UNIVERSITY OF CALIFORNIA,
IRVINE

Asynchronous Transmission in Multiuser Networks

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Electrical Engineering

by

Mehdi Ganji

Thesis Committee:
Chancellor's Professor, Hamid Jafarkhani, Chair
Professor, Syed A. Jafar
Professor, Ender Ayanoglu

2018

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Prof. Hamid Jafarkhani for the continuous support of my research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me a lot in the time of research and writing of this thesis.

I would like to thank my friends for accepting nothing less than excellence from me. Last but not the least, I would like to thank my family: my parents and to my brothers and sister for supporting me spiritually throughout my research studies and my life in general.

ABSTRACT OF THE THESIS

Asynchronous Transmission in Multiuser Networks

By

Mehdi Ganji

Master of Science in Electrical Engineering

University of California, Irvine, 2018

Chancellor's Professor, Hamid Jafarkhani, Chair

Time asynchrony inherently exists in many wireless communication systems, specially in multiuser scenarios where the signals come from different locations. Different locations and paths impose different delays on the received signals, resulting in asynchronous reception at the receiver. In most of the works in the literature, perfect synchronization is a common presumption. However, it might be impossible to synchronize all the nodes even if an ideal infrastructure with signal overheads is considered. For example, if the receiver encompasses multiple receive antennas or there are multiple distributed base stations, then, the synchronization can be realized at one of them at most. Thus, it is of great importance to investigate the effect of the time asynchrony in the wireless systems. One natural question is that how to eliminate the time asynchrony and make all the received signals aligned at the receiver. This question is analyzed under the notion of time synchronization. There are many methods in literature trying to achieve this goal. However, the other question which is atypical but even more important is that, is synchronizing the received signals necessary? does the time asynchrony degrade the performance?

When the receiver is designed with the presumption of having perfect synchronization,

YES, the time asynchrony will degrade the performance. Nevertheless, what if we design the system and the receiver structure with TIME ASYNCHRONY in mind. Does the system that is designed based on the time asynchrony provide worse performance compared to the synchronous one? We will thoroughly investigate this question in this thesis. In a nutshell, we show that by investigating inherent time delays between different users in a multiuser/multi-antenna scenario, we are able to improve the performance. By using proper transmission and receiver design, time asynchrony provides additional degrees of freedom in a time limited communication settings which can be exploited to improve the performance.

In Chapter one, more details about time synchronization and the problem of time asynchrony are presented. Then, we introduce the proper structure for exploiting time asynchrony and present the resulting system model. We also explain the reason behind possible advantages of time asynchrony. In Chapter two, we show implementation of different detection methods based on the asynchronous system model. We include different methods like maximum likelihood sequence detection (MLSD), successive interference cancellation (SIC) and Zero Forcing (ZF). In Chapter three, we analytically analyze the achievable performance by the asynchronous transmission and compare it with the conventional synchronous transmission. Two performance criteria are considered. One is the bit error rate (BER) in a fading channel. The other one is the achievable rate in an additive white Gaussian noise (AWGN) channel. The diversity gain of the BER performance and the multiplexing gain of the achievable rates are also derived. Finally, in Chapter 4, simulation results, some discussions and future work are presented.

Chapter 1

System Model

In this section, we present the problem statement which is the main motivation to consider asynchronous transmission. Then, some insightful concepts are provided to further support the use of asynchronous transmission. Next, the derivation of the system model are presented. At the end, some features of the resulting system model are analyzed.

1.1 Problem Statement

Time and Carrier Synchronization are inseparable ingredients of wireless communications. Timing synchronization is the process by which the proper sampling instants are acquired by the receiver. If the receiver cannot grasp the right times to sample the incoming signal, the resulting performance will be degraded. On the other hand, the carrier synchronization is defined as the process by which the receiver node adjusts the phase and frequency of its local oscillator to those of the received signal [36]. In this

thesis, the main focus is time synchronization and carrier synchronization is part of the future work, although the underlying concepts are more or less the same.

Generally, in wireless systems including single-input-single-output (SISO), multiple-input-multiple-output (MIMO), cooperative relaying, and multi-user interference networks, the transmitted signals are distorted by changes in amplitude, phase and timing which are priori unknown to the receiver node. Usually, the channel parameters are estimated in the channel estimation process using pilot based or blind methods. For example, in SISO systems, the received signal at the receiver antenna is affected by a single timing offset (TO). The receiver has to estimate this parameter and compensate for its effects in order to decode the transmitted symbols. The receiver may or may not have the knowledge of channel state information (CSI). In the case of no CSI availability, the receiver has to carry out channel estimation (CE) in addition to TO estimation. In a multi-antenna system, data is transmitted across different channels. The received signal at an antenna is given by a linear combination of the distorted and delayed version of the data symbols transmitted from different transmit antennas. In multi-antenna systems, multiple signal streams arrive at a receive antenna from different transmit antennas resulting in multiple timing offsets (MTOs) [33].

The synchronization challenge in distributed MIMO systems and up-link of the Multi-user systems is to jointly estimate and compensate for the effect of multiple TOs in order to mitigate inter-user interference (ISI). For example, in LTE standards, synchronization is achieved through periodically transmitted primary and secondary synchronization signals from the base station. Any user who has not yet acquired the up-link synchronization can use the primary and secondary synchronization signals to first achieve synchronization in the down-link. Next, compensation for the propagation loss is made as part of the up-link random access procedure. However, another

synchronization challenge is that the correction of one users timing at the receiver can misalign those of the other users [10]. Hence, a huge effort is made to receive all the transmitted signals from possibly distributed antennas synchronized at all the receive antennas. Besides from the additional overhead signaling and processing, the main question is that whether it is necessary to enforce perfect time synchronization and align all the received signals or asynchrony can, in fact, be beneficial. There are some works in the literature which used the phenomena of time asynchrony to improve the system performance [11, 5, 6, 43, 1, 41, 18].

1.2 Multi-user Systems: Inter-user Interference

From another perspective, the other issue in Multi-user systems is inter-user interference (IUI) which is the result of sharing a common medium to transmit data. There are many applications where multiple users share a common channel to transmit data to a receiver. Numerous examples of multi-access communication include up-link transmission of a single cell in a cellular system, a group of twisted-pair copper subscriber lines transmitting data to the same switching office, multiple ground stations communicating with a satellite and interactive cable television networks. The key challenge in these types of networks is interference from unwanted signals. Over several decades, many methods have been introduced to address this problem [24, 46, 29, 8]. Most of these methods are based on assigning orthogonal dimensions to different users to be able to separate them and prevent interference. For example, time division multiple access (TDMA) protocols allocate different time slots to different users to mitigate interference. The same concept can be applied by partitioning the frequency spectrum among different users, which is called frequency division multiple access (FDMA). Code

division multiple access (CDMA) is another scheme used to surpass IUI in which users are multiplexed by distinct codes rather than by orthogonal frequency bands, or by orthogonal time slots [34]. More recently, multiple receive antennas are utilized at the receive side to take advantage of the spatial domain in order to cancel interference [35], [28].

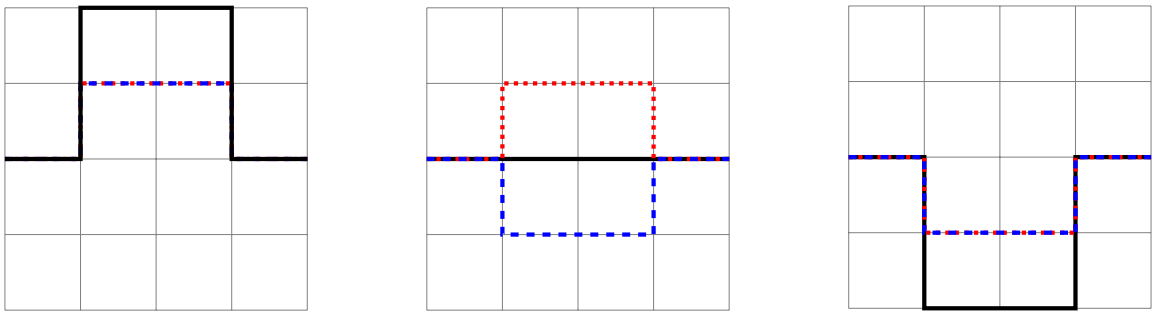
However, all the aforementioned works in the literature assume perfect time synchronization which is practically challenging as explained before. In this thesis, we investigate the timing asynchrony between users as an additional resource to address the problem of IUI. By exploiting time delays between users and employing an appropriate sampling method, we design detection methods which not only cancel the interference effectively, but also outperform the synchronous ones. When timing mismatch is used to cancel the IUI, resources like frequency spectrum, time and receive antenna can be employed to improve the performance. Next, we will provide some simple examples to show the concepts behind the timing asynchrony.

1.3 The Concept Behind Asynchrony

In this section, we try to provide some insights into possible benefits of asynchronous transmission. First, we provide a toy example to demonstrate the advantages of asynchronous transmission. Assume that we have two single-antenna users, namely, U_1 and U_2, trying to transmit their corresponding symbols from alphabet $\{-1, 1\}$ to a common single-antenna receiver at a same time and same frequency. Assuming AWGN channel, the received signal at the receiver can be denoted as:

$$y = s_1 + s_2 + n. \tag{1.1}$$

Considering Rectangular pulse shape and perfectly synchronized scenario, the possible outcomes for different permutations of input symbols are shown in Fig. 1.1. As it can be seen, no matter how small the additive noise is, it is impossible to distinguish between transmission of $\{1, -1\}$ and $\{-1, 1\}$. As a consequence of IUI, we are unable to decode the transmitted symbols in synchronous fashion. However, assume that either we add intentional time delay to transmission of one of the users or channel itself impose different delays to the transmitted signals. Then, the resulting combination of transmitted symbols can be one of the options in Fig. 1.2 .

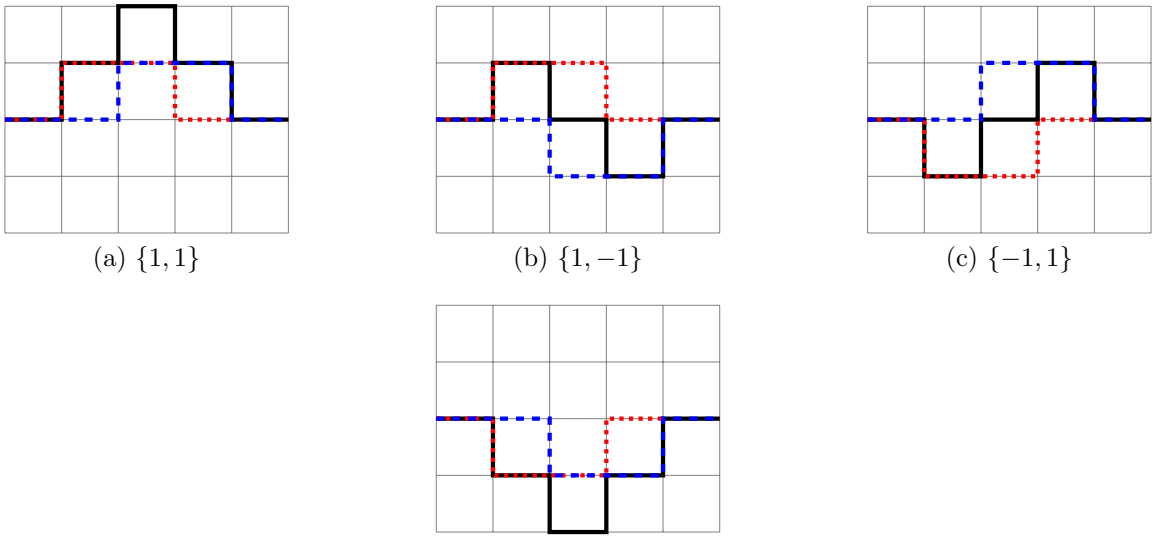


(a) $\{1, 1\}$ (b) $\{-1, 1\}, \{1, -1\}$ (c) $\{-1, -1\}$

Figure 1.1: Possible outcomes for synchronous transmission

Thanks to the timing offset between the two received signals, all possible combinations of the transmitted symbols are distinguishable and, hence, decodable. However, if we follow conventional sampling methods like sampling the middle point or calculate the integral over the range of transmission, then the benefits of asynchrony will be lost and the performance might even be worse. Thus, the proper sampling method is of great importance in order to take advantage of asynchronous transmission. Although in this simple example, there is 150% increase in the time interval which results in the reduced rate, by using sufficiently long frames, the loss due to adding time delays will be negligible.

The metric to numerically express "distinguishability" is the minimum distance be-



(a) $\{1, 1\}$ (b) $\{1, -1\}$ (c) $\{-1, 1\}$

(d) $\{-1, -1\}$

Figure 1.2: Possible outcomes for asynchronous transmission

tween the resulting signals. The distance between two continuous and integrable signals, e.g., $f(t)$ and $g(t)$, is defined as:

$$D(f, g) = \int_{-\infty}^{\infty} |f(t) - g(t)|^2 dt \tag{1.2}$$

Based on the aforementioned example, the minimum distance between the resulting signals in the synchronous scenario is zero while in the asynchronous scenario, it depends on the introduced delay and is equal to $\min\{\frac{2T}{T+\tau}, \frac{4\tau}{T+\tau}\}$ where T and τ are the symbol length and the time delay, respectively.(the result is normalized by the transmission interval). Assume that the time delay is a portion of the symbol length, i.e., $\tau = \kappa T$, then, the minimum distance can be shown with respect to κ as follows:

Maximizing the the minimum distance between the received signals results in minimizing the BER. Therefore, roughly speaking, we can interpret from Fig. 1.3 that the timing offset equal to half of the symbol length results in the best BER performance. These concepts are explained in more details in Chapter 3.

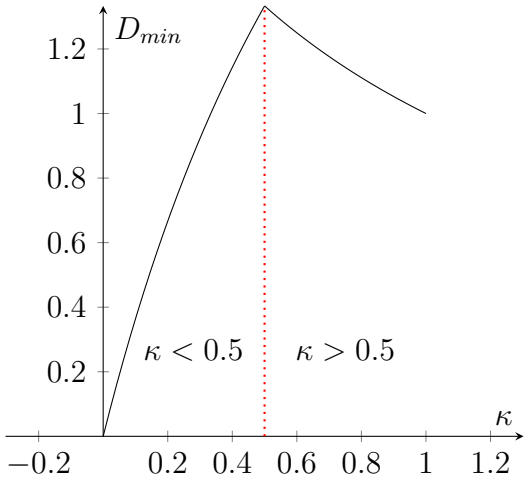


Figure 1.3: D_{min} with respect to $\kappa = \frac{\tau}{T}$

Next, we provide another example using the truncated Raised Cosine (R.C.) pulse shape. Assume that two symbols modulated by truncated R.C. are transmitted through two different channels, namely, h_1, h_2. Then, the received signal is equal to:

$$y(t) = h_1 s_1(t) + h_2 s_2(t) + n(t). \tag{1.3}$$

where $s_i(t) = s_i p(t)$, $p(t)$ being the truncated R.C. pulse shape. If the two signals are

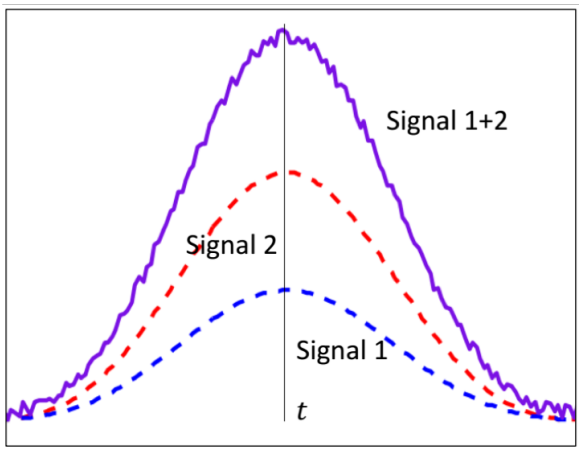


Figure 1.4: Illustration of synchronous reception

perfectly aligned, then the received signal is sampled at the peak point as shown in

Fig. 1.4. Assuming a normalized pulse shape, the discrete sample can be written as $y_1 = h_1 s_1 + h_2 s_2 + n_1$. Note that any extra sample at any other instant, e.g., t' will result in $y_2 = \alpha h_1 s_1 + \alpha h_2 s_2 + n_2$ where α is the amplitude of the pulse shape at the new sampling point t'. Now, assume that there is a time delay between the received signals. Then, based on the sampling method shown in Fig. 1.5, the two samples are obtained as $y_1 = h_1 s_1 + \beta h_2 s_2 + n_1$ and $y_2 = \gamma h_1 s_1 + h_2 s_2 + n_2$ where β and γ are the amplitude of the normalized pulse shape at the times t_1 and t_2, respectively. If we put

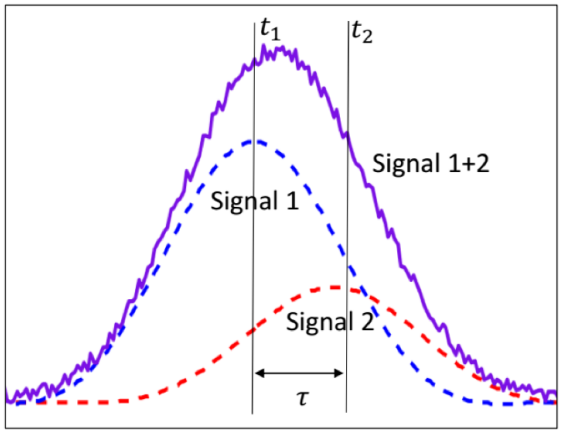


Figure 1.5: Illustration of asynchronous reception

the samples in the matrix form, for both synchronous and asynchronous scenarios, we will have:

$$\boldsymbol{y_{synch}} = \begin{bmatrix} h_1 & h_2 \\ \alpha h_1 & \alpha h_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \boldsymbol{n_{synch}}, \qquad \boldsymbol{y_{asynch}} = \begin{bmatrix} h_1 & \beta h_2 \\ \gamma h_1 & h_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \boldsymbol{n_{asynch}}. \tag{1.4}$$

$$= \boldsymbol{H_{synch}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \boldsymbol{n_{synch}}, \qquad = \boldsymbol{H_{asynch}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \boldsymbol{n_{asynch}}. \tag{1.5}$$

The main difference between the synchronous and asynchronous transmissions is in

the structure of the resulting mixing matrices denoted as $\boldsymbol{H_{synch}}$ and $\boldsymbol{H_{asynch}}$. One critical characteristic of the mixing matrix is its rank which is usually called the degree of freedom or multiplexing gain of the channel. In other words, by using proper processing, any channel matrix $\boldsymbol{H}$ can be divided into $rank(\boldsymbol{H})$ number of parallel sub-channels. It can be easily shown that the rank of matrices $\boldsymbol{H_{synch}}$ and $\boldsymbol{H_{asynch}}$ are 1 and 2, respectively, i.e., $rank(\boldsymbol{H_{synch}}) = 1,\ rank(\boldsymbol{H_{asynch}}) = 2$ (channel coefficients are assumed to be independent). Therefore, synchronous transmission results in one degree of freedom while the asynchronous transmission provides two degrees of freedom. Again, these concepts will be shown in more details in Chapter 3.

