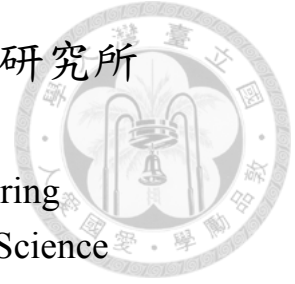國立臺灣大學電機資訊學院電信工程研究所
碩士論文
Graduate Institute of Communication Engineering
College of Electrical Engineering and Computer Science
National Taiwan University
Master Thesis

擴散式分子通訊之同步與通道估計
Synchronization and Channel Estimation in Diffusion-Based
Molecular Communications

許博凱
Bo-Kai Hsu

指導教授：葉丙成教授
Advisor: Ping-Cheng Yeh, Ph.D.

中華民國 104 年 7 月
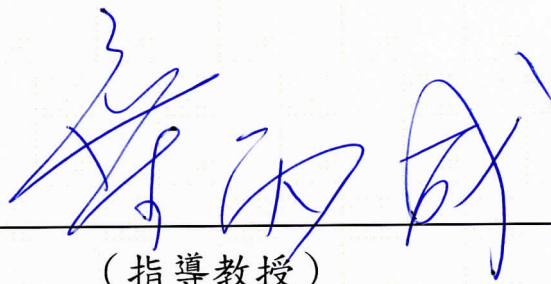July, 2015

國立臺灣大學
電信工程研究所

碩士論文

擴散式分子通訊之同步與通道估計

許博凱　撰

104
7

國立臺灣大學 碩士學位論文
口試委員會審定書

擴散式分子通訊之同步與通道估計
Synchronization and Channel Estimation in Diffusion-Based Molecular Communications

本論文係許博凱君（R02942058）在國立臺灣大學電信工程學研究所完成之碩士學位論文，於民國 104 年 8 月 18 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

_____（簽名）

（指導教授）

_____　_____

_____　_____

所　　長　_____（簽名）

致謝

中文摘要

奈米科技發展於實現在生醫、產業、環境及軍事上的應用，最近越來越多的研究人員關注奈米尺度的裝置間合作的模式以建構更複雜的系統，在奈米尺度下的通訊中最被看好的一種就是分子通訊，奈米機器人藉由在流體中釋放與接收分子來交換訊息，這種通訊跟傳統電磁波通訊有很大的不同，分子通訊不受到天線尺寸的限制，也與生物系統有較佳的相容性，而且以分子擴散的形式傳遞訊息是一種相當節省能量的機制，然而，分子擴散的隨機性會在通道中產生雜訊，很多研究員紛紛提出通訊系統的設計以對抗通道中的雜訊及干擾來增進分子通訊的品質，大多數在文獻中提出的設計都需要傳送端與接收端的同步，在這篇論文中我們主要探討擴散式分子通訊下的同步問題，我們考慮兩種非同步效應包括時鐘偏移與時鐘偏斜，在訓練式同步下，我們提出的疊代線性估計式比起其他方法只需要相當低的計算量，在盲同步下，我們應用費雪計分演算法來對抗符元間干擾並且有效的更新時鐘偏移估計，這兩種方法的估計效率都蠻接近最低均方誤差的比特曼估計，此外，即使在沒有通道資訊的情況下，我們也可以藉由疊代估計通道與時鐘偏移量來達到同步，這種方法的均方誤差也相當接近在沒有通道資訊的克拉馬羅下限。
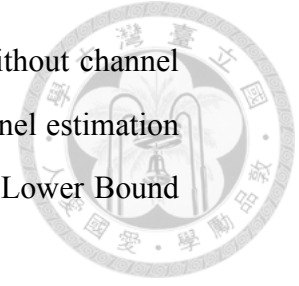
Abstract

Nanotechnology has been developing to fulfill applications in biomedicine, industry, environment, and military. Recently, more and more researchers have discussed on the cooperation between nano-scale devices to form a more complicated system. One of most promising solutions to construct nano-scale communications is molecular communications. In a fluid medium, nano-scale devices exchange messages by releasing and capturing molecules. This type of communications is quite different from the conventional microwave communications. Unrestricted by the antenna size, molecular communications has a good compatibility with biological systems at the nanoscale. Moreover, molecular diffusion is really a good power saving mechanism to propagate information. However, the randomness in the diffusion process arises noises in molecular communication channel. Many researchers have designed communication systems against noises and interferences to improve communication quality. Among most systems proposed in literature, synchronization between transmitter and receiver nanomachines is inevitable. In this thesis, we deal with synchronization problem in diffusion-based molecular communications. Non-synchronous effects including clock offset and clock skew have been considered. In training-based synchronization, we proposed the Iterative Linear Estimator (ILE) with much lower complexity than other methods. In blind synchronization, we applied Fisher's Score Algorithm (FSA) to efficiently update clock offset estimation against the Inter Symbol Interference (ISI) effect. The efficiencies of ILE and FSA are close to lower bound of

Mean Square Error (MSE) by the Pitman estimator. Even without channel information, we proposed a method to iteratively update channel estimation and clock offset estimation. Its MSE is close to Cramer Rao Lower Bound (CRLB) without channel information.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Combining with nano-technology, nano-scale communication is a promising technology in information and communication field. Recently, more and more researchers have discussed on the cooperation between nano-scale devices to form a functional nano-network. Inspired by the biological systems in nature, molecular communications are thought as the most popular solution to construct nano-scale communications. By diffusing across a fluid medium, molecules as messenger barriers are used to propagate information. Most molecular communication systems proposed in literature require synchronization between nano-scale nodes to communicating with each other. In our thesis, we discuss on synchronization problem in molecular diffusion channel. We begin with an overview of diffusion-based molecular communications and then introduce some related works and our main contributions on synchronization in molecular communications.

1.1 Overview

In recent years, nanotechnology has been developing rapidly. Beginning with manipulating atoms and molecules precisely, nanotechnology generally includes all the technology in the scale from 1 to 100 nanometers. In such small size, it is possible to manufacture an artificial device in scale of molecules, nano-machine. A nano-machine refers to a most basic unit in nano-scale but can perform specific function such as computing, sensing or data storage [1]. Because of the size, the computational capability and the memory of a

nano-machine are limited. To reach a more complicated system, we need to form a nano-network by communicating and cooperating between nano-machines. A nano-network consists of a collection of nano-machines. They link together and transfer information with each other to form a functional system. Developing a nano-network is an attractive topic in nano-technology.

Nano-network can be applied to many aspects of applications like biomedicine, industry, environment, and military [2]. The blood pressure monitoring and drug delivering system [3] are examples in biomedicine. Besides, nano-network are possibly used in water or food inspection and air pollution control. For this prospect in the future, more and more researchers in communication and information field are interesting in nono-network.

The communication mechanisms between nano-scale devices are diverse, including electromagnetic, acoustic, mechanical, and molecular communication [2]. Due to the antenna size, the traditional electromagnetic communication are limited. Moreover, much higher carrier frequency in nano-scale antenna requires computation beyond the capability of nano-machine. Inspired by cellular communication, one of promising communication mechanism between nano-machines is molecular communication. In molecular communication, the information is propagated by transmitting and receiving messenger molecules. Unlike mechanical communication require physical contacts, molecular communication could use the messenger molecules as barriers embed information and communicate indirectly. The advantage of molecular communication make it suitable in developing nano-network.

Molecular communications are widely used in biological systems. The distance of information propagation ranges from hundreds of nano-meters to some centimeters or meters. In long range molecular communications, pheromone is an example [4]. The effect of pheromone propagates from interspecies to one body by circulatory system. In short range, neurons control their electrical activity by the concentration of calcium ions signal [5]. Messenger molecules like proteins, ions, or DNAs are transported in short range molecular communications by motor, gap junction, or diffusion. For intracellular transport, motor proteins carry messenger molecules along microtubule rails [6]. For

inter-cellular communications, gap junction connects two adjacent cells as a channel or messenger molecules like calcium ions diffuses across a free space. In this thesis, we focus on designing mechanisms in the diffusion-based molecular communications in short range.

In Diffusion-based Molecular Communications (DMC), molecules diffuse across a fluid medium from regions of high to low concentration. This process can be modeled by Fick's laws of diffusion and Brownian motion process [7]. In this field, many papers design a good modulation and detection to improve the quality of molecular communication. For example, the paper [8] considers a time-slotted diffusion-based molecular communication with information embedded in different quantity levels. However, most literature assume perfect synchronization between the transmitter and the receiver. In reality, how to form a time-slotted system in DMC is still a problem. To solve this problem, we analyze the timing synchronization problem in this thesis.

1.2   Related Works On Timing Synchronization

Generally, transmitter converts information bit stream into a sequence of symbols. Then, transmitter assigns a period of time called *symbol duration* to transmit each symbol. However, with non-synchronous clocks of transmitter and receiver, a clock offset which is constant but unknown for receiver exists between these two clocks. Accordingly, receiver may not identify the beginning of each symbol duration. This is the problem of *timing synchronization* or *symbol synchronization* [9].

In conventional electromagnetic communications, Orthogonal Frequency Division Multiplexing (OFDM) system is vulnerable to symbol synchronization error. In molecular communications, the propagation of diffusion is much slower than light speed, so a longer symbol duration than conventional communications is inevitable. Due to the long and random propagation delay, timing synchronization problem in DMC is difficult [10]. Besides, the paper [11] proposed a sampling-based detection requires a perfect synchronization system between transmitter nano-machine and receiver nano-machine. The performance of sampling-based detection affected by timing synchronization error. Accordingly, solving

synchronization problem in DMC is worth investigating. Analyzing the characteristic of propagation delay could be a solution [11]. In this thesis, we try to estimate the random delay in DMC.
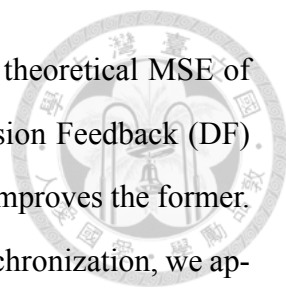
Recently, some papers in literature start discussing the synchronization problem in molecular communication because many people notice that it is necessary to keep the clock synchronized among nanomachines. For example, inspired from biological mechanism, the paper [12] surveys the Quorum Sensing between Bacteria to reach synchronization in a cluster of nodes of a nanonetwork. They use some kinds of molecules called autoinducers. When sensing the concentration of autoinducers over a threshold level, the node in the nanonetwork will activated and releasing autoinducers as positive feedback. This kind of mechanism causes chain effect so that the authors claim that the mechanism is efficient and the delay among nanomachines is acceptable.

For timing synchronization problem in concentration-based molecular communication, the first blind synchronization algorithm has been proposed in [13]. This paper use the concentration single measured by receiver to efficiently estimate the propagation delay of transmission. But our system model is different from this paper's. The situation in our transmission is releasing a few number of molecules, which is less than the level to form a concentration single. Then, when molecules arrive to the position of receiver, they will be captured one by one and receiver can measure the arrival times of molecules. These two types of system model of communication in DMC has been studied in literature [8, 14].

1.3 Major Contributions

Our main work is to design an efficient estimator of clock offset and clock skew so that RN has an estimated time slotted system, which synchronizes with TN's. The key contributions of this thesis are listed as below.

First, we compare the Mean Square Error (MSE) and computational complexity of three methods in *clock offset estimation*, Maximum Likelihood Estimation (MLE), Linear Estimation (LE), and Iterative LE. Among them, we proposed the best one, Iterative LE, with lower complexity and its MSE reaches almost the same efficient level with the

others. On the other hand, in *blind synchronization*, we analyze the theoretical MSE of two methods, LE for the arrival time of the first molecule and Decision Feedback (DF) method. Furthermore, we give a sufficient condition when the latter improves the former. When dealing with Inter Symbol Interference (ISI) effect in blind synchronization, we apply Fisher Score Algorithm (FSA) to improve the initial estimation of clock offset. The MSE of FSA in blind synchronization is much closer to the optimal location parameter estimator, Pitman estimator, than Iterative LE proposed in training-based synchronization.

Second, in channel estimation, we proposed initial estimation of clock offset τ and channel estimation μ and λ. To make the MSE converge to zero overtime, we proposed a method iteratively updating τ, μ, and λ. The MSE of this method is close to the CRLB with unknown μ and λ.

Third, in clock skew estimation, we proposed Iterative LE for clock skew rate with channel information and Quasi-likelihood clock skew rate without channel information. The latter is better than the former when the number of symbols is large enough.

The following structure of our thesis begins with system model and problem formulation in Chapter. 2. We will explain what Additive Inverse Gaussian Noise (AIGN) channel and quantity-based modulation are in DMC. Then, for clock offset estimation, we discuss training-based and blind synchronization in Chapter. 3. The channel estimation is considered in Chapter. 4 when dealing with clock offset estimation. Moreover, we discuss the clock skew estimation with and without channel information in Chapter. 5. Finally, the conclusion and future work are discussed in Sec. 6.

Chapter 2

Synchronization Problem Formulation In Additive Inverse Gaussian Noise Channel Model

In this chapter, we describe our system model in diffusion-based molecular communications. The channel is modeled by Additive Inverse Gaussian Noise channel, which was investigated by former researchers. Under this model, we discuss on the synchronization problem for quantity-based modulation. In this type of synchronization problem, we consider both clock offset and clock skew effect.

2.1 System Model

We consider an end-to-end communication in a volume with fluid medium. The transmitter is a nano-machine, and so is the receiver. We call them Transmitter Nano-machine (TN) and Receiver Nano-machine (RN). They communicate with each other by releasing and capturing molecules. The channel between them is molecular diffusion based on Brownian motion to propagate information message.

As the model described in [15], TN is in a small scale of molecules, but RN is in a large scale of the distance apart from TN. Because of the scale, TN is approximately a point in three-dimensional space but RN is approximately an infinite plane for the molecules

released from TN. Besides, on the surface of RN is full of receptors to capture the arrival molecules .The Brownian motion is isotropic for each direction, so we only consider the one-dimensional movement perpendicular to the infinite plane of RN. All molecules and receptors are identical. When a molecule diffuses across the fluid medium and captured by a receptor, it is completely removed from the space. This absorbing process is different from concentration-based molecular communications [16].

The movements of molecules are affected by two kinds of effect. One is the Brownian motion from the particle collision in the fluid medium. The other one is the drift velocity towards RN. Combine these two kinds of effect, we model the movement of molecules by the Wiener process with drift. Moreover, because all molecules are in nano-scale and diffuse in three dimensional space, the collision of two molecules is almost impossible. We model each movement of molecule is independent. Based on [17], a one-dimensional molecular diffusion can be described by one-dimensional Brownian motion and the first hitting time to a specific position follows the Inverse Gaussian distribution. This channel in our model is called Additive Inverse Gaussian Noise channel model as in [18].

2.1.1 Additive Inverse Gaussian Noise Channel

We define the spatial location by a one-dimensional extent with the origin as TN's position. Let $d > 0$ denote the position of RN apart from TN. When a molecule is released from TN at time x, it will act as one-dimensional Brownian motion with positive drift velocity $v > 0$, which is the direction to RN. Because of positive drift, in a sufficiently long period of time, this molecule must arrive in the RN's position and be captured by RN. However, the arrival time of this molecule is random. Based on [17], we describe the arrival time as $x + T$, where the additive first hitting time T is the random variable with Inverse Gaussian distribution.

$$f_T(t) = \begin{cases} \sqrt{\frac{\lambda}{2\pi t^3}} \exp\{-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\} & \text{if } t > 0, \\ 0 & \text{if } t \leq 0. \end{cases} \qquad (2.1)$$

8

with parameters as follows:

$$\mu = \frac{d}{v} \ , \ \lambda = \frac{d^2}{2D} \text{ and, } D = \frac{k_B T_a}{6\pi\eta r},$$

where k_B is the Boltzmann constant, T_a is the absolute temperature, η is the viscosity of the fluid medium, and r is the radius of molecule.

2.1.2 Quantity-based Modulation

A modulation called quantity-based modulation have been discussed on the paper [8] in AIGN channel. As shown in Fig. 2.1, quantity-based modulation embeds information message in the quantities of molecules, which is similar to amplitude modulation in the traditional wireless communications. Before communication, TN and RN shared a common period of time, called symbol duration denoted by T_s. Information message is conveyed by a sequence of symbols in consecutive symbol durations as shown in Fig. 2.1. For each symbol, TN assign the corresponding quantities of molecules by the information message and release these molecules at once in the begging of the symbol duration. In binary case, TN release L_0 and L_1 unlabeled molecules representing a binary zero and a binary one, respectively. All molecules are in the same type, so RN cannot distinguish them. Then, RN will accumulate the quantity of arrival molecules during symbol duration T_s and demodulate this symbol by some detection methods.
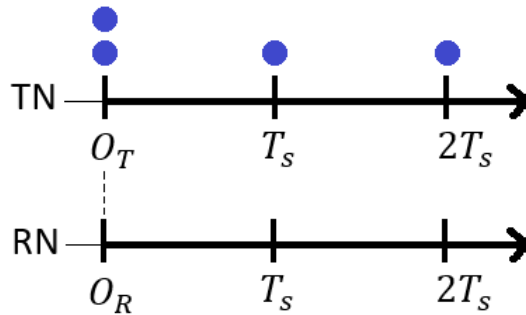
Figure 2.1: Time-slotted system in quantity-based modulation under perfect synchronization.

An optimal detection in quantity-based modulation is proposed in [8] to improve quality of communication. This kind of modulation and detection rely on a synchronous time-

slotted system, as shown in Fig. 2.1. RN needs to share a perfect synchronous time-slotted system with TN, otherwise it causes the error to increase. The main work in this thesis is to meet the non-synchronous gap in quantity-based modulation.
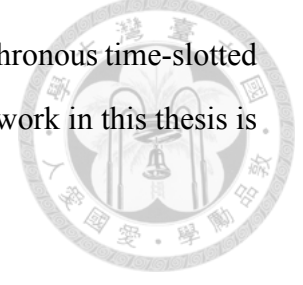
2.1.3 Clock Offset Estimation

We have shown the time-slotted system in quantity-based modulation under perfect synchronization in Fig. 2.1. However, in the beginning of communication, TN and RN may have non-synchronous clocks with each other, so the starting point of their clock may be different. We denote the starting point of TN's and RN's clock by O_T and O_R respectively. As shown in Fig. 2.2, a clock offset $\tau := O_T - O_R$ exists generally. We define the clock offset as the subtraction of O_R from O_T, which may be negative. Consequently, when RN receive a molecule, the arrival time measured by RN includes not only the random delay T, but also the clock offset τ. Remind that the first hitting time T is random and follows Inverse Gaussian distribution caused by AIGN channel. The clock offset τ is constant but unknown for RN caused by the non-synchronous phenomenon. The problem is how to efficiently estimate the clock offset τ by the sequence of arrival time, which is random, measured by RN to reach synchronization with TN.
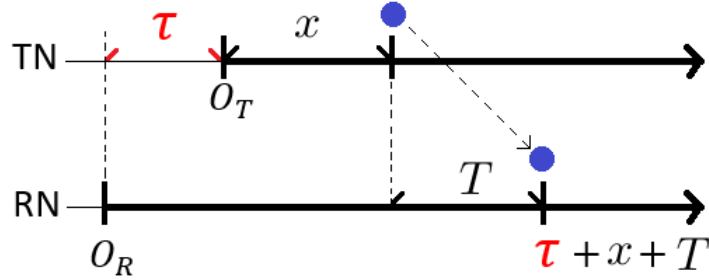
Figure 2.2: Phenomenon of clock offset.

In quantity-based modulation, TN releases consecutive symbols and each symbol carry an amount of molecules. All of the molecules are unlabeled and in the same type. They are captured by RN in a sequence of time according to the ascending order. Therefore, one molecule released in a later time may be captured earlier than another molecule released in a previous time. We call it crossover effect. The crossover effect is natural in AIGN channel. RN cannot distinguish whether the observation affect by crossover effect.
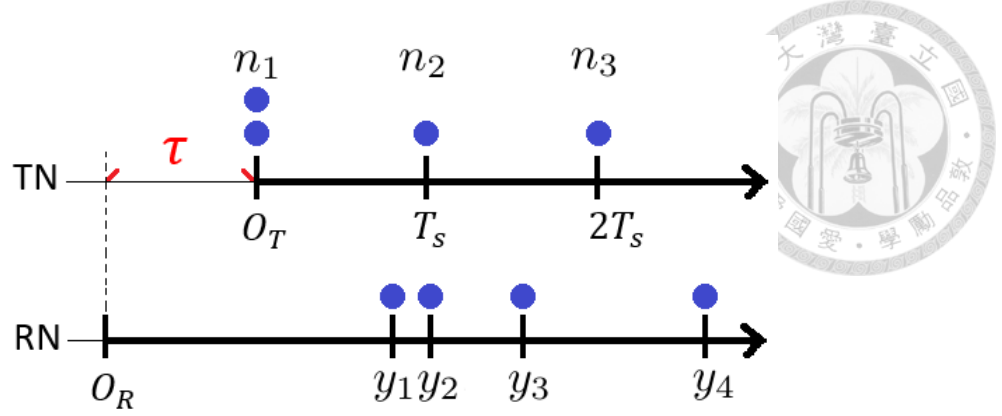
Figure 2.3: Clock offset estimation for quantity-based modulation.

As shown in Fig. 2.3, TN releases $n_1, n_2, ..., n_K$ molecules in the beginning of consecutive K symbols. In binary case, $n_k \in \{L_0, L_1\}$. What RN observe is a sequence of arrival times of molecules denoted by $y_1, y_2, ..., y_N$ according to the ascending order of timing, where $N = n_1 + n_2 + ... + n_K$. The problem is how to design an efficient estimator $\hat{\tau}(\mathbf{y})$ based on the observation $\mathbf{y} = [y_1, y_2, ..., y_N]$. We want to make the Mean Square Error (MSE) defined by $E[(\tau - \hat{\tau}(\mathbf{y}))^2]$ as small as possible, where the function $E[.]$ denotes the expectation of some random variable. This is the first problem we deal with called *clock offset estimation*.

2.1.4 Clock Skew Estimation

The second problem we deal with is the *clock skew estimation* as shown in Fig. 2.4. In real scenarios, RN's clock may tick in a different rate from TN's, that is called *clock drift* or *clock skew*. Under clock skew effect, RN could count symbols on a different timing duration, maybe longer or shorter than TN's. A bit error of clock skew will accumulate overtime, causing non-synchronous phenomena gradually. The accumulated error is more severe in molecular communications than in traditional wireless communications, because molecules diffuse with much loner delay and the whole system require much longer communication duration.