Before delving into details of asynchronous transmission, let us discuss Hilbert Space of signals, briefly. Hilbert space generalizes Euclidean space of real numbers to finite energy signals. Each finite energy signal can be represented by a vector in Hilbert space with each coordinate given by an inner product with the corresponding orthonormal basis functions. In more details, any finite energy signal like $x(t)$ can be written as a linear combination of the orthonormal basis functions as:

$$x(t) = \sum_{n \in \Gamma} x[n] b_n(t) \tag{1.6}$$

where $b_n(t)$ is an orthonormal basis function, i.e., $\langle b_n(t).b_m(t) \rangle = \delta[n-m]$ and $x[n]$ is the corresponding coefficient in the direction of the basis function $b_n(t)$ which is obtained by the following inner product.

$$x[n] = \langle x(t).b_n(t) \rangle \tag{1.7}$$

The Hilbert space representation is particularly useful as it allows a complete description of time-continuous signals $x(t)$ by a set of discrete values $x[n]$. If we further

constrain our finite energy signals to strictly band-limited ones, then Hilbert Space is called Paley-Wiener space. The Nyquist sampling theorem states that any signal in Paley-Wiener space whose Fourier transform is supported on $f \in (-W\ W)$ can be written as the linear combination of some Sinc pulses, i.e.,

$$x(t) = \sum_{n=-\infty}^{\infty} x[n] \left(\sqrt{2W} sinc(2W(t - nT)) \right) \tag{1.8}$$

where T is the Nyquist interval, i.e., $T = \frac{1}{2W}$ [12]. Due to the completeness of the of Nyquist rate Sinc pulses, every band-limited signal, even Sinc pulses that do not lie at integer multiples of T, e.g., their shifted version, i.e., $sinc(2W(t - nT - \tau))$, still lie completely in Paley-Wiener space.

Assume that, in Eq. (1.8), $x[n]$ is the transmitted symbol modulated on Sinc pulse $b_n(t) = \sqrt{2W} sinc(2W(t-nT))$. In practice, pulses spanning an unlimited time domain are not feasible, hence, they are usually truncated within a desired interval. Assume that the transmission interval is truncated into NT seconds, then we are capable of transmitting approximately $2WNT$ symbols by using the Nyquist rate signaling. In other words, Hilbert space offers 2WNT dimensions by the Nyquist rate signaling in the case of finite-time transmissions [17]. However, due to the truncation, the finite set of Nyquist rate Sinc pulses, i.e.:

$$S_{Nyquist} \equiv \left\{ b_n(t) = \sqrt{2W} sinc(2W(t - nT)), n = 0, \cdots, 2WNT - 1 \right\}$$

is not complete anymore and does not span the whole signal space. Therefore, we can insert additional pulses to exploit more signaling dimensions which leads to higher data throughput.

For example, defining $b_{2WNT} = \sqrt{2W}sinc(2W(t-\tau))$ and applying the well-known Gram-Schmidt orthogonalization process, provide us an orthonormal basis function with size $2WNT+1$. The newly formed basis function exploits an additional signaling dimension. We can continue this procedure and take advantage of the rest of available signaling dimensions. In fact, it is shown in the literature that the available degree of freedom in a time-limited channel is unbounded [1]. We will use asynchronous transmission to fully exploit the available degree of freedom inherent in such a channel. So far, we have provided some basic examples and insights to show the underlying benefits behind asynchronous transmission. However, as explained before, taking advantage of asynchronous transmission requires proper sampling and processing at the receiver. In the next section, the proper sampling and the resulting system model will be presented.

1.4 Sampling Method

First, we explain the system model for a point to point AWGN channel. Then, the generalization to Multi-user/MIMO fading channels is provided. In the conventional point to point Nyquist-rate transmission, after performing encoding and modulation, the modulated symbols are shaped with appropriate waveforms suited to the communication channel. Let the block length be equal to N, then the transmitted signal will be:

$$x(t) = \sum_{n=1}^{N} x[n]p(t-(n-1)T) \tag{1.9}$$

where $p(t)$ is the pulse shape, e.g., root raised cosine, which is truncated and its length is denoted by T_p, i.e., $p(t) = 0,\ t \notin [0, T_p]$. T is the symbol interval and is usually equal

to $\frac{1}{2W}$ where W is the occupied bandwidth in which most of the transmitted power is concentrated. To incorporate asynchrony in the transmission, assume that instead of N symbols, we transmit KN symbols using shifted versions of the pulse shapes. Then the transmitted signal will be:

$$x(t) = \sum_{k=1}^{K} \sum_{n=1}^{N} x_k[n] p(t - (n-1)T - \tau_k) \tag{1.10}$$

It can be interpreted as the superimpose of K sub-streams with the corresponding time shift of $\tau_k,\ 0 \leq \tau_1 < \tau_2 < \cdots < \tau_K < T$. We call K the over-signaling factor. By assuming AWGN channel, the received signal is described as:

$$y(t) = \sum_{k=1}^{K} \sum_{n=1}^{N} x_k[n] p(t - (n-1)T - \tau_k) + n(t) \tag{1.11}$$

To detect the transmitted symbols $x_k[n]$, instead of working with the continuous random process $y(t)$, we use a set of statistics, $Z_j = r_j(y(t)), j = 1, \cdots, J$ that are sufficient for the detection of the transmitted symbols. Intuitively, $Z_1, Z_2, \cdots, Z_J$ are jointly sufficient statistics if the statistician who knows the values of $Z_1, Z_2, \cdots, Z_J$ can do just as good a job of estimating the transmitted symbols as the statistician who knows the entire random process $y(t)$. We use the well-known factorization theorem to find the sufficient statistics.

Theorem 1.1. *Let $Y_1, \cdots, Y_n$ be random variables with joint density $f(y_1, y_2, \cdots, y_n|\theta)$. The statistics*

$$Z_j = r_j(Y_1, Y_2, \cdots, Y_n), j = 1, \cdots, J \tag{1.12}$$

are jointly sufficient to estimate θ if and only if the joint density can be factored as

follows:

$$f(y_1, y_2, \cdots, y_n|\theta) = u(y_1, y_2, \cdots, y_n).v(r_1(y_1, y_2, \cdots, y_n), \cdots, r_J(y_1, y_2, \cdots, y_n), \theta)$$

where u and v are non-negative functions [32].

The density of $y(t)$ given the transmitted symbols is calculated as follows:

$$f(y(t)|\{x_k[n]\}) = c \exp\left[\int_{-\infty}^{\infty} [z(t)]^2 dt\right]$$

where $z(t) = y(t) - \sum_{k=1}^{K}\sum_{n=1}^{N} x_k[n]p(t-(n-1)T-\tau_k)$ and c is a constant value. By expanding the likelihood function, it can be observed that the transmitted symbol $x_k[n]$ is related to $y(t)$ through the value of $\int_{-\infty}^{\infty} y(t)p(t-(n-1)T-\tau_k)dt$. Therefore, using the factorization theorem stated above, we can conclude that sufficient statistics for detecting the transmitted symbols are:

$$y_l[m] = \int_{-\infty}^{\infty} y(t)p(t-(m-1)T-\tau_l)dt \tag{1.13}$$

which is known as the matched filter in the literature and can also be implemented using convolution followed by a sampler, i.e.,:

$$y_l[m] = y(t) * p(t)|_{T_p+(m-1)T+\tau_l} \tag{1.14}$$

Denoting $p(t) * p(t)$ as $g(t)$, the sufficient statistics can be represented as:

$$y_l[m] = \sum_{k=1}^{K}\sum_{n=1}^{N} x_k[n]g(T_p + (m-n)T + (\tau_l - \tau_k)) + n_l[m] \tag{1.15}$$

where $n_l[m] = n(t) * p(t)|_{T_p+(m-1)T+\tau_l}$. The NK obtained samples can be put to-

gether in two different ways. If we define $\boldsymbol{y[m]} = [y_1[m], \cdots, y_K[m]]^T$ and $\boldsymbol{x[m]} = [x_1[m], \cdots, x_K[m]]^T$, the input-output relation of the system can be presented in matrix form as follows:

$$\begin{bmatrix} \boldsymbol{y[1]} \\ \boldsymbol{y[2]} \\ \vdots \\ \boldsymbol{y[N]} \end{bmatrix} = \begin{bmatrix} \boldsymbol{R_{11}} & \boldsymbol{R_{12}} & \cdots & \boldsymbol{R_{1N}} \\ \boldsymbol{R_{21}} & \boldsymbol{R_{22}} & \cdots & \boldsymbol{R_{2K}} \\ \vdots & \ddots & \ddots & \vdots \\ \boldsymbol{R_{N1}} & \boldsymbol{R_{N2}} & \cdots & \boldsymbol{R_{NN}} \end{bmatrix} \begin{bmatrix} \boldsymbol{x[1]} \\ \boldsymbol{x[2]} \\ \vdots \\ \boldsymbol{x[N]} \end{bmatrix} + \begin{bmatrix} \boldsymbol{n[1]} \\ \boldsymbol{n[2]} \\ \vdots \\ \boldsymbol{n[N]} \end{bmatrix} \tag{1.16}$$

where $\boldsymbol{R_{mn}}$ is the $K \times K$ constructing sub-block whose elements are defined as:

$$\boldsymbol{R_{mn}}(l,k) = g(T_p + (m-n)T + (\tau_l - \tau_k)) \tag{1.17}$$

Matrix $\boldsymbol{R}$ is a Hermitian block-Toeplitz matrix, i.e., $\boldsymbol{R_{ij}} = \boldsymbol{R_{i'j'}} \quad if \quad i-j = i'-j'$, whose sub-blocks are not necessarily Toeplitz. Two examples of matrix $\boldsymbol{R'}$ are provided for Rectangular and R.R.C pulse shapes, $K = 3, \boldsymbol{\tau} = [0, 0.2, 0.7]$ and $N = 3$.

$$\boldsymbol{R}_{Rect.} = \left[\begin{array}{ccc|ccc|ccc} 1 & 0.8 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.8 & 1 & 0.3 & 0.2 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.3 & 1 & 0.7 & 0.5 & 0 & 0 & 0 & 0 \\ \hline 0 & 0.2 & 0.7 & 1 & 0.8 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.8 & 1 & 0.3 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.3 & 1 & 0.7 & 0.5 & 0 \\ \hline 0 & 0 & 0 & 0 & 0.2 & 0.7 & 1 & 0.8 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0.8 & 1 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.3 & 1 \end{array}\right] \tag{1.18}$$

$$\boldsymbol{R}_{R.R.C.} = \left[\begin{array}{ccc|ccc|ccc} 1 & 0.93 & 0.33 & 0 & -0.11 & -0.07 & 0 & 0.02 & 0.01 \\ 0.93 & 1 & 0.6 & 0.2 & 0 & -0.12 & -0.04 & 0 & 0.02 \\ 0.33 & 0.6 & 1 & 0.84 & 0.6 & 0 & -0.13 & -0.12 & 0 \\ \hline 0 & 0.2 & 0.84 & 1 & 0.93 & 0.33 & 0 & -0.11 & -0.07 \\ -0.11 & 0 & 0.6 & 0.93 & 1 & 0.6 & 0.2 & 0 & -0.12 \\ -0.07 & -0.12 & 0 & 0.33 & 0.6 & 1 & 0.84 & 0.6 & 0 \\ \hline -0 & -0.04 & -0.13 & 0 & 0.2 & 0.84 & 1 & 0.93 & 0.33 \\ 0.02 & 0 & -0.12 & -0.11 & 0 & 0.6 & 0.93 & 1 & 0.6 \\ 0.01 & 0.017 & 0 & -0.07 & -0.12 & 0 & 0.33 & 0.6 & 1 \end{array}\right] \tag{1.19}$$

On the other hand, if we define $\boldsymbol{y_l} = [y_l[1], \cdots, y_l[N]]^T$ and $\boldsymbol{x_k} = [x_k[1], \cdots, x_k[N]]^T$, the input-output relation of the system can be presented in matrix form as follows:

$$\begin{bmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \\ \vdots \\ \boldsymbol{y_K} \end{bmatrix} = \begin{bmatrix} \boldsymbol{R'_{11}} & \boldsymbol{R'_{12}} & \cdots & \boldsymbol{R'_{1K}} \\ \boldsymbol{R'_{21}} & \boldsymbol{R'_{22}} & \cdots & \boldsymbol{R'_{2K}} \\ \vdots & \ddots & \ddots & \vdots \\ \boldsymbol{R'_{K1}} & \boldsymbol{R'_{K2}} & \cdots & \boldsymbol{R'_{KK}} \end{bmatrix} \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \vdots \\ \boldsymbol{x_K} \end{bmatrix} + \begin{bmatrix} \boldsymbol{n_1} \\ \boldsymbol{n_2} \\ \vdots \\ \boldsymbol{n_K} \end{bmatrix} \tag{1.20}$$

where $\boldsymbol{R'_{lk}}$ is the $N \times N$ constructing sub-block whose elements are defined as:

$$\boldsymbol{R'_{lk}}(m,n) = g(T_p + (m-n)T + (\tau_l - \tau_k)) \tag{1.21}$$

Matrix $\boldsymbol{R'}$ is a Hermitian matrix whose sub-blocks, i.e., $\boldsymbol{R'_{lk}}$ are banded Toeplitz blocks of order u, where $u = \frac{T_p}{T}$. Two examples of matrix $\boldsymbol{R'}$ are provided for Rectangular

and R.R.C. pulse shapes, $K = 3, \boldsymbol{\tau} = [0, 0.2, 0.7]$, and $N = 3$.

$$\boldsymbol{R'}_{Rect.} = \left[\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & 0.8 & 0 & 0 & 0.3 & 0 & 0 \\ 0 & 1 & 0 & 0.2 & 0.8 & 0 & 0.7 & 0.3 & 0 \\ 0 & 0 & 1 & 0 & 0.2 & 0.8 & 0 & 0.7 & 0.3 \\ \hline 0.8 & 0.2 & 0 & 1 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.8 & 0.2 & 0 & 1 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0.8 & 0 & 0 & 1 & 0 & 0.5 & 0.5 \\ \hline 0.3 & 0.7 & 0 & 0.5 & 0.5 & 0 & 1 & 0 & 0 \\ 0 & 0.3 & 0.7 & 0 & 0.5 & 0.5 & 0 & 1 & 0 \\ 0 & 0 & 0.3 & 0 & 0 & 0.5 & 0 & 0 & 1 \end{array}\right] \tag{1.22}$$

$$\boldsymbol{R'}_{R.R.C} = \left[\begin{array}{ccc|ccc|ccc} 1 & 0 & 0 & 0.93 & -0.11 & 0.02 & 0.33 & -0.07 & 0.01 \\ 0 & 1 & 0 & 0.2 & 0.93 & -0.11 & 0.84 & 0.33 & -0.07 \\ 0 & 0 & 1 & -0.04 & 0.2 & 0.93 & -0.13 & 0.84 & 0.33 \\ \hline 0.93 & 0.2 & -0.04 & 1 & 0 & 0 & 0.6 & -0.12 & 0.02 \\ -0.11 & 0.93 & 0.2 & 0 & 1 & 0 & 0.6 & 0.6 & -0.12 \\ 0.02 & -0.11 & 0.93 & 0 & 0 & 1 & -0.12 & 0.6 & 0.6 \\ \hline 0.33 & 0.84 & -0.13 & 0.6 & 0.6 & -0.12 & 1 & 0 & 0 \\ -0.07 & 0.33 & 0.84 & -0.12 & 0.6 & 0.6 & 0 & 1 & 0 \\ 0.01 & -0.07 & 0.33 & 0.02 & -0.12 & 0.6 & 0 & 0 & 1 \end{array}\right] \tag{1.23}$$

If the distance between the introduced time delays are equal, i.e., $\tau_k = (k-1)\frac{T}{K},\ k = 1, \cdots, K$, then the second structure turns into a block-Toeplitz matrix which each blocks is also Toeplitz. On the other hand, the first structure will turn into a Toeplitz

matrix. Therefore, these two structures are used interchangeably based on the application. Note that, in fact, the uniformly distributed set of time delays provide the best performance, which will be shown in Chapter 3.

Now, let us adapt the point to point model to a fading multi-user model. We consider a system with K users, transmitting data to a common receiver simultaneously, which can have one receive antenna or multiple ones. Here, we assume one receive antenna but having multiple receive antennas also follows the same structure. The signal transmitted from User k is described by:

$$x_k(t) = \sum_{n=1}^{N} x_k[n]p(t-(n-1)T) \tag{1.24}$$

Due to the different physical locations of the users, their signal is received with various time delays. The kth transmitted signal is received with a relative delay of τ_k and a channel path gain of h_k. Then, the received signal can be represented by:

$$y(t) = \sum_{k=1}^{K} h_k x_k(t-\tau_k) + n(t) \tag{1.25}$$

where K is the number of users and $n(t)$ is the white noise with variance of σ_n^2. Without loss of generality, we assume that $0 = \tau_1 < \tau_2 < \cdots < \tau_K < T$. The channel coefficients are assumed to be fixed during each frame. Again, using the factorization theorem stated above, we can obtain sufficient statistics and put them in the matrix form as:

$$\begin{bmatrix} \boldsymbol{y}[1] \\ \boldsymbol{y}[2] \\ \vdots \\ \boldsymbol{y}[N] \end{bmatrix} = \boldsymbol{R} \begin{bmatrix} \boldsymbol{H} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{H} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}[1] \\ \boldsymbol{x}[2] \\ \vdots \\ \boldsymbol{x}[N] \end{bmatrix} + \begin{bmatrix} \boldsymbol{n}[1] \\ \boldsymbol{n}[2] \\ \vdots \\ \boldsymbol{n}[N] \end{bmatrix} \tag{1.26}$$

or

$$\begin{bmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \\ \vdots \\ \boldsymbol{y_K} \end{bmatrix} = \boldsymbol{R'} \begin{bmatrix} \boldsymbol{H'_1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H'_2} & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{H'_K} \end{bmatrix} \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \vdots \\ \boldsymbol{x_K} \end{bmatrix} + \begin{bmatrix} \boldsymbol{n_1} \\ \boldsymbol{n_2} \\ \vdots \\ \boldsymbol{n_K} \end{bmatrix} \tag{1.27}$$

where $\boldsymbol{H} = \text{diag}[h_1, h_2, \ldots, h_K]$ and $\boldsymbol{H'_k} = h_k \boldsymbol{I_N}$. The elements of matrices $\boldsymbol{R}$ and $\boldsymbol{R'}$ are defined as before.