To reach synchronous with TN, RN needs to adapt symbol duration corresponding to its own clock skew. Here, we assume the skew is constant overtime but unknown for RN. Nevertheless, RN perceives the arrival time sequence $\mathbf{y}$ by its own clock, so RN could
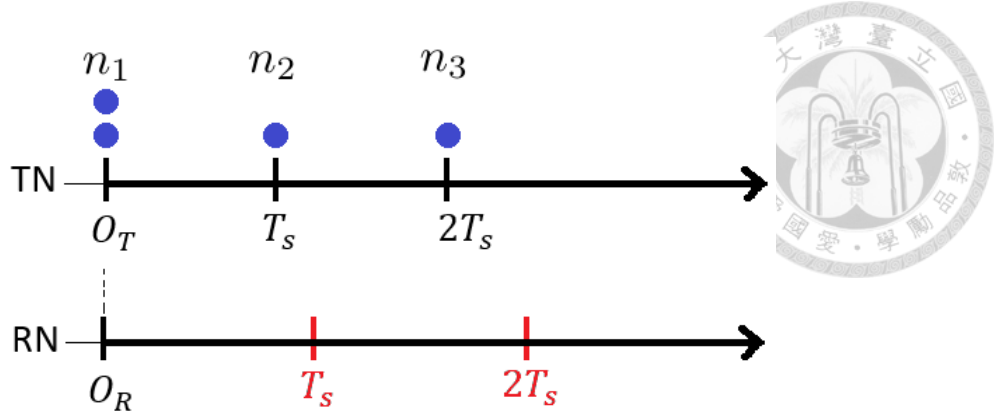
Figure 2.4: Clock skew estimation for quentity-based modulation.

extract the pattern in $\mathbf{y}$ to estimate $T_s \times R$, where R is the clock skew rate. Then, TN and RN could count symbols by T_s and $T_s \times \hat{R}$ respectively with almost the same timing duration to reach synchronization. For example, if $T_s = 10$ (sec) and RN's clock ticks in a twice rate from TN's, that is $R = 2$, then RN can compensate the clock skew effect by adapting its symbol duration to $T_s \times \hat{R} \approx 20$ (sec). Even though TN counts symbols by $T_s = 10$ (sec) and RN counts symbols by $T_s \times \hat{R} \approx 20$ (sec), they actually count symbols with almost the same duration because RN's closk ticks in a twice rate from TN's.

Combine with above two non-synchronous effect, RN needs to estimate the clock offset τ and the unknown clock skew rate R simultaneously. For TN's time slotted system, $\tau + (k-1)T_s$ is the boundary between $(k-1)$-th symbol and the k-th symbol, where $k = 1, 2, ..., K$. However, becasue the clock skew rate R for RN's clock from TN's clcok, the true boundary between $(k-1)$-th symbol and the k-th symbol for RN's time slotted system is $R[\tau + (k-1)T_s]$, denoted by S_k which is the true parameters we want to estimate. The target is to synchronize the sampling time $\hat{S}_k(\mathbf{y}) = \hat{R}(\mathbf{y})[\hat{\tau}(\mathbf{y}) + (k-1)T_s]$ as close to $S_k = R[\tau + (k-1)T_s]$ as possible. By minimum mean square error (MMSE) criteria, we try to make $E[(S_k - \hat{S}_k(\mathbf{y}))^2]$ as small as possible.

2.2 Synchronization Problem Formulation

In AIGN channel model, we consider many kinds of synchronization problem for quantity-based modulation. The following, we describe our overall system model by plotting the whole system diagram. Moreover, our system diagram shows all main parameters

in our model. With different assumption on these parameters, we formulate many kinds of synchronization problem described in a table.

2.2.1 System Diagram

Fig. 2.5 shows an overall system diagram for quantity-based modulation. In this system, TN modulates information bit stream I_k into the sequence of molecular amounts for each symbol $\{n_k | 1 \leq k \leq K\}$. Then, TN releases molecules based on TN's clock. We denote the molecular releasing time sequence as $\mathbf{x} = [x_1, x_2, ..., x_N]$, where $N = n_1 + n_2 + ... + n_K$. For quantity-based modulation, $x_i = (j-1)T_s$ if $n_1 + ... + n_{j-1} < i \leq n_1 + ... + n_j$ for $i \in \{1, 2, ...N\}$. For example, if $K = 3$ and $n_1 = 2, n_2 = 1, n_3 = 1$, as shown in Fig. 2.1, we have $\mathbf{x} = [0, 0, T_s, 2T_s]$.
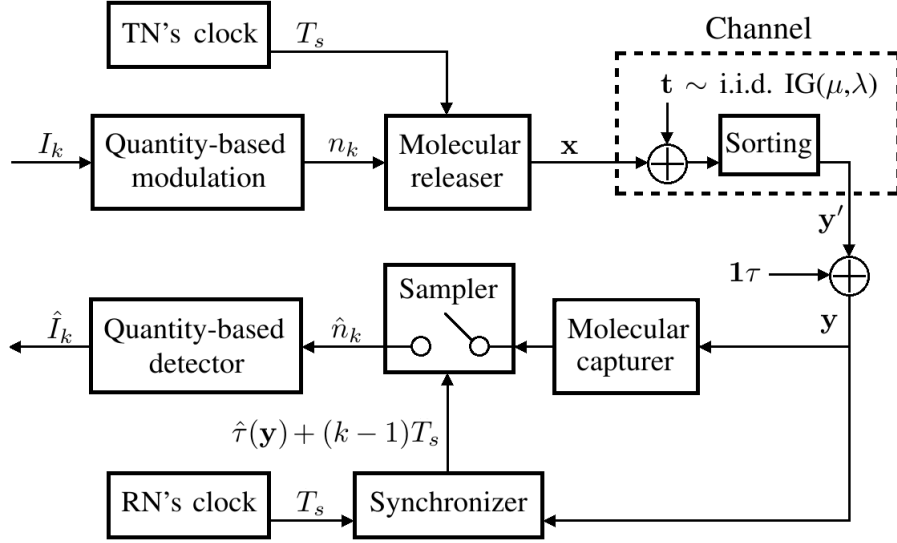
Figure 2.5: System diagram of clock offset synchronization for quantity-based modulation in AIGN channel.

After passing through AIGN channel, $\mathbf{x}$ becomes $\mathbf{y}' = \text{sort}(\mathbf{x} + \mathbf{t})$, where $\mathbf{t}$ is a random sample from Inverse Ganssion distribution. We model the channel effect with an additive Inverse Gaussain Noise t_i and the crossover effect by sorting the arrival time sequence. Moreover, because of the non-synchronous clock between TN and RN, the actual arriving time measured by RN $\mathbf{y}$ includes the unknown clock offset τ. According to $\mathbf{y}$, we design an efficient estimator $\hat{\tau}(\mathbf{y})$ to estimate the unknown clock offset τ. This is the synchronization problem only with clock offset.

13
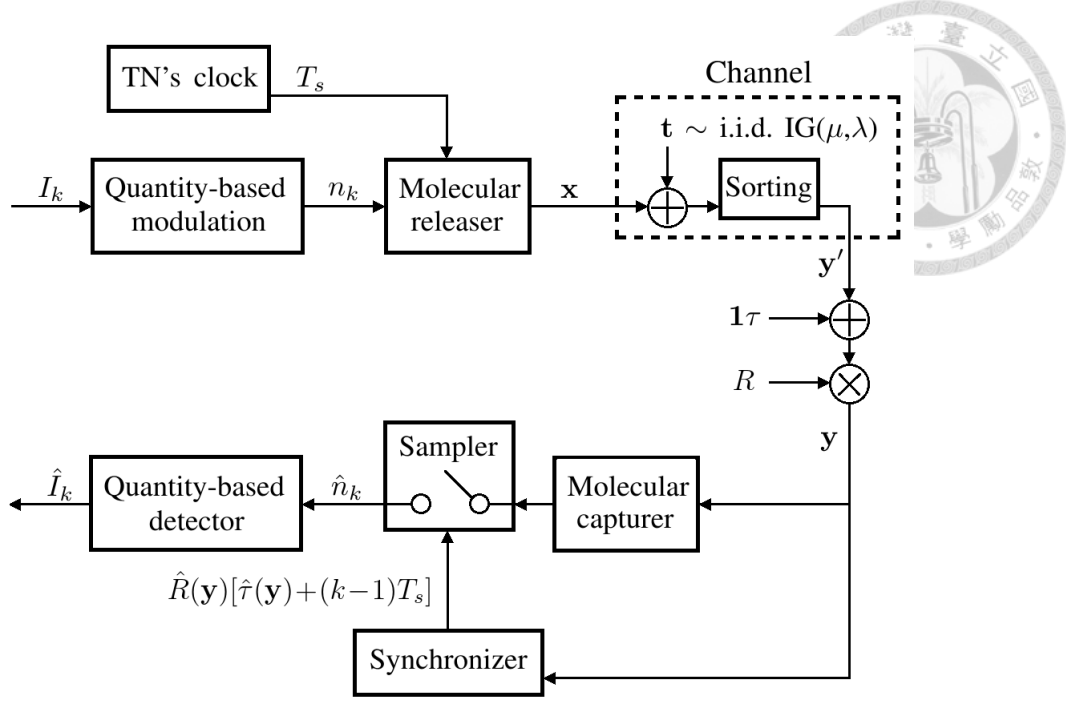
Figure 2.6: System diagram of clock offset and clock skew synchronization for quantity-based modulation in AIGN channel.

Fig. 2.6 shows an overall system diagram combining clock offset and clock skew. After passing through AIGN channel, $\mathbf{y}'$ not only add clock offset τ but also multiple clock skew rate R, where both τ and R are unknown for RN. According to $\mathbf{y} = R[\tau\mathbf{1} + \mathbf{y}']$, we design two efficient estimators $\hat{\tau}(\mathbf{y})$ and $\hat{R}(\mathbf{y})$ to make the sampling time $\hat{S}_k = \hat{R}(\mathbf{y})[\hat{\tau}(\mathbf{y}) + (k-1)T_s]$ synchronize with TN's time slotted system.

2.2.2 Training-based And Blind Synchronization

There are two types of synchronization: training-based synchronization and blind synchronization.

In training-based synchronization, each communication includes tow phases. The first one is the training phase. In training-phase, TN transmits the training sequence which RN has already known. After training signal passes through the channel, RN can estimate the channel information or synchronize with TN by receiving the observations in training-phase. After training-phase, the other one is communicating phase. In communicating phase, TN transmits the information message. This time, RN focuses on detecting the information message with the parameters was fixed before.

14

In our work, when we discuss on the training-based synchronization, we want to design the mechanism of synchronization in training phase. For quantity-based modulation, RN knows the training sequence I_k in Fig. 2.6. Then, the quantities of molecules released by TN n_k for each symbol is known for RN. For simplification, we often discuss the constant training sequence. That is, each symbol has the same quantities of molecules $n = n_1 = n_2 = ... = n_K$. TN transmits total $K \times n$ molecules during K symbol durations.

In contrast, when we discuss on the blind synchronization, we want to design the mechanism of synchronization in communicating phase. That is, even TN transmit information message, RN has to synchronize with TN and detect the message simultaneously. For quantity-based modulation, TN transmit the information message I_k randomly. Then the quantities of molecules n_k for each symbol is random for RN. Traditionally, we often use the M-ary with equal *a priori* probability. That is n_k is an discrete uniform distribution over $\{l_0, l_1, ..., l_{M-1}\}$, where l_i is the i-th level of quantities.

2.2.3 Channel Information

Besides the information transmitted by TN, the channel information is also important for RN. For AIGN channel, the additive random delay T follows inverse gaussian distribution with two parameters μ and λ. These tow parameters depend on some channel characteristics like the distance between TN and RN d, the drift velocity v, and the diffusion coefficient D of some types of molecules. Nevertheless, μ and λ determine all information about an AIGN channel as shown in Fig.2.6. Therefore, it is enough for RN to evaluate the values of μ and λ.

We model the AIGN channel with clock offset τ and clock skew rate R as

$$\mathbf{y} = R[\text{sort}(\mathbf{x} + \mathbf{t}) + \mathbf{1}\tau], \tag{2.2}$$

where $\mathbf{x}$ is the input releasing time sequence and $\mathbf{y}$ is the output arrival time sequence and $\mathbf{t}$ is the random sample from the inverse gaussian distribution. Specifically, for an arrival

cases	assume known	fix but unknown
clock offset estimation	μ, λ, R	τ
channel estimation	R	τ, μ, λ
clock skew estimation	μ, λ	τ, R
joint estimation		μ, λ, τ, R

Table 2.1: List 4 kinds of problems with different assumptions about parameters in stochastic model.

time y_i of a molecule which was released at $x_j = k_j T_s$, that is at the k_j-th symbol,

$$y_i = R[x_j + t_i + \tau] = R[k_j T_s + t_i + \tau]. \tag{2.3}$$

Note that i might not equal j because the sorting effect changes the order of timing. Naturally, we model the crossover effect by sorting the arrival time sequence.

Overall our system model, we have a stochastic model for an arrival time y_i observed by RN as

$$y_i \sim f_T(\frac{y_i}{R} - \tau - k_j T_s | \mu, \lambda) = f(y_i | \mu, \lambda, \tau, R, k_j). \tag{2.4}$$

It is clear that we have 5 main parameters, μ, λ, τ, R, and k_j, in this stochastic model. τ and R are clock offset and clock skew rate, respectively, which are our target unknown parameters. μ and λ are the AIGN channel parameters. k_j is related with x_j depends on the signal transmitted by TN. We have discussed that n_k is known for training-based synchronization and n_k is random for blind synchronization in Sec. 2.2.2. Moreover, different assumptions about μ, λ, τ, and R lead to many kinds of problems listed as in Table 2.1.

First of all, we consider a simple case, clock offset estimation, in Chapter. 3. We assume the channel information μ, λ, and the clock skew rate R are all known for RN. The observations $\mathbf{y}$ have been compensated by clock skew rate R as in system diagram Fig. 2.5. Our target is to estimate the unknown clock offset τ. In clock offset estimation, we discuss on both training-based synchronization and blind synchronization.

Second, if we have a short symbol duration and overall communication period does not last too long, the effect of clock skew rate is negligible. The overall system model is described as in Fig. 2.5. The arrival times of molecules $\mathbf{y}$ observed by RN does not

short ⟋ channel estimation (Chp. 4)

symbol duration ⟨ insensitive ⟋ clock skew estimation (Chp. 5.1)

long ⟍ delay ⟨

sensitive ⟍ joint estimation (Chp. 5.2)
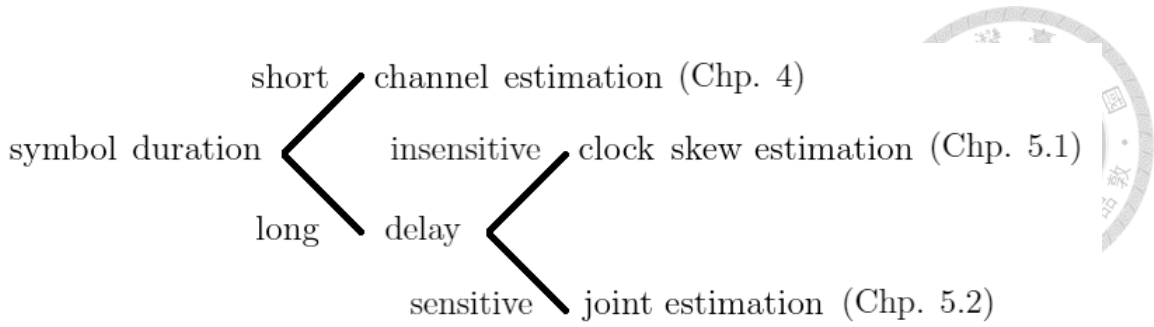
Figure 2.7: Three kinds of situation by the short or long symbol duration and whether delay is sensitive or not.

multiple by clock skew rate R. This way, what we really focus on is the unknown channel information. We call this situation is channel estimation problem. For channel estimation, we discuss on the case when RN has no idea about channel information but without clock skew effect in Chapter. 4. The target estimation parameter is still clock offset τ with minimum MSE.

Third, if we have a long symbol duration or overall communication period lasts too long, the clock skew effect cannot be ignored. Estimating the clock skew rate R is inevitable. For simplification, we begin with the case RN knows the channel information μ and λ to estimate τ and R simultaneously. In realistic, this is in the situation with delay insensitive communication, so non-synchronous channel estimation methods like [16] can be applied before synchronization. Then, RN can synchronize with TN under channel information. We call this problem clock skew estimation and discuss it in Sec. 5.1. Otherwise, in delay sensitive application like health monitoring, RN needs to do synchronization and channel estimation simultaneously. This is the toughest problem we deal with, joint estimation. In joint estimation, we considers the case when these 4 parameters are all unknown. Our target is estimate the unknown clock offset τ and clock skew rate R so that the sampling time S_k of RN can be synchronized with TN's time slotted system in Sec. 5.2.

Besides the simple case, clock offset estimation, we divided the other three kinds of situation by short or long symbol duration and delay sensitive or insensitive communication. If we have a short symbol duration and overall communication period does not last too long, we focus on the channel estimation but the clock skew effect. Otherwise, if

17

in an urgent communication like health monitoring, we have no time to estimate channel parameters before synchronization, so joint estimation is required. Otherwise, clock skew estimation is the case only consider the clock skew effect with known channel information.
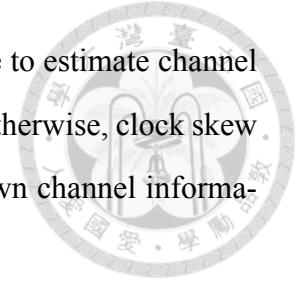
Chapter 3

Clock Offset Estimation In Training-based And Blind Synchronization

The first problem we deal with is clock offset estimation problem. If we have a short symbol duration and overall communication period does not last too long, the effect of clock skew rate is negligible. The overall system model is described as in Fig. 2.5. The arrival times of molecules $\mathbf{y}$ observed by RN does not multiple by clock skew rate R. This way, we only focus on estimating the unknown clock offset τ.

3.1 Training-based Synchronization

In this section, we focus on the training-based synchronization. In training-based synchronization, there is a training phase to synchronize between TN and RN before they transmit and receive information message. In training phase, TN transmit a pilot signal which RN have already known, so $\mathbf{x}$ is constant for RN.

In clock offset estimation, because the clock skew effect has been compensated perfect, the evaluation reduce to the MSE of clock offset τ for any k. That is, the error of all boundary of symbols are the same. RN's time slotted system is a shift version of TN's

time slotted system.

$$E[(S_k - \hat{S}_k)^2] = E[(\tau + (k-1)T_s - \hat{\tau} - (k-1)T_s)^2]$$
$$= E[(\tau - \hat{\tau})^2] \qquad (3.1)$$

The following, We propose some methods to estimate τ under the assumption that RN knows $\mathbf{x}$.

3.1.1 Maximum Likelihood Estimation

In classic point estimation problem, Maximum Likelihood Estimation (MLE) can be applied if the joint probability distribution of observations is known. The following, we derive the probability distribution of RN's observations (the arrival times of molecules) under non-synchronous situation.

Based on [14], when $\tau = 0$, that is perfect synchronization, the joint probability density function (pdf) of observation denoted by $\mathbf{y}'$ given releasing time sequence $\mathbf{x}$ has been derived as below:

$$f(\mathbf{y}'|\mathbf{x}) = \sum_{\mathbf{u} \in \mathbf{P}(\mathbf{y}')} f(\mathbf{u}|\mathbf{x})$$
$$= \begin{cases} \sum_{\mathbf{u} \in \mathbf{P}(\mathbf{y}')} \prod_{i=1}^{N} f_T(u_i - x_i), & \text{if } \mathbf{y}' = \text{sort}(\mathbf{y}'); \\ 0, & \text{otherwise.} \end{cases} \qquad (3.2)$$

where $\mathbf{P}(\mathbf{y}')$ is the set of all possible permutation of $\mathbf{y}'$ and the function sort$(\mathbf{y}')$ sorts $\mathbf{y}'$ according to ascending order.

In our work, because of the non-synchronous phenomenon, the observation of RN is $\mathbf{y} = \mathbf{y}' + \tau\mathbf{1}_N$, where $\mathbf{1}_N$ is the $1 \times N$ vector with all value are equal to 1. Therefore, we

can derive likelihood function as below:

$$f(\mathbf{y}|\mathbf{x}, \tau) = \begin{cases} \sum_{\mathbf{u} \in \mathbf{P}(\mathbf{y})} \prod_{i=1}^{N} f_T(u_i - x_i - \tau), & \text{if } \mathbf{y} = \text{sort}(\mathbf{y}); \\ 0, & \text{otherwise.} \end{cases} \quad (3.3)$$
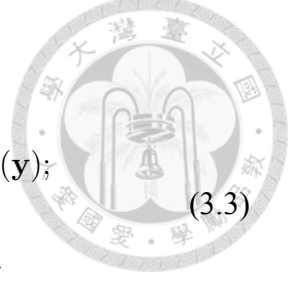
Then, we can apply MLE and denote the estimator by $\hat{\tau}_{\text{MLE}}$.

$$\hat{\tau}_{\text{MLE}} := \arg\max_{\tau} f(\mathbf{y}|\mathbf{x}, \tau). \quad (3.4)$$

However, the time complexity grows rapidly by factorial on N because of the permutation of $\mathbf{y}$. In reality, it beyond the computational capability of nano-machine.

3.1.2   Linear Estimation

By considering the complexity issue, we try to apply Linear Estimation (LE). Assume $\hat{\tau}_{\text{LE}} := \mathbf{a}_N \mathbf{y}^\top + b$ for some constant $\mathbf{a}_N = [a_1, a_2, ..., a_N]$ and b such that MSE is minimal.

We have $\mathbf{y} = \mathbf{y}' + \tau \mathbf{1}_N$. By derivation above, we have the joint pdf of $\mathbf{y}'$ given $\mathbf{x}$, so we can derived the mean and covariance of $\mathbf{y}'$, which are used to derive the coefficient $\mathbf{a}_N$ and b.

Because the parameter τ is constant but unknown, we set the constraint on $\mathbf{a}_N \mathbf{1}_N^\top = 1$ to eliminate τ as below:

$$\begin{aligned} E[(\tau - \hat{\tau}_{\text{LE}})^2] &= E[(\tau - \mathbf{a}_N \mathbf{y}^\top - b)^2] \\ &= E[(\tau - \mathbf{a}_N \mathbf{1}^\top \tau - \mathbf{a}_N \mathbf{y}'^\top - b)^2] \\ &= E[(\mathbf{a}_N \mathbf{y}'^\top + b)^2]. \end{aligned} \quad (3.5)$$

According to [19], applying the solution in Non-random parameter estimation, we can

make MSE reach the minimal value when

$$\mathbf{a}_N = \mathbf{1}_N \mathbf{C}_{\mathbf{y}'}^{-1} \{\mathbf{1}_N \mathbf{C}_{\mathbf{y}'}^{-1} \mathbf{1}_N^\top\}^{-1} \text{ and}$$

$$b = -E[\mathbf{a}_N \mathbf{y}'^\top] \tag{3.6}$$

,where $\mathbf{C}_{\mathbf{y}'}$ is the covariance matrix of the random vector $\mathbf{y}'$.

Besides eliminating unknown τ, setting $\mathbf{a}_N \mathbf{1}^\top = 1$ make the unbiased property possible, so we actually apply LE under the unbiased constraint.

$$
\begin{aligned}
E[\hat{\tau}_{\text{LE}}] &= E[\mathbf{a}_N \mathbf{y}^\top + b] \\
&= E[\mathbf{a}_N \mathbf{1}_N^\top \tau + \mathbf{a}_N \mathbf{y}'^\top + b] \\
&= E[\tau] + E[\mathbf{a}_N \mathbf{y}'^\top] - E[\mathbf{a}_N \mathbf{y}'^\top] = \tau.
\end{aligned} \tag{3.7}
$$

The last step in (3.7) follows by $b = -E[\mathbf{a}\mathbf{y}'^\top]$ in (3.6) and the unknown τ is constant.
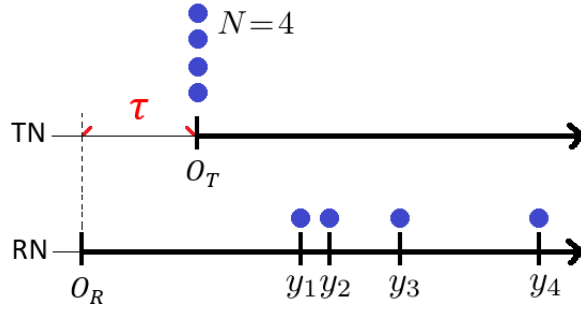
Figure 3.1: An example for LE when $\mathbf{x} = \mathbf{0}_4$.

Let's give an example when $K = 1$, $n_1 = N = 4$, and $\mathbf{x} = \mathbf{0}_N$ as a zero vector with length N. As shown in Fig. 3.1, if TN release N molecules at the beginning of communication. In this case, the random vector $\mathbf{y}' = [T_{(1)}, T_{(2)}, ..., T_{(N)}]$, where $T_{(i)}$ is the i-th order statistic of N independent and identical distribution (iid) random variables with generic Inverse Gaussian distribution $f_T(t)$.

For simplification, we only take the first two components y_1 and y_2 as example. Ap-

plying the solution of LE in this case, we get the linear estimator as below:

$$\hat{\tau}_{\mathrm{LE}} := \mathbf{a}_2 \mathbf{y}_2^\top + b$$

$$= a_1 y_1 + a_2 y_2 - E[a_1 T_{(1)} + a_2 T_{(2)}]$$

$$= a_1(y_1 - \mu_1) + (1 - a_1)(y_2 - \mu_2), \tag{3.8}$$

where $\mu_1 = E[T_{(1)}]$, $\mu_2 = E[T_{(2)}]$, and

$$a_1 = \frac{Var[T_{(2)}] - E[(T_{(1)} - \mu_1)(T_{(2)} - \mu_2)]}{E[\{(T_{(1)} - \mu_1) - (T_{(2)} - \mu_2)\}^2]}$$

$$= \frac{Var[T_{(2)}] - Cov[T_{(1)}, T_{(2)}]}{Var[T_{(1)} - T_{(2)}]}, \tag{3.9}$$

where the function $Var[.]$ and $Cov[.,.]$ denotes, respectively, the variance of some random variable and the covariance of two random variables.

In (3.8), it is clear that $\hat{\tau}_{\mathrm{LE}}$ is actually a weighted mean of each unbiased estimator $y_i - E[y_i]$. We use the information in the arrival time of each molecule and take weighted average to minimum MSE.

In this simple example, the theoretical MSE can be derived as below:

$$E[(\tau - \hat{\tau}_{\mathrm{LE}})^2] = Var[\mathbf{a}_2 \mathbf{y}_2^\top]$$

$$= \frac{Var[T_{(1)}]Var[T_{(2)}] - Cov[T_{(1)}, T_{(2)}]^2}{Var[T_{(1)} - T_{(2)}]}. \tag{3.10}$$

The theoretical MSE is used to verify our simulation results in the following section.

3.1.3 Iterative Linear Estimation For τ

In the above, we proposed LE, which reduces the complexity of MLE, in training-based synchronization. The LE only needs linear time computation on N after we have the weighted value $\mathbf{a}_N$. However, in the algorithm of LE, the complexity of computing the weighted value $\mathbf{a}_N$ and mean vector $\mathbf{u}$ beforehand still grows rapidly by factorial on N, which is incomputable when N is large. For example, if TN transmit $K = 10$ symbols

and each symbol release $n_i = 8$ molecules for all $1 \leq i \leq 10$, then we have to offline compute 80 means and weighted value by 80×80 covariance matrix $\mathbf{C}_{\mathbf{y}'}$. It costs much. To simplify the computation of $\mathbf{a}_N$ and make linear estimation for large N possible, we rewrite the algorithm to iterative form, which called Iterative LE (ILE) by us.

First of all, let's recall the linear combination in LE in (3.11). We denote the estimator of LE for the first N molecules by $\hat{\tau}_N$ and the corresponding coefficient by $\mathbf{a}_N$. Then, we introduce a constant vector $\mathbf{u}$ to represent the expected value of the random vector $\mathbf{y}'$. This way, it is clear that $\mathbf{a}_N$ is actually the weighted value of N unbiased estimators $y_i - u_i$ for $i = 1, 2, ..., N$.

$$\hat{\tau}_N := \mathbf{a}_N \mathbf{y}^\top + b = \mathbf{a}_N (\mathbf{y} - \mathbf{u})^\top, \tag{3.11}$$

where $\mathbf{u} = E[\mathbf{y}']$ is the mean vector.

Assume TN transmits total K symbols with the same quantity of n molecules. That is, the training sequence $\{n_k | 1 \leq k \leq K\}$ is a constant sequence, so $n_1 = n_2 = ... = n_K = n$. When RN receives k symbols, we can apply LE for the first $N = kn$ molecules. When RN receives $k + 1$ symbols, we can apply LE for the first $N = (k + 1)n$ molecules. We compare these two estimators to derive the iterative form of ILE.

$$\hat{\tau}_{kn} = \mathbf{a}_{kn} (\mathbf{y}^{[1:kn]} - \mathbf{u}^{[1:kn]})^\top,$$

$$\hat{\tau}_{(k+1)n} = \mathbf{a}_{(k+1)n} (\mathbf{y}^{[1:(k+1)n]} - \mathbf{u}^{[1:(k+1)n]})^\top \tag{3.12}$$

,where $\mathbf{y}^{[i_1:i_2]}$ denote the i_1-th component to the i_2-th component of vector $\mathbf{y}$ and so does $\mathbf{u}^{[i_1:i_2]}$.

In (3.12), the first kn components of $\mathbf{y}_{kn}$ and $\mathbf{y}_{(k+1)n}$ are the same, so are $\mathbf{u}_{kn}$ and $\mathbf{u}_{(k+1)n}$. Therefore, $\mathbf{y}_{(k+1)n} = [\mathbf{y}_{kn}, \mathbf{y}^{[kn+1:(k+1)n]}]$ and $\mathbf{u}_{(k+1)n} = E[\mathbf{y}_{(k+1)n}] = [\mathbf{u}_{kn}, \mathbf{u}^{[kn+1:(k+1)n]}]$ However, the first kn components of $\mathbf{a}_{kn}$ and $\mathbf{a}_{(k+1)n}$ are different. We need to derive the relationship with them.

Recall that $\mathbf{a}_N = \mathbf{1}_N \mathbf{C}_{\mathbf{y}'}^{-1} \{\mathbf{1}_N \mathbf{C}_{\mathbf{y}'}^{-1} \mathbf{1}_N^\top\}^{-1}$ from (3.6). The inverse of covariance matrix $\mathbf{C}_{\mathbf{y}'}^{-1}$ has the following two properties.
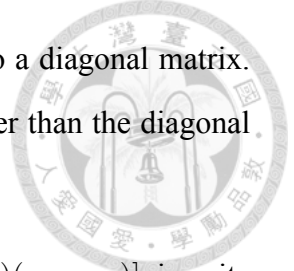
Approximate diagonal property : The matrix $\mathbf{C}_{\mathbf{y}'}^{-1}$ is similar to a diagonal matrix. That is, the entries outside the main diagonal are significantly smaller than the diagonal entries.

This property results from the covariance of y_i' and y_j', $E[(y_i - u_i)(y_j - u_j)]$, is quite small when $|i - j|$ is large. Accordingly, the matrix $\mathbf{C}_{\mathbf{y}'}$ is similar to a diagonal matrix, and so is its inverse matrix $\mathbf{C}_{\mathbf{y}'}^{-1}$.

Approximate repetition property : The diagonal entries of matrix $\mathbf{C}_{\mathbf{y}'}^{-1}$ repeats by the period of n except the first submatrix. That is, the matrix $\mathbf{C}_{\mathbf{y}'}^{-1}$ approximately shows as below.

$$
\mathbf{C}_{\mathbf{y}'}^{-1} \approx
\begin{bmatrix}
\mathbf{A}_{n\times n} & \mathbf{0} & \mathbf{0} & \dots & & \mathbf{0} \\
\mathbf{0} & \mathbf{B}_{n\times n} & \mathbf{0} & & & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{B}_{n\times n} & & & \vdots \\
\vdots & & & \ddots & & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \dots & & \mathbf{0} & \mathbf{B}_{n\times n}
\end{bmatrix}
$$

This repetition property results from $y_{i+(k-1)n}'$ is similar to y_{i+kn}' for $i = 1, 2, ..., n$ and $2 \leq k \leq K - 1$ because they all affect by similar level of ISI effect. We assume the crossover effect over two or more symbols can be ignored. Under this assumption, the effect of ISI on the second symbol is the same as the third one, and so are the following symbols. On the other hand, the first symbol without ISI causes the exception.

The above two properties of $\mathbf{C}_{\mathbf{y}'}^{-1}$ is useful when we find the relationship between $\mathbf{a}_{kn}$ and $\mathbf{a}_{(k+1)n}$. When $N = (k + 1)n$, based on the properties of $\mathbf{C}_{\mathbf{y}'}^{-1}$, we can derive $\mathbf{a}_{(k+1)n}$ as

$$
\begin{aligned}
\mathbf{a}_{(k+1)n} &= \frac{\mathbf{1}_N \mathbf{C}_{\mathbf{y}'}^{-1}}{\mathbf{1}_N \mathbf{C}_{\mathbf{y}'}^{-1} \mathbf{1}_N^\top} = \frac{\mathbf{1}_N \mathbf{C}_{\mathbf{y}'}^{-1}}{\mathbf{1}_n \mathbf{A}_{n\times n} \mathbf{1}_n^\top + k\mathbf{1}_n \mathbf{B}_{n\times n} \mathbf{1}_n^\top} \\
&= \frac{\mathbf{1}_{(k+1)n} \mathbf{C}_{\mathbf{y}'}^{-1}}{A + kB},
\end{aligned}
\tag{3.13}
$$

where $A = \mathbf{1}_n \mathbf{A}_{n\times n} \mathbf{1}_n^\top$ and $B = \mathbf{1}_n \mathbf{B}_{n\times n} \mathbf{1}_n^\top$. Moreover, when $N = kn$, we can derive

$\mathbf{a}_{kn}$ as

$$\mathbf{a}_{kn} = \frac{\mathbf{1}_{kn}\mathbf{C}_{\mathbf{y}'}^{-1}}{A + (k-1)B}. \tag{3.14}$$

In the same way, we split $\mathbf{a}_{(k+1)n}$ into two parts, $\mathbf{a}^{[1:kn]}$ and $\mathbf{a}^{[kn+1:(k+1)n]}$, where $\mathbf{a}^{[1:kn]}$ is the first kn components of $\mathbf{a}_{(k+1)n}$ and $\mathbf{a}^{[kn+1:(k+1)n]}$ is n components from the $(kn+1)$-th one to the $(k+1)n$-th one. In (3.13), the denominator is a constant $A + kB$, and the nominator is a row vector which is the sum of all row vectors of $\mathbf{C}_{\mathbf{y}'}^{-1}$. Accordingly, we can derive the relationship as below.

$$\begin{aligned}
\mathbf{a}^{[1:kn]} &= \frac{\mathbf{1}_{kn}\mathbf{C}_{\mathbf{y}'}^{-1} + \mathbf{0}}{A + kB} = \frac{\mathbf{1}_{kn}\mathbf{C}_{\mathbf{y}'}^{-1}}{A + (k-1)B} \frac{A + (k-1)B}{A + kB} \\
&= \mathbf{a}_{kn}\frac{A + (k-1)B}{A + kB}
\end{aligned} \tag{3.15}$$

$$\begin{aligned}
\mathbf{a}^{[kn+1:(k+1)n]} &= \frac{\mathbf{0} + ... + \mathbf{0} + \mathbf{1}_n\mathbf{B}_{n\times n}}{A + kB} = \frac{\mathbf{1}_n\mathbf{B}_{n\times n}}{B} \frac{B}{A + kB} \\
&= \mathbf{w}_n\frac{B}{A + kB},
\end{aligned} \tag{3.16}$$

where $\mathbf{w}_n = \frac{\mathbf{1}_n\mathbf{B}_{n\times n}}{B}$.

Then, we split $\mathbf{a}_{(k+1)n}$ and replace $\mathbf{a}^{[1:kn]}$ and $\mathbf{a}^{[kn+1:(k+1)n]}$ by (3.15) and (3.16). The estimator of LE for the first $N = (k+1)n$ molecules can be derived from the estimator of LE for the first $N = kn$ molecules by the iterative form as below.

$$\begin{aligned}
\hat{\tau}_{(k+1)n} &:= \mathbf{a}_{(k+1)n}(\mathbf{y}^{[1:(k+1)n]} - \mathbf{u}^{[1:(k+1)n]})^\top \\
&= \mathbf{a}^{[1:kn]}(\mathbf{y}^{[1:kn]} - \mathbf{u}^{[1:kn]})^\top \\
&\quad + \mathbf{a}^{[kn+1:(k+1)n]}(\mathbf{y}^{[kn+1:(k+1)n]} - \mathbf{u}^{[kn+1:(k+1)n]})^\top \\
&= \frac{A + (k-1)B}{A + kB}\mathbf{a}_{kn}(\mathbf{y}^{[1:kn]} - \mathbf{u}^{[1:kn]})^\top \\
&\quad + \frac{B}{A + kB}\mathbf{w}_{n_1}(\mathbf{y}^{[kn+1:(k+1)n]} - \mathbf{u}^{[kn+1:(k+1)n]})^\top \\
&= \frac{A + (k-1)B}{A + kB}\hat{\tau}_{kn} + \frac{B}{A + kB}\hat{\tau}_{\text{new}}^{(k+1)},
\end{aligned} \tag{3.17}$$

where $\hat{\tau}_{\text{new}}^{(k+1)} = \mathbf{w}_n(\mathbf{y}^{[kn+1:(k+1)n]} - \mathbf{u}^{[kn+1:(k+1)n]})^\top$.

For simplification, in ILE, we denote the previous estimator $\hat{\tau}_{kn}$ and the next estimator $\hat{\tau}_{(k+1)n}$ in LE by $\hat{\tau}_k$ and $\hat{\tau}_{k+1}$, respectively. Moreover, another parameter $\alpha = \frac{A}{B}$ represents the importance of the new estimator $\hat{\tau}_{\text{new}}^{(k+1)}$ with respect to the previous estimator $\hat{\tau}_k$.

$$\hat{\tau}_{k+1} = \frac{\alpha + k - 1}{\alpha + k}\hat{\tau}_k + \frac{1}{\alpha + k}\hat{\tau}_{\text{new}}^{(k+1)}, \tag{3.18}$$

where $\alpha = \frac{A}{B}$. In the case when $\alpha = 1$, which means $A = B$, that is T_s is large enough so that all symbols are almost independent with each other. Then, we treat $\hat{\tau}_{\text{new}}^{(k+1)}$ with the same importance with $\hat{\tau}_k$ in this case. On the other hand, in the case when $\alpha > 1$, which means $A > B$, that is the following symbols are influenced by the ISI effects. Then, we reduce the importance of $\hat{\tau}_{\text{new}}^{(k+1)}$ with respect to $\hat{\tau}_k$.

Moreover, because of the periodic property of $\mathbf{y}'$, $u_{i+(k-1)n} - u_{i+kn}$ is close to T_s, which is constant, for $i = 1, 2, ..., n$ and $1 \leq k \leq K-1$. As a result, $\mathbf{u}^{[kn+1:(k+1)n]} = kT_s\mathbf{1}_n + \mathbf{m}_n$ for some constant vector $\mathbf{m}_n = [m_1, m_2, ..., m_n]$.

$$\begin{aligned}\hat{\tau}_{\text{new}}^{(k+1)} &= \mathbf{w}_n(\mathbf{y}^{[kn+1:(k+1)n]} - \mathbf{u}^{[kn+1:(k+1)n]})^\top \\ &= \mathbf{w}_n(\mathbf{y}^{[kn+1:(k+1)n]} - \mathbf{m}_n)^\top - kT_s\end{aligned} \tag{3.19}$$

To sum up, the algorithm of ILE is described as below. First, we compute $\mathbf{a}_n$ and $\mathbf{u}^{[1:n]}$ beforehand to initialize the first estimator $\hat{\tau}_1$. Second, we compute α, $\mathbf{w}_n$, and $\mathbf{m}_n$ beforehand to iteratively update the previous estimator $\hat{\tau}_k$ for $1 \leq k \leq K-1$.

$$\hat{\tau}_1 = \mathbf{a}_n(\mathbf{y}^{[1:n]} - \mathbf{u}^{[1:n]})^\top \tag{3.20}$$

$$\hat{\tau}_{k+1} = \frac{\alpha + k - 1}{\alpha + k}\hat{\tau}_k + \frac{1}{\alpha + k}[\mathbf{w}_n(\mathbf{y}^{[kn+1:(k+1)n]} - \mathbf{m}_n)^\top - kT_s] \tag{3.21}$$

The algorithm of ILE only needs the statistics of the first symbol without ISI, $\mathbf{a}_n$ and $\mathbf{u}_n$, the second or third symbol with ISI, $\mathbf{w}_n$ and $\mathbf{m}_n$, and the ratio of importance between them, α. By these information, it is enough to iteratively derive LE for large amount of molecules N and large index of symbol k without too much performance lost. Unlike

$\mathbf{a}_n$	the weighted value of each arrival time without ISI (derived from covariance matrix of the arrival time).
$\mathbf{u}_n$	the mean vector of the arrival time without ISI.
$\mathbf{w}_n$	the weighted value of each arrival time with ISI (derived from covariance matrix of the arrival time).
$\mathbf{m}_n$	the mean vector of the arrival time with ISI.
α	the weighted value between the previous estimation and the new estimation.

Table 3.1: The parameters of the Iterative LE for clock offset τ.

LE, whose complexity grows rapidly when N and k increases, ILE only needs a constant amount of computational complexity when offline computing the weighted values and means.

3.1.4 Lower Bound Without Crossover

In this part, we derive the theoretical lower bound of MSE without ISI as a benchmark to evaluate the performance of our estimator proposed above.

Because we want to derive the lower bound of MSE, the ideal case under some unrealistic condition can be considered. Here, we just consider the situation without crossover effect. That is, if TN transmit K symbols, these K symbols are without ISI. Moreover, if n_k molecules are released in the k-th symbol, these n_k arrival times of molecules are i.i.d. To sum up, All arrival times of molecules $\mathbf{y}$ are a random sample with the sample size $N = n_1 + ... + n_K$.

According to the classic estimation theory, the classic lower bound of MSE for all unbiased estimators is Cramer-Rao Lower Bound (CRLB). The fisher information number of inverse gaussian distribution with respect to τ is

$$
\begin{aligned}
I_0 &= -E[\frac{\partial^2}{\partial \tau^2}\{\ln f_T(t-\tau)\}] \\
&= -E[\{\frac{3}{2(t-\tau)^2} - \frac{\lambda}{(t-\tau)^3}\}] \\
&= -E[\frac{3}{2(T)^2} - \frac{\lambda}{(T)^3}] \\
&= \frac{\lambda}{\mu^3} + \frac{9}{2\mu^2} + \frac{21}{2\mu\lambda} + \frac{21}{2\lambda^2}.
\end{aligned}
\tag{3.22}
$$

Moreover, we assume **y** are a random sample with the sample size $N = n_1 + \ldots + n_K$, so the CRLB is

$$\frac{1}{NI_0} = \frac{1}{I_0 \sum_{k=1}^{K} n_k}, \tag{3.23}$$
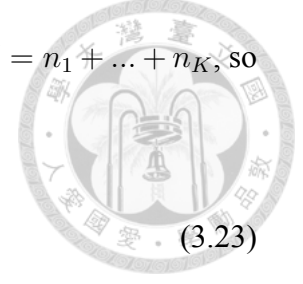
However, the parameter we want to estimate is the location parameter τ. The support of $f_T(t - \tau)$ is $(\tau, \infty]$, which depends on the parameter τ. Accordingly, the assumption of Cramer-Rao inequality does not hold, so we cannot claim CRLB is the lower bound of MSE for all unbiased estimator.