In the conventional transmission scheme, the set of modulated signals, hence, the set of matched filters are orthogonal, which results in independent noise samples. However, by introducing timing offsets, the set of matched filters are not orthogonal anymore. Due to the additional signaling, the noise samples are not independent anymore and their covariance is not identity matrix. Remember that:

$$n_l[m] = \int_{-\infty}^{\infty} n(t) p(t - (m-1)T - \tau_l) dt \tag{1.28}$$

Thus, the covariance between $n_l[m]$ and $n_k[n]$, denoted as δ_{lk}^{mn}, will be equal to:

$$\begin{aligned} \delta_{lk}^{mn} &= E[n_l[m] n_k^*[n]] \\ &= E\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} n(t) p(t - (m-1)T - \tau_l) n^*(s) p(s - (n-1)T - \tau_k) dt ds \right] \\ &= \sigma_n^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(t-s) p(t - (m-1)T - \tau_l) p(s - (n-1)T - \tau_k) dt ds \\ &= \sigma_n^2 \int_{-\infty}^{\infty} p(t - (m-1)T - \tau_l) p(t - (n-1)T - \tau_k) dt \\ &= \sigma_n^2 p(t) * p(t)|_{T_p + (m-n)T + (\tau_l - \tau_k)} \\ &= \sigma_n^2 g(T_p + (m-n)T + (\tau_l - \tau_k)) \end{aligned} \tag{1.29}$$

Therefore, the covariance matrix of noise vectors $\boldsymbol{n} = [\boldsymbol{n[1]}^T, \cdots, \boldsymbol{n[N]}^T]^T$ and $\boldsymbol{n'} = [\boldsymbol{n_1}^T, \cdots, \boldsymbol{n_K}^T]^T$ are equal to $\boldsymbol{R}\sigma_n^2$ and $\boldsymbol{R'}\sigma_n^2$, respectively. However, we can use a different set of matched filters to produce independent noise samples which circumvent the noise whitening procedure involving Cholesky decomposition. For example, one way is to break down the integrals corresponding to the sampling in Fig. 1.6 to define a new sampling method as shown in Fig. 1.7 [38, 3]. The corresponding output samples are written as follows where τ_{k+1} is an auxiliary variable equal to T_p.

$$y_l[m] = \int_{\tau_l+(m-1)T}^{\tau_{(l+1)}+(m-1)T} \sum_{k=1}^{K}\sum_{n=1}^{N} x_k[n]p(t-(m-1)T-\tau_l)p(t-(n-1)T-\tau_k)h_k dt$$
$$+ \int_{\tau_l+(m-1)T}^{\tau_{(l+1)}+(m-1)T} n(t)p(t-(m-1)T-\tau_l)dt \qquad 1 \le l \le K,\ 1 \le m \le N+1$$

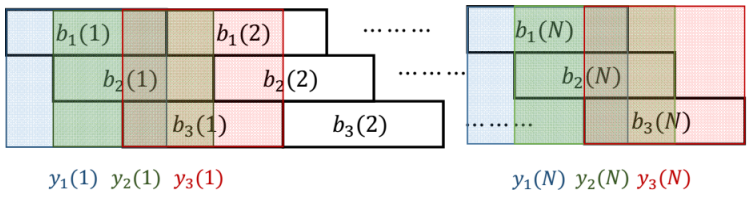


Figure 1.6: Sampling method with large sampling interval

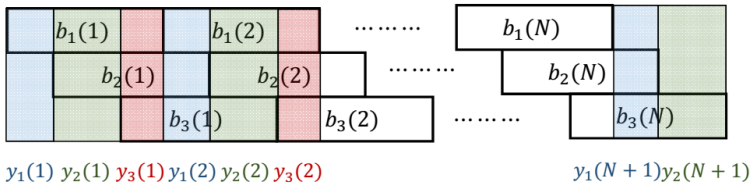


Figure 1.7: Sampling method with small sampling interval

By defining intermediate variables $u_{mn}(l,k)$ and noise samples $v_l[m]$ as follows:

$$u_{mn}(l,k) = \int_{\tau_l+(m-1)T}^{\tau_{(l+1)}+(m-1)T} p(t-(m-1)T-\tau_l)p(t-(n-1)T-\tau_k)dt \quad (1.30)$$

$$v_l[m] = \int_{\tau_l+(m-1)T}^{\tau_{(l+1)}+(m-1)T} n(t)p(t-(m-1)T-\tau_l)dt \quad (1.31)$$

, respectively, we can write the output samples in a more compact way:

$$y_l[m] = \sum_{k=1}^{K} \sum_{n=1}^{N} x_k[n] u_{mn}(l,k) h_k + v_l[m] \tag{1.32}$$

$$m = 1, \ldots, N+1 \quad \text{sampling time index}$$

$$l = 1, \ldots, K \quad \text{index of the matched user}$$

Defining $\boldsymbol{y[m]} = [y_1[m], y_2[m], \ldots, y_K[m]]^T$ and $\boldsymbol{x[n]} = [x_1[n], x_2[n], \ldots, x_K[n]]^T$, then, $\boldsymbol{y[m]}$ for different values of m can be written as:

$$\boldsymbol{y[m]} = \sum_{n=1}^{N} \boldsymbol{U_{mn} H x[n]} + \boldsymbol{v[m]} \quad 1 \leq m \leq N+1 \tag{1.33}$$

where $\boldsymbol{H} = \text{diag}[h_1, h_2, \ldots, h_K]$, $\boldsymbol{v[m]} = [v_1(j), v_2(j), \ldots, v_K(j)]^T$ and $\boldsymbol{U_{mn}}$ is a $K \times K$ matrix whose elements are defined as $\boldsymbol{U_{mn}}(l,k) = u_{mn}(l,k)$. The next step is to put all vectors of $\boldsymbol{y[m]}$ together and define $\boldsymbol{y}$ as $[\boldsymbol{y[1]}^T, \boldsymbol{y[2]}^T, \ldots, \boldsymbol{y[N+1]}^T]^T$. Then, $\boldsymbol{y}$ can be written as:

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{U_{11}} & \boldsymbol{U_{12}} & \boldsymbol{U_{13}} & \ldots & \boldsymbol{U_{1N}} \\ \boldsymbol{U_{21}} & \boldsymbol{U_{11}} & \boldsymbol{U_{12}} & \ldots & \boldsymbol{U_{1(N-1)}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{U_{(N-1)1}} & \ldots & \boldsymbol{U_{21}} & \boldsymbol{U_{11}} & \boldsymbol{U_{12}} \\ \boldsymbol{U_{N1}} & \ldots & \boldsymbol{U_{31}} & \boldsymbol{U_{21}} & \boldsymbol{U_{11}} \\ \boldsymbol{U_{(N+1)1}} & \ldots & \boldsymbol{U_{41}} & \boldsymbol{U_{31}} & \boldsymbol{U_{21}} \end{bmatrix} \begin{bmatrix} \boldsymbol{H} & \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H} & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{H} & \boldsymbol{0} \\ \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{H} \end{bmatrix} \begin{bmatrix} \boldsymbol{x[1]} \\ \boldsymbol{x[2]} \\ \vdots \\ \boldsymbol{x[N]} \end{bmatrix} + \boldsymbol{v}$$

$$\boldsymbol{y} = \boldsymbol{U}\bar{\boldsymbol{H}}\boldsymbol{x} + \boldsymbol{\nu} \tag{1.34}$$

Block Toeplitz structure of $\boldsymbol{U}$ originates from the fact that $u_{(j+m)(i+m)}(l,k) = u_{ji}(l,k)$. This can be verified by a change of variable in Eq. (1.30). Based on the relation

between T_p and T, different numbers of adjacent symbols interfere with each other. For example, for Rectangular pulse shapes, i.e., $T_p = T$, at each instant only current and the previous symbol cause interference. In other words, only $\boldsymbol{U_{11}}$ and $\boldsymbol{U_{21}}$ are nonzero. Without loss of generality, we assume that $T = 1$, therefore $\boldsymbol{U_{11}}$ and $\boldsymbol{U_{21}}$ are defined as follows:

$$\boldsymbol{U_{11}} = \begin{bmatrix} \tau_2 - \tau_1 & 0 & \dots & 0 \\ \tau_3 - \tau_2 & \tau_3 - \tau_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \tau_K - \tau_{K-1} & \dots & \tau_K - \tau_{K-1} & 0 \\ 1 - \tau_K & \dots & 1 - \tau_K & 1 - \tau_K \end{bmatrix} \tag{1.35}$$

$$\boldsymbol{U_{21}} = \begin{bmatrix} 0 & \tau_2 - \tau_1 & \dots & \tau_2 - \tau_1 \\ 0 & 0 & \dots & \tau_3 - \tau_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \tau_{K-1} - \tau_K \\ 0 & \dots & 0 & 0 \end{bmatrix} \tag{1.36}$$

Hence, for Rectangular pulse shapes, the system model simplifies to:

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{U_{11}} & \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{U_{21}} & \boldsymbol{U_{11}} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \dots & \boldsymbol{U_{21}} & \boldsymbol{U_{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{U_{21}} & \boldsymbol{U_{11}} \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{U_{21}} \end{bmatrix} \begin{bmatrix} \boldsymbol{H} & \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{H} & \boldsymbol{0} \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{H} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}[\boldsymbol{1}] \\ \boldsymbol{x}[\boldsymbol{2}] \\ \vdots \\ \boldsymbol{x}[\boldsymbol{N}] \end{bmatrix} + \boldsymbol{v}$$

The important fact about this sampling method is that the covariance matrix of noise samples is diagonal. With a small abuse of notation, we denote $Diag(\boldsymbol{U_{11}})$ as a diagonal matrix including diagonal elements of $\boldsymbol{U_{11}}$. Then, it can be shown that $E[\boldsymbol{v}\boldsymbol{v}^H]$ is equal to $\sigma^2(\boldsymbol{I_{N+1}} \otimes Diag(\mathbf{U_{11}}))$, where $\boldsymbol{I_n}$ is an $n \times n$ identity matrix and $(\otimes)$ is Kronecker product.

Since the statistically sufficient samples in Fig. 1.6 can be created from samples in Fig. 1.7, the samples in Fig. 1.7, i.e. Eq. (1.37), are sufficient statistics too. Both of these sampling methods introduce intentional ISI and impose memory on the system; however, they have some differences:

1. Sampling intervals in Fig. 1.7 are smaller and need faster sampler.

2. Since sampling intervals are disjoint in Fig. 1.7, noise samples are independent. However, due to sampling overlap, the noise samples in Fig. 1.6 are correlated.

3. The sampling in Fig. 1.7 results in an overdetermined system, while the number of output samples in Fig. 1.6 is equal to the number of input symbols.

Note that, the notion of increasing signaling dimension by using shifted versions of the pulse shapes, introduced ISI in the system. Hence, the benefit of having more degree of freedom is obtained in the expense of dealing with ISI which necessitates the need for designing low complexity receiver architectures. The question of whether the mixing matrix, i.e., $\boldsymbol{R}/\boldsymbol{R'}$, is invertible or not and how it behaves asymptotically as the block length N tends to infinity will have important consequences in the performance of system. Hence, we will investigate this question in the next section.

1.5 Features of Mixing Matrix

To understand the asymptotic behavior of matrix $\boldsymbol{R}/\boldsymbol{R'}$, we will utilize the Szego theorem which states that [20]:

Theorem 1.2. *Let* $\boldsymbol{T_N} = [t_{k-j}; k, j = 0, 1, 2, \cdots, N-1]$ *be a sequence of Hermitian Toeplitz matrices whose generating functions is defined as* $f(w) = \sum_{k=-\infty}^{\infty} t_k e^{ikw}, w \in [0, 2\pi]$. *Also,* $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{N-1}$ *are the sorted eigenvalues of matrix* $\boldsymbol{T_N}$. *Then, for any function* F *that is continuous on the range of* f*:*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} F(\lambda_k) = \frac{1}{2\pi} \int_0^{2\pi} F(f(w))dw \tag{1.37}$$

In addition, the largest and smallest eigenvalues asymptotically converge to:

$$\lim_{N\to\infty} \lambda_0 = \max_w f(w)$$

$$\lim_{N\to\infty} \lambda_{N-1} = \min_w f(w)$$

The proof comes from asymptotic equivalence of sequences of Hermitian Toeplitz matrices and their corresponding circulant versions which result in asymptotic convergence of the their eigenvalues [21]. Therefore, this theorem can be utilized to specify the eigenvalues of matrix $\boldsymbol{R}$ when the time delays are uniformly distributed. As N tends to infinity, the asymptotic eigenvalues of matrix $\boldsymbol{R}$ with uniformly distributed time delays approach the equispaced samples of the $\frac{K}{T}$-folded spectrum of the pulse shape. $\frac{K}{T}$-folded spectrum is defined the same as the conventional folded spectrum except that the frequency shifts are $\frac{K}{T}$ [37, 22].

Eigenvalues of matrix $\boldsymbol{R}$ with uniform time delays and different pulse shapes including

Sinc, R.R.C. and Rectangular are shown in Figs. 1.8, 1.9 and 1.10, respectively. As you can see, for the conventional transmission, i.e., $K = 1$, all the eigenvalues are equal to one for all the three mentioned pulse shapes. The reason is that because of Nyquist no-ISI condition, independent of the used pulse shape, the folded spectrum is always flat. Due to the time limited transmission (infinite-time pulse shapes can not be practically realized), the spectrum is theoretically non-zero for every frequency, thus the extra eigenvalues provided by the asynchronous transmission are always positive. However, the used pulse shape has a huge impact on the magnitude of the extra eigenvalues. For example, for Sinc function, adding asynchrony is not very beneficial, because the extra eigenvalues have very small values. However, for other pulse shapes which have some nonzero spectrum outside the $\frac{1}{T}$ bandwidth, like, Rectangular and R.R.C. pulse shapes, adding asynchrony provides additional nonzero eigenvalues. For example, for R.R.C. pulse shape, β percent additional eigenvalues are available to exploit. For Rectangular pulse shape, due to having unlimited spectrum, there are unlimited number of eigenvalues to exploit, however, by increasing K, the additional eigenvalues get closer to zero. In summary, the benefits of the asynchronous transmission depends on the utilized pulse shape and its $\frac{K}{T}$-folded spectrum.

Fortunately, the Szego Theorem can also be extended to Toeplitz block matrices like $\boldsymbol{R'}$ [22]. The generalized Szego Theorem relates the collective behavior of the eigenvalues to the generalized generating function, $\boldsymbol{R'(w)}$, which is defined as:

$$\boldsymbol{R'(w)} = \begin{bmatrix} f_{11}(w) & f_{12}(w) & \cdots & f_{1K}(w) \\ f_{21}(w) & f_{22}(w) & \cdots & f_{2K}(w) \\ \vdots & \ddots & \ddots & \vdots \\ f_{K1}(w) & f_{K2}(w) & \cdots & f_{KK}(w) \end{bmatrix} \tag{1.38}$$

where $f_{lk}(w)$ is the generating function for the corresponding Toeplitz block $\boldsymbol{R'_{lk}}$. The generalized Szego Theorem states that for any continuous function F [37]:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{NK-1}F[\lambda_k(\boldsymbol{R'})]=\frac{1}{2\pi}\int_0^{2\pi}\sum_{j=0}^{K-1}F[\lambda_j(\boldsymbol{R'(w)})]dw \tag{1.39}$$

In particular, for $F(x)=x$,

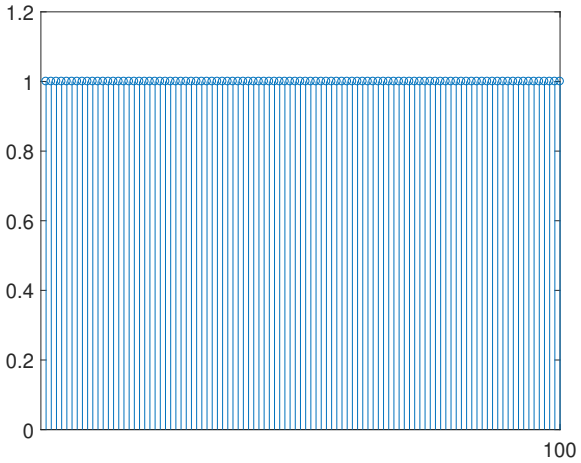
$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{NK-1}\lambda_k(\boldsymbol{R'})=\frac{1}{2\pi}\int_0^{2\pi}\sum_{j=0}^{K-1}\lambda_j(\boldsymbol{R'(w)})dw \tag{1.40}$$

Moreover, same convergence results can be obtained for the largest and smallest eigenvalues, i.e,:

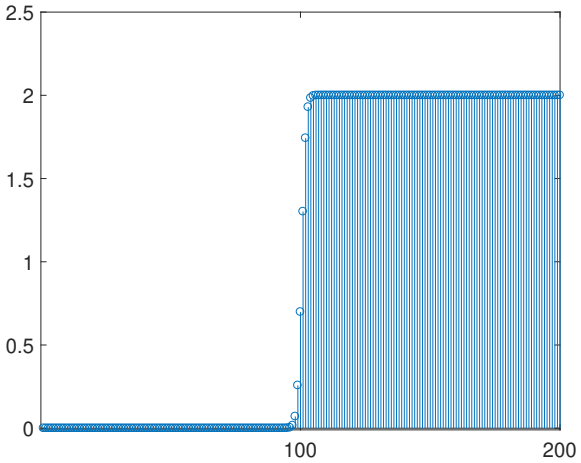
$$\lim_{N\to\infty}\lambda_0(\boldsymbol{R'})=\max_w\lambda_0(\boldsymbol{R'(w)})$$

$$\lim_{N\to\infty}\lambda_{NK-1}(\boldsymbol{R'})=\min_w\lambda_{K-1}(\boldsymbol{R'(w)})$$

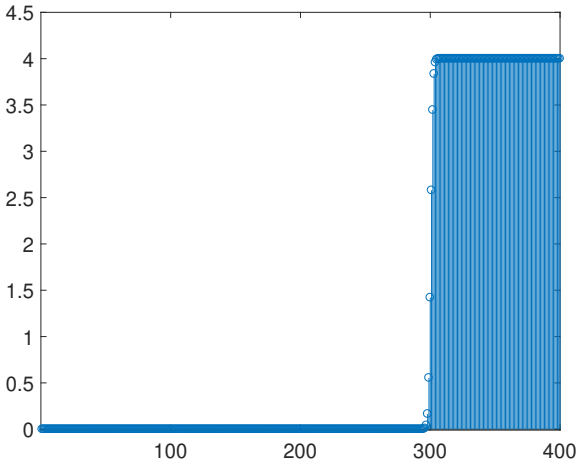
Therefore, some of the properties of matrix $\boldsymbol{R'}$ can be deducted by properties of matrix $\boldsymbol{R'(w)}$ when the block length is sufficiently large. For example, it is proved in the literature that for time limited transmission, i.e., finite value of u, matrix $\boldsymbol{R'(w)}$ is positive definite with bounded eigenvalues, thus, eigenvalues of matrix $\boldsymbol{R'}$ are nonzero and bounded. On the other hand, when the pulse shapes are strictly band-limited, matrix $\boldsymbol{R'(w)}$ is singular which results in singularity of $\boldsymbol{R'}$ [44]. In the next section, we adapt well-known receiver algorithms to the asynchronous transmission.