Nevertheless, the minimum MSE estimator for location parameter was investigated by Pitman [20]. For all estimators $\hat{\tau}(\mathbf{y})$ satisfy the invariant property as

$$\hat{\tau}(\mathbf{y} + c\mathbf{1}) = \hat{\tau}(\mathbf{y}) + c, \tag{3.24}$$

for any constant c, then Pitman estimator $\hat{\tau}_P(\mathbf{y})$ has the minimum MSE.

$$\hat{\tau}_P(\mathbf{y}) := \frac{\int_{-\infty}^{\infty} \theta f(\mathbf{y}|\mathbf{x}, \tau = \theta) \, \mathrm{d}\theta}{\int_{-\infty}^{\infty} f(\mathbf{y}|\mathbf{x}, \tau = \theta) \, \mathrm{d}\theta} \tag{3.25}$$

The joint pdf $f(\mathbf{y}|\mathbf{x}, \tau = \theta)$ described in (3.3) is too complicated to integral. Under the crossover effect, we have to consider all the permutation of **y** in (3.3), so $\mathbf{P}(\mathbf{y})$ has $N!$ terms need to be considered. The factorial $N!$ grows so rapidly that the computational complexity beyonds the capacity of nano-machines. This is why we need to investigate the other estimator with lower complexity and without too much MSE loss.

Nevertheless, under the condition without crossover effect, we treat **y** is a random

sample with sample size N. We can simplify the Pitman estimator as

$$\hat{\tau}_P(\mathbf{y}) := \frac{\int_{-\infty}^{\infty} \theta \prod_{i=1}^{N} f_T(y_i - 0 - \theta) \, \mathrm{d}\theta}{\int_{-\infty}^{\infty} \prod_{i=1}^{N} f_T(y_i - 0 - \theta) \, \mathrm{d}\theta} \tag{3.26}$$

$$= \frac{\int_{-\infty}^{y_1} \theta \prod_{i=1}^{N} f_T(y_i - \theta) \, \mathrm{d}\theta}{\int_{-\infty}^{y_1} \prod_{i=1}^{N} f_T(y_i - \theta) \, \mathrm{d}\theta} \tag{3.27}$$

$$= \frac{\int_{\infty}^{0} [y_1 - \alpha] \prod_{i=1}^{N} f_T(y_i - y_1 + \alpha) \, \mathrm{d}(y_1 - \alpha)}{\int_{\infty}^{0} \prod_{i=1}^{N} f_T(y_i - y_1 + \alpha) \, \mathrm{d}(y_1 - \alpha)} \tag{3.28}$$

$$= y_1 - \frac{\int_{0}^{\infty} \theta \prod_{i=1}^{N} f_T(y_i - y_1 + \theta) \, \mathrm{d}\theta}{\int_{0}^{\infty} \prod_{i=1}^{N} f_T(y_i - y_1 + \theta) \, \mathrm{d}\theta} \tag{3.29}$$

The integral in (3.29) has no close form, so we numerically compute the value in our simulation.

3.1.5 Simulation Results

We simulate the quantity-based modulation in AIGN channel with non-synchronous effect and try to synchronize by MLE, LE, and ILE proposed in training-based synchronization. Moreover, we numerically plot the CRLB and Pitman estimator shown as benchmark to evaluate the efficiency of these estimators.

For the parameters of AIGN channel, we set $T_a = 298(25°C)$, $\eta = 8.9 \times 10^{-4}$(water in 25°C), $r = 10^{-8}$(10nm), $d = 2 \times 10^{-5}$(20μm), and $v = 2 \times 10^{-6}$(2μm/sec), so we use the random variable $T \sim \mathrm{IG}(\mu,\lambda)$ with $\mu = 10$ and $\lambda \approx 8.1955$ in the whole thesis.

The following, we present two simulation results in one case when $K = 1$, which is just considering the first symbol, and another case when $K \geq 2$, which is considering multi-symbol with ISI effect.

In Fig. 3.2, we only consider the first symbol with $n_1 = N$ molecules. When we use LE for n_1 arrival times to estimate τ, the MSE can reach as small as using MLE, but the complexity of LE reduce to linear time on N, which is quite lower than MLE. Considering the efficiency, the MSE of LE is closer to that of Pitman estimator than MLE when the quantity of molecules n_1 is small. This is reasonable because of the asymptotic efficiency property of MLE. Moroever, we find out that the MSE of both two methods are close to CRLB when the quantity of molecules N is large enough.
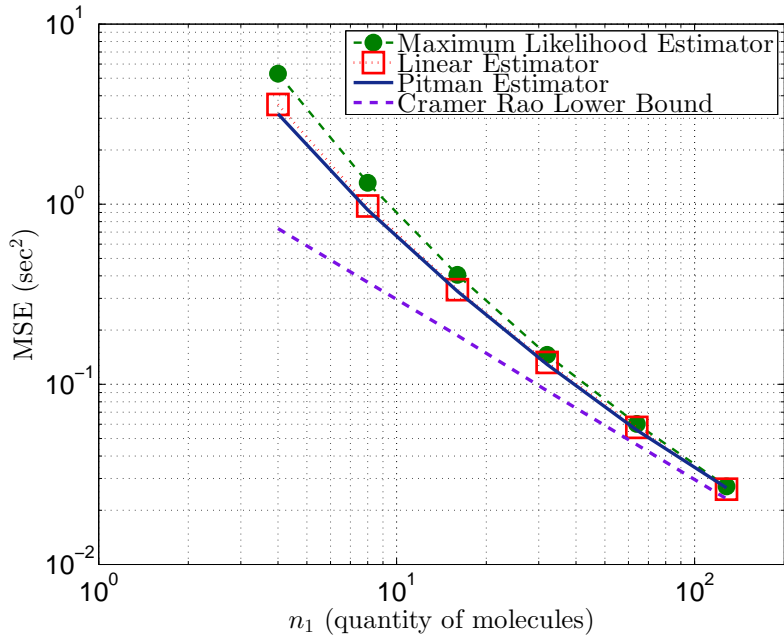
Figure 3.2: The MSE of MLE and LE with CRLB and Pitman as benchmark in case when $K = 1$, $n_1 = N$, and $\mathbf{x} = \mathbf{0}_{n_1}$.

The experiment in Fig. 3.2 only simulates the special case when we consider the first symbol, which is $K = 1$. In general, when we use multi-symbol to estimate τ, that is $K \geq 2$, the ISI effect will affect our result. In this case, we simulate LE for all arrival N molecules when TN releases K symbols with n_1 molecules per symbol, that is $n_1 = n_2 = ... = n_K$. Because ISI effect, the symbol duration T_s affects the performance of estimation.

We simulate $K = 6$ symbols in Fig. 3.3. When T_s is closed to $\mu := E[T]$, that is ISI effect is severe, the MSE of our estimators increase. On the other hand, when T_s is large enough, all symbols become almost independent, so all symbols act as the first symbol without ISI effect. The MSE of our estimators are close to the lower bound, Pitman or CRLB without ISI effect. Moreover, the MSE of Iterative LE is really close to that of LE, especially when T_s is large, which verify that we do not lose too much information when we reform the iterative process in order to reduce the complexity.
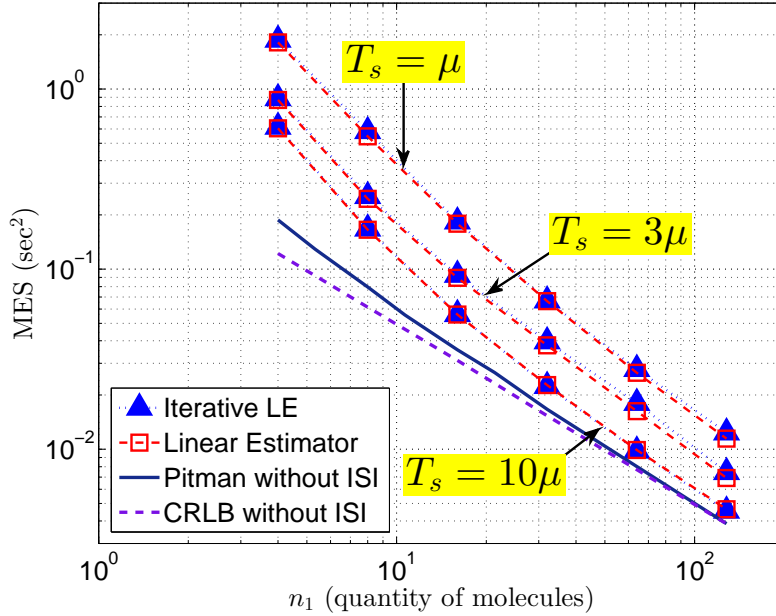
Figure 3.3: The MSE of LE with multi-symbol in case when $K = 6$, and $n_1 = n_2 = ... = n_6$, where μ is the mean of T.

3.2 Blind Synchronization

In this section, we discuss on clock offset estimation when $\{n_k\}$ for $1 \leq k \leq K$ is not constant but random for RN, because of the message embedded in the quantity level of molecules for each symbol. Therefore, different from above discussion, this case has no training phase anymore before communication.

Following the paper [8], we consider M-ary quantity-based modulation in general, that is $n_k \in \{L_0, L_1, ..., L_{M-1}\}$ as M hypotheses and $L_i > 0$ for $0 \leq i < M$. Accordingly, RN need to use the information in $\mathbf{y}$ to do both synchronization (estimating τ) and demodulation (detecting n_k).

3.2.1 Non-decision-directed Parameter Estimation

Under the uncertainty of $\{n_k\}$, the releasing time sequence $\mathbf{x}$ is also random for RN. We can rewrite the joint pdf of observation $\mathbf{y}'$ when $\tau = 0$ by averaging the conditional joint pdf over the probability of $\mathbf{x}$. Here, we assume *a priori* probability of $\mathbf{x}$ is known for

RN, so the joint pdf of $\mathbf{y}'$ can be derived as follows:

$$f(\mathbf{y}') = \sum_{\mathbf{x}} \Pr\{\mathbf{x}\} f(\mathbf{y}'|\mathbf{x}). \tag{3.30}$$

For simplification, we just apply Linear Estimation (LE) proposed in Sec. 3.1.2 only for the first molecule, that is $\hat{\tau}_{\mathrm{LE}} := y_1 - E[y_1']$,which is a simple and efficient estimator. Obviously, we just use the first molecule in the first symbol duration to estimate τ. In general case when T_s is large enough, the probability of crossover happen to all the molecules from the first symbol is quite small, so we almost can assume the first arrival molecule is from the first symbol. As a result, what we really care about is the uncertainty of n_1, because y_1' is almost the first order statistic of n_1 iid random variables with generic distribution $f_T(t)$.

$$\begin{aligned} f(y_1') &= \sum_{l=L_0}^{L_{M-1}} \Pr\{n_1 = l\} f(y_1'|\mathbf{x} = \mathbf{0}_l) \\ &= \sum_{j=0}^{M-1} p_j f_{T_{(1)}^{(L_j)}}(y_1'), \end{aligned} \tag{3.31}$$

where $T_{(i)}^{(l)}$ is the i-th order statistic of l iid random variables with generic distribution $f_T(t)$ and p_j is *a priori* probability of $n_1 = L_j$. Accordingly, we can rewrite the LE only for the first molecule as below:

$$\hat{\tau}_{\mathrm{LE}} := y_1 - \sum_{j=0}^{M-1} p_j E[T_{(1)}^{(L_j)}]. \tag{3.32}$$

With this estimator, the theoretical MSE can be derived as below:

$$E[(\tau - \hat{\tau}_{\mathrm{LE}})^2] = Var[y_1'] = \sum_{j=0}^{M-1} p_j \sigma_j^2 + \sum_{j=0}^{M-1} \sum_{i=0}^{j} (v_j - v_i)^2 p_i p_j, \tag{3.33}$$

where $\sigma_j^2 = Var[T_{(1)}^{(L_j)}]$, and $v_j = E[T_{(1)}^{(L_j)}]$ for $j = 0, 1, ..., (M-1)$.
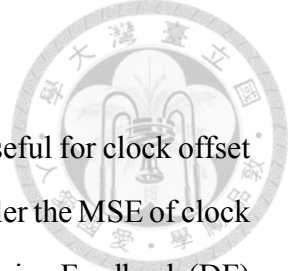
3.2.2 Decision Feedback

Because of blind synchronization, the output of demodulation is useful for clock offset estimation. That is, the better the performance of detection is, the smaller the MSE of clock offset estimation will be, and vice versa. As a result, we can apply Decision Feedback (DF) method to improve LE only for the first molecule.

The steps are listed as follows. First, get $\hat{\tau}_1 := y_1 - E[y_1']$ by LE only for the first molecule, and then demodulate the first symbol to get $\tilde{n}_1$ based on $\hat{\tau}_1$, where $\tilde{n}_1$ is the detected value of the molecular quantity of the first symbol n_1. Finally, get $\hat{\tau}_{\text{DF}} := y_1 - E[T_{(1)}^{(\tilde{n}_1)}]$ by DF method, where $E[T_{(1)}^{(\tilde{n}_1)}]$ is a derived random variable on the domain of $\{v_0, v_1, ..., v_{M-1}\}$.

To analyze the performance of DF, let $q_{i,j}$ denote the probability when RN detects n_1 as L_j condition on $n_1 = L_i$, that is $q_{i,j} := \Pr\{\tilde{n}_1 = L_j | n_1 = L_i\}$. We can derive the theoretical MSE as below:

$$E[(\tau - \hat{\tau}_{\text{DF}})^2] = \sum_{j=0}^{M-1} p_j \sigma_j^2 + \sum_{j=0}^{M-1} \sum_{i=0}^{j} (v_j - v_i)^2 (p_i q_{i,j} + p_j q_{j,i}). \qquad (3.34)$$

If we compare the (3.34) with (3.33), the only difference is the coefficient of $(v_j - v_i)^2$ term. Therefore, applying DF will improve performance when $p_i q_{i,j} + p_j q_{j,i} < p_i p_j$ for all $0 \leq i < j \leq M - 1$, that is the average crossover error probability of level i and j is less than $p_i p_j$. Moreover, the best performance of DF can make MSE reach $\sum_{j=0}^{M-1} p_j \sigma_j^2$ when $q_{ij} = 0$ for $i \neq j$, that is in the absence of demodulation errors, $\tilde{n}_1 = n_1$, which known as Decision-directed Parameter Estimation.

3.2.3 Fisher's Scoring Algorithm For Symbols With ISI

The above methods to estimate clock offset τ only use the first symbol without ISI. In blind synchronization, how to use the information in the following symbols with ISI is still a problem. We introduce a method to iteratively update the initial estimation $\hat{\tau}_{\text{LE}}$ or $\hat{\tau}_{\text{DF}}$. This way, in one communication, RN can both receive information message and synchronize with TN overtime. We expect that the updated clock offset will be more and
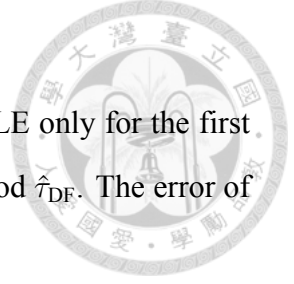
more accurate overtime.

First, we denote the initial estimation by $\hat{\tau}_{(0)}$, which may be the LE only for the first molecule in blind synchronization or use the decision feedback method $\hat{\tau}_{\text{DF}}$. The error of initial estimation is defined by

$$\epsilon := \hat{\tau}_{(0)} - \tau. \tag{3.35}$$

If ϵ is positive, it means that our initial estimation is overestimated. Otherwise, our initial estimation is underestimated. The following, we want to use the symbols with ISI to detect whether $\hat{\tau}_{(0)}$ is overestimated or underestimated. Then updating to the new clock offset estimation by compensating the estimated error $\hat{\epsilon}$.

To overcome the ISI effect, we try to derive the new observations, which affected by ISI. We define a function $g(x)$ as below.

$$g(x) := \begin{cases} T_s, & \text{if } \frac{x}{T_s} \in \mathbb{Z}; \\ x - \left\lfloor \frac{x}{T_s} \right\rfloor T_s, & \text{otherwise.} \end{cases}$$
$$= \sum_{k=0}^{\infty} [x - (k-1)T_s] \mathbb{1}\{(k-1)T_s < x \le kT_s\} \tag{3.36}$$

As shown in Fig. 3.4, the domain of $g(x)$ is $(-\infty, \infty)$ and the range of $g(x)$ is $(0, Ts]$. To short, the function $g(x)$ is the residue of x divided by T_s if x is not a multiple of T_s, otherwise $g(kT_s) = T_s$ for some $k \in \mathbb{Z}$. $g(x)$ is a many-to-one function, which maps the values from $(-\infty, \infty)$ to $(0, Ts]$.
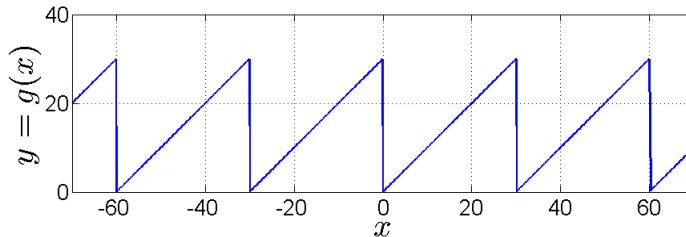
Figure 3.4: Function $g(x)$ from the domain $(-\infty, \infty)$ to the range $(0, T_s]$ with $T_s = 30$.

In AIGN channel, the random delay T follows inverse gaussian distribution. If the

error of initial estimation is zero, that is, $\hat{\tau}_{(0)} = \tau$ and $\epsilon = 0$, then $T > T_s$ causes ISI effect. $g(T)$ is the derived random variable, which means the residue of T divided by T_s. In RN's time slotted system, if an arrival time y_i of a molecule falls in the k_i-th symbol duration of RN, that is $\tau + (k_i - 1)T_s < y_i \leq \tau + k_i T_s$, then $t_i' := y_i - \tau - (k_i - 1)T_s$ follows the distribution of $g(T)$. As shown in Fig. 3.5, the meaning of t_i' is the duration from the previous boundary of the k_i-th symbol to the arrival time y_i.
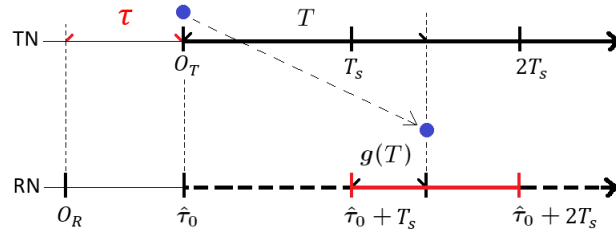
Figure 3.5: The variable $g(T)$ is the residue of random delay T divided by T_s in RN's time slotted system under perfect synchronization.

However, $\epsilon \neq 0$ in general. To derive the random quantity $t_i' := y_i - \hat{\tau}_{(0)} - (k_i - 1)T_s$ with error, we define the derived random variable T^{ISI} from the inverse gaussian random variable T by

$$T^{\text{ISI}} := g(T - \epsilon) \tag{3.37}$$

The meaning of T^{ISI} is shown in Fig. 3.6. RN derives the initial estimation $\hat{\tau}_{(0)}$ when receiving the first symbol. Then, according to $\hat{\tau}_{(0)}$, RN constructs its own time slotted system $\{\hat{\tau}_{(0)}, \hat{\tau}_{(0)} + T_s, \hat{\tau}_{(0)} + 2T_s, ...\}$ assume clock skew has been compensated perfect. RN receives information message by this time slotted system and simultaneously observes the random quantities $t_i' = y_i - \hat{\tau}_{(0)} - (k_i - 1)T_s$ if the arrival time y_i falls in the k_i-th symbol, where t_i' is a realization of random variable T^{ISI} follows the pdf $f_{T^{\text{ISI}}}(t')$.

The following, we want to derive the distribution of T^{ISI} by the inverse gaussian dis-
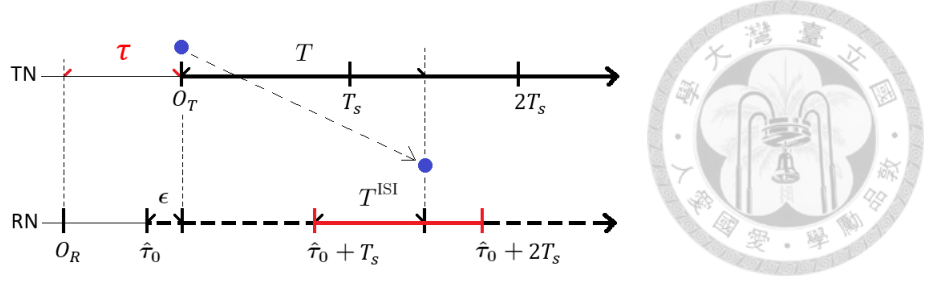
Figure 3.6: The meaning of T^{ISI} in RN's time slotted system.

tribution $f_T(t)$. Let us begin with a simple case when $\epsilon = 0$, then

$$F_{T^{\text{ISI}}}(t') = Pr(T^{\text{ISI}} \le t') = Pr(g(T) \le t') = \int_0^\infty \mathbb{1}\{g(t) \le t'\}f_T(t)\,\mathrm{d}t \qquad (3.38)$$

$$= \sum_{k=0}^\infty \int_{(k-1)T_s}^{kT_s} \mathbb{1}\{t - (k-1)T_s \le t'\}f_T(t)\,\mathrm{d}t \qquad (3.39)$$

$$= \sum_{k=0}^\infty \int_{(k-1)T_s}^{(k-1)T_s+t'} f_T(t)\,\mathrm{d}t \qquad (3.40)$$

The equation (3.40) shows the cumulative distribution function (cdf) of T^{ISI} under perfect synchronization. Take derivation with respect to t' to get the pdf of T^{ISI}.

$$f_{T^{\text{ISI}}}(t') = \sum_{k=0}^\infty \frac{\mathrm{d}}{\mathrm{d}t'} \int_{(k-1)T_s}^{(k-1)T_s+t'} f_T(t)\,\mathrm{d}t = \sum_{k=0}^\infty f_T(t' + (k-1)T_s), \qquad (3.41)$$

where $0 < t' \le T_s$. As shown in (3.41), the pdf of T^{ISI} under perfect synchronization accumulates the inverse gaussian distribution $f_T(t)$ with the period T_s and shifts each period to the support $(0, T_s]$.

In general, if $\epsilon \ne 0$, then

$$F_{T^{\text{ISI}}}(t') = Pr(g(T - \epsilon) \le t') = \int_0^\infty \mathbb{1}\{g(t - \epsilon) \le t'\}f_T(t)\,\mathrm{d}t \qquad (3.42)$$

$$= \sum_{k=0}^\infty \int_{(k-1)T_s}^{kT_s} \mathbb{1}\{t - \epsilon - (k-1)T_s \le t'\}f_T(t)\,\mathrm{d}t \qquad (3.43)$$

$$= \sum_{k=0}^\infty \int_{\epsilon+(k-1)T_s}^{\epsilon+(k-1)T_s+t'} f_T(t)\,\mathrm{d}t \qquad (3.44)$$

Similarly, let us take derivation with respect to t' to both side.

$$f_{T^{\text{ISI}}}(t') = \sum_{k=0}^\infty \frac{\mathrm{d}}{\mathrm{d}t'} \int_{\epsilon+(k-1)T_s}^{\epsilon+(k-1)T_s+t'} f_T(t)\,\mathrm{d}t = \sum_{k=0}^\infty f_T(t' + \epsilon + (k-1)T_s) \qquad (3.45)$$

where $0 < t' \leq T_s$. Note that the support of T^{ISI} is still $(0, T_s]$, even though $\epsilon \neq 0$. We see the pdf of T^{ISI} as a distribution function with two parameters, ϵ and T_s as in (3.46).
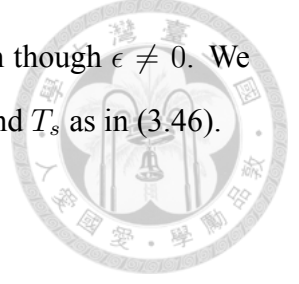
$$f_{T^{\text{ISI}}}(t'|\epsilon, T_s) = \sum_{k=0}^{\infty} f_T(t' + \epsilon + (k-1)T_s)\mathbb{1}\{0 < t' \leq T_s\} \tag{3.46}$$

The support of T^{ISI} only depends on T_s and is independent on the parameter ϵ which we want to estimate.

According to the observations t'_i affected by ISI and its distribution $f_{T^{\text{ISI}}}(t'|\epsilon, T_s)$, we want to apply Fisher's Scoring Algorithm to find the MLE of the true error $\epsilon_0 = \hat{\tau}_{(0)} - \tau$ of initial estimation.

Let $t'_1, t'_2, ..., t'_n$ be a random sample from the generic random variable T^{ISI} follows $f_{T^{\text{ISI}}}(t'|\epsilon, T_s)$ distribution. Assume T_s is known, so the support is independent on the parameter ϵ. Denote the MLE for the parameter ϵ by $\hat{\epsilon}$ as:

$$\hat{\epsilon} := \arg\max_{\epsilon} \sum_{i=1}^{n} \log f_{T^{\text{ISI}}}(t'_i|\epsilon, T_s) \tag{3.47}$$

To find $\hat{\epsilon}$, we need to find the root of score statistics $s_{T^{\text{ISI}}}(t'_i + \epsilon)$ defined by the derivation of log likelihood with respect to ϵ. Note that $f_{T^{\text{ISI}}}(t'_i|\epsilon, T_s)$ is a function of $t'_i + \epsilon$, so is the score function.

$$0 = \sum_{i=1}^{n} \frac{\partial}{\partial \epsilon} \log f_{T^{\text{ISI}}}(t'_i|\hat{\epsilon}, T_s) = \sum_{i=1}^{n} s_{T^{\text{ISI}}}(t'_i + \hat{\epsilon}) \tag{3.48}$$

To simply the score statistic, Let us apply chain rule and substitute (3.46) into (3.48).

$$s_{T^{\text{ISI}}}(t_i' + \epsilon) := \frac{\partial}{\partial \epsilon} \log f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s) = \frac{1}{f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s)} \frac{\partial f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s)}{\partial \epsilon} \tag{3.49}$$

$$= \frac{1}{f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s)} \sum_{k=0}^{\infty} \frac{\partial}{\partial \epsilon} f_T(t' + \epsilon + (k-1)T_s) \tag{3.50}$$

$$= \frac{1}{f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s)} \sum_{k=0}^{\infty} f_T(t_i' + \epsilon + (k-1)T_s) \frac{\partial \log f_T(t_i' + \epsilon + (k-1)T_s)}{\partial \epsilon}$$

$$\tag{3.51}$$

$$= \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k-1)T_s) \frac{\partial \log f_T(t' + \epsilon + (k-1)T_s)}{\partial \epsilon} \tag{3.52}$$

$$= \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k-1)T_s) s_T(t_i' + \epsilon + (k-1)T_s) \tag{3.53}$$

where the weighted statistic $w(t_i' + \epsilon + (k-1)T_s)$ and the score statisic of inverse gaussian $s_T(t + \epsilon)$ is defined as below. In (3.51), we apply chain rule again.

$$w(t_i' + \epsilon + (k-1)T_s) := \frac{f_T(t_i' + \epsilon + (k-1)T_s)}{f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s)} \tag{3.54}$$

$$= \frac{f_T(t_i' + \epsilon + (k-1)T_s)}{\sum_{k'=0}^{\infty} f_T(t' + \epsilon + (k'-1)T_s)} \tag{3.55}$$

In (3.55), it is obvious that $w(t_i' + \epsilon + (k-1)T_s)$ is the $(k-1)$-th weighted value of inverse gaussian pdf $f_T(t_i' + \epsilon + (k-1)T_s)$ when $t_i' + \epsilon$ and T_s is fixed. Note that the denominator of (3.55) must converge because of the integral test and the integral of a pdf is equal to 1.

$$s_T(t + \epsilon) := \frac{\partial}{\partial \epsilon} \log f_T(t + \epsilon) = \frac{1}{f_T(t + \epsilon)} \frac{\partial f_T(t + \epsilon)}{\partial \epsilon} \tag{3.56}$$

$$= \frac{1}{f_T(t + \epsilon)} \frac{\partial}{\partial \epsilon} \left\{ \sqrt{\frac{\lambda}{2\pi(t + \epsilon)^3}} \exp\left\{-\frac{\lambda(t + \epsilon - \mu)^2}{2\mu^2(t + \epsilon)}\right\} \right\} \tag{3.57}$$

$$= \frac{1}{f_T(t + \epsilon)} \left\{ -\frac{3}{2(t + \epsilon)} - \frac{\lambda}{2\mu^2} + \frac{\lambda}{2(t + \epsilon)^2} \right\} f_T(t + \epsilon) \tag{3.58}$$

$$= \frac{\lambda}{2(t + \epsilon)^2} - \frac{3}{2(t + \epsilon)} - \frac{\lambda}{2\mu^2} \tag{3.59}$$

It is not surprise that the score statistic of inverse gaussian is simple as (3.59) becasue inverse gaussian distribution is one of the exponential family.

The equation (3.53) points out that the score of T^{ISI} is the weighted mean of the score of T, $s_T(t'_i + \epsilon + (k-1)T_s)$, and the weighed value is proportional to the probability density $f_T(t'_i + \epsilon + (k-1)T_s)$. The meaning of (3.53) is intuitive. When RN observes an arrival time with ISI t'_i, this molecule could be released before the random delay t_i equals $t'_i + \epsilon + 0$, $t'_i + \epsilon + T_s$, or $t'_i + \epsilon + 2T_s$ ..., and t_i follows inverse gaussian distribution and the chance of each possible delay is proportional to the probability density $f_T(t'_i + \epsilon + (k-1)T_s)$.

To reduce the complexity when computing $s_{T^{\mathrm{ISI}}}(t'_i + \epsilon)$, which is an infinite weighted mean, we approximate it by $\tilde{s}_{T^{\mathrm{ISI}}}(t'_i + \epsilon)$ with just the first two terms. The approximate is reasonable if lever-1 crossover probability dominates lever-2 or more crossover probability, especially in the case $T_s > E[T] = \mu$.

$$s_{T^{\mathrm{ISI}}}(t'_i + \epsilon) \approx w(t_0)s_T(t_0) + (1 - w(t_0))s_T(t_0 + T_s) = \tilde{s}_{T^{\mathrm{ISI}}}(t'_i + \epsilon), \qquad (3.60)$$

where $w(t_0)$ and $s_T(t_0)$ are the first non-zero terms of weighted value and score value in the infinite summation (3.53). $s_T(t_0 + T_s)$ is the second non-zero term of score value.

The first non-zero term might be $k = 0$, 1, or 2 in the following three cases, respectively. In most cases, $t'_i + \epsilon \in (0, T_s]$ falls in the support of $f_{T^{\mathrm{ISI}}}(t')$, as shown in Fig. 3.7, then the first non-zero term is $k = 1$ in (3.53), that is, $t_0 = t'_i + \epsilon$ However, sometimes the initial estimation $\hat{\tau}_0$ is too underestimated. It causes $t'_i + \epsilon \leq 0$ as shown in Fig. 3.8. In this case, the first non-zero term is $k = 2$ in (3.53), that is, $t_0 = t'_i + \epsilon + T_s$. On the other hand, sometimes the initial estimation $\hat{\tau}_0$ is too overestimated. It causes $t'_i + \epsilon > T_s$ as shown in Fig. 3.9. In this case, the first non-zero term is $k = 0$ in (3.53), that is $t_0 = t'_i + \epsilon - T_s$.

Normal Case :If $0 < t'_i + \epsilon \leq T_s$ for some t'_i, then $f_T(t'_i + \epsilon - T_s) = 0$ and $f_T(t'_i + \epsilon) > 0$. Therefore, the first non-zero term is $k = 1$ in (3.53). We approximate the score statistic by $\tilde{s}_{T^{\mathrm{ISI}}}(t'_i + \epsilon)$ with $w(t_0) = \frac{f_T(t'_i + \epsilon)}{f_{T^{\mathrm{ISI}}}(t'_i | \epsilon, T_s)}$, $s_T(t_0) = \frac{\lambda}{2(t'_i + \epsilon)^2} - \frac{3}{2(t'_i + \epsilon)} - \frac{\lambda}{2\mu^2}$, and $s_T(t_0 + T_s) = \frac{\lambda}{2(t'_i + \epsilon + T_s)^2} - \frac{3}{2(t'_i + \epsilon + T_s)} - \frac{\lambda}{2\mu^2}$.

Too Underestimated Case :If $\epsilon < 0$ and $t'_i + \epsilon \leq 0$ for some t'_i, then $f_T(t'_i + \epsilon - T_s) = f_T(t'_i + \epsilon) = 0$ and $f_T(t'_i + \epsilon + T_s) > 0$. Therefore, the first non-zero term is $k = 2$ in (3.53). We approximate the score statistic by $\tilde{s}_{T^{\mathrm{ISI}}}(t'_i + \epsilon)$ with $w(t_0) = \frac{f_T(t'_i + \epsilon + T_s)}{f_{T^{\mathrm{ISI}}}(t'_i | \epsilon, T_s)}$, $s_T(t_0) =$
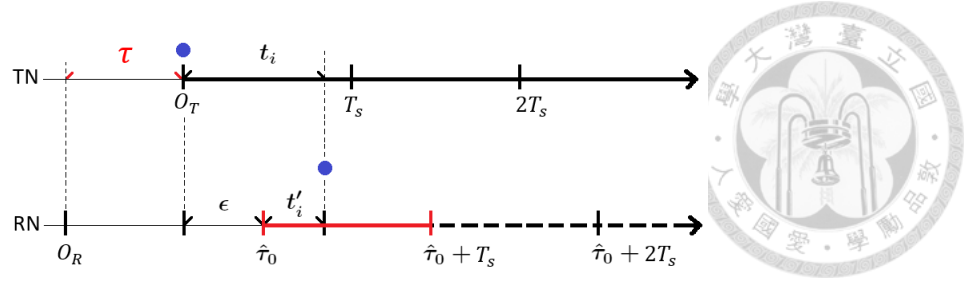
Figure 3.7: It is the normal case when $t_i' + \epsilon$ falls in the support $(0, T_s]$.

$\frac{\lambda}{2(t_i'+\epsilon+T_s)^2} - \frac{3}{2(t_i'+\epsilon+T_s)} - \frac{\lambda}{2\mu^2}$, and $s_T(t_0 + T_s) = \frac{\lambda}{2(t_i'+\epsilon+2T_s)^2} - \frac{3}{2(t_i'+\epsilon+2T_s)} - \frac{\lambda}{2\mu^2}$. Because the initial estimation is underestimated, it is possible that the arrival molecules actually was released from the previous symbol. The situation is significant, so our consideration have to include this part.
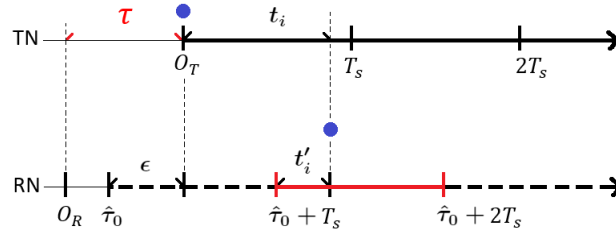
Figure 3.8: In too Underestimated case when $t_i' + \epsilon \leq 0$, the probability of crossover from the past symbol is significant.

Too Overestimated Case : If $0 < \epsilon$ and $T_s < t_i' + \epsilon$ for some t_i', then $f_T(t_i' + \epsilon - T_s) > 0$. Therefore, the first non-zero term is $k = 0$ in (3.53). We approximate the score statistic by $\tilde{s}_{T^{\mathrm{ISI}}}(t_i' + \epsilon)$ with $w(t_0) = \frac{f_T(t_i'+\epsilon-T_s)}{f_{T^{\mathrm{ISI}}}(t_i'|\epsilon, T_s)}$, $s_T(t_0) = \frac{\lambda}{2(t_i'+\epsilon-T_s)^2} - \frac{3}{2(t_i'+\epsilon-T_s)} - \frac{\lambda}{2\mu^2}$, and $s_T(t_0 + T_s) = \frac{\lambda}{2(t_i'+\epsilon)^2} - \frac{3}{2(t_i'+\epsilon)} - \frac{\lambda}{2\mu^2}$. Because the initial estimation is overestimated, it is possible that the arrival molecules actually was released from the next symbol. The situation is significant, so our consideration have to include this part.

Figure 3.9: In too overestimated case when $t_i' + \epsilon > T_s$, the probability of crossover from the future symbol is significant.

By Fisher's Score Algorithm (FSA), we iteratively update the estimated error $\hat{\epsilon}_{(k+1)}$

by

$$\hat{\epsilon}_{(k+1)} = \hat{\epsilon}_{(k)} + \frac{1}{nI_0(\epsilon_0)} \sum_{i=1}^{n} s_{T^{\text{ISI}}}(t'_i + \hat{\epsilon}_{(k)}), \qquad (3.61)$$

where $I_0(\epsilon_0)$ is the fisher information number defined by
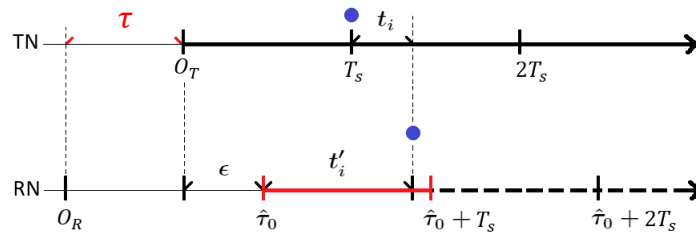
$$E[-\frac{\partial^2 \log f_{T^{\text{ISI}}}(T^{\text{ISI}}|\epsilon_0, T_s)}{\partial \epsilon^2}|\epsilon_0] \qquad (3.62)$$

$$= E[-\frac{\partial s_{T^{\text{ISI}}}(T^{\text{ISI}} + \epsilon_0)}{\partial \epsilon}|\epsilon_0], \qquad (3.63)$$

and ϵ_0 is the true error of the initial estimation.

The actual fisher information number of T^{ISI} is hard to derive, so we approximate it by $\tilde{I}_0$ which is in the case without ISI. When T_s is large enough with respect to $E[T] = \mu$, $f_T(t'_i + \epsilon + (k-1)T_s)$ decays quickly over k, so $w(t_0) \approx 1$ and

$$s_{T^{\text{ISI}}}(T^{\text{ISI}} + \epsilon_0) \approx s_T(T^{\text{ISI}} + \epsilon_0) = \frac{\lambda}{2(T^{\text{ISI}} + \epsilon_0)^2} - \frac{3}{2(T^{\text{ISI}} + \epsilon_0)} - \frac{\lambda}{2\mu^2}. \qquad (3.64)$$

Moreover, because the crossover effect is negligible, so the residue effect of $g(x)$ is also negligible, that is $g(x) \approx x$. Therefore, $T^{\text{ISI}} + \epsilon_0 = g(T - \epsilon_0) + \epsilon_0 \approx T - \epsilon_0 + \epsilon_0 = T$. This way, $\tilde{I}_0$ is simple as

$$I_0(\epsilon_0) \approx \tilde{I}_0 := E[-\frac{\partial s_T(T^{\text{ISI}} + \epsilon_0)}{\partial \epsilon}|\epsilon_0] = E[\frac{\lambda}{(T^{\text{ISI}} + \epsilon_0)^3} - \frac{3}{2(T^{\text{ISI}} + \epsilon_0)^2}] \qquad (3.65)$$

$$= E[\frac{\lambda}{T^3} - \frac{3}{2T^2}] = \frac{\lambda}{\mu^3} + \frac{9}{2\mu^2} + \frac{21}{2\mu\lambda} + \frac{21}{2\lambda^2}. \qquad (3.66)$$

The fisher information number is the value after taking expectation, so it only depends on the parameters of AIGN channel but the observations. Moreover, the fisher information number is always positive because it is the variance of score statistic.

Finally, the approximated FSA is

$$\hat{\epsilon}_{(k+1)} = \hat{\epsilon}_{(k)} + \frac{1}{n\tilde{I}_0} \sum_{i=1}^{n} \tilde{s}_{T^{\text{ISI}}}(t_i' + \hat{\epsilon}_{(k)}) \tag{3.67}$$

$$= \hat{\epsilon}_{(k)} + \frac{1}{n\tilde{I}_0} \sum_{i=1}^{n} \{w(t_0)s_T(t_0) + (1 - w(t_0))s_T(t_0 + T_s)\} . \tag{3.68}$$

The iteration keeps going until the absolute value of $\tilde{s}_{T^{\text{ISI}}}(t_i' + \epsilon)$ is small enough.

3.2.4 Simulation Results

In blind synchronization, we present the numerical MSE of LE with the first symbol, Decision Feedback (DF) method, and Fisher's Score Alogrithm (FSA) propsed above. Also, we verify our results by matching with theoretical curve and compare it with MSE of Pitman estimator in training-based methods as benchmark.

The parameters for M-ary quantity-based modulation is described as below; we set $L_j = \frac{(2j+1)n_1}{M}$ and $p_j = \frac{1}{M}$ so that $\sum_{j=0}^{M-1} p_j L_j = \bar{L} = n_1$ for $M = 2, 4, 8$. Moreover, when $M = 1$, the case reduce to training-based synchronization, which $n_1 = L_0$ is constant.
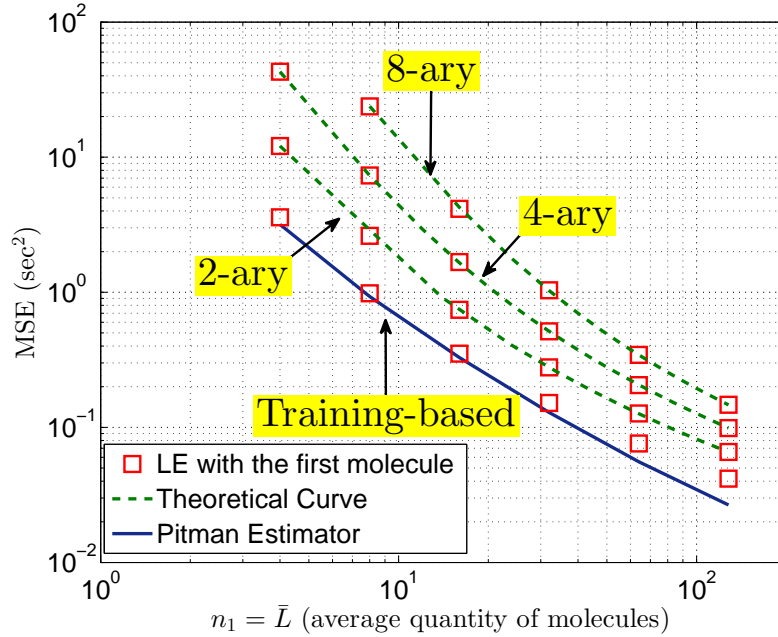

Figure 3.10: The MSE of LE with the first molecule and its theoretical curve in case when $K = 1$, $T_s = 3\mu$, where μ is the mean of T.

Fig. 3.10 shows that the performance of LE with the first molecule is worse when M

(a) Pitman Estimation　　　　(b) Initial Estimation
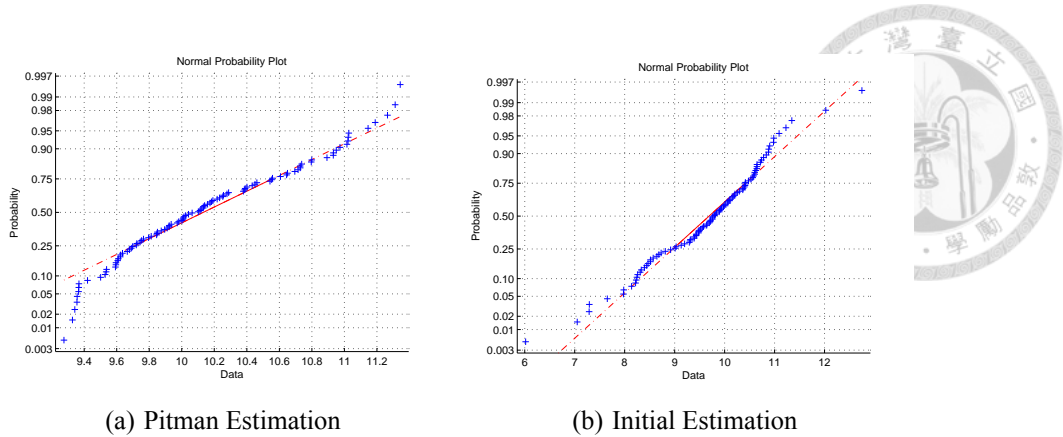
Figure 3.11: Normplot of Pitman Estimation and Initial Estimation of clock offset τ.

increases, which can be intuitively interpreted as the more random the message is, the harder we can estimate clock offset efficiently. For $M = 2$, 4, and 8, the numerical MSE matches the theoretical curve followed by (3.33). For $M = 1$, training-based synchronization performs better than blind synchronization.

In Fig. 3.11, we plot some estimations of clock offset τ and fit it with normal distribution. The estimation points are close to a line so that we use normal distribution to show the relationship of MSE and BER as below.