(a) $N = 100$ and $K = 1$

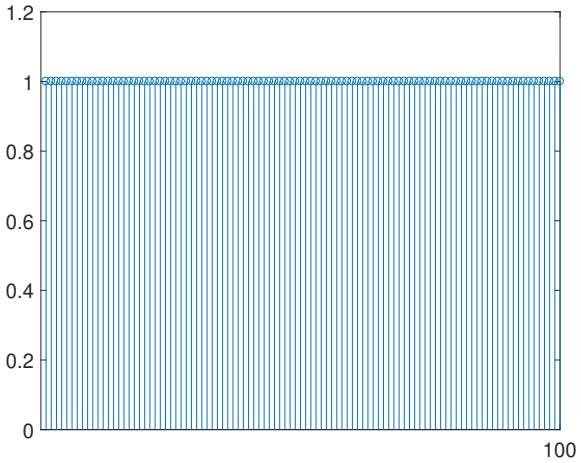


(b) $N = 100$ and $K = 2$

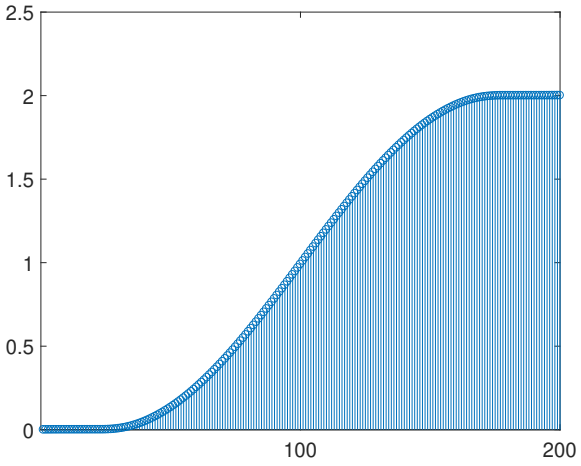


(c) $N = 100$ and $K = 4$

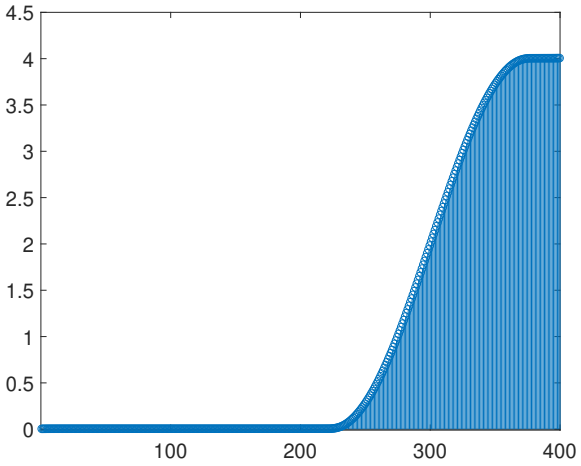
Figure 1.8: Eigenvalues of matrix $\boldsymbol{R}$ with Sinc pulse shape



(a) $N = 100$ and $K = 1$

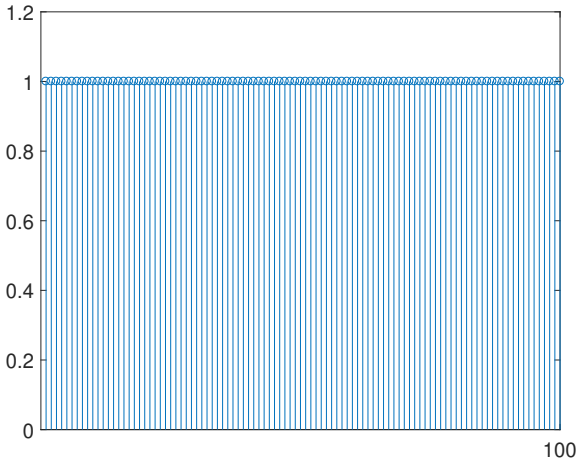


(b) $N = 100$ and $K = 2$

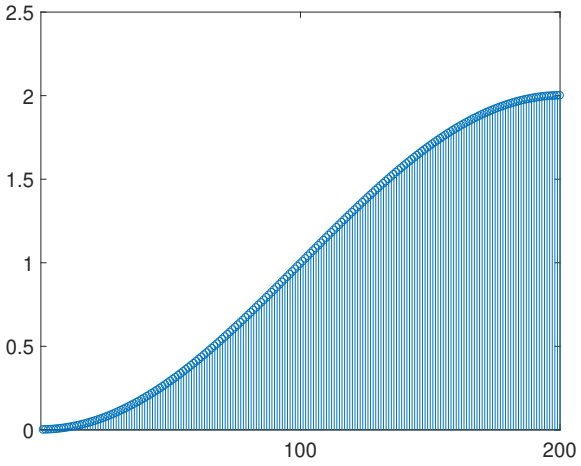


(c) $N = 100$ and $K = 4$

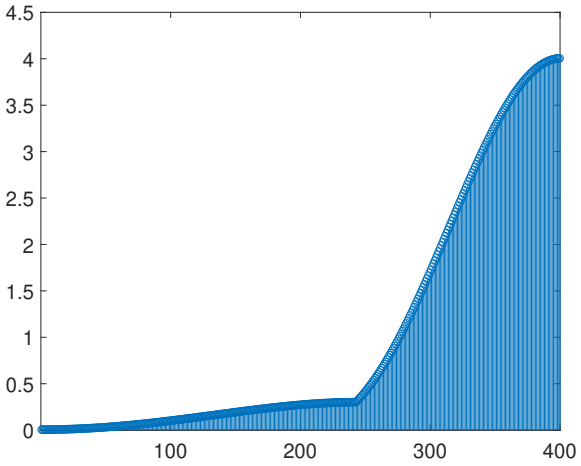
Figure 1.9: Eigenvalues of matrix $\boldsymbol{R}$ with R.R.C. pulse shape ($\beta = 0.75$)



(a) $N = 100$ and $K = 1$



(b) $N = 100$ and $K = 2$



(c) $N = 100$ and $K = 4$

Figure 1.10: Eigenvalues of matrix $\boldsymbol{R}$ with Rectangular pulse shape

Chapter 2

Receiver Design

In this section, we introduce different detection methods which take advantage of asynchronous transmission. In more detail, the problem is to recover the vector $\boldsymbol{b} \in \mathbb{R}^{NK}$ from an observation vector in the form of:

$$\boldsymbol{y} = \boldsymbol{U_{tot}}\boldsymbol{x} + \boldsymbol{\nu} \tag{2.1}$$

where $\boldsymbol{U_{tot}} = \boldsymbol{U}\bar{\boldsymbol{H}}$ is the known channel matrix and $\boldsymbol{\nu} \sim N(0, \sigma^2\boldsymbol{\Sigma_n})$. Thanks to the new sampling method introduced in the previous section, the covariance matrix of noise vector is a diagonal matrix. For the sake of notational simplicity, we can multiply the obtained vector with ${\boldsymbol{\Sigma_n}}^{-1/2}$ to equalize the variance of noise for different sub-channels. Thus, we will have $\hat{\boldsymbol{y}} = \hat{\boldsymbol{U}}\boldsymbol{x} + \hat{\boldsymbol{\nu}}$ where $\hat{\boldsymbol{U}} = {\boldsymbol{\Sigma_n}}^{-1/2}\boldsymbol{U}\bar{\boldsymbol{H}}$ and $\boldsymbol{\nu} \sim N(0, \sigma^2\boldsymbol{I_{(N+1)K}})$. The elements of $\boldsymbol{x}$ belong to a finite alphabet Ω of size $|\Omega|$. Therefore, there are $|\Omega|^{NK}$ possible vectors of $\boldsymbol{x}$. Detecting $\boldsymbol{x}$ in the maximum-likelihood (ML) sense is equivalent

to:

$$\min_{\boldsymbol{x}\in\Omega^{NK}} \left\| \hat{\boldsymbol{y}} - \hat{\boldsymbol{U}}\boldsymbol{x} \right\| \tag{2.2}$$

Eq. (2.2) is a finite-alphabet-constrained least-squares (LS) problem, which is known to be nondeterministic polynomial-time (NP)-hard. The complicating factor is of course the constraint $\boldsymbol{x} \in \Omega^{NK}$, otherwise it would be just the classical LS regression [30]. Due to the distinct feature of the sampling method shown in Fig. 1.7 , i.e., $\hat{\boldsymbol{U}}(i,j) = 0, j > i$, the metric in Eq. (2.2) can be rewritten as:

$$\min_{\boldsymbol{x}\in\Omega^{NK}} \{f_1(x_1) + f_2(x_1, x_2) + \cdots + f_{NK}(x_1, \cdots, x_{NK})\} \tag{2.3}$$

where

$$f_n(x_1, \cdots, x_n) = \left(\hat{y}_n - \sum_{m=1}^{n} \hat{\boldsymbol{U}}(n,m) x_m \right)^2 \tag{2.4}$$

The new representation can be visualized as a decision tree with NK+1 layers, $|\Omega|$ branches emanating from each node, and $|\Omega|^{NK}$ leaf nodes. To any branch, we associate a hypothetical decision on x_n , and the branch metric $f_n(x_1, \cdots, x_n)$. Also, to any node, we associate the cumulative metric which is just the sum of all branch metrics accumulated when traveling to that node from the root. Finally, to each node, we associate the symbols $\{x_1, x_2 \cdots, x_n\}$ it takes to reach there from the root.

Clearly, a naive but valid way of solving the minimization in Eq. (2.3) would be to traverse the entire tree to find the leaf node with the smallest cumulative metric. However, such a brute-force search is extremely inefficient, since there are $|\Omega|^n$ leaf nodes to examine. We will now review some efficient, popular, but approximate solutions for

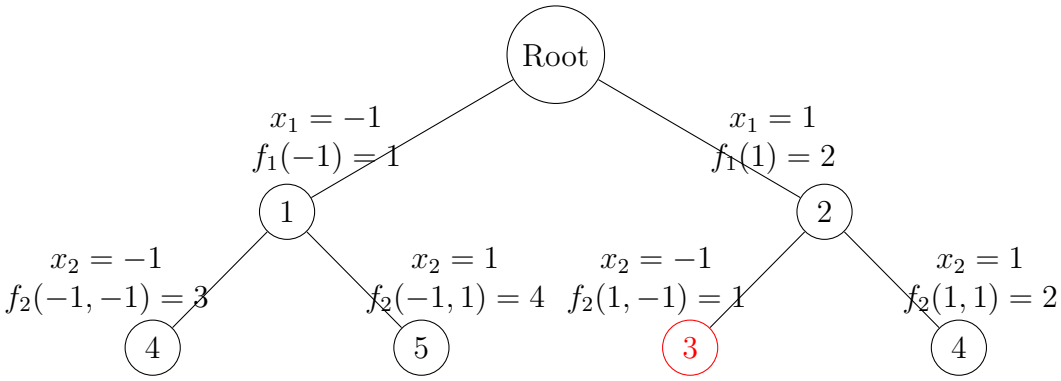


Figure 2.1: Tree representation of the minimization problem

this problem setting [48, 2].

2.1 Maximum-Likelihood Sequence Detection

Simply stated, Maximum-Likelihood Sequence Detection (MLSD) finds the sequence through the trellis that looks most like the received output sequence, or, in other words, minimizes the relation in Eq. (2.3) [16]. As a simple example, the highlighted path in Fig. 2.2 is the MLSD decision for the sequence that best matches the received outputs. The receiver need only wait until the entire sequence is received and then compare it against all possible encoder sequences.

The concept is simple, but the complexity grows exponentially with length of the sequence. The Viterbi Algorithm reduces this complexity through a recursive elimination procedure where trellis paths that are worse than other paths are eliminated early from further consideration. At each stage of the trellis, the decoder keep only the best surviving path into each state, eliminating all the rest into that state at that time. Thanks to the memory introduced in the system by the time delays, we can implement MLSD by using Viterbi algorithm. As a result, the complexity order of $\mathcal{O}(|\Omega|^{NK})$ offered by

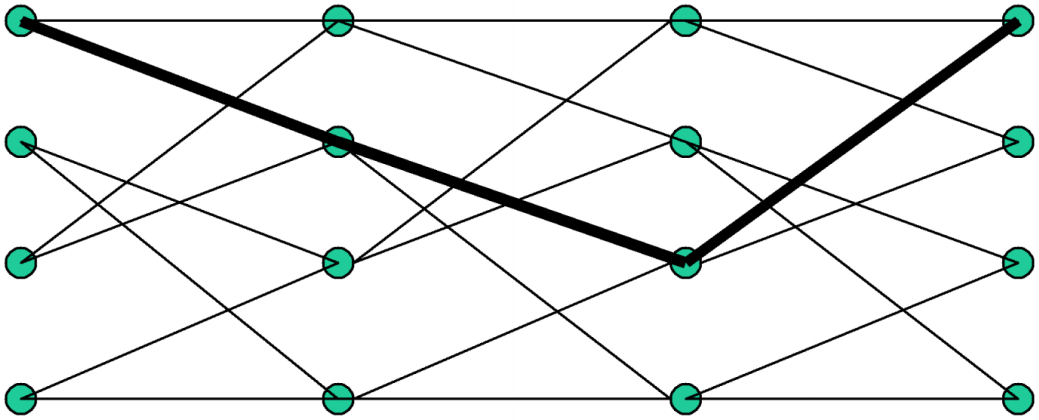


Figure 2.2: Illustration of the concept of MLSD in a trellis

exhaustive search reduces to complexity order of $\mathcal{O}(|\Omega|^K)$[39].

Based on the recursive relation between the input and output which is described as:

$$\hat{\boldsymbol{y}}[\boldsymbol{j}] = \hat{\boldsymbol{U}}_{\boldsymbol{11}}\boldsymbol{x}[\boldsymbol{j}] + \hat{\boldsymbol{U}}_{\boldsymbol{21}}\boldsymbol{x}[\boldsymbol{j}-\boldsymbol{1}] + \hat{\boldsymbol{\nu}}[\boldsymbol{j}] \qquad 2 \leq j \leq N$$

the trellis diagram of the system includes $|\Omega|^K$ states with $|\Omega|^K$ outgoing paths to the next states, and $|\Omega|^K$ incoming paths from previous states. To calculate the metric for each path, we need to calculate the likelihood function as follows:

$$\begin{aligned} Pr(\hat{\boldsymbol{y}}[\boldsymbol{j}]|\boldsymbol{x}[\boldsymbol{j}], \boldsymbol{x}[\boldsymbol{j}-\boldsymbol{1}]) =& Pr(\hat{\boldsymbol{\nu}}[\boldsymbol{j}] = \hat{\boldsymbol{y}}[\boldsymbol{j}] - \hat{\boldsymbol{U}}_{\boldsymbol{11}}\boldsymbol{x}[\boldsymbol{j}] - \hat{\boldsymbol{U}}_{\boldsymbol{21}}\boldsymbol{x}[\boldsymbol{j}-\boldsymbol{1}]) \\ =& \frac{1}{\sqrt{(2\pi)^K \sigma_n^2}} \exp\left(-\frac{1}{2\sigma_n^2} \boldsymbol{s_j}^H \boldsymbol{s_j}\right) \end{aligned}$$

where $\boldsymbol{s_j} = \hat{\boldsymbol{y}}[\boldsymbol{j}] - \hat{\boldsymbol{U}}_{\boldsymbol{11}}\boldsymbol{x}[\boldsymbol{j}] - \hat{\boldsymbol{U}}_{\boldsymbol{21}}\boldsymbol{x}[\boldsymbol{j}-\boldsymbol{1}]$. By discarding common terms and simple calculations, the metric for each path can be defined as $\sum_{i=1}^{K} |\boldsymbol{s_j}(i)|^2$. After calculating the path metrics, the final goal is to find the surviving path and trace it back to decode the transmitted symbols.

2.2 Zero Forcing (ZF)

The ZF detector first solves the problem mentioned in Eq. (2.2) neglecting the finite alphabet constraint.

$$\tilde{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \left\| \hat{\boldsymbol{y}} - \hat{\boldsymbol{U}}\boldsymbol{x} \right\| \tag{2.5}$$

$$= \hat{\boldsymbol{U}}^{-1}\hat{\boldsymbol{y}} \tag{2.6}$$

Note that $\hat{\boldsymbol{U}}^{-1}$ does not need to be explicitly calculated. For example, one way is to use Gaussian elimination, i.e.,

$$\tilde{x}_1 = \frac{\hat{y}_1}{\hat{\boldsymbol{U}}(1,1)} \tag{2.7}$$

$$\tilde{x}_2 = \frac{\hat{y}_2 - \hat{\boldsymbol{U}}(2,1)\tilde{x}_1}{\hat{\boldsymbol{U}}(2,2)} \tag{2.8}$$

$$\vdots \tag{2.9}$$

$$\tilde{x}_n = \frac{\hat{y}_n - \sum_{m=1}^{n-1} \hat{\boldsymbol{U}}(n,m)\tilde{x}_m}{\hat{\boldsymbol{U}}(n,n)} \tag{2.10}$$

ZF detector, then approximate the answers by projecting each $\tilde{x}_n$ onto the closet point in the constellation point, i.e.,

$$\bar{x}_n = \underset{x_k \in \Omega}{\operatorname{argmin}} \|x_k - \tilde{x}_k\| \tag{2.11}$$

It can be seen that, $\tilde{\boldsymbol{x}} = \boldsymbol{x} + \hat{\boldsymbol{U}}^{-1}\hat{\boldsymbol{\nu}}$, meaning that all the ISI has been completely removed. This is, in fact, how ZF got its name. However, unfortunately ZF works poorly unless $\hat{\boldsymbol{U}}$ is well conditioned. The reason is that the correlation between the

noise samples in the projection phase is neglected which can be high when the matrix $\hat{\boldsymbol{U}}$ is ill-conditioned. This can be improved by using MMSE estimate but it does not overcome the fundamental problem of the approach [30].

There is a fundamental difference between synchronous and asynchronous ZF. In the synchronous ZF, we need more number of receive antennas than transmit antennas to be able to cancel interference completely. However, in the asynchronous ZF, even with one receive antenna, we are able to remove all the interference completely. The required dimension at the receiver signal space is provided by asynchronous transmission and oversampling. Denoting M as the number of receiver antenna, it is well-known in the literature that, synchronous ZF provides $M - K + 1$ diversity orders when K independent symbols are transmitted. $K - 1$ diversity orders are wasted to remove the interference [23]. In more details, transmitting S ($S \leq K \leq M$) independent symbols will result in $(M-S+1)(K-S+1)$ diversity orders [40]. On the other hand, asynchronous ZF provides additional degrees of freedom by means of oversampling which enable us to enjoy full diversity of M while transmitting K independent symbols. These claims are mathematically proved in Chapter 3.

Here, we present the system model when multiple receive antennas are used at the receiver. By stacking output samples of all receive antennas together we can represent the system model as follows:

$$\begin{bmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \\ \vdots \\ \boldsymbol{y_M} \end{bmatrix} = \begin{bmatrix} \boldsymbol{U} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U} & \dots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{U} \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{H}}_1 \\ \bar{\boldsymbol{H}}_2 \\ \vdots \\ \bar{\boldsymbol{H}}_M \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} \boldsymbol{\nu_1} \\ \boldsymbol{\nu_2} \\ \vdots \\ \boldsymbol{\nu_M} \end{bmatrix}$$

$$\boldsymbol{y_{tot}} = \boldsymbol{U_{tot}} \boldsymbol{H_{tot}} \boldsymbol{x} + \boldsymbol{\nu_{tot}}$$

$$\boldsymbol{y_{tot}} = \boldsymbol{L_{tot}} \boldsymbol{x} + \boldsymbol{\nu_{tot}}$$

where M is number of receive antennas. Then, the zero-forcing detector is defined as:

$$\tilde{\boldsymbol{y}} = (\boldsymbol{L_{tot}}^H \boldsymbol{\Sigma_{tot}}^{-1} \boldsymbol{L_{tot}})^{-1} \boldsymbol{L_{tot}}^H \boldsymbol{\Sigma_{tot}}^{-1} \boldsymbol{y_{tot}} = \boldsymbol{x} + \tilde{\boldsymbol{\nu}} \tag{2.12}$$

where $\boldsymbol{\Sigma_{tot}} = \boldsymbol{I_M} \otimes \boldsymbol{\Sigma}$, and $\boldsymbol{\Sigma}$ is a diagonal matrix representing the covariance matrix of noise samples $\boldsymbol{\nu_i}$, $1 \leq i \leq M$. Note that, because the receive antennas are co-located the matrix $\boldsymbol{U}$ which corresponds to the timing offsets is the same at all the receive antennas. The noise enhancement factor is $(\boldsymbol{L_{tot}}^H \boldsymbol{\Sigma_{tot}}^{-1} \boldsymbol{L_{tot}})^{-1}$, which affects the receiver performance and will be studied in Chapter 3.

2.3 Successive Interference Cancellation with Hard Decision Passing

Successive interference cancellation (SIC) detection that takes a serial approach to cancel interference can be used to reduce complexity. SIC is sometimes called ZF detector with decision feedback (ZF-DF). ZF-DF detector performs Gaussian elimination method mentioned before to detect the transmitted symbol. However, it performs the Gaussian elimination with the modification that it projects the symbols onto the con-

stellation Ω in each step of the Gaussian elimination, rather than afterwards. In more details:

$$\bar{x}_1 = \underset{x_1 \in \Omega}{\operatorname{argmin}} f_1(x_1) \tag{2.13}$$

$$= \left[\frac{\hat{y}_1}{\hat{\boldsymbol{U}}(1,1)} \right] \tag{2.14}$$

where [.] means the projection onto the nearest point in the constellation. For the next symbols, the effect of previously detected symbols are removed by subtracing them from the corresponding samples, i.e.,

$$\bar{b}_n = \underset{x_n \in \Omega}{\operatorname{argmin}} \{ f_n(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_{n-1}, x_n) \} \tag{2.15}$$

$$= \left[\frac{\hat{y}_n - \sum_{m=1}^{n-1} \hat{\boldsymbol{U}}(n,m) \bar{x}_m}{\hat{\boldsymbol{U}}(n,n)} \right] \tag{2.16}$$

In the decision-tree perspective, ZF-DF can be considered as just examining one single path down from the root. When deciding on x_n, it considers $x_1, x_2, \cdots, x_{n-1}$ known and takes the x_n that corresponds to the smallest branch metric. Clearly, after NK steps we end up at one of the leaf nodes, but not necessarily in the one with the smallest cumulative metric. For example, in Fig. 2.1, considering the branch metrics for x_1 will result in detecting $x_1 = -1$. Next, deciding on x_2 given that $x_1 = -1$ will yield the detected vector as $(-1, -1)$ which is not the optimal answer, i.e., $(1, -1)$.

Thus, the problem with ZF-DF is error propagation. If, due to noise, an incorrect decision is taken in any of the steps, then this error will propagate and many of the subsequent symbols are likely to be detected wrong as well. The detection order can be optimized to minimize the effects of error propagation. It is best to start with

the symbol for which ZF produces the most reliable result, i.e., the sub-channel with smallest noise variance, and then proceed to less and less reliable symbols. However, even with the optimal ordering, error propagation severely limits the performance [4].

Using the sampling method in Fig. 1.7, this serial approach can be either a forward SIC initiated from the first transmitted symbol, i.e., $x_1[1]$, or a backward SIC started from the last transmitted symbol, i.e., $x_K[N]$. For example, for forward processing, $x_1[1]$ can be decoded by using $y_1[1]$ without interference, then $x_2[1]$ can be decoded by canceling the interference of $x_1[1]$ from $y_2[1]$, and so on. The same procedure can be performed backwards. One can also combine forward and backward operations. However, when hard decisions are used, such a combination will not result in a noticeable gain. On the other hand, by using soft decisions, combining the forward and backward operations will improve the results as explained in the following section.

2.4 Forward Backward Belief Propagation Detection

In the previous section we introduced an SIC method which was performed by passing hard decisions of previously decoded symbols to cancel the interference. In this section, we introduce a similar detection method which passes likelihood values.

In practice, each symbol x_k typically is composed of information-carrying bits. It is then of interest to take decisions on the individual bits, and often, also to quantify how reliable these decisions are. Such reliability information about a bit is called a soft decision, and is typically expressed via the probability ratio. Hard decisions does not distinguish between two events of $p(x = 0|y) = 0.49, p(x = 0|y) = 0.51$ and

$p(x=0|y)=0.01, p(x=0|y)=0.99$, while, the soft decisions take this difference into account. Therefore, soft decisions carry more information with respect to the hard decisions. By using likelihood values, instead of hard decisions, performance can be improved. Additionally, this method provides the opportunity to exploit benefits of backward processing as well. We explain the strategy of decoding for BPSK modulation and $K=2$, but it can be also generalized to other modulations and other values of K. We also assume that transmitted symbols have the same prior probabilities and calculate the conditional probabilities as follows:

$$\begin{cases} a = P(y_1[1]|x_1[1]=0) = \frac{1}{\sqrt{2\pi\rho_1\sigma_n^2}} \exp\left(-\frac{|y_1[1]+h_1\rho_1|^2}{2\rho_1\sigma_n^2}\right) \\ b = P(y_1[1]|x_1[1]=1) = \frac{1}{\sqrt{2\pi\rho_1\sigma_n^2}} \exp\left(-\frac{|y_1[1]-h_1\rho_1|^2}{2\rho_1\sigma_n^2}\right) \\ P_0^{fw}(x_1[1]) = P(x_1[1]=0|y_1[1]) = \frac{a}{a+b} \\ P_1^{fw}(x_1[1]) = P(x_1[1]=1|y_1[1]) = \frac{b}{a+b} \end{cases}$$

$$\begin{cases} c = P(y_2[1]|x_2[1]=0, y_1[1]) = \\ P_0^{fw}(x_1[1]) \frac{1}{\sqrt{2\pi\rho_2\sigma_n^2}} \exp\left(-\frac{|y_2[1]+h_1\rho_1+h_2\rho_2|^2}{2\rho_2\sigma_n^2}\right) + P_1^{fw}(x_1[1]) \frac{1}{\sqrt{2\pi\rho_2\sigma_n^2}} \exp\left(-\frac{|y_2[1]-h_1\rho_1+h_2\rho_2|^2}{2\rho_2\sigma_n^2}\right) \\ d = P(y_2[1]|x_2[1]=1, y_1[1]) = \\ P_0^{fw}(x_1[1]) \frac{1}{\sqrt{2\pi\rho_2\sigma_n^2}} \exp\left(-\frac{|y_2[1]+h_1\rho_1-h_2\rho_2|^2}{2\rho_2\sigma_n^2}\right) + P_1^{fw}(x_1[1]) \frac{1}{\sqrt{2\pi\rho_2\sigma_n^2}} \exp\left(-\frac{|y_2[1]-h_1\rho_1-h_2\rho_2|^2}{2\rho_2\sigma_n^2}\right) \\ P_0^{fw}(x_2[1]) = P(x_2[1]=0|y_1[1], y_2[1]) = \frac{c}{c+d} \\ P_1^{fw}(x_2[1]) = P(x_2[1]=1|y_1[1], y_2[1]) = \frac{d}{c+d} \end{cases}$$

where $\rho_i = \boldsymbol{U_{11}}(i,i)$. Using these successive calculations, $P_0^{fw}(x_k[n])$ and $P_1^{fw}(x_k[n])$ can be found for all values of $1 \le n \le N$ and $1 \le k \le K$. As explained before, due to the structure of the sampling method in Fig. 1.7, the last transmitted symbol can also be detected without interference and the same procedure can be applied backward to find $P_0^{bw}(x_k[n])$ and $P_1^{bw}(x_k[n])$. Using either of these likelihood sets as a detection

metric will result in an improvement over the hard-decision SIC method that was presented in the previous section. Moreover, the performance can be furthered improved if we use forward and backward operations together and define the detection metric as:

$$P_0(x_k[n]) = P_0^{fw}(x_k[n])P_0^{bw}(x_k[n])$$
$$P_1(x_k[n]) = P_1^{fw}(x_k[n])P_1^{bw}(x_k[n])$$

Simulation results are presented in Chapter 4 .

Chapter 3

Performance Analysis

In this section, the achievable performance by the asynchronous transmission is analyzed. Different criteria can be considered to show the efficiency of the asynchronous transmission. In this work, we use BER and the achievable rate. In the analysis of BER, one important factor is diversity gain which specifies the reliability of the transmission. In the analysis of the achievable rate, the asymptotic metric in high SNR is called the multiplexing gain. In this section, we analyze the diversity and multiplexing gain provided by the asynchronous transmission.

3.1 Diversity

Unlike the AWGN channel, the fading channel suffers from sudden declines in the power. This is due to the destructive addition of multi-path signals in the propagation media. Therefore, the effective signal-to-noise ratio (SNR) at the receiver can go through deep fades and be dropped dramatically. Usually we can assume a threshold

for the received SNR in which the receiver can reliably detect and decode the transmitted signal. If the received SNR is lower than such a threshold, a reliable recovery of the transmitted signal is impossible. and this event is called an outage. The probability of outage, which is the probability of having a received power lower than the given threshold, can be calculated based on the statistical model of the channel or based on the actual measurements of the channel. One way to combat the outage events is to provide different replicas of the transmitted signal to the receiver, which is the main idea behind diversity. If these different replicas fade independently, it is less probable to have all copies of the transmitted signal in deep fade simultaneously. Therefore, the receiver can reliably decode the transmitted signal using these received signals. To define diversity quantitatively, we use the relationship between the received SNR, denoted by δ , and the probability of error, denoted by P_e. A tractable definition of the diversity, or diversity gain, is

$$D = -\lim_{\delta \to \infty} \frac{\log P_e}{\log \delta} \tag{3.1}$$

where P_e is the error probability at an SNR equal to δ . In other words, diversity is the slope of the error probability curve in terms of the received SNR in a log-log scale [26].

As mentioned before, the conventional ZF detector is unable to fully exploit the available diversity gain which is M. However, we show that by using asynchronous transmission, all the interference can be canceled by one receive antenna and therefore, diversity gain of M can be achieved by existing diversity combining methods. In what follows, we show analytically that asynchronous ZF is able to provide full diversity of M where M is the number of receive antennas. The system represented in Eq. (2.12)

consists of NK subchannels, each of them having SNR of $\frac{E[|x_k[i]|^2]}{\boldsymbol{COV}_{\tilde{\boldsymbol{\nu}}}(i,i)}$, $1 \leq i \leq NK$, where $\boldsymbol{COV}_{\tilde{\boldsymbol{\nu}}}$ can be calculated as:

$$\boldsymbol{COV}_{\tilde{\boldsymbol{\nu}}} = E[\tilde{\boldsymbol{\nu}}\tilde{\boldsymbol{\nu}}^H] = \sigma_n^2(\boldsymbol{L_{tot}}^H \boldsymbol{\Sigma_{tot}}^{-1} \boldsymbol{L_{tot}})^{-1} \tag{3.2}$$

$$= \sigma_n^2(\sum_{i=1}^{M} \boldsymbol{L_i}^H \boldsymbol{\Sigma}^{-1} \boldsymbol{L_i})^{-1} \tag{3.3}$$

$$= \sigma_n^2(\sum_{i=1}^{M} \boldsymbol{H_i}^* \boldsymbol{R} \boldsymbol{H_i})^{-1} \tag{3.4}$$

where $\boldsymbol{L_i} = \bar{\boldsymbol{U}}\boldsymbol{H_i}$. In the derivation of $\boldsymbol{COV}_{\tilde{\boldsymbol{\nu}}}$, Eq. (3.3) is found by some matrix manipulation and Eq. (3.4) is obtained by using the fact that $\bar{\boldsymbol{U}}^H \boldsymbol{\Sigma}^{-1} \bar{\boldsymbol{U}} = \boldsymbol{R}$. Unfortunately, due to the complex structure of $\left(\sum_{i=1}^{M} \boldsymbol{H_i}^* \boldsymbol{R} \boldsymbol{H_i}\right)^{-1}$ for $M \geq 1$, finding the exact expression of BER for $M \geq 1$ is not easy. We derive an upper bound on BER by finding an upper bound on the diagonal elements of $\boldsymbol{COV}_{\tilde{\boldsymbol{\nu}}}$ and show that full diversity is achieved. Because $\boldsymbol{R}$ is positive definite, for every $1 \leq i \leq M$, $\boldsymbol{H_i^* R H_i}$ is also positive definite. Therefore, we can apply the following lemma.

Lemma 3.1. *For n positive definite matrices $\boldsymbol{A_i}, 1 \leq i \leq n$, we have:*

$$(\sum_{i=1}^{n} \boldsymbol{A_i})^{-1} \leq \sum_{i=1}^{n} \boldsymbol{A_i}^{-1} \tag{3.5}$$

where $\boldsymbol{B} \leq \boldsymbol{C}$ means that $\boldsymbol{C} - \boldsymbol{B}$ is positive semidefinite.

Proof. This lemma is a straightforward result of the following inequality, which can be found in [25].

$$(\boldsymbol{A} + \boldsymbol{B})^{-1} \leq \boldsymbol{A}^{-1} \quad \boldsymbol{A}, \boldsymbol{B} : \text{ positive definite matrices}$$

□

As a result, we can conclude that $\boldsymbol{COV_{\tilde{\nu}}} \leq \sigma_n^2 \sum_{i=1}^{M} (\boldsymbol{H_i}^* \boldsymbol{R} \boldsymbol{H_i})^{-1}$. This inequality implies that the diagonal elements of the covariance matrix of noise are upper bounded as follows:

$$\boldsymbol{COV_{\tilde{\nu}}}(i,i) \leq \frac{\sigma_n^2 \boldsymbol{R}^{-1}(i,i)}{\sum_{j=1}^{M} |h_{(1+(i-1)modK),j}|^2} \quad 1 \leq i \leq NK \tag{3.6}$$

where $h_{k,m}$ is the channel coefficient between User k and Receive Antenna m.
The BER expression for an AWGN channel with average transmit power of $E[|x_k[i]|^2]$ and noise variance of $\frac{\sigma_n^2 \boldsymbol{R}^{-1}(i,i)}{\sum_{j=1}^{M} |h_{((1+(i-1)modK),j)}|^2}$ is equal to:

$$p_i = \frac{\sqrt{\frac{\delta_0 2}{\pi \boldsymbol{R}^{-1}(i,i)}}}{2\left(1+\frac{\delta_0 2}{\boldsymbol{R}^{-1}(i,i)}\right)^{M+\frac{1}{2}}} \frac{\Gamma(M+\frac{1}{2})}{\Gamma(M+1)} \times {}_2F_1(1, M+\frac{1}{2}; M+1; \frac{1}{1+\frac{\delta_0 2}{\boldsymbol{R}^{-1}(i,i)}}) \tag{3.7}$$

where $\delta_0 = \frac{E[|x_k[i]|^2]}{\sigma_n^2}$. The details of derivation can be found in Appendix A. Due to having the same average transmit power and a lower noise variance, we conclude that BER for each subchannel is upper bounded by p_i, i.e., $P_i \leq p_i$. If we define $D_i = -\lim_{\delta_0 \to \infty} \frac{\log P_i}{\log \delta_0}$ and $d_i = -\lim_{\delta_0 \to \infty} \frac{\log p_i}{\log \delta_0}$, it is clear that $D_i \geq d_i$. By using the fact that the hypergeometric function of form ${}_2F_1(1, m+\frac{1}{2}; m+1; \frac{1}{1+c})$ converges to one as c grows large [47], we can calculate that $d_i = M$. Therefore, the diversity of the ith subchannel is greater than or equal to M. On the other hand, M is the maximum available diversity for this system, which completes the proof of achieving full diversity, i.e. $D_i = M$.

3.1.1 Effect of Time Delays on Performance

In this section, we calculate the optimal values of delays for the ZF detection in order to achieve the lowest average BER with one receive antenna at high SNR. Because for $M = 1$ the inequality in Eq. (3.6) turns into equality, the exact BER expression for each subchannel can be obtained as:

$$P_i = \frac{\sqrt{\frac{\delta_0 2}{\pi \boldsymbol{R}^{-1}(i,i)}}}{2\left(1+\frac{\delta_0 2}{\boldsymbol{R}^{-1}(i,i)}\right)^{3/2}} \frac{\Gamma(3/2)}{\Gamma(2)}\ {}_2F_1(1,3/2;2;\frac{1}{1+\frac{\delta_0 2}{\boldsymbol{R}^{-1}(i,i)}})$$

Approximating P_{avg} at high SNR for one receive antenna results in: (see Appendix B for more details)

$$\widetilde{P_{avg}} = \frac{1}{4\sqrt{\pi}NK}\frac{\Gamma(3/2)}{\Gamma(2)} \times \frac{\sum_i \boldsymbol{R}^{-1}(i,i)}{\delta_0},$$

For a fixed number of users and frame length, in order to minimize $\widetilde{P_{avg}}$, we need to minimize the trace($\boldsymbol{R}^{-1}$) which is related to time delays between different users. In what follows, we derive the relationship between the trace($\boldsymbol{R}^{-1}$) and time delays, and consequently find optimum time delays.

Lemma 3.2. *the sum of the diagonal elements of the inverse of matrix* $\boldsymbol{R}$ *is equal to:*

$$trace(\boldsymbol{R}^{-1}) = \frac{(N-1)(N+1)}{3(1+\tau_1-\tau_K)} + \frac{2N+1}{3(N+1+\tau_1-\tau_K)} + \frac{N(N+2)}{3}\sum_{i=1}^{K-1}\frac{1}{\tau_{i+1}-\tau_i} \tag{3.8}$$

The proof is presented in Appendix C.

Theorem 3.1. *The optimum time delays which result in the lowest average BER for*

ZF detection at high SNR are: (τ_1 is assumed to be zero)

$$\tau_{i-1} = \frac{i-2}{i-1} \times \tau_i \quad 3 \leq i \leq K \tag{3.9}$$

Also τ_K is found by solving the following equation:

$$A\tau_K^4 + B\tau_K^3 + C\tau_K^2 + D\tau_K + E = 0 \tag{3.10}$$

where

$$A = (1-(K-1)^2)\frac{(N+2)}{3}.$$

$$B = \frac{-2}{3}(1-(K-1)^2)N^2 + 2(4(K-1)^2-1)\frac{(N+1)}{3}.$$

$$C = \frac{1}{3}(1-(K-1)^2)N^3 + \frac{2}{3}(1-4(K-1)^2)N^2 - 2(K-1)^2(3N+2).$$

$$D = \frac{2}{3}(K-1)^2(N^3+5N^2+8N+4).$$

$$E = -\frac{1}{3}(K-1)^2(N^3+4N^2+5N+2).$$

The proof is easily obtained by taking the derivation of Eq. (3.8) with respect to time delays.

For $K = 2$, A will be zero and Eq. (3.10) is a polynomial of degree 3 which has a closed-form solution as follows:

$$\tau_{opt} = \frac{N+2-\sqrt[3]{N^3+1.5N^2-1.5N-1}}{3} \tag{3.11}$$

where N is the block length. However, for other values of K, Eq. (3.10) should be solved numerically. After finding τ_K, the remaining time delays are calculated recursively using Eq. (3.9). The optimum delay values for different K and N values

are reported in Tables 3.1 and 3.2. Optimum time delays approach uniform time

Table 3.1: Optimum Time Delays when $K = 2$

Case	N=10	N=32	N=64	N=128	N $\to \infty$
K=2	0.5240	0.5077	0.5039	0.5019	0.5

Table 3.2: Optimum delays when $N = 128$

Case	N=128
K=4	[0.2505,0.5010,0.7514]
K=6	[0.1669,0.3338,0.5006,0.6675,0.8344]
K=8	[0.1251,0.2502,0.3754,0.5004,0.6256,0.7507,0.8758]

delays, i.e, $\tau_k = \frac{k-1}{K}$, $2 \leq k \leq K$, as N increases. The effects of time delay values on the performance are studied numerically in the following section.

In this section, we showed that unlike synchronous ZF, asynchronous ZF can provide full diversity of M. If we assume $M > K$, then transmitting independent symbols from different users and using ML detection can provide diversity order of M. However, conventional ZF detection wastes $K - 1$ diversity orders to null the interferences resulting in diversity order of $M - K + 1$. On the other hand, if $M < K$, we are only able to transmit maximum number of M symbols which is apprehend as the multiplexing gain. At the receiver side, M diversity gain is available which can be exploited by ML detection, however, the ZF detection is unable to decode the transmitted symbols because the null space of the channel matrix is empty. These diversity results are summarize in Table 3.3

In summary, a MIMO system can provide two types of gains: diversity gain and spatial multiplexing gain. Given a MIMO channel, both gains can, in fact, be simultaneously obtained, but there is a fundamental trade-off between how much of each type of gain any coding scheme can extract: higher multiplexing gain comes at the price of

Table 3.3: Diversity Gain results

Diversity gain	$M > K$	$M < K$
synchronous ML	M	M
synchronous ZF	$M - K + 1$	unable to decode
Asynchronous	M	M

sacrificing diversity [50]. However, by means of asynchronous transmission, another domain for exploiting multiplexing gain is provided which let us enjoy full diversity provided by spatial domain without sacrificing the multiplexing gain. Multiplexing gain is particularly important in the high-SNR regime where the system is degree-of-freedom limited and is defined as:

$$G_m = \lim_{\delta \to \infty} \frac{R}{\log \delta} \tag{3.12}$$

where R is the achievable rate. In the next section, the results for multiplexing gain are provided.

3.2 Achievable Rate and Multiplexing Gain

In this section, we consider AWGN channels to simplify the analysis. However, the results can be extended to fading channels. For the conventional synchronous transmission systems, the achievable rate for AWGN channels is:

$$C_{synch} = log_2 \left(1 + \frac{P}{\sigma_n^2}\right) \tag{3.13}$$

where P is the transmit power and σ_n^2 is the variance of noise. This rate is achieved by Gaussian symbols with variance of P and coding over infinite time slots. The ratio of

$\frac{P}{\sigma_n^2}$ is understood as signal to noise ratio (SNR), i.e., δ and thus the multiplexing gain of the channel is $\lim_{\delta\to\infty} \frac{C}{\log \delta} = 1$.

The system model for an asynchronous transmission with K shifted sub-streams in AWGN channels can be written as:

$$\boldsymbol{y'}_{NK\times 1} = \boldsymbol{R'}_{NK\times NK}\boldsymbol{x}_{NK\times 1} + \boldsymbol{n'}_{NK\times 1} \tag{3.14}$$

where $\boldsymbol{R'}$ is a block-Toeplitz matrix and depends on time delays and $\boldsymbol{n'}$ is the noise vector with a covariance matrix of $\boldsymbol{R'}$. As shown in Chapter 1, matrix $\boldsymbol{R'}$ is a positive definite matrix for all time-limited pulse shapes. This system model can be regarded as a point to point system where the sub-streams are superimposed (P2P) or a multi-node system where shifted sub-streams are transmitted from separate nodes (MP2P). In the P2P case, $x(t) = \sum_{k=1}^{K} x_k(t)$ is transmitted from a single node to the receiver node. In the MP2P case, each node is transmitting one of the shifted sub-streams. For example, Node 1 transmits $x_1(t)$, Node 2 transmits $x_2(t)$ and so forth. Both of these scenarios result in the same system model at the receiver side, denoted in Eq. (3.14). The achievable rate for the system in Eq. (3.14) can be written as [19]:

$$\begin{aligned} C_{asynch} &= \lim_{N\to\infty} \frac{1}{N+1} \max_{\boldsymbol{Q}} log_2 \left(\left| \left(\boldsymbol{R'}\sigma_n^2 + \boldsymbol{R'Q}\boldsymbol{R'}^H \right) \frac{\boldsymbol{R'}^{-1}}{\sigma_n^2} \right| \right) \\ &= \lim_{N\to\infty} \frac{1}{N+1} \max_{\boldsymbol{Q}} log_2 \left(\left| \boldsymbol{I_{NK}} + \frac{1}{\sigma_n^2}\boldsymbol{R'Q} \right| \right) \end{aligned} \tag{3.15}$$

Where $\boldsymbol{Q}$ is the covariance matrix of the transmitted vector and the maximization is over all possible covariance matrices which satisfy the power constraint. However, the two aforementioned scenarios, i.e., P2P and MP2P, result in different power constraints. For a fair comparison, let us assume that for the both cases, the total transmit power during N time slots is NP.

3.2.1 P2P case:

Denoting the sub-stream corresponding to the time shift τ_k as $x_k(t)$, the power constraint for the P2P case can be written as:

$$E\left[\int_{-\infty}^{\infty}\left(\sum_{k=1}^{K}x_k(t)\right)\left(\sum_{k=1}^{K}x_k(t)\right)^{*}dt\right]\leq NP \tag{3.16}$$

The power constraint can be simplified in the following steps:

$$E\left[\int_{-\infty}^{\infty}\left(\sum_{k=1}^{K}x_k(t)\right)\left(\sum_{k=1}^{K}x_k(t)\right)^{*}dt\right]=$$
$$E\left[\int_{-\infty}^{\infty}\left(\sum_{k=1}^{K}\sum_{n=1}^{N}x_k[n]p(t-(n-1)T-\tau_k)\right)\left(\sum_{k=1}^{K}\sum_{n=1}^{N}x_k[n]p(t-(n-1)T-\tau_k)\right)^{*}dt\right]=$$
$$\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}E\left[x_{k_1}[n_1]x_{k_2}[n_2]\right]\int_{-\infty}^{\infty}p(t-(n_1-1)T-\tau_{k_1})p(t-(n_2-1)T-\tau_{k_2})dt=$$
$$\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}trace(\boldsymbol{R'_{k_1k_2}}COV[\boldsymbol{x_{k_1}},\boldsymbol{x_{k_2}}])=trace(\boldsymbol{R'}COV[\boldsymbol{x}])$$

Thus, the power constraint for the P2P case can be written as:

$$trace(\boldsymbol{R'Q})\leq NP \tag{3.17}$$

Then, the corresponding achievable rate can be written as:

$$C_{asynch}^{P2P}=\lim_{N\to\infty}\frac{1}{N+1}\max_{\boldsymbol{Q}}log_2|\boldsymbol{I_{NK}}+\frac{1}{\sigma_n^2}\boldsymbol{R'Q}|$$
$$\boldsymbol{s.t.}\quad trace(\boldsymbol{R'Q})\leq NP \tag{3.18}$$

Matrix $\boldsymbol{R'}$ is a Hermitian positive definite matrix, thus its SVD decomposition can be written as:

$$\boldsymbol{R'} = \boldsymbol{U}_{R'} \begin{bmatrix} \lambda_0 & 0 & \cdots & 0 \\ 0 & \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{NK-1} \end{bmatrix} \boldsymbol{U}_{R'}^H \tag{3.19}$$

$$= \boldsymbol{U}_{R'} \boldsymbol{\Lambda} \boldsymbol{U}_{R'}^H \tag{3.20}$$

where $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{NK-1}$ are the eigenvalues of matrix $\boldsymbol{R'}$ and $\boldsymbol{U}_R$ is Unitary matrix. Because the time delays are known at the transmitter, the matrix $\boldsymbol{R'}$ is known at the transmitter. Hence, the transmitted symbols can be beam-formed in the direction of eigen-vectors of matrix $\boldsymbol{R'}$ and the powers to each sub-channel can be assigned based on the strength of the corresponding eigen-value. Consequently, the SVD decomposition of matrix $\boldsymbol{Q}$ can be denoted as $\boldsymbol{Q} = \boldsymbol{U}_{R'} \boldsymbol{P} \boldsymbol{U}_{R'}^H$ where $\boldsymbol{P}$ is the diagonal matrix containing the assigned powers to each symbol, i.e., $\boldsymbol{P} = diag[P_0, P_2, \cdots, P_{NK-1}]$. Then, the product of $\boldsymbol{R'Q}$ is equal to:

$$\boldsymbol{R'Q} = \boldsymbol{U}_{R'} \begin{bmatrix} P_0\lambda_0 & 0 & \cdots & 0 \\ 0 & P_1\lambda_1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{NK-1}\lambda_{NK-1} \end{bmatrix} \boldsymbol{U}_{R'}^H \tag{3.21}$$

Therefore, the achievable rate can be further simplified to:

$$C_{asynch}^{P2P} = \lim_{N\to\infty} \frac{1}{N+1} \max_{\boldsymbol{P}} \sum_{n=0}^{NK-1} log_2\left(1+\frac{P_i\lambda_i}{\sigma_n^2}\right)$$
$$\boldsymbol{s.t.} \quad \sum_{n=0}^{NK-1} P_i\lambda_i \leq NP \tag{3.22}$$

The optimal answer for the $P2P$ maximization problem is very simpler and is obtained by assigning P_is such that $P_0\lambda_0 = P_1\lambda_1 = \cdots = P_{NK-1}\lambda_{NK-1} = \frac{P}{K}$. As a result, C_{asynch}^{P2P} will be:

$$C_{asynch}^{P2P} = Klog_2\left(1+\frac{P}{K\sigma_n^2}\right)$$

Now, the multiplexing gain of the P2P case can be easily calculated as:

$$G_m^{P2P} = \lim_{\delta\to\infty} \frac{C_{asynch}^{P2P}}{\log\delta} \tag{3.23}$$
$$= \lim_{\delta\to\infty} \frac{Klog_2(1+\frac{\delta}{K})}{\log\delta} = K \tag{3.24}$$

As mentioned before, despite synchronous transmission which only provides one degree of freedom, by means of asynchronous transmission K degrees of freedom can be achieved where K is the number of shifted sub-streams with distinct timing offsets. Next, the achievable rate in the MP2P case is analyzed.

3.2.2 MP2P case:

Denoting the sub-stream corresponding to the kth node as $x_k(t)$, the power constraint can be written as:

$$\sum_{k=1}^{K} E\left[\int_{-\infty}^{\infty} x_k(t)x_k(t)^* dt\right] \leq NP \tag{3.25}$$

Each of the summands can be simplified to:

$$
\begin{aligned}
E\left[\int_{-\infty}^{\infty} x_k(t)x_k(t)^* dt\right] &= E\left[\int_{-\infty}^{\infty}\left(\sum_{n=1}^{N} x_k[n]p(t-(n-1)T)\right)\left(\sum_{n=1}^{N} x_k[n]p(t-(n-1)T)\right)^* dt\right] \\
&= \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} E[x_k[n_1]x_k[n_2]^*]\int_{-\infty}^{\infty} p(t-(n_1-1)T)p(t-(n_2-1)T)dt \\
&= \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} E[x_k[n_1]x_k[n_2]^*]\delta[n_1-n_2] \\
&= \sum_{n=1}^{N} E[x_k[n]x_k[n]^*] \\
&= trace\left(COV(\boldsymbol{x_k})\right)
\end{aligned}
$$

Denoting $COV(\boldsymbol{x_k})$ as $\boldsymbol{Q_k}$, we can write the power constraint as:

$$\sum_{k=1}^{K} trace(\boldsymbol{Q_k}) \leq NP \quad or \quad trace(\boldsymbol{Q}) \leq NP \tag{3.26}$$

Then, the corresponding achievable rate can be written as:

$$
\begin{aligned}
C_{asynch}^{MP2P} = \lim_{N\to\infty} \frac{1}{N+1} \max_{\boldsymbol{Q}} log_2\left(\left|\boldsymbol{I_{NK}} + \frac{1}{\sigma_n^2}\boldsymbol{R'Q}\right|\right) \\
\boldsymbol{s.t.} \quad trace(\boldsymbol{Q}) \leq NP
\end{aligned} \tag{3.27}
$$

Similar to the P2P case, by using the knowledge of eigen-vectors of matrix $\boldsymbol{R'}$ and proper beam-forming, the achievable rate can be simplified as:

$$
\begin{aligned}
C_{asynch}^{MP2P} = \lim_{N\to\infty} \frac{1}{N+1} \max_{\boldsymbol{P}} \sum_{n=0}^{NK-1} log_2\left(1+\frac{P_i\lambda_i}{\sigma_n^2}\right) \\
\boldsymbol{s.t.} \quad \sum_{n=0}^{NK-1} P_i \leq NP
\end{aligned} \tag{3.28}
$$

The optimal power assignment for this maximization problem is obtained by the well-known water filling algorithm. The intuitive behind the water filling algorithm is to assign more power to more reliable sub-channels. It can be described mathematically as follows.

$$P_i = \begin{cases} \mu - \frac{\sigma_n^2}{\lambda_i} & \mu \geq \frac{\sigma_n^2}{\lambda_i} \\ 0 & \mu \leq \frac{\sigma_n^2}{\lambda_i} \end{cases} \tag{3.29}$$

where μ is some cutoff value to make sure the total power constraint is met [45]. In low SNR regime, the capacity achieving method of water filling forces us to drop all the sub-channels except the strongest one and take advantage of only one degree of freedom. However, as available transmit power increases, the algorithm becomes less strict on choosing affordable sub-channels. In other words, in high SNR regime, i.e, $\frac{P}{\sigma_n^2} \to \infty$, all NK sub-channels will be utilized and also be assigned the same power. To calculate the degree of freedom provided by the MP2P method and analyze C_{asynch}^{MP2P}, let us look at the water-filling algorithm from another perspective.

For simplicity, assume that σ_n^2 is equal to one. As described before, power is assigned to each eigenvector based on the reciprocal of the corresponding eigenvalue. Fig. 3.1 also illustrates the concept of water-filling algorithm. For example, if μ is between $\frac{1}{\lambda_{n-1}}$ and $\frac{1}{\lambda_n}$, then power is assigned only to n largest eigenvalues in a way that:

$$\sum_{i=0}^{n-1} \left(\mu - \frac{1}{\lambda_i} \right) = NP \tag{3.30}$$

which results in $\mu = \frac{1}{n} \left(NP + \sum_{i=0}^{n-1} \frac{1}{\lambda_i} \right)$. Hence, if P is between $\frac{1}{N} \left(\frac{n-1}{\lambda_{n-1}} - \sum_{i=0}^{n-2} \frac{1}{\lambda_i} \right)$

and $\frac{1}{N}\left(\frac{n}{\lambda_n} - \sum_{i=0}^{n-1}\frac{1}{\lambda_i}\right)$, then the achievable rate will be equal to:

$$\begin{aligned} C &= \frac{1}{N+1}\sum_{i=0}^{n-1}\log(1+P_i\lambda_i) \\ &= \frac{n}{N+1}\log\mu + \frac{1}{N+1}\log\left(\prod_{i=0}^{n-1}\lambda_i\right) \\ &= \frac{n}{N+1}\log\left(\frac{NP}{n}+a\right) + \frac{n}{N+1}\log b \end{aligned} \tag{3.31}$$

where $a = \frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{\lambda_i}$ and $b = \sqrt[n]{\prod_{i=0}^{n-1}\lambda_i}$.

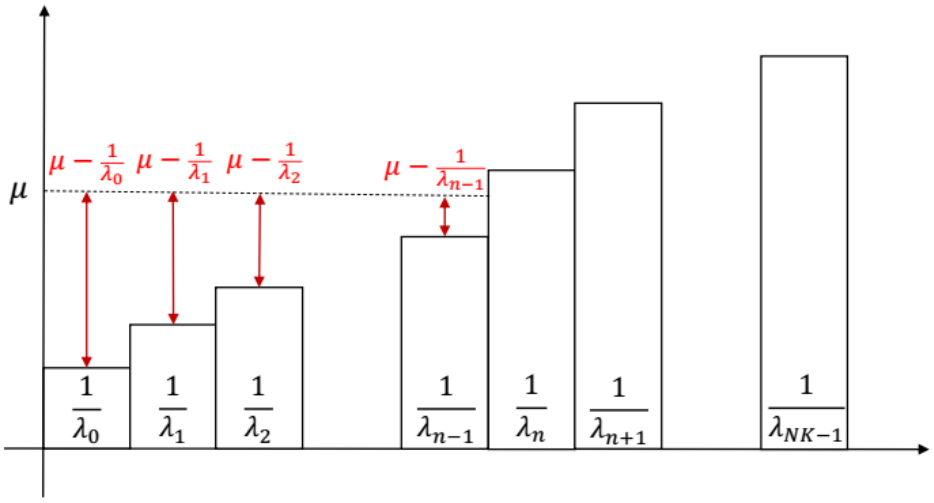


Figure 3.1: Water-filling Algorithm

These results can be generalized as follows:

$$\begin{cases} P \le t_1 \Rightarrow C_1(P) = \frac{1}{N+1}\log\left(NP + \frac{1}{\lambda_0}\right) + \frac{1}{N+1}\log\lambda_0 \\ t_{n-1} \le P \le t_n \Rightarrow C_n(P) = \frac{n}{N+1}\log\left(\frac{NP}{n} + a_n\right) + \frac{n}{N+1}\log b_n \\ P \ge t_{NK-1} \Rightarrow C_{NK}(P) = \frac{NK}{N+1}\log\left(\frac{P}{K} + a_{NK}\right) + \frac{NK}{N+1}\log b_{NK} \end{cases} \tag{3.32}$$

where $a_n = \frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{\lambda_i}$, $b_n = \sqrt[n]{\prod_{i=0}^{n-1}\lambda_i}$ and $t_n = \frac{1}{N}\left(\frac{n}{\lambda_n} - \sum_{i=0}^{n-1}\frac{1}{\lambda_i}\right)$. Therefore, the

degrees of freedom provided by the MP2P case depends on the available transmit power. Let us consider two extreme cases. First, assume that, $P \leq t_n$ where $\lim_{N\to\infty} \frac{n}{N} = 0$, then $C_{asynch}^{MP2P} = 0$. The reason is that when the transmit power is low such that among infinite number of available sub-channels only a few of them (not comparable with N) are used, then the total rate approaches zero. However, as P increases, the water-filling method assigns power to more sub-channels resulting in higher degree of freedom. For a sufficiently large P, we will have $C_{asynch}^{MP2P} = K \log\left(\frac{P}{K\sigma_n^2} + A\right) + KB$ where A and B are defined as follows:

$$A = \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{f(w)} dw \tag{3.33}$$

$$B = \frac{1}{2\pi} \int_0^{2\pi} \log_2 f(w) dw \tag{3.34}$$

where $f(w)$ is the $\frac{K}{T}$-folded spectrum of the pulse shape defined in Chaper 1. The derivation of this result is based on the simple use of the Szego theorem stated in Chapter 1. If we let $P \to \infty$, we can easily show that the MP2P case also provides a multiplexing gain of K, i.e.,:

$$G_m^{MP2P} = \lim_{\delta\to\infty} \frac{C_{asynch}^{MP2P}}{\log \delta} \tag{3.35}$$

$$= \lim_{\delta\to\infty} \frac{K \log\left(\frac{\delta}{K} + A\right) + KB}{\log \delta} = K \tag{3.36}$$

Therefore, asynchronous transmission also provides a multiplexing gain of K in the MP2P cases. Note that due to use of the water-filling algorithm, the performance of MP2P depends on the eigenvalues of matrix $\boldsymbol{R'}$, while the performance of the P2P case is independent of the eigenvalues as long as they are nonzero. Also note that the concept of superimposing shifted sub-streams can be applied to the MP2P case. Thus, a system with J points transmitting to a common receiver and each of them

superimposing K sub-streams will result in JK multiplexing gain as long as the timing delays are distinct. The summary of results for multiplexing gain are provided in Table 3.4.

Table 3.4: Multiplexing Gain results

Multiplexing gain	P2P	MP2P
synchronous ML	1	1
synchronous ZF	1	unable to decode
Asynchronous	K	K

In the next section, the simulation results for BER and achievable rates, some insightful discussion and future works are presented.

Chapter 4

Simulation Results and Discussions

In this section, we provide simulation results in order to validate our theoretical results and compare different methods. We present the results in two sections. First, the BER results in the Rayleigh fading channel are presented and the diversity gain is analyzed. In the second part, the achievable rate results for the AWGN channel are presented and the multiplexing gain is analyzed.

4.1 BER Results: Diversity Gain Analysis

In this part, we present the BER results in the Rayliegh fading channel. In all simulations, channel coefficients are independent Rayliegh fadings with variance one, fixed during the block and changing independently for each block. All users have the same average power of one and variance of noise (σ_n^2) is equal to $10^{\frac{-\text{SNR}}{10}}$ where SNR is in dB. To avoid inter-block interference, the last symbol of each block should be idle for asynchronous methods. This will reduce spectral efficiency, but it is negligible for

large block lengths. In all simulations, the block length is 128 and the time delays are uniform except in the case where we report the time delays to study their effects on the performance. The number of users and the number of receive antennas is denoted by K and M, respectively. When M is not specified, the assumption is that only one receive antenna is used. Transmitted symbols are chosen from BPSK modulation and the comparing criterion is the average bit error rate among all the users.

In Fig. 4.1,we compare the performance of the asynchronous MLSD method with that of the synchronous ML. Asynchronous MLSD outperforms synchronous ML detection with similar complexity. Fig. 4.1 also includes the single-user bound for a better comparison. As can be seen in the figure, asynchronous MLSD for $K = 2$ achieves performance of the single user system at high SNR.

Fig. 4.2 shows the performance of different SIC methods presented in Section 2.3.