How the synchronization error affects the whole communication system evaluated by Bit Error Rate (BER)? Fig. 3.12 shows the BER with $L_0 = 32$ and $L_1 = 96$ versus synchronization error which follows normal distribution with mean from 0.4 to 1.2 and variance from 0.1 to 0.5. We set the demodulation threshold at the middle of two adjacent quantity levels, that is $\tilde{n}_k := \arg\min_{L_j} |L_j - l_1|$, where l_1 is the quantity of molecules in the range of $(\hat{\tau}_1 + (k-1)T_s, \hat{\tau}_1 + kT_s]$. Note that the minimal BER happens to the synchronization error follows normal distribution with mean near 0.9 and variance near 0 instead of the perfect synchronization with zero synchronization error. This is reasonable because a bit overestimation of clock offset τ is helpful due to the propagation delay. The optimal error ϵ_{op} satisfies $f_T(\epsilon_{\text{op}}) = f_T(\epsilon_{\text{op}} + T_s)$. Moreover, the effect of synchronization error to BER is not symmetric with respect to zero. This means that 1 second overestimation does not equivalent to 1 second underestimation.

As shown in Fig. 3.13, compared with Pitman Estimator, LE with the first molecule loses a bit MSE when the number of molecules $n_1 = \bar{L}$ is large. However, the inaccuracy
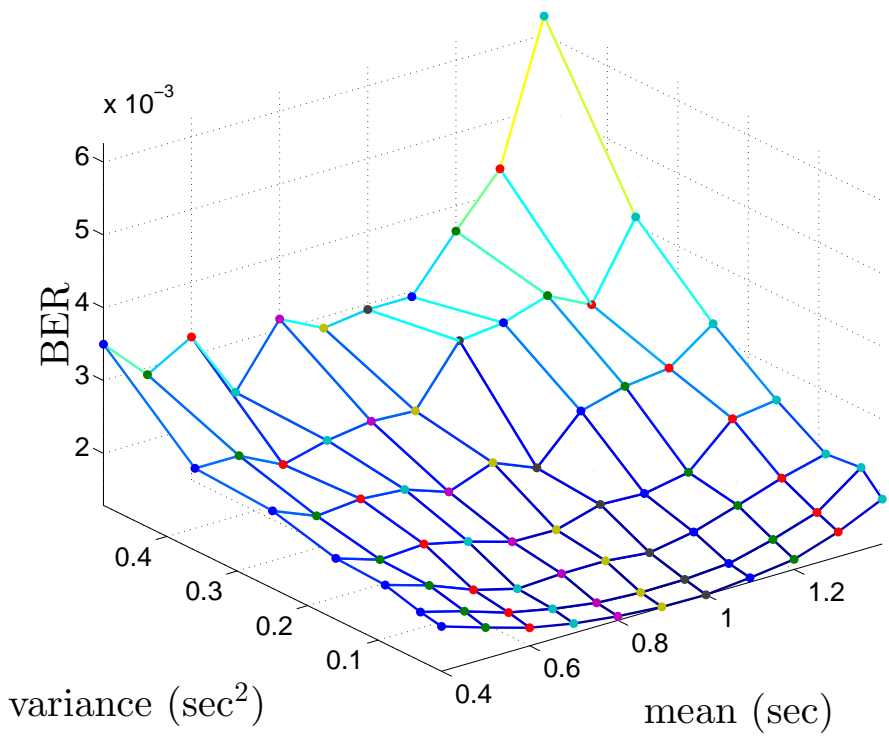
Figure 3.12: The BER versus the synchronization error which follows normal distribution with mean from 0.4 to 1.2 and variance from 0.1 to 0.5.
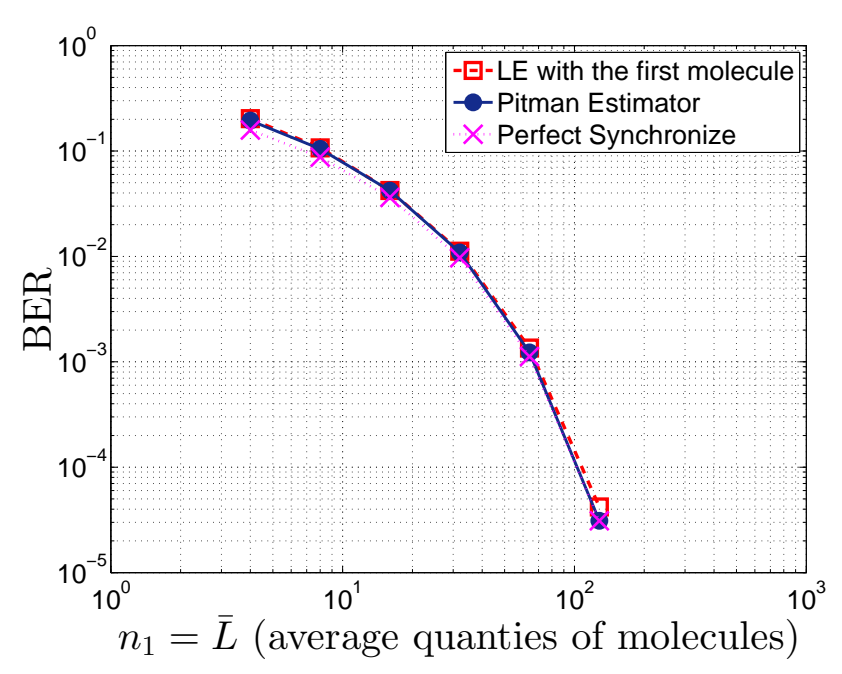
Figure 3.13: The BER of LE with the first molecule and Pitman Estimator with $T_s = \mu$.

of LE with the first molecule does not affect much BER loss for overall communication system. In this simulation, all estimation has been compensated by the optimal error ϵ_{op}.
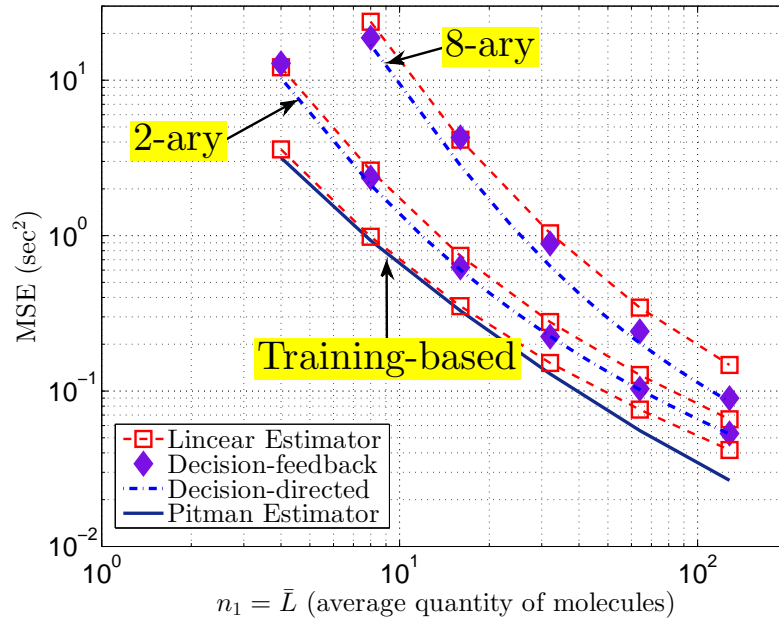
Figure 3.14: The MSE of Decision Feedback and theoretical Decision-directed Parameter Estimation in case when $K = 1, T_s = 3\mu$.
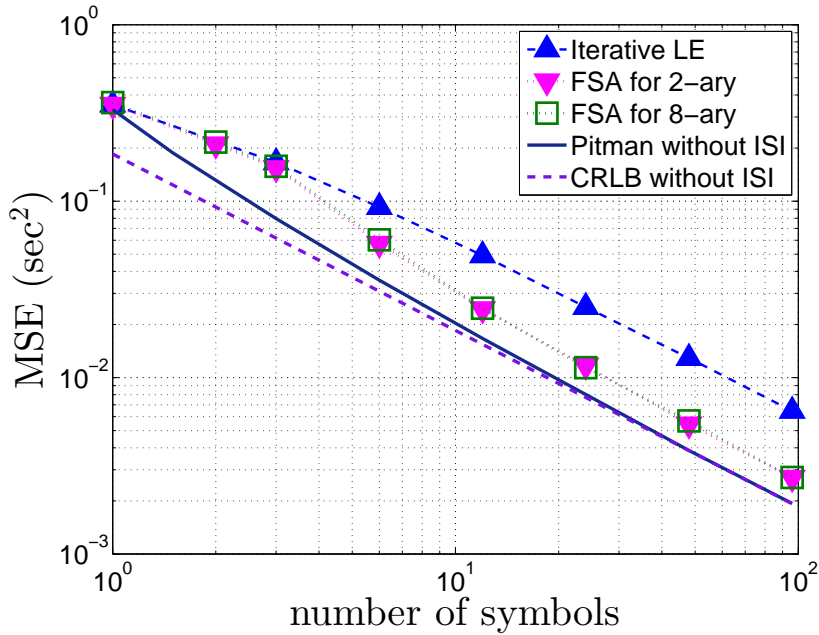
Figure 3.15: The MSE of FSA $\tau_{(K)}$ with M-ary versus the number of symbols K when $T_s = 3\mu$ and $n_1 = \bar{L} = 16$.

In Fig. 3.14, we use the same detection threshold to apply DF method. The numerical result shows that DF improves the MSE of LE with the first molecule and is closed to Decision-directed Parameter Estimation for this detection. Moreover, when M is large, the improvement of DF method becomes notable.

In Fig. 3.15, We compare Iterative LE (ILE) proposed in training-based synchronization with Fisher's Scoring Algorithm (FSA) proposed in blind synchronization. When number of symbols is 1, that is initial estimation, the MSE of ILE and FSA are almost the same. However, when we use the information in the following symbols with ISI and iteratively update clock offset estimation over K symbols, FSA is quite better than ILE. The MSE of FSA with ISI effect is close to Pitman estimator and CRLB without ISI. This means that FSA is close to the optimal estimator in the sense with minimum MSE. Moreover, the performance of FSA is independent on M for M-ary quantity-based modulation.

Chapter 4

Channel Estimation In Training-based Synchronization

In this chapter, we still focus on clock offset estimation problem in AIGN channel. We have proposed some approaches in both training-based synchronization and blind synchronization. Those approaches are applied only when RN knows perfect information about AIGN channel. In AIGN channel, the random delay T has two parameters, $\mu = \frac{d}{v}$ and $\lambda = \frac{d^2}{2D}$, which depend on the distance d, drift velocity v, and diffusion coefficient D of molecule. However, in reality, the distance d of RN apart from TN is unknown for RN. Some channel estimation approaches must apply before synchronization, such as [16] propose solutions about distance estimation in diffusion-based molecular communications.

Here, we want to release the assumption that RN knows the perfect information about channel, such as distance d. Moreover, instead of d, the target parameters used in synchronization are μ and λ. Therefore, we want to design some approaches such that RN can estimate the channel parameters μ and λ and clock offset τ simultaneously, base on the observed arrival time sequence $\mathbf{y}$.

4.1   Initial Estimation

In our system model, RN received symbols with ISI except for the first symbol. The estimations for the symbols with and without ISI are quite different, so we separate our

49

estimation into two parts. First, according to the first symbol without ISI, RN estimate all parameters with initial estimation, $\hat{\tau}_{(0)}$, $\hat{\mu}_{(0)}$, and $\hat{\lambda}_{(0)}$. Then, by the following symbols with ISI, RN update all estimations with $\hat{\tau}_{(k)}$, $\hat{\mu}_{(k)}$, and $\hat{\lambda}_{(k)}$ for $k > 1$. Note that the channel parameters μ and λ are nuisance parameters, our target is the MSE of $\hat{\tau}_{(k)}$ as small as possible over k.

Let y_1, y_2, ..., y_{n_1} be the arrival times of molecules in the first symbol. RN knows that TN transmits n_1 molecules in the first symbol. Assume the ISI effect from the second symbol is negligible, then y_1, y_2, ..., y_{n_1} are i.i.d follows the three parameters inverse gaussian distribution.

$$f_Y(y|\tau, \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi(y-\tau)^3}} \exp\{-\frac{\lambda(y-\tau-\mu)^2}{2\mu^2(y-\tau)}\} \mathbb{1}\{y > \tau\}. \tag{4.1}$$

Note that the support of three parameters inverse gaussian distribution depend on the location parameter τ, so this is not one of the exponential family distribution.

According to [21], the estimators of three parameters inverse gaussian distribution have been investigated. The MLE of τ, μ, and λ are computed by numerically applying Newton-Raphson iteration in [21]. Because of the computational complexity issue, we do not apply iteration in the initial estimation, so the initial estimator of τ is

$$\hat{\tau}_{(0)} = y_1 - \frac{(\bar{y} - y_1)^3}{2s^2 \log n}, \tag{4.2}$$

where $\bar{y} = \frac{1}{n_1} \sum_{i=1}^{n_1} y_i$ is the sample mean, and $s^2 = \frac{1}{n_1-1} \sum_{i=1}^{n_1} (y_i - \bar{y})^2$ is the sample variance. We leave the iteration of $\hat{\tau}_{(0)}$ to the following update with ISI symbols.

With the initial estimation $\hat{\tau}_{(0)}$, the MLE of inverse gaussian parameters μ and τ are well known as:

$$\hat{\mu}_{(0)} = \frac{1}{n_1} \sum_{i=1}^{n_1} (y_i - \hat{\tau}_{(0)}) = \bar{y} - \hat{\tau}_{(0)} \tag{4.3}$$

, and

$$\frac{1}{\hat{\lambda}_{(0)}} = \frac{1}{n_1} \sum_{i=1}^{n_1} \frac{1}{y_i - \hat{\tau}_{(0)}} - \frac{1}{\hat{\mu}_{(0)}}. \tag{4.4}$$

The sample mean is used to estimate the mean parameter μ. On the other hand, the sample moment with order -1 is used to estimate the shape parameter λ.

After receiving the first symbol, RN keeps updating all parameters by the information in the following symbols with ISI. The process is to iteratively update τ, μ, and λ. That is, RN fixes the previous estimation $\hat{\mu}_{(K-1)}$ and $\hat{\lambda}_{(K-1)}$ to update $\hat{\tau}_{(K)}$. Then, update $\hat{\mu}_{(K)}$ by the new estimation $\hat{\tau}_{(K)}$. Finally, update $\hat{\lambda}_{(K)}$ by the new estimation $\hat{\tau}_{(K)}$ and $\hat{\mu}_{(K)}$. After receiving the $(K+1)$-th symbol, RN keeps repeating the iterative update process. Our goal is to make the MSE of $\hat{\tau}_{(K)}$ as small as possible. Besides, we hope that the MSE converges to zero when K approaches infinity.

In (3.68), we have an approximated FSA to estimate the initial error $\epsilon_0 := \hat{\tau}_{(0)} - \tau$ when fixing the channel parameters μ and λ. In channel estimation, we still apply this method in our iterative update process by

$$\hat{\tau}_{(K)} = \hat{\tau}_{(0)} - \hat{\epsilon}_{(K)}, \tag{4.5}$$

where $\hat{\epsilon}_{(K)}$ is computed by (3.68) with the previous channel estimation $\hat{\mu}_{(K-1)}$ and $\hat{\lambda}_{(K-1)}$ for $K = 1, 2, \dots$. This way, the problem remains how to update $\hat{\mu}_{(K)}$ by $\hat{\tau}_{(K)}$ and how to update $\hat{\lambda}_{(K)}$ by $\hat{\tau}_{(K)}$ and $\hat{\mu}_{(K)}$.

4.2 Iterative update $\hat{\mu}_{(K)}$

Although the second or later symbols have the ISI effect, this effect only sorts the arrival times of molecules but the value of arrival times. Therefore, the summation over all arrival times $y_1, y_2, \dots, y_N$ is independent on the ISI effect, that is, $\sum_{i=1}^{N} y_i' = \sum_{i=1}^{N} (x_i + t_i)$ because $\mathbf{y}' = \text{sort}(\mathbf{x} + \mathbf{t})$. For this reason, the sample mean statistic is quite good against

the ISI effect, so we have

$$\mathbf{y}\mathbf{1}_N^\top = \mathbf{y}'\mathbf{1}_N^\top + N\tau = \mathbf{x}\mathbf{1}_N^\top + \mathbf{t}\mathbf{1}_N^\top + N\tau. \tag{4.6}$$

Then, divided by N to get the sample mean of all arrival times $\mathbf{y}$, that is,

$$\bar{y} - \tau = \bar{x} + \bar{t} \tag{4.7}$$

where $\bar{y} = \frac{1}{N}\mathbf{y}\mathbf{1}_N^\top$, and so are $\bar{x}, \bar{t}$. Moreover, we derive the relation between $\bar{x}$ and T_s,

$$\bar{x} = \frac{1}{N}\sum_{i=1}^N x_i = \frac{1}{N}\sum_{k=1}^K n_k(k-1)T_s = \bar{n}_K T_s, \tag{4.8}$$

where $\bar{n}_K = \frac{1}{N}\sum_{k=1}^K n_k(k-1)$ is constant. In channel estimation problem, we assume that RN knows the training sequence n_k, that is in training-based synchronization. If TN transmit a constant training sequence with $n_1 = n_2 = ... = n_K$, then $\bar{n}_K = \frac{K-1}{2}$.

Because the first hitting time t_i of every molecule is i.i.d. and follows generic Inverse Gaussian distribution, by Central Limit Theory,

$$\bar{t} = \frac{1}{N}\sum_{i=1}^N t_i \sim \mathcal{N}(E[T], \frac{Var[T]}{N}) = \mathcal{N}(\mu, \frac{\mu^3}{N\lambda}). \tag{4.9}$$

We use the sample mean $\bar{t}$ to estimate the mean parameter of AIGN channel μ. After receiving the K-th symbols with ISI, RN can estimate μ by $\hat{\mu}_{(K)}$ as

$$\hat{\mu}_{(K)} := \bar{y}_K - \hat{\tau}_{(K)} - \bar{n}_K T_s, \tag{4.10}$$

where $\bar{y}_K$ is the sample mean of all arrival times during K symbols.

4.3 Iterative update $\hat{\lambda}_{(K)}$

Unlike the mean parameter μ could be estimated by sample mean, the shape parameter λ is hard to estimate under ISI effect. In this section, we apply Maximum Conditional

Likelihood (MCL) to estimate λ under the initial error ϵ_0 is fixed.

Recall that we have derived the distribution of arrival times with ISI effect in (3.46). In channel estimation, the unknown parameters of this distribution include the channel information μ and λ, so

$$f_{T^{\text{ISI}}}(t'|\epsilon, T_s, \mu, \lambda) = \sum_{k=0}^{\infty} f_T(t' + \epsilon + (k-1)T_s|\mu, \lambda)\mathbb{1}\{0 < t' \leq T_s\} \tag{4.11}$$

where $f_T(t|\mu, \lambda)$ is the inverse gaussian distribution with two unknown parameters μ and λ.

We define the conditional likeilhood function of μ and λ under the initial error ϵ is fixed by

$$L(\mu, \lambda|\mathbf{t}', \epsilon) := \sum_{i=1}^{n} \log f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s, \mu, \lambda). \tag{4.12}$$

The MCL of μ and λ is a function of ϵ as

$$[\hat{\mu}(\epsilon), \hat{\lambda}(\epsilon)] := \arg\max_{\mu, \lambda} L(\mu, \lambda|\mathbf{t}', \epsilon). \tag{4.13}$$

To solve this maximization problem, let us take first order derivation to both parameters and try to find the simultaneous solution $[\hat{\mu}(\epsilon), \hat{\lambda}(\epsilon)]$,

$$\begin{cases} 0 = \sum_{i=1}^{n} \frac{\partial}{\partial \mu} \log f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s, \hat{\mu}(\epsilon), \hat{\lambda}(\epsilon)), \\ 0 = \sum_{i=1}^{n} \frac{\partial}{\partial \lambda} \log f_{T^{\text{ISI}}}(t_i'|\epsilon, T_s, \hat{\mu}(\epsilon), \hat{\lambda}(\epsilon)). \end{cases} \tag{4.14}$$

To simplify the notation, we denote $\hat{\mu}(\epsilon)$ and $\hat{\lambda}(\epsilon)$ by $\hat{\mu}$ and $\hat{\lambda}$, respectively. They are still the functions of the fix parameter ϵ.

The score function of inverse gaussian distribution with respect to μ is

$$
\begin{aligned}
\frac{\partial}{\partial \mu} \log f_T(t + \epsilon | \mu, \lambda) &= \frac{1}{f_T(t + \epsilon)} \frac{\partial f_T(t + \epsilon)}{\partial \mu} \\
&= \frac{1}{f_T(t + \epsilon)} \frac{\partial}{\partial \mu} \left\{ \sqrt{\frac{\lambda}{2\pi(t + \epsilon)^3}} \exp\left\{ -\frac{\lambda(t + \epsilon - \mu)^2}{2\mu^2(t + \epsilon)} \right\} \right\} \\
&= \frac{1}{f_T(t + \epsilon)} \left\{ \frac{\lambda(t + \epsilon)}{\mu^3} - \frac{\lambda}{\mu^2} \right\} f_T(t + \epsilon) \\
&= \frac{\lambda(t + \epsilon)}{\mu^3} - \frac{\lambda}{\mu^2}.
\end{aligned}
\tag{4.15}
$$

Similar to the derivation of the score function of ϵ in (3.53), the score function under ISI effect is a weighted mean of the score function of original inverse gaussian distribution. Accordingly, (4.15) is derived for the original score function of inverse gaussian distribution with respect to μ and substitute in (4.16).

$$
\begin{aligned}
0 &= \sum_{i=1}^{n} \frac{\partial}{\partial \mu} \log f_{T^{\text{ISI}}}(t_i' | \epsilon, T_s, \hat{\mu}, \hat{\lambda}) \\
&= \sum_{i=1}^{n} \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k - 1)T_s) \frac{\partial}{\partial \mu} \log f_T(t_i' + \epsilon + (k - 1)T_s | \hat{\mu}, \hat{\lambda}) \\
&= \sum_{i=1}^{n} \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k - 1)T_s) \left\{ \frac{\hat{\lambda}(t_i' + \epsilon + (k - 1)T_s)}{\hat{\mu}^3} - \frac{\hat{\lambda}}{\hat{\mu}^2} \right\}
\end{aligned}
\tag{4.16}
$$

In (4.16), eliminate the scale constant $\hat{\lambda}$ and $\hat{\mu}^3$ to get

$$
\hat{\mu}(\epsilon) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k - 1)T_s | \hat{\mu}(\epsilon), \hat{\lambda}(\epsilon))(t_i' + \epsilon + (k - 1)T_s).
\tag{4.17}
$$

Note that the weighted value $w(t_i' + \epsilon + (k - 1)T_s | \hat{\mu}(\epsilon), \hat{\lambda}(\epsilon))$ is proportional to $f_T(t_i' + \epsilon + (k - 1)T_s | \mu, \lambda)$ which depends on the channel parameters μ and λ.

Similarly, the same derivation is applied with respect to the shape parameter λ.

$$
\begin{aligned}
\frac{\partial}{\partial \lambda} \log f_T(t+\epsilon | \mu, \lambda) &= \frac{1}{f_T(t+\epsilon)} \frac{\partial f_T(t+\epsilon)}{\partial \lambda} \\
&= \frac{1}{f_T(t+\epsilon)} \frac{\partial}{\partial \lambda} \left\{ \sqrt{\frac{\lambda}{2\pi(t+\epsilon)^3}} \exp\left\{ -\frac{\lambda(t+\epsilon-\mu)^2}{2\mu^2(t+\epsilon)} \right\} \right\} \\
&= \frac{1}{f_T(t+\epsilon)} \left\{ \frac{1}{2\lambda} - \frac{t+\epsilon}{2\mu^2} + \frac{1}{\mu} - \frac{1}{2(t+\epsilon)} \right\} f_T(t+\epsilon) \\
&= \frac{1}{2\lambda} - \frac{t+\epsilon}{2\mu^2} + \frac{1}{\mu} - \frac{1}{2(t+\epsilon)}
\end{aligned}
\tag{4.18}
$$

The equation (4.18) is derived for the original score function of inverse gaussian distribution with respect to λ and substitute in (4.19).

$$
\begin{aligned}
0 &= \sum_{i=1}^{n} \frac{\partial}{\partial \lambda} \log f_{T^{\text{ISI}}}(t_i' | \epsilon, T_s, \hat{\mu}, \hat{\lambda}) \\
&= \sum_{i=1}^{n} \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k-1)T_s) \frac{\partial}{\partial \lambda} \log f_T(t_i' + \epsilon + (k-1)T_s | \hat{\mu}, \hat{\lambda}) \tag{4.19} \\
&= \sum_{i=1}^{n} \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k-1)T_s) \\
&\quad \left\{ \frac{1}{2\hat{\lambda}} - \frac{t_i' + \epsilon + (k-1)T_s}{2\hat{\mu}^2} + \frac{1}{\hat{\mu}} - \frac{1}{2(t_i' + \epsilon + (k-1)T_s)} \right\} \tag{4.20}
\end{aligned}
$$

There are four terms in (4.20). The first term is $\frac{n}{2\hat{\lambda}}$ and the third term is $\frac{n}{\hat{\mu}}$ because they are independent on k and i. Moreover, the second term includes the solution $\hat{\mu}(\epsilon)$ in (4.17). We try to solve the simultaneous solutions $[\hat{\mu}(\epsilon), \hat{\lambda}(\epsilon)]$, so we could substitute (4.17) in (4.20).

$$
\begin{aligned}
0 &= \frac{n}{2\hat{\lambda}} - \frac{n\hat{\mu}}{2\hat{\mu}^2} + \frac{n}{\hat{\mu}} - \sum_{i=1}^{n} \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k-1)T_s) \left\{ \frac{1}{2(t_i' + \epsilon + (k-1)T_s)} \right\} \\
&= \frac{n}{2\hat{\lambda}} + \frac{n}{2\hat{\mu}} - \frac{1}{2} \sum_{i=1}^{n} \sum_{k=0}^{\infty} \left\{ \frac{w(t_i' + \epsilon + (k-1)T_s)}{t_i' + \epsilon + (k-1)T_s} \right\} \tag{4.21}
\end{aligned}
$$

Then, eliminate the scale constant $\frac{1}{2}$ to get

$$\frac{1}{\hat{\lambda}(\epsilon)} = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=0}^{\infty} \left\{ \frac{w(t_i' + \epsilon + (k-1)T_s | \hat{\mu}(\epsilon), \hat{\lambda}(\epsilon))}{t_i' + \epsilon + (k-1)T_s} \right\} - \frac{1}{\hat{\mu}(\epsilon)} \tag{4.22}$$

$$\approx \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{w(t_0 | \hat{\mu}(\epsilon), \hat{\lambda}(\epsilon))}{t_0} + \frac{(1 - w(t_0 | \hat{\mu}(\epsilon), \hat{\lambda}(\epsilon)))}{t_0 + T_s} \right\} - \frac{1}{\hat{\mu}(\epsilon)}. \tag{4.23}$$

Therefore, the solution of MCL $[\hat{\mu}(\epsilon), \hat{\lambda}(\epsilon)]$ is in (4.17) and (4.23). Similar to MLE of μ and λ in original inverse gaussian distribution, $\hat{\mu}(\epsilon)$ is estimated by the sample mean and $\hat{\lambda}(\epsilon)$ is estimated by the sample moment with order -1. Moreover, because the observations y_i' is affected by the ISI effect, we need to take weighted average to all possible score value with the weighted value $w(t_i' + \epsilon + (k-1)T_s)$ proportional to the probability density $f_T(t_i' + \epsilon + (k-1)T_s)$.

The same with (3.60), the equation (4.23) is approximated by the first and second terms of summation over k. The approximation is reasonable because the crossover effect over two symbol durations is negligible, that is, $w(t_0 + 2T_s)$ and $w(t_0 + 3T_s)$... are dominated by $w(t_0)$ and $w(t_0 + T_s)$. Moreover, the score value $\frac{1}{t_0 + (k-1)T_s}$ decreases over k, so we only consider the first and second terms.

However, we cannot approximate (4.17) by the first and second term. The reason is even the weighted value concentrates on the first and second term, the score value increases over k. Moreover, $t_0 + (k-1)T_s$ approaches to infinity when k approaches infinity. Therefore, the value of $\hat{\mu}(\epsilon)$ is hard to evaluate.

$$\hat{\mu}(\epsilon) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=0}^{\infty} w(t_i' + \epsilon + (k-1)T_s | \hat{\mu}(\epsilon), \hat{\lambda}(\epsilon))(t_i' + \epsilon + (k-1)T_s)$$

$$\not\approx \frac{1}{n} \sum_{i=1}^{n} w(t_0)t_0 + w(t_0 + T_s)(t_0 + T_s). \tag{4.24}$$

Nevertheless, we still can estimated μ by sample mean as described in (4.10). For this reason, We replace the MCL of $\hat{\mu}$ in (4.23) by the sample mean estimator $\hat{\mu}_{(K)}$ in (4.10)

to iteratively update $\hat{\lambda}_{(K)}$.

$$\frac{1}{\hat{\lambda}_{(K)}} := \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{w(t_0|\hat{\mu}_{(K)}, \hat{\lambda}_{(K-1)})}{t_0} + \frac{(1 - w(t_0|\hat{\mu}_{(K)}, \hat{\lambda}_{(K-1)}))}{t_0 + T_s} \right\} - \frac{1}{\hat{\mu}_{(K)}}. \quad (4.25)$$

where t_0 uses the information of the new estimation $\epsilon_{(K)}$, and the weighted value uses the information of the previous estimation $\hat{\mu}_{(K)}$ and $\hat{\lambda}_{(K-1)}$.

To sum up, the whole channel estimation process includes three initial estimations and three iterative update estimations. The initial estimations of τ, μ, and λ are $\hat{\tau}_{(0)}$, $\hat{\mu}_{(0)}$, and $\hat{\lambda}_{(0)}$, respectively, as described in (4.2), (4.3), and (4.4). The iterative update estimation of τ, μ, and λ are $\hat{\tau}_{(K)}$, $\hat{\mu}_{(K)}$, and $\hat{\lambda}_{(K)}$, respectively, as described in (4.5), (4.10), and (4.25).

4.4   Lower Bound with unknown μ and λ

Different from Sec. 3.1.4, the channel information μ and λ are unknown for channel estimation problem. Now, we have three unknown parameters $\theta = [\tau, \mu, \lambda]$ to be estimated.

Similar with Sec. 3.1.4, let us derive the Cramer-Rao Lower Bound of τ under μ and λ are unknown. Without crossover effect, all the arrival times of molecules $\mathbf{y}$ are a random sample from a generic distribution $f_T(y_i - \tau|\mu, \lambda)$ with the sample size N. For this three

parameters model $\theta = [\tau, \mu, \lambda]$, the fisher information matrix is

$$\mathbf{I}_0 := -E\left\{\begin{bmatrix} \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \tau^2} & \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \tau \partial \mu} & \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \tau \partial \lambda} \\ \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \mu \partial \tau} & \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \mu^2} & \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \mu \partial \lambda} \\ \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \lambda \partial \tau} & \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \lambda \partial \mu} & \frac{\partial^2 \log f_T(y-\tau|\mu,\lambda)}{\partial \lambda^2} \end{bmatrix}\right\} \qquad (4.26)$$

$$= -E\left\{\begin{bmatrix} \frac{3}{2T^2} - \frac{\lambda}{T^3} & \frac{\lambda}{\mu^3} & \frac{1}{2T^2} - \frac{1}{2\mu^2} \\ \frac{\lambda}{\mu^3} & -\frac{3\lambda T}{\mu^4} + \frac{2\lambda}{\mu^3} & \frac{T}{\mu^3} - \frac{1}{\mu^2} \\ \frac{1}{2T^2} - \frac{1}{2\mu^2} & \frac{T}{\mu^3} - \frac{1}{\mu^2} & -\frac{1}{2\lambda^2} \end{bmatrix}\right\} \qquad (4.27)$$

$$= \begin{bmatrix} \tilde{I}_0 & -\frac{\lambda}{\mu^3} & -\frac{3}{2\mu\lambda} - \frac{3}{2\lambda^2} \\ -\frac{\lambda}{\mu^3} & \frac{\lambda}{\mu^3} & 0 \\ -\frac{3}{2\mu\lambda} - \frac{3}{2\lambda^2} & 0 & \frac{1}{2\lambda^2} \end{bmatrix} \qquad (4.28)$$

The inverse of the fisher information matrix for $\theta = [\tau, \mu, \lambda]$ is

$$\mathbf{I}_0^{-1} = \frac{1}{\det(\mathbf{I}_0)} \begin{bmatrix} \frac{1}{2\mu^3\lambda} & \frac{1}{2\mu^3\lambda} & \frac{3}{2\mu^4} + \frac{3}{2\lambda\mu^3} \\ \frac{1}{2\mu^3\lambda} & \frac{1}{2\mu^3\lambda} + \frac{3}{4\mu\lambda^3} + \frac{3}{\lambda^4} & \frac{3}{2\mu^4} + \frac{3}{2\lambda\mu^3} \\ \frac{3}{2\mu^4} + \frac{3}{2\lambda\mu^3} & \frac{3}{2\mu^4} + \frac{3}{2\lambda\mu^3} & (\tilde{I}_0 - \frac{\lambda}{\mu^3})\frac{\lambda}{\mu^3} \end{bmatrix}, \qquad (4.29)$$

where $\det(\mathbf{I}_0) = \frac{1}{2\mu^3\lambda}\left(\frac{3}{2\mu\lambda} + \frac{6}{\lambda^2}\right)$. By Cramer-Rao inequality, for any unbiased estimator $\hat{\theta} = [\hat{\tau}, \hat{\mu}, \hat{\lambda}]$, its covariance matrix has the property that $Cov[\hat{\theta}] - \mathbf{I}_0^{-1}$ is a positive semidefinite matrix.

$$\mathbf{x}[Cov[\hat{\theta}] - \mathbf{I}_0^{-1}]\mathbf{x}^\top \geq 0, \qquad (4.30)$$

for any three dimensional vector $\mathbf{x} \in \mathbb{R}^3$. For $\mathbf{x} = [1, 0, 0]$, we get the lower bound of the variance $\hat{\tau}$ with unknown μ and λ as

$$\frac{1}{N} \frac{1}{\det(\mathbf{I}_0)} \frac{1}{2\mu^3\lambda} = \frac{1}{N\left(\frac{3}{2\mu\lambda} + \frac{6}{\lambda^2}\right)} \qquad (4.31)$$

The CRLB of clock offset τ with unknown μ and λ is in (4.31), but we still cannot claim the lower bound property because the support depends on the parameter τ. Moreover, the
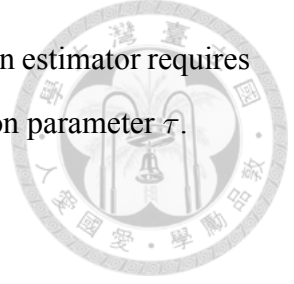
Pitman estimator cannot applied because μ and λ are unknown. Pitman estimator requires the information of whole distribution $f_T(t|\mu, \lambda)$ except for the location parameter τ.

4.5 Simulation Results

In this section, we simulate the MSE of some methods proposed above and plot lower bound to evaluate the performance of our estimators. The parameters of AIGN channel are the same with Sec. 3.1.5.

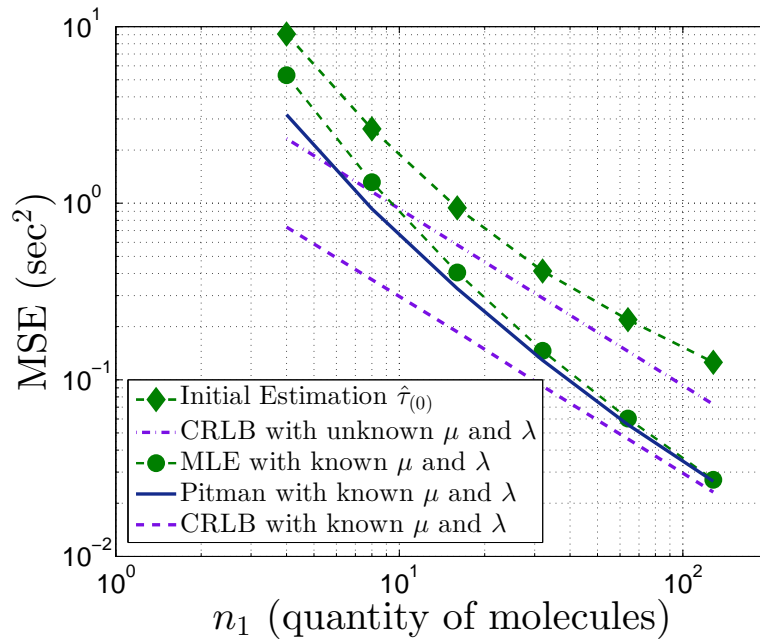Figure 4.1: Initial Estimation $\hat{\tau}_{(0)}$ and CRLB with unknown μ and λ.

In Fig. 4.1, the MSE of our initial estimator for clock offset τ as in (4.2) is worse than the MLE proposed in Sec. 3.1.1. This is reasonable because the latter use more channel information than the former. Moreover, the MSE of our initial estimator for clock offset τ is close to the CRLB with unknown channel information μ and λ as derived in (4.31).

In Fig. 4.2, the MSE of FSA with updating $\mu_{(K)}$ and $\lambda_{(K)}$ proposed above has been plotted. When we only use the first symbol ,that is the number of symbols $K = 1$, the MSE is the initial estimator proposed in (4.2). Moreover, when we use the information in the following symbols with ISI effect and iteratively update our estimator of clock offset τ and channel information μ and λ, respectively, proposed in (4.10) and (4.25), the MSE
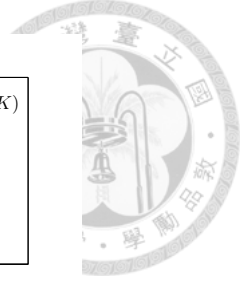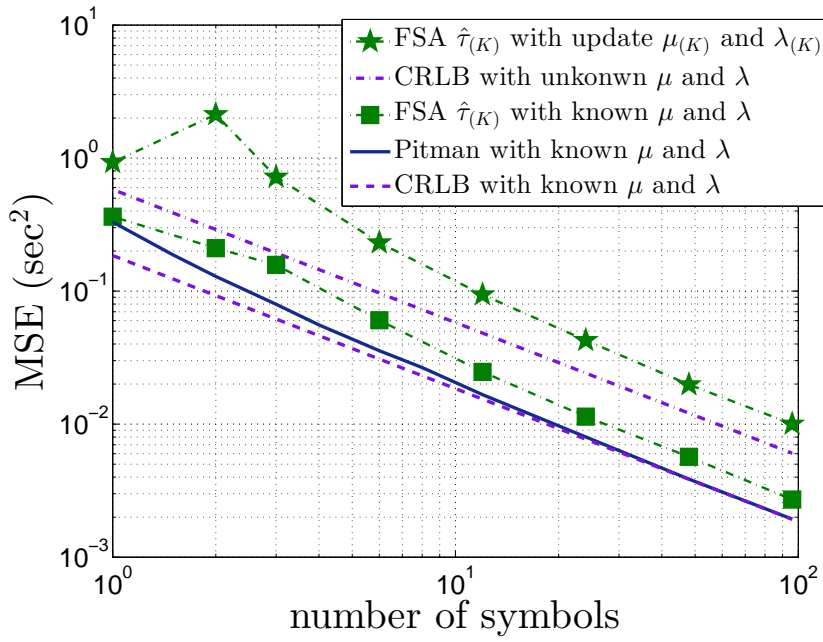
Figure 4.2: The MSE of FSA $\tau_{(K)}$ with update $\mu_{(K)}$ and $\lambda_{(K)}$ versus the number of symbols K when $T_s = 3\mu$ and $n_1 = \bar{L} = 16$.

improves over the number of symbols K except for the case when K is small. The reason why the updated clock offset when K is small worse than $K = 1$ is that the sample size is not enough for FSA to find a quite good statistical estimation. Besides, compare with FSA with known μ and λ proposed in Sec. 3.2.3, we lose a bit accuracy than before. Nevertheless, its MSE is close to the CRLB with unknown μ and λ as derived in (4.31).

Chapter 5

Clock Skew Estimation With And Without Channel Information

In this chapter, we consider a more tough problem with clock skew effect. In Fig. 2.6, the sampling time is $\hat{S}_k = \hat{R}[\hat{\tau} + (k-1)T_s]$. Before this chapter, we assume the clock skew rate R is known and fix, so $\hat{R}(\mathbf{y})$ is correct without any error. This way, the error of all starting point of symbols $\hat{S}_k$ are the same and equal to the error of clock offset estimation $\hat{S}_1 = \hat{\tau}$. However, in real scenarios, RN's clock may tick in a different rate from TN's, that's called *clock drift* or *clock skew*. Under clock skew effect, RN needs to adapt its symbol duration T_s corresponding to its own clock skew rate R so that they can count symbols on the same time duration. How to estimate $\hat{R}(\mathbf{y})$ from the pattern in the observation signal $\mathbf{y}$ received by RN is still a problem. We deal with this problem in this chapter.

5.1 Training-based Synchronization With Perfect Channel Information

For a solid communication system, the effect of clock skew is more severe than that of clock offset because the error of clock skew would accumulate over time. If we don't keep tracing the starting points of every symbol, the accumulated error will make our system

gradually become non-synchronous, as shown in Fig. 2.4. A good methodology to solve this problem should make the error of sampling time S_k converges to zero over time, that is, we can make RN's time slotted system locked with TN's time slotted system in steady state.

Consider the case without clock offset error, then the MSE of sampling time S_k becomes

$$E[(\hat{S}_k - S_k)^2] = E[(\hat{R}[\tau + (k-1)T_s] - R[\tau + (k-1)T_s])^2]$$
$$= [\tau + (k-1)T_s]^2 E[(\hat{R}(\mathbf{y}) - R)^2] \qquad (5.1)$$

The equation (5.1) presents that the error of clock skew rete R will cause its $[\tau + (k-1)T_s]^2$ times error of sampling time and the factor increases over time. This is the effect of accumulated error of clock skew. To conquer this effect, we require a much efficient estimator to reduce the error over $[\tau + (k-1)T_s]^2$ times. Otherwise, it is impossible that the error of sampling time S_k converges to zero over time

For clock skew estimation, we only consider the training-based synchronization and the training sequence is constant for all symbols. That is, TN transmits total K symbols and each symbol has the same quantity of molecules, so the training sequence is $n = n_1 = n_2 = \cdots = n_K$. Then, the arrival timing sequence of molecules $\mathbf{y}$ observed by RN has some periodic patterns of cycle n. We can use the information of the period to estimate the unknown clock skew rate R

5.1.1 Iterative Linear Estimation For Clock Skew Rate

In Sec. 3.1.3, we have derived two kind of weight, $\mathbf{a}_n$ without ISI and $\mathbf{w}_n$ with ISI. For the second and later symbols with ISI, we use the wighted value $\mathbf{w}_n$ and mean value $\mathbf{m}_n$ to estimate the clock offset τ as (3.19). However, In (3.19), we assume the clock skew rate R is known and has been compensated in $\mathbf{y}^{[1:n]}$. However, in the situation with unknown

clock skew rate R, (3.19) should be rewritten as

$$\hat{\tau}_{\text{new}}^{(k+1)} = \mathbf{w}_n\Big(\frac{\mathbf{y}^{[kn+1:(k+1)n]}}{R} - \mathbf{m}_n\Big)^\top - kT_s \tag{5.2}$$

In Sec. 3.1.3, we estimate the unknown clock offset τ with correct clock skew rate. Reversely, we could estimate the unknown clock skew rate R with correct clock offset τ as

$$\tau = \mathbf{w}_n\Big(\frac{\mathbf{y}^{[kn+1:(k+1)n]}}{\hat{R}} - \mathbf{m}_n\Big)^\top - kT_s, \tag{5.3}$$

where $1 \leq k \leq K - 1$. For $k = 1$,

$$\tau = \mathbf{w}_n\Big(\frac{\mathbf{y}^{[n+1:2n]}}{\hat{R}} - \mathbf{m}_n\Big)^\top - T_s. \tag{5.4}$$

Let us subtract (5.4) from (5.3), then we get

$$(k - 1)T_s = \frac{\mathbf{w}_n\big(\mathbf{y}^{[kn+1:(k+1)n]} - \mathbf{y}^{[n+1:2n]}\big)^\top}{\hat{R}}. \tag{5.5}$$

The notation $\mathbf{y}^{[kn+1:(k+1)n]}$ is the $(k + 1)$-th symbol received by RN. We want to derive the estimator of R for the k-th symbol, so we replace $(k + 1)$ with k in (5.5).Then, we can get an estimator of R by the n molecules in the k-th symbol received by RN.

$$\hat{R}^{(k)} := \frac{\mathbf{w}_n\big(\mathbf{y}^{[(k-1)n+1:kn]} - \mathbf{y}^{[n+1:2n]}\big)^\top}{(k - 2)T_s}, \tag{5.6}$$

where $3 \leq k \leq K$. In (5.6), $\hat{R}^{(2)} = \frac{0}{0}$ has no definition, so k begins with 3. This way, we lost the information in the first symbol $\mathbf{y}^{[1:n]}$ because the first symbol without ISI use different weighted values $\mathbf{a}_n$ from the other symbols with ISI. To include the information in the first symbol $\mathbf{y}^{[1:n]}$, we rewrote the form (5.6) to

$$\hat{R}^{(k)} \approx \frac{\mathbf{w}_n\big(\mathbf{y}^{[(k-1)n+1:kn]}\big)^\top - \mathbf{a}_n\big(\mathbf{y}^{[1:n]}\big)^\top}{(k - 1)T_s}. \tag{5.7}$$

where $2 \leq k \leq K$. This way, our estimator $\hat{R}^{(k)}$ begins with $k = 2$. Even though the first symbol $\mathbf{a}_n(\mathbf{y}^{[1:n]})^\top$ has a bit error with $\mathbf{w}_n(\mathbf{y}^{[n+1:2n]})^\top + T_s$, the error is divided by $k - 1$ and is eliminated gradually.
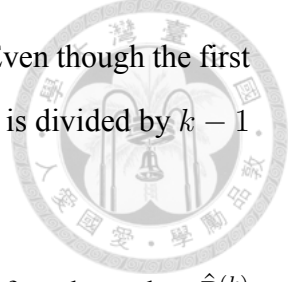
For every symbol, $2 \leq k \leq K$, we use (5.7) to estimate R. We found out that $\hat{R}^{(k)}$ is more accurate as k increases because the estimator divided by $k - 1$ can reduce the variance. Specifically,

$$Var[\hat{R}^{(k)}] = \frac{Var[\hat{R}^{(2)}]}{(k-1)^2}. \tag{5.8}$$

Accordingly, when using overall estimators by all symbols, we give more weighted value on the latter estimator.

$$\hat{R}_{\text{LE}}^{(K)} := \frac{1\hat{R}^{(2)} + 4\hat{R}^{(3)} + ... + (K-1)^2 \hat{R}^{(K)}}{\frac{(K-1)K(2K-1)}{6}} \tag{5.9}$$

We call (5.9) Linear Estimation (LE) for the clock skew rate R. Actually, $\hat{R}_{\text{LE}}^{(K)}$ is a linear combination of all arrival times of molecules $y_1, y_2, ..., y_N$ with some coefficients.