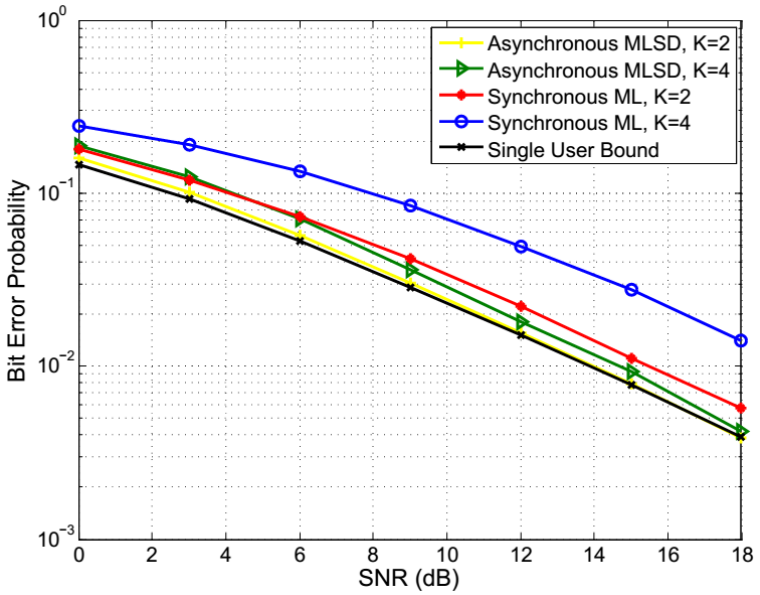


Figure 4.1: Comparing asynchronous MLSD and synchronous ML

Our new forward backward belief propagation method using the sampling method in

Fig. 1.7 improves the performance of traditional SIC method by about 3 dB.

Fig. 4.3 compares the performance of the synchronous and asynchronous ZF detec-

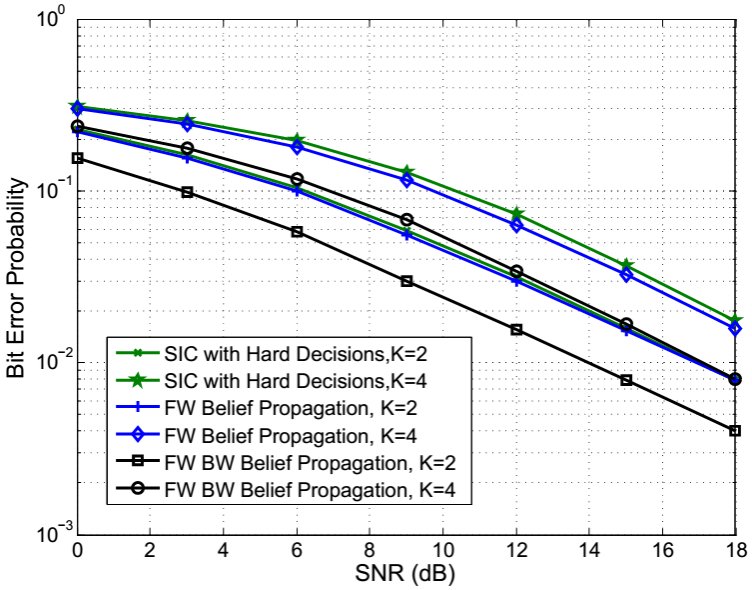


Figure 4.2: Performance of SIC method with hard decisions and soft decisions

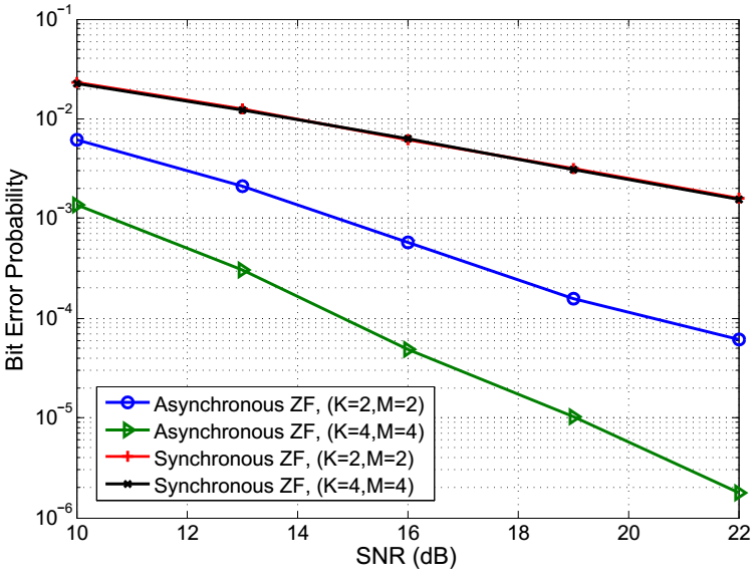


Figure 4.3: Comparing synchronous and asynchronous ZF

tors. Although asynchronous ZF is even possible with one receive antenna, for fair

comparison, we consider the cases where the number of receive antennas and users are the same. Since all users are assumed to have the same transmit power, synchronous ZF for $(K = 2, M = 2)$ and $(K = 4, M = 4)$ provides the same performance and both of them have diversity of one. However, for asynchronous ZF detection, diversity of 2 and 4 is achieved for $(K = 2, M = 2)$ and $(K = 4, M = 4)$, respectively.

We study the effects of time delay values on the performance of a ZF system with $K = 4$ users and one receive antenna in Fig. 4.4. Note that a synchronous ZF solution does not exist in this case as we need at least $M = 4$ receive antennas. We show the results for six different sets of time delays. For optimum time delays we use the result of Section 3.1.1 as reported in Table 3.2. The curve associated with random time delays represents the average performance over uniformly distributed random time delays. The remaining sets of time delays are specified in the figure. The optimum time delays and time delays of $[0.01, 0.1, 0.9]$ have the best and worst performances, respectively. They also have the lowest and the highest trace$(\boldsymbol{R}^{-1})$, respectively, which are presented along with other sets of time delays in Table 4.1. As can be seen, a lower trace$(\boldsymbol{R}^{-1})$ results in a better performance. This observation is in line with the analysis in Section 3.1.1 where trace$(\boldsymbol{R}^{-1})$ was introduced as a criterion to compare the performance of different time delays.

Finally, to compare different methods with each other, we include the performance

Table 4.1: Comparing trace$(\boldsymbol{R}^{-1})$ for different time delays in Fig. 4.4

Time delays	trace$(\boldsymbol{R}^{-1})$
$[0.2505, 0.5010, 0.7514]$	8.8404×10^4
$[0.4, 0.6, 0.8]$	9.6639×10^4
$[0.1, 0.4, 0.7]$	1.1065×10^5
$[0.1, 0.2, 0.9]$	1.7347×10^5
$[0.01, 0.1, 0.9]$	6.7784×10^5

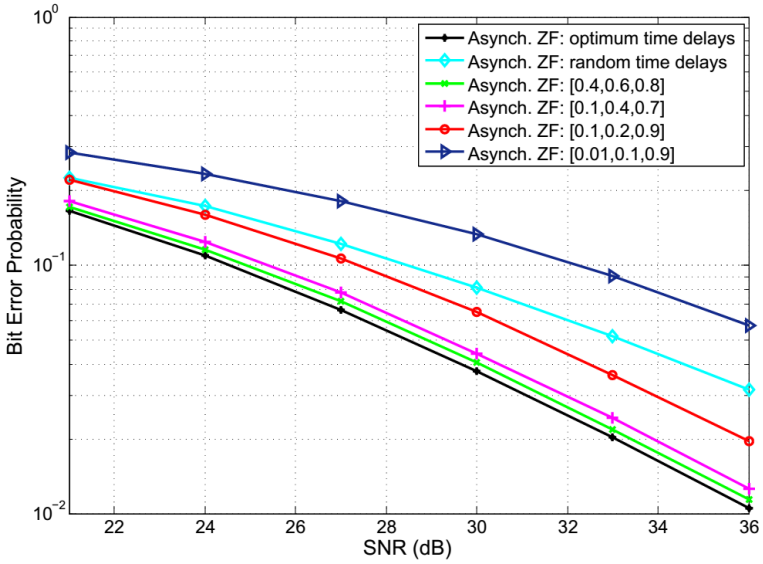


Figure 4.4: Effect of time delays in asynchronous ZF detection for $K = 4$

of all detection methods for $K = 2$ in Fig. 4.5. Both MLSD and forward-backward BP detection methods not only outperform the synchronous ML detection, but also achieve the performance of the single user system. In addition, the low complexity method of SIC with hard decisions also provides good performance.

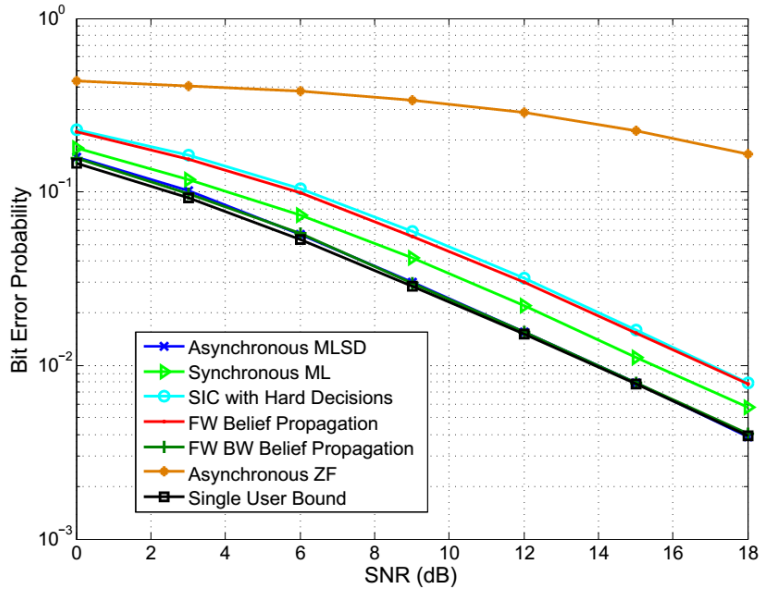


Figure 4.5: Comparison of all detection methods for $K = 2$

4.2 Achievable Rate: Multiplexing Gain Analysis

In this part, the achievable rate results are provided. The achievable rates by synchronous transmission, asynchronous transmission in P2P case and asynchronous transmission in MP2P case are compared. Different pulse shapes including Rectangular, R.R.C. and Sinc pulse shapes are considered to show the effect of pulse shapes in the performance of the asynchronous transmission. Note that the achievable rate depends on the occupied spectrum, therefore comparing different pulse shapes like Rectangular pulse shape and R.R.C. pulse shape is not fair. However, here, our main goal is comparing the synchronous and asynchronous transmission using the same pulse shape of which the bandwidth is identical. Hence, the notion of occupied bandwidth is neglected. Asynchronous results are provided for three different values of K, i.e., 2, 4 and 6. As you can see, in all the figures with different pulse shapes, the asynchronous transmissions in both P2P and MP2P cases outperform the synchronous transmission. In the P2P case where the shifted signals are superimposed, the performance is independent of the pulse shape but in the MP2P case, the performance depends on the pulse shape. Wide range of SNR is considered to effectively show the difference in the multiplexing gain provided by different methods.

In Fig. 4.6, Rectangular pulse shape is used. The synchronous transmission provides the multiplexing gain of one which is expected. The asynchronous transmissions in both P2P and MP2P cases provides higher multiplexing gain. It can be observed from the slope of the figures that the provided multiplexing gain for the asynchronous transmission with $K = 2, 4$ and 6 are about $2, 4$ and 6 times the synchronous one. Although the slopes for the P2P and MP2P curves are the same, at low SNR, the MP2P case outperforms the P2P case, while at the high SNR, P2P provides higher

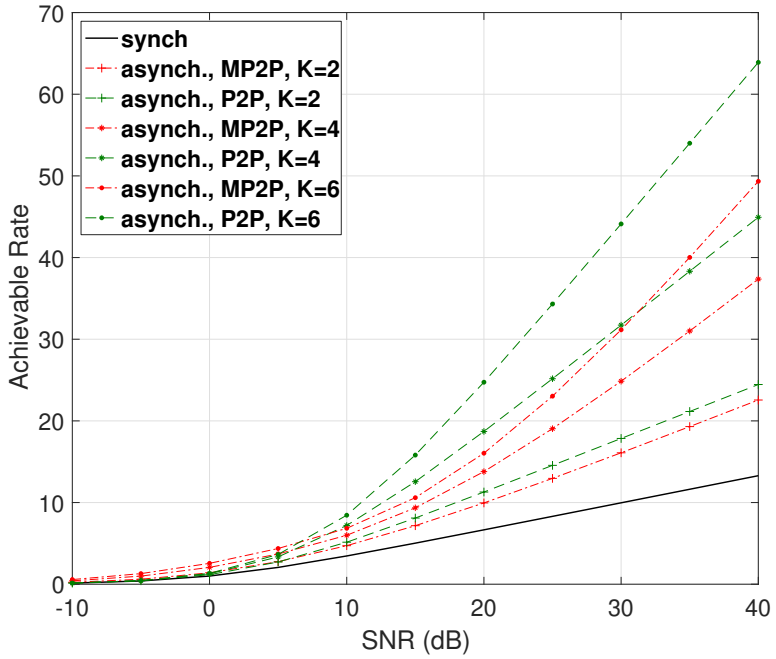


Figure 4.6: Comparison of the achievable rates obtained by synchronous and asynchronous transmission using Rectangular pulse shape

achievable rate.

As explained before, the reason behind the superiority of asynchronous transmission is the non-zero spectrum outside the Nyquist bandwidth, i.e., $\frac{1}{T}$. Rectangular pulse shape has infinite bandwidth, which provides the best opportunity for the asynchronous transmission. In the conventional synchronous transmission using Nyquist pulse shapes, the folded spectrum is a flat function independent of the pulse shape which ignores the out-of-bound spectrum. However, in the asynchronous transmission, the $\frac{K}{T}$-folded spectrum depends on the pulse shape and the more residual out-of-bound spectrum the more degree of freedom is available to exploit. On the other hand, if the spectrum is strictly band-limited, the asynchronous transmission superiority becomes marginal in the MP2P case.

In Fig. 4.7, Sinc pulse shape is considered. Sinc function is a perfect Nyquist pulse

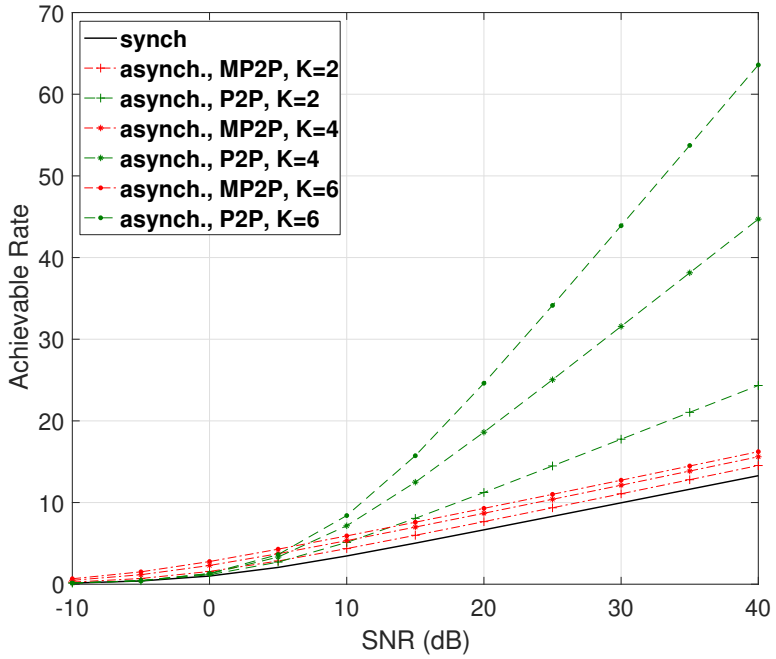


Figure 4.7: Comparison of the achievable rates obtained by synchronous and asynchronous transmission using Sinc pulse shape

shape meaning that it is strictly band-limited to the $[-\frac{1}{T}, \frac{1}{T}]$ interval. In practice, even Sinc function can only be realized in a time-limited fashion, thus the spectrum is always non-zero, however, the nonzero values outside $[-\frac{1}{T}, \frac{1}{T}]$ are very small. Hence, the additional eigenvalues provided by the asynchronous transmission are very close to zero but still positive. In the P2P case, having nonzero eigenvalues is sufficient to exploit higher degrees of freedom because the power assignments help to compensate for the small eigen-values. Thus, asynchronous transmission in the P2P case is able to provide higher multiplexing gain even with the band-limited pulse shapes. However, in the MP2P scenario where the powers are assigned by using the water filling algorithm, the small eigen-values are neglected, thus the multiplexing gain is exactly the same as the synchronous transmission. Still some marginal improvement is provided in the P2P case which is because of the mentioned small eigen-values. Asymptotically, if an infinite transmit power is available, even the very small eigen-values are utilized by the

water-filling algorithm resulting in the same multiplexing gain as the P2P case.

In Figs 4.8 and 4.9, the R.R.C pulse shape is used. R.R.C. pulse shape is parametrized by β which is called the roll-of factor. The roll-of factor can be any number in the interval $[0, 1]$. The higher the roll-of factor the more out-of-bound spectrum exists. One can already guess that asynchronous transmission in the $MP2P$ case can provide higher achievable rate when β is higher which is verified by the simulation results.

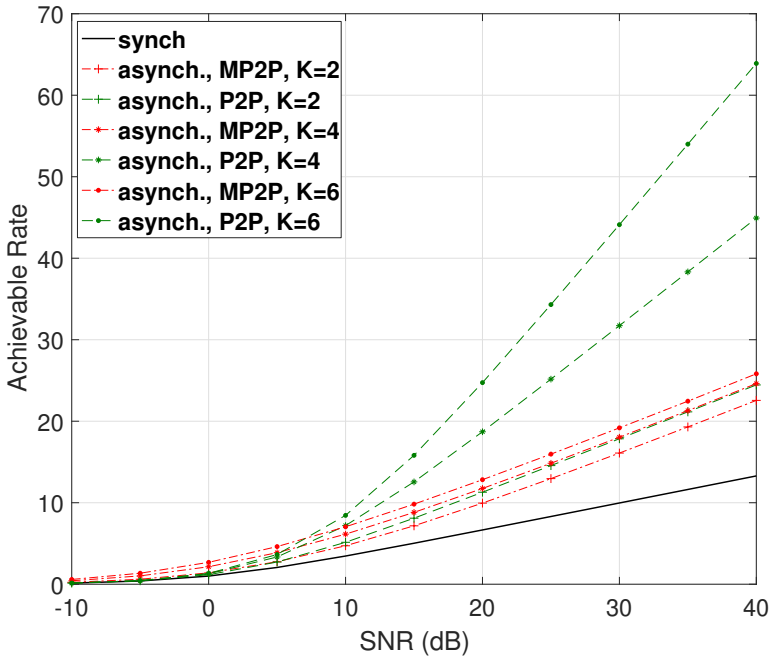


Figure 4.8: Comparison of the achievable rates obtained by synchronous and asynchronous transmission using R.R.C pulse shape ($\beta = 1$)

In Fig. 4.8, β is 1 and the MP2P case provides a multiplexing gain of 2. This is perfectly aligned with our expectation because when $\beta = 1$, the out-of-bound spectrum is exactly equal to the Nyquist bandwidth which makes the whole spectrum interval twice the one in the synchronous transmission. In the MP2P case, the mentioned out-of-bound spectrum is utilized by using $K = 2$. However, further increase of K does not change the multiplexing gain. On the other hand, in the $P2P$ case, where having

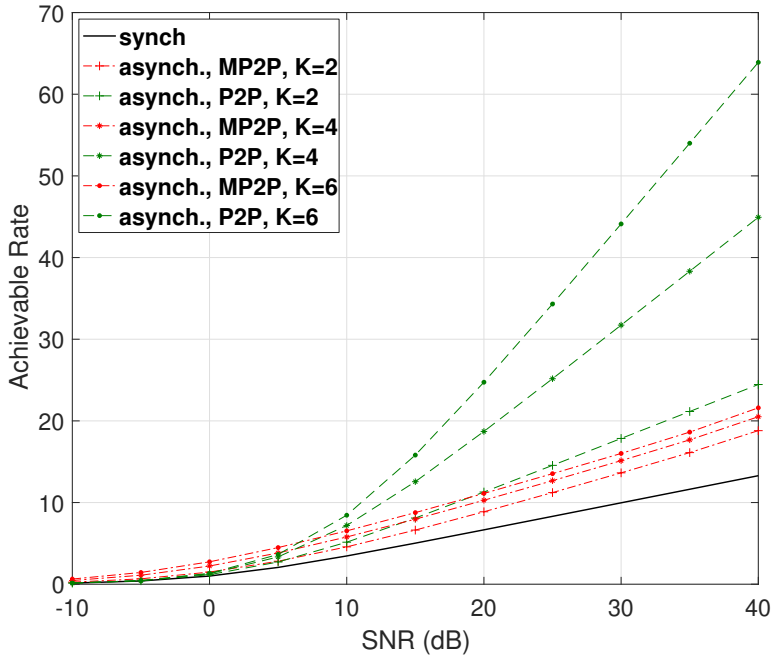


Figure 4.9: Comparison of the achievable rates obtained by synchronous and asynchronous transmission using R.R.C pulse shape ($\beta = 0.5$)

nonzero out-of-bound spectrum is sufficient, a multiplexing gain of K is achieved like the previous pulse shapes. In Fig. 4.9, β is 0.5, thus the multiplexing gain in the MP2P case is around 1.5 which is the same for all values of K. Note the multiplexing gain is theoretically defined as the ratio of the achievable rate to the SNR while SNR goes to infinity. Sometimes, we misused the multiplexing term to just show the aforementioned ratio at some finite SNRs. As shown in the mathematical analysis, asynchronous transmission provides a multiplexing gain of K in both P2P and MP2P cases. However, to see it in the MP2P case we literally need infinite transmit power because the extra eigen-values are very close to zero. Thus, with band-limited pulse shapes and in the practical range of SNR, the full potential of the asynchronous transmission is exploited only in the P2P case.

4.3 Discussions and Future work

In this work, we studied the benefits of asynchrony when multiple users are sending data simultaneously to a common receiver. Instead of treating asynchrony as a disruptive factor, we exploited it as an additional resource to cancel interference. We have shown that asynchrony between data streams adds a favorable ISI which makes interference cancellation possible. It also introduces memory to the system which can be exploited by methods like maximum-likelihood sequence detection. In addition to MLSD, a novel forward-backward belief propagation detection method was presented and this method outperforms synchronous ML detection. Exact BER expression for ZF detection was derived and it was verified that a diversity equal to the number of receive antennas is achievable by asynchronous transmission.

It is also shown that asynchronous transmission can provide higher degrees of freedom with respect to synchronous transmission. The concept of asynchrony can be used either with superimposing shifted time delays called P2P or consider the time delays across different users/transmit antennas called MP2P. These two scenarios are analyzed in this work. We analytically showed that both of these structures are able to provide a multiplexing gain of K where K is the number of distinct time delays in the system. It is shown that in the P2P scenario the multiplexing gain of K is realized even in the practical range of SNR and independent of the used pulse shape. However, when band-limited pulse shapes like Sinc and R.R.C pulse shapes are utilized , the MP2P case is unable to exploit the full potential of asynchrony in the practical range of SNR. Nevertheless, analytically speaking, it still provides a multiplexing gain of K because the term multiplexing is defined when $P \to \infty$.

In this work, we considered fairly simple system models to demonstrate the potentials of

asynchronous transmission. For example, we considered a flat fading model to analyze the diversity gain and an AWGN model to analyze the multiplexing gain. Thus, further work is required to analyze these concepts under more complicated and more practical models like frequency selective channels. Frequency selectivity will provide multiple replicas with different time delays at the receiver which is handled by using equalizers. The structure of equalizers can be complex in the time domain, which motivates the use of multi-carrier transmission like Orthogonal Frequency Division Multiplexing (OFDM) easing the equalization process [31]. Thus, it is vital to analyze the asynchronous transmission in the conjunction with time domain equalizers and also in the multi-carrier systems. Some rough analysis have shown the same advantages in the use of frequency offsets in the multi-carrier domain. However, more detailed analysis is required which is the topic of our future work.

In addition, each of the mentioned methods has its own issues. For example, in the P2P case where the multiplexing gain of K is achieved, the shifted streams are superimposed resulting in high Peak to Average Power Ratio (PAPR) which needs to be addressed. There are many work on this concept within the OFDM framework [27], thus the applicability of those works to our system model needs to be examined. On the other hand, MP2P method is free of PAPR drawback, while suffer from dependency on the used pulse shape. Thus, further analysis is required to design the proper pulse shapes to maximize the achieved gain. The other issue is the complexity of the receiver which can handle the intentional ISI and detect the transmitted symbols efficiently. Using optimum Receivers can be very complicated. We used some sub-optimum receiver methods like ZF and SIC receivers, however, further research is required to adapt more efficient receiver designs. For example, low complexity algorithms designed for large scale MIMO systems [9, 14] can be good candidates for the asynchronous transmission

system model. Designing proper matched filters at the receiver is also very crucial to the performance and also complexity of the receiver. For example, we introduced two different set of matched filter where one of them resulted in correlated noise samples and the other one created independent noise samples easing the detection process. However, realizing these matched filter with very small durations might be practically inefficient, which highlight the need for exploring new sets of match filters. For the analysis of the achievable rate, we assumed that the transmitter knows the time delays in order to fully exploit the available degrees of freedom. This assumption is reasonable when the time delays are intentionally added by the transmitter and the channel does not impose any time delay. However, this might not be the case in most of the practical channels. Thus, the receiver needs to send some feedback information about the time delays to the transmitter. Efficient design of the aforementioned feedback process and also the achieved gain in such system can be a topic of the future research.

Lastly, we used the concept of asynchrony mainly in the Multiuser scenario where multiple users/antennas transmit data to a common receiver. All of these concepts can also be applied to other networks like relay networks [49], broadcast networks and cognitive radio networks [42].

Bibliography

[1] Capacity of MIMO systems with asynchronous PAM, author=Barman, Kishor and Dabeer, Onkar, journal=IEEE Transactions on Communications, volume=57, number=11, year=2009, publisher=IEEE.

[2] H. Artés, D. Seethaler, and F. Hlawatsch. Efficient detection algorithms for MIMO channels: A geometrical approach to approximate ML detection. *IEEE transactions on signal processing*, 51(11):2808–2820, 2003.

[3] M. R. Avendi and H. Jafarkhani. Differential distributed Space-Time coding with imperfect synchronization in frequency-selective channels. *IEEE Transactions on Wireless Communications*, 14(4):1811–1822, April 2015.

[4] L. Bai, J. Choi, and Q. Yu. *Low complexity MIMO receivers*. Springer, 2014.

[5] S. Barghi and H. Jafarkhani. Exploiting asynchronous Amplify-and-Forward relays to enhance the performance of IEEE 802.11 networks. *IEEE/ACM Transactions on Networking*, 23(2):479–490, April 2015.

[6] S. Barghi, H. Jafarkhani, and H. Yousefi'zadeh. MIMO-assisted MPR-aware MAC design for asynchronous WLANs. *IEEE/ACM Transactions on Networking*, 19(6):1652–1665, Dec 2011.

[7] D. S. Bernstein. *Matrix mathematics: theory, facts, and formulas*. Princeton University Press, 2009.

[8] V. R. Cadambe and S. A. Jafar. Interference alignment and degrees of freedom of the k-user interference channel. *IEEE Transactions on Information Theory*, 54(8):3425–3441, 2008.

[9] A. Chockalingam and B. S. Rajan. *Large MIMO systems*. Cambridge University Press, 2014.

[10] E. Dahlman, S. Parkvall, and J. Skold. *4G: LTE/LTE-advanced for mobile broadband*. Academic press, 2013.

[11] A. Das and B. D. Rao. MIMO systems with intentional timing offset. *EURASIP Journal on Advances in Signal Processing*, 2011(1):1–14, 2011.

[12] L. Debnath and P. Mikusiński. *Hilbert spaces with applications*. Academic press, 2005.

[13] M. Dow. Explicit inverses of Toeplitz and associated matrices. *ANZIAM Journal*, 44:185–215, 2008.

[14] A. Elghariani and M. Zoltowski. Low complexity detection algorithms in large-scale MIMO systems. *IEEE Transactions on Wireless Communications*, 15(3):1689–1702, 2016.

[15] T. Eng and L. B. Milstein. Coherent DS-CDMA performance in Nakagami multipath fading. *IEEE Transactions on Communications*, 43(2/3/4):1134–1143, 1995.

[16] G. Forney. Maximum-likelihood sequence estimation of digital sequences in the presence of inter-symbol interference. *IEEE Transactions on Information theory*, 18(3):363–378, 1972.

[17] R. G. Gallager. *Principles of digital communication*, volume 1. Cambridge University Press Cambridge, UK:, 2008.

[18] M. Ganji and H. Jafarkhani. Interference mitigation using asynchronous transmission and sampling diversity. In *Global Communications Conference (GLOBECOM), 2016 IEEE*, pages 1–6. IEEE, 2016.

[19] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *IEEE Journal on selected areas in Communications*, 21(5):684–702, 2003.

[20] R. Gray. On the asymptotic eigenvalue distribution of Toeplitz matrices. *IEEE Transactions on Information Theory*, 18(6):725–730, 1972.

[21] R. M. Gray et al. Toeplitz and circulant matrices: A review. *Foundations and Trends® in Communications and Information Theory*, 2(3):155–239, 2006.

[22] J. Gutiérrez-Gutiérrez and P. M. Crespo. Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: Applications to MIMO systems. *IEEE Transactions on Information Theory*, 54(12):5671–5680, 2008.

[23] A. Hedayat and A. Nosratinia. Outage and diversity of linear receivers in flat-fading MIMO channels. *IEEE Transactions on Signal Processing*, 55(12):5868–5873, 2007.

[24] M. L. Honig. *Advances in multiuser detection*. Wiley Online Library, 2009.

[25] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 2012.

[26] H. Jafarkhani. *Space-time coding: theory and practice*. Cambridge university press, 2005.

[27] T. Jiang and Y. Wu. An overview: Peak-to-average power ratio reduction techniques for OFDM signals. *IEEE Transactions on broadcasting*, 54(2):257–268, 2008.

[28] J. Kazemitabar and H. Jafarkhani. Multiuser interference cancellation and detection for users with more than two transmit antennas. *IEEE Transactions on Communications*, 56(4):574–583, 2008.

[29] J. Kazemitabar and H. Jafarkhani. Performance analysis of multiple antenna multi-user detection. In *Information Theory and Applications Workshop, 2009*, pages 150–159. IEEE, 2009.

[30] E. G. Larsson. MIMO detection methods: How they work [lecture notes]. *IEEE signal processing magazine*, 26(3), 2009.

[31] Y. G. Li and G. L. Stuber. *Orthogonal frequency division multiplexing for wireless communications*. Springer Science & Business Media, 2006.

[32] J. M. Mendel. *Lessons in estimation theory for signal processing, communications, and control*. Pearson Education, 1995.

[33] H. Meyr, M. Moeneclaey, and S. A. Fechtel. *Digital communication receivers: synchronization, channel estimation, and signal processing*. Wiley Online Library, 1998.

[34] S. Moshavi. Multi-user detection for DS-CDMA communications. *IEEE Communications Magazine,*, 34(10):124–136, 1996.

[35] A. F. Naguib, N. Seshadri, and A. R. Calderbank. Applications of space-time block codes and interference suppression for high capacity and high data rate wireless systems. *Record of the 32nd Asilomar Conference on Signals, Systems & Computers.*, 2:1803–1810, 1998.

[36] A. A. Nasir, S. Durrani, H. Mehrpouyan, S. D. Blostein, and R. A. Kennedy. Timing and carrier synchronization in wireless communication systems: a survey and classification of research in the last 5 years. *EURASIP Journal on Wireless Communications and Networking*, 2016(1):180, Aug 2016.

[37] M. Oudin and J. P. Delmas. Asymptotic generalized eigenvalue distribution of block Toeplitz matrices and application to space-time beamforming. In *2007 15th European Signal Processing Conference*, pages 2449–2453, Sept 2007.

[38] S. Poorkasmaei and H. Jafarkhani. Asynchronous orthogonal differential decoding for multiple access channels. *IEEE Transactions on Wireless Communications*, 14(1):481–493, Jan 2015.

[39] J. G. Proakis. *Intersymbol interference in digital communication systems*. Wiley Online Library, 2001.

[40] E. Sengul, E. Akay, and E. Ayanoglu. Diversity analysis of single and multiple beamforming. *IEEE Transactions on Communications*, 54(6):990–993, 2006.

[41] S. Shao, Y. Tang, T. Kong, K. Deng, and Y. Shen. Performance analysis of a modified V-BLAST system with delay offsets using zero-forcing detection. *IEEE Transactions on Vehicular Technology*, 56(6):3827–3837, 2007.

[42] S. Sodagari and H. Jafarkhani. Enhanced spectrum sharing and cognitive radio using asynchronous primary and secondary users. *IEEE Communications Letters*, PP(99):1–1, 2018.

[43] M. Torbatian. Communication over asynchronous networks: Signaling and rate-reliability analysis. 2011.

[44] Torbatian, Mehdi. *Communication over Asynchronous Networks: Signaling and Rate-Reliability Analysis*. PhD thesis, 2011.

[45] D. Tse and P. Viswanath. *Fundamentals of wireless communication*. Cambridge university press, 2005.

[46] S. Verdu. *Multiuser detection*. Cambridge University Press, 1998.

[47] R. Xu and F. Lau. Performance analysis for MIMO systems using zero forcing detector over fading channels. *IEEE Proceedings-Communications*, 153(1):74–80, 2006.

[48] S. Yang and L. Hanzo. Fifty years of MIMO detection: The road to large-scale MIMOs. *IEEE Communications Surveys & Tutorials*, 17(4):1941–1988, 2015.

[49] X. Zhang, M. Ganji, and H. Jafarkhani. Exploiting asynchronous signaling for multiuser cooperative networks with analog network coding. In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, March 2017.

[50] L. Zheng and D. N. C. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels. *IEEE Transactions on information theory*, 49(5):1073–1096, 2003.

Appendices

A Derivation of Bit Error Rate (BER) Expression

For an AWGN channel with an average transmit power of $E[|x_k[i]|^2]$ and noise variance of $\frac{\sigma_n^2 \boldsymbol{R}^{-1}(i,i)}{\sum_{j=1}^{M} |h_{(1+(i-1)modK),j}|^2}$, the post SNR at the receiver can be expressed as:

$$\delta_i = \frac{\delta_0 \sum_{j=1}^{M} |h_{(1+(i-1)modK),j}|^2}{\boldsymbol{R}^{-1}(i,i)} \tag{A.1}$$

where $\delta_0 = \frac{E[|x_k[i]|^2]}{\sigma_n^2}$. We know that $|h_{i,j}|^2$ follows a chi-squared distribution with two degrees of freedom for all is and js. Therefore, $\sum_{j=1}^{M} |h_{(1+(i-1)modK),j}|^2$ is chi-squared distributed with $2M$ degrees of freedom. As a result, the distribution of δ_i can be calculated as follows:

$$P_{\delta_i}(\delta) = \frac{\boldsymbol{R}^{-1}(i,i)}{\delta_0} \frac{\left(\frac{\boldsymbol{R}^{-1}(i,i)}{\delta_0}\delta\right)^{M-1} \exp\left(-\frac{\boldsymbol{R}^{-1}(i,i)}{\delta_0 2}\delta\right)}{2^M \Gamma(M)} \tag{A.2}$$

where $\Gamma(.)$ is the Gamma function. For a specific value of SNR, BER varies according to the modulation. We assume that BPSK is used, however, extension to other modulations is straightforward. Based on this assumption, the BER for a given value of

SNR, e.g., δ is equal to $Q(\sqrt{2\delta})$. The next step is to calculate the following integral:

$$p_i = \int_0^\infty Q(\sqrt{2\delta}) P_{\delta_i}(\delta) d\delta$$

The integral of $\frac{a^m}{\Gamma(m)} \int_0^\infty \exp(-az) z^{m-1} Q(\sqrt{bz}) dz$ has a closed-form of:

$$\frac{\sqrt{b/2\pi a}}{2\left(1+\frac{b}{2a}\right)^{m+1/2}} \frac{\Gamma(m+1/2)}{\Gamma(m+1)} {}_2F_1(1, m+\frac{1}{2}; m+1; \frac{1}{1+\frac{b}{2a}})$$

where ${}_2F_1(q, w; e; r)$ is the hypergeometric function [15]. Therefore, the bit error rate, i.e., p_i is equal to:

$$p_i = \frac{\sqrt{\frac{\delta_0 2}{\pi \boldsymbol{R}^{-1}(i,i)}}}{2\left(1+\frac{\delta_0 2}{\boldsymbol{R}^{-1}(i,i)}\right)^{M+\frac{1}{2}}} \times \frac{\Gamma(M+\frac{1}{2})}{\Gamma(M+1)} \times {}_2F_1(1, M+\frac{1}{2}; M+1; \frac{1}{1+\frac{\delta_0 2}{\boldsymbol{R}^{-1}(i,i)}}) \tag{A.3}$$

B Average BER and Its Approximation at High SNR

In Eq. (A.3), p_i depends on $\boldsymbol{R}^{-1}(i,i)$ which varies for different values of i, and therefore each subchannel has a different BER. This is unlike the synchronous ZF, where all resulting subchannels have the same performance. In order to evaluate the performance of the entire system, we define the average BER performance as follows:

$$p_{avg} = \frac{\sum_{i=1}^{NK} p_i}{NK} \tag{B.4}$$

Since p_{avg} is not tractable, we approximate it at high SNR, using the fact that ${}_2F_1(1, m+\frac{1}{2}; m+1; \frac{1}{1+c})$ converges to one as c grows large [47]. Hence, at high SNR, p_{avg} can be approximated as follows:

$$\widetilde{p_{avg}} = \text{Const } \times \frac{\sum_{i=1}^{NK} (\boldsymbol{R}^{-1}(i,i))^M}{\delta_0^M} \tag{B.5}$$

where the constant value is equal to $\frac{1}{2^{(M+1)}NK\sqrt{\pi}} \frac{\Gamma(M+\frac{1}{2})}{\Gamma(M+1)}$.

C Proof of Lemma 3.2

When the frame length is N, we denote $\boldsymbol{R}$ by $\boldsymbol{R^N}$. Then, we prove by induction that, for all $N \in \mathbb{Z}_+$,

$$\text{trace}((\boldsymbol{R^N})^{-1}) = \frac{(N-1)(N+1)}{3(1+\tau_1-\tau_K)} + \frac{2N+1}{3(N+1+\tau_1-\tau_K)} + \frac{N(N+2)}{3} \sum_{i=1}^{K-1} \frac{1}{\tau_{i+1}-\tau_i}$$

Base case: When $N = 1$, $\boldsymbol{R^1}$ is equal to $\boldsymbol{R_{11}}$ which can be written as a generalized Fiedlers matrix whose inverse is given by [13]:

$$\boldsymbol{R_{11}}^{-1} = -\frac{1}{2} \begin{pmatrix} d_1 & \frac{1}{\tau_2-\tau_1} & ... & 0 & f \\ \frac{1}{\tau_2-\tau_1} & d_2 & \frac{1}{\tau_3-\tau_2} & ... & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & ... & \frac{1}{\tau_{K-1}-\tau_{K-2}} & d_{K-1} & \frac{1}{\tau_K-\tau_{K-1}} \\ f & 0 & ... & \frac{1}{\tau_K-\tau_{K-1}} & d_K \end{pmatrix}$$

where f and d_is are defined as:

$$f = \frac{1}{\tau_K - \tau_1 - 2} \tag{C.6}$$

$$d_1 = \frac{1}{\tau_1 - \tau_2} - \