To reduce nano-machine's computational complexity, we rewrite LE for the clock skew rate to iterative form as

$$
\begin{aligned}
\hat{R}_{\text{LE}}^{(2)}(\mathbf{y}) &= \hat{R}^{(2)} \\
&= \frac{\mathbf{w}_n(\mathbf{y}^{[n+1:2n]})^\top - \mathbf{a}_n(\mathbf{y}^{[1:n]})^\top}{T_s} \\
\hat{R}_{\text{LE}}^{(k)}(\mathbf{y}) &= \left[1 - \frac{6(k-1)}{k(2k-1)}\right]\hat{R}_{\text{LE}}^{(k-1)}(\mathbf{y}) + \frac{6(k-1)}{k(2k-1)}\hat{R}^{(k)} \\
&= \left[1 - \frac{6(k-1)}{k(2k-1)}\right]\hat{R}_{\text{LE}}^{(k-1)}(\mathbf{y}) \\
&\quad + \frac{6\left[\mathbf{w}_n(\mathbf{y}^{[(k-1)n+1:kn]})^\top - \mathbf{a}_n(\mathbf{y}^{[1:n]})^\top\right]}{k(2k-1)T_s},
\end{aligned}
$$

$$ (5.10) $$
$$ (5.11) $$

where $2 < k \leq K$, and $\hat{R}_{\text{LE}}^{(k-1)}(\mathbf{y})$ is the previous estimator only using the previous $(k-1)$ symbols to estimate R.

5.2 Training-based Synchronization Without Channel Information

In this section, we discuss on joint estimation. Both clock offset τ and clock skew rate R are unknown and are considered in our channel model as in Fig.2.6 without channel information μ and λ.

Remind that (2.3) describes the observations by RN in our stochastic model. Actually, y_i includes three terms, $R\tau$, $RT_s k_j$, and Rt_i as

$$y_i = R\tau + RT_s k_j + Rt_i, \tag{5.12}$$

where R and τ are unknown and our target parameters. The additive first hitting time t_i follows inverse gaussian distribution with channel parameters μ and λ, which are both unknown. The symbol duration T_s is known and shared with TN before communication and y_i is released at the k_j-th symbols. Because of the ISI effect, RN actually is not sure what k_j is. However, ISI effect only affects the arrival times of molecules at the margin of symbol duration. Most arrival times of molecules are without ISI effect, so we actually can estimate k_j by RN's time slotted system with a bit detection error.

Compare our stochastic model in (5.12) with the classical linear regression model. The sample y_i with covarites x_i follow the stochastic model as

$$y_i = \beta_{00} + \beta_{01} x_i + \epsilon_i, \tag{5.13}$$

for $1 \leq i \leq N$ with the sample size N. The error term $\epsilon_i \sim$ i.i.d. $\mathcal{N}(0, \sigma^2)$ with zero mean and unknown but constant variance σ^2. The linear model $y = \beta_{00} + \beta_{01} x$ with the intercept β_{00} and slope β_{01}. The estimator $\hat{\beta} = [\hat{\beta}_{00}, \hat{\beta}_{01}]^\top$ can be derived by Least Square (LS) criterion as

$$\hat{\beta} = (Z^\top Z)^{-1} Z^\top Y \tag{5.14}$$

where $Z = [Z_1, Z_2, ..., Z_n]^\top$, $Z_i = [1, x_i]^\top$ and $Y = [y_1, y_2, ..., y_n]^\top$. The model in (5.13) is similar to our model in (5.12). Let the unknown intercept β_{00} is $R\tau$, the unknown slope β_{01} is RT_s and the error term ϵ_i is Rt_i, respectively.
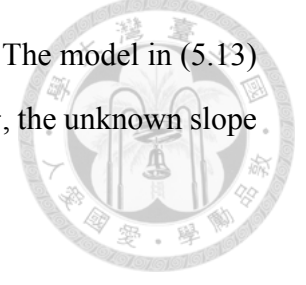
However, tow stochastic models have something different. First, the pairs (y_i, x_i) are clear in classical linear regression model, but (y_i, k_j) are not clear in our model because of the crossover effect. Nevertheless, with constant training sequence , RN knows that TN transmit n molecules each symbol. RN can estimate k_j by the quotient of i divided by n because the arrival time sequence **y** has a cycle with n molecules. Second, the error term Rt_i does not follow normal distribution with zero mean but the scaled inverse gaussian distribution with parameters R, μ, and λ as

$$Rt_i \sim \frac{1}{R} f_T(\frac{t}{R}|\mu, \lambda). \tag{5.15}$$

Nevertheless, generalized linear regression model can deal with the model with error terms follows a more general distribution. By quasi-likelihood estimator, we only need the mean structure and variance structure of our model, then the estimator $\hat{\beta}$ can be derived as

$$0 = \sum_{i=1}^{n} \frac{\partial \mu(Z_i, \hat{\beta})}{\partial \beta} \frac{y_i - \mu(Z_i, \hat{\beta})}{v(\mu(Z_i, \hat{\beta}))}, \tag{5.16}$$

where $\mu(Z_i, \hat{\beta})$ is the mean structure and $v(\mu(Z_i, \hat{\beta}))$ is the variance structure. In our model, we have linear mean structure $\mu(Z_i, \hat{\beta}) = Z_i\hat{\beta}$ and constant variance structure $v(\mu(Z_i, \hat{\beta})) = Var[Rt_i] = \frac{R^2\mu^3}{\lambda}$. Then, the quasi-likelihood estimator $\hat{\beta}$ in our model reduces to (5.14) because

$$0 = s(\hat{\beta}) = \sum_{i=1}^{n} \frac{\partial \mu(Z_i, \hat{\beta})}{\partial \beta} \frac{y_i - \mu(Z_i, \hat{\beta})}{v(\mu(Z_i, \hat{\beta}))}$$
$$= \sum_{i=1}^{n} Z_i^\top \frac{y_i - Z_i^\top \hat{\beta}}{Var[Rt_i]} = Var[Rt_i]^{-1} Z^\top(Y - Z\hat{\beta}). \tag{5.17}$$

Note that $E[Rt_i] = R\mu \neq 0$. The non-zero mean property causes the estimated intercept

$\hat{\beta}_{00}$ to be shifted. Nevertheless, we still can estimate clock skew rate by

$$\hat{R}_q := \frac{\hat{\beta}_{01}}{T_s}, \tag{5.18}$$

because the symbol duration T_s is known and the slope $\hat{\beta}_{01}$ is derived in (5.16). We call (5.18) by Quasi-likelihood clock skew estimator and denoted by $\hat{R}_q$

By the asymptotic efficiency of MLE, we know that $Var[\hat{\beta}_{01}]$ approaches to Cramer Rao Lower Bound. Let us derive the fisher information matrix of $\beta = [\beta_{00}, \beta_{01}]^\top$ in this model as

$$
\begin{aligned}
F(\beta) = E[-\frac{\partial s(\beta)}{\partial \beta}] &= Z^\top Var[Rt_i]^{-1} Z \\
&= \frac{1}{Var[Rt_i]} \begin{bmatrix} N & \sum_{i=1}^{N} k_i \\ \sum_{i=1}^{N} k_i & \sum_{i=1}^{N} k_i^2 \end{bmatrix} \\
&= \frac{N\lambda}{R^2 \mu^3} \begin{bmatrix} 1 & \bar{k} \\ \bar{k} & \bar{k}_2 \end{bmatrix},
\end{aligned}
\tag{5.19}
$$

where $\bar{k} = \frac{1}{N} \sum_{i=1}^{N} k_i$ and $\bar{k}_2 = \frac{1}{N} \sum_{i=1}^{N} k_i^2$. For $\mathbf{k}_N = [0\mathbf{1}_n, 1\mathbf{1}_n, ..., (K-1)\mathbf{1}_n]$ and $N = Kn$, the inverse matrix of $F(\beta)$ is

$$
\begin{aligned}
F(\beta)^{-1} &= \frac{R^2 \mu^3}{N\lambda(\bar{k}_2 - \bar{k}^2)} \begin{bmatrix} \bar{k}_2 & -\bar{k} \\ -\bar{k} & 1 \end{bmatrix} \\
&= \frac{\mu^3 R^2}{\lambda n K (K-1)(K+1)} \begin{bmatrix} 2(K-1)(2K-1) & -6(K-1) \\ -6(K-1) & 12 \end{bmatrix}.
\end{aligned}
\tag{5.20}
$$

The variance of $\hat{R}_q$ equals $\frac{Var[\hat{\beta}_{01}]}{T_s^2}$, which approaches

$$[0,1]F(\beta)^{-1}[0,1]^\top = \frac{12\mu^3 R^2}{\lambda n K (K-1)(K+1)T_s^2}. \tag{5.21}$$

The equation (5.21) is an asymptotic MSE of $\hat{R}_q$ by the asymptotic efficiency property of MLE.

5.3 Simulation Results

In this section, we simulate the MSE of clock skew rate R for Quasi-likelihood estimator and Iterative LE proposed above. The parameters of AIGN channel are the same with Sec. 3.1.5.

In Fig. 5.1, the MSE of Quasi-likelihood clock skew estimator $\hat{R}_q$ approaches to the asymptotic variance as in (5.21). The reason why the simulated MSE is lower than CRLB is that the error term in our stochastic model Rt_i actually affected by ISI effect. If a realization of error Rt_i is larger than the symbol duration T_s, it causes ISI effect. Therefore, the actually variance of Rt_i is smaller than $R^2 Var[T]$ because the data affected by ISI effect is like censored data.
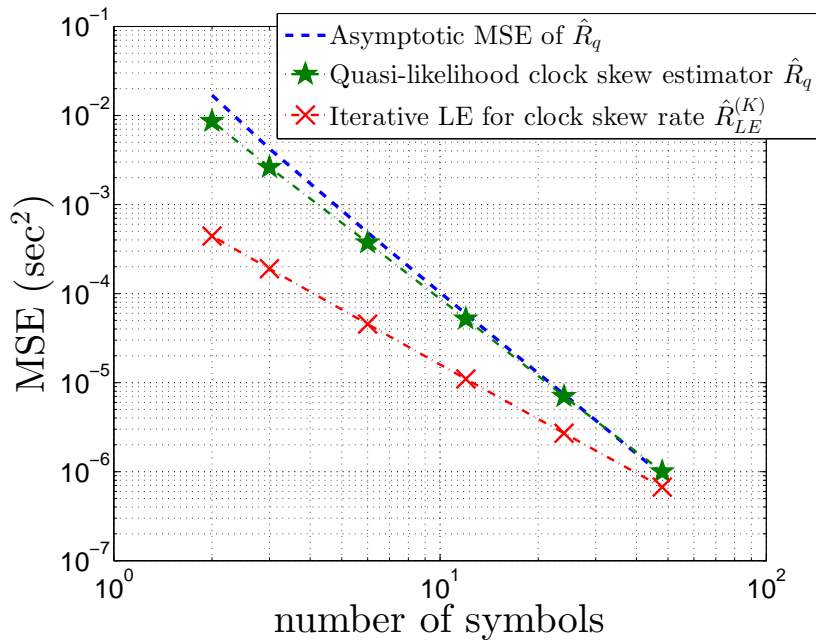

Figure 5.1: The MSE of clock skew rate R for Quasi-likelihood estimator and Iterative LE with $R = 1$, $T_s = 3\mu$ and $n_1 = \bar{L} = 16$.

On the other hand, the MSE of Iterative LE for clock skew rate $\hat{R}_{LE}^{(K)}$ is lower than that of Quasi-likelihood clock skew estimator $\hat{R}_q$. This is reasonable because the former uses more channel information than the latter. However, the MSE of the latter decreases more quickly than that of the former. We expect that Quasi-likelihood clock skew estimator $\hat{R}_q$ becomes better when the number of symbols is large enough.
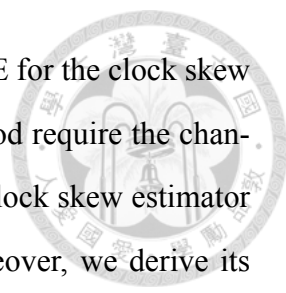
Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this work, as far as we know, we first discuss on the timing synchronization problem for quantity-based modulation in Additive Inverse Gaussian Channel.

In training-based synchronization for clock offset estimation problem, we have proposed Iterative LE, whose computational complexity is much lower than LE and MLE. Moreover, its MSE reaches almost the same efficient level with LE and MLE.

In blind synchronization for clock offset estimation problem, we compare the theoretical MSE of LE with that of DF, and give a sufficient condition when the latter improves the former. Besides the initial estimator, we proposed FSA to update the clock offset estimation under ISI effect. The MSE of FSA is close to the optimal location parameter estimator, Pitman estimator. Accordingly, our FSA is nearly optimal in the minimum MSE sense. Moreover, FSA is suitable in blind synchronization and the accuracy is independent on M for M-ary quantity-based modulation.

In channel estimation, we take the clock offset estimator without channel information μ and λ proposed in [21] as our initial estimator. The MSE of initial estimator is close to the CRLB with unknown μ and λ. Moreover, we use the MLE of μ and λ as our initial channel estimation. Then, the method updating channel estimation $\hat{\mu}_{(K)}$ and $\hat{\lambda}_{(K)}$ with ISI symbols was proposed. Combining with FSA proposed in Sec. 3.2.3, its MSE is close to the CRLB with unknown μ and λ.

When considering the clock skew effect, we proposed Iterative LE for the clock skew rate R and its complexity is linear time on N. However, these method require the channel information. In joint estimation, we proposed Quasi-likelihood clock skew estimator to estimate the clock skew rate without channel information. Moreover, we derive its asymptotic MSE. When the number of symbols are large enough, the Quasi-likelihood clock skew estimator is better than Iterative LE for clock skew rate.

6.2   Future work

The next question we face is how accurate we need to estimate the clock offset and the clock skew. To answer this question, the Bit Error Rate (BER) has to be considered for future work. This analysis depends on the whole modulation and detection scheme we choose, which is more complicated and difficult to extent to general situation.
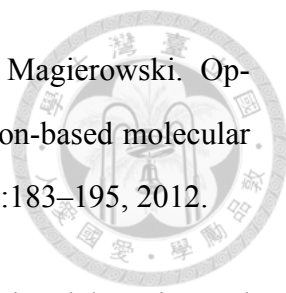
In our thesis, we use MSE as criterion to evaluate our estimators. Because the square loss function is symmetric to the true parameter, overestimation has the same loss with underestimation. However, if the BER of the whole communication system is considered, overestimation will perform worse than underestimation. By our investigation, the reason is that overestimation will cause ISI from the future symbols, which affects BER significantly. Intuitively, when considering BER as our evaluation, we prefer to design the mechanism tends to underestimation.
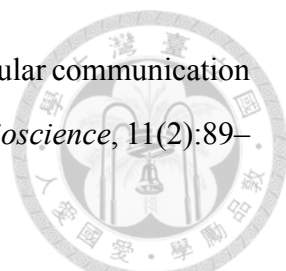
In our thesis, we only discuss on blind synchronization in clock offset estimation. The clock skew estimation or channel estimation in blind synchronization might be the next topic. Furthermore, our channel model only consider one-dimensional first hitting and molecular capture one by one rather than concentration-based molecular communications, which has been a popular model recently in molecule communications. In concentration-based synchronization, the estimation from a random process as observations need to be investigated rather than our analysis in random sample as observations.

Bibliography

[1] Tatsuya Suda, Michael Moore, Tadashi Nakano, Ryota Egashira, Akihiro Enomoto, Satoshi Hiyama, and Yuki Moritani. Exploratory research on molecular communication between nanomachines. In *Genetic and Evolutionary Computation Conference (GECCO), Late Breaking Papers*, volume 25, page 29. Citeseer, 2005.

[2] Ian F. Akyildiz, Fernando Brunetti, and Cristina Blázquez. Nanonetworks: A new communication paradigm. *Computer Networks (Elsevier) Journal*, 52:2260–2279, Aug. 2008.

[3] Ignacio Llatser, Eduard Alarcón, and M Pierobon. Diffusion-based channel characterization in molecular nanonetworks. In *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 467–472, Apr. 2011.

[4] Lluís Parcerisa Giné and Ian F Akyildiz. Molecular communication options for long range nanonetworks. *Computer Networks*, 53(16):2753–2766, Nov. 2009.

[5] T. Nakano, T. Suda, M. Moore, R. Egashira, A. Enomoto, and K. Arima. Molecular communication for nanomachines using intercellular calcium signalling. In *Proc. 5th IEEE Conference on Nanotechnology*, volume 2, pages 478–481, July 2005.

[6] Akihiro Enomoto, Michael Moore, Tadashi Nakano, Ryota Egashira, Tatsuya Suda, Atsushi Kayasuga, Hiroaki Kojima, Hitoshi Sakakibara, and Kazuhiro Oiwa. A molecular communication system using a network of cytoskeletal filaments. In *Proceedings of the 2006 NSTI Nanotechnology Conference*, 2006.

[7] Hoda ShahMohammadian, Geoffrey G. Messier, and Sebastian Magierowski. Optimum receiver for molecule shift keying modulation in diffusion-based molecular communication channels. *Nano Communication Networks*, 3(3):183–195, 2012.

[8] Wei-An Lin, Yen-Chi Lee, Ping-Cheng Yeh, and Chia-Han Lee. Signal detection and ISI cancellation for quantity-based amplitude modulation in diffusion-based molecular communications. In *Proc. IEEE GLOBECOM*, pages 4362–4367, Dec. 2012.

[9] Ma Shaodan, Pan Xinyue, Yang Guang-Hua, and Ng Tung-Sang. Blind symbol synchronization based on cyclic prefix for ofdm systems. *Vehicular Technology, IEEE Transactions on*, 58(4):1746–1751, 2009.

[10] Ignacio Llatser, Albert Cabellos-Aparicio, and Eduard Alarcon. Networking challenges and principles in diffusion-based molecular communication. *Wireless Communications, IEEE*, 19(5):36–41, 2012.

[11] D. Makrakis M.U. Mahfuz and H. Mouftah. On the detection of binary concentration-encoded unicast molecular communication in nanonetworks. In *4th International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS-2011)*, pages 446–449, 26-29 January 2011.

[12] S. Abadal and I.F. Akyildiz. Bio-inspired synchronization for nanocommunication networks. In *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, pages 1–5, Dec 2011.

[13] H. ShahMohammadian, G. Messier, and S. Magierowski. Blind synchronization in diffusion-based molecular communication channels. *Communications Letters, IEEE*, PP(99):1–4, 2013.

[14] Andrew W. Eckford. Nanoscale communication with Brownian motion. In *Proc. International Conference on Information Sciences and Systems*, pages 160–165, Mar. 2007.

[15] Sachin Kadloor, Raviraj S Adve, and Andrew W Eckford. Molecular communication using brownian motion with drift. *IEEE Transactions on NanoBioscience*, 11(2):89–99, June 2012.

[16] Jiun-Ting Huang, Hsin-Yu Lai, Yen-Chi Lee, Chia-Han Lee, and Ping-Cheng Yeh. Distance estimation in concentration-based molecular communications. to appear in *IEEE Global Communications Conference (GLOBECOM)*, Dec. 2013.

[17] Raj S. Chhikara and J. Leroy Folks. *The inverse gaussian distribution: theory, methodology, and applications*. Marcel Dekker, Inc., 1989.

[18] K. V. Srinivas, R. S. Adve, and Andrew W. Eckford. Molecular communication in fluid media: The additive inverse gaussian noise channel. *IEEE Transactions on Information Theory*, 58(7):4678–4692, 2012.

[19] Cagatay Candan. Notes on non-random parameter estimation. 2011.

[20] EJG Pitman. The estimation of the location and scale parameters of a continuous population of any given form. *Biometrika*, pages 391–421, 1939.

[21] RCH Cheng and NAK Amin. Maximum likelihood estimation of parameters in the inverse gaussian distribution, with unknown origin. *Technometrics*, 23(3):257–263, 1981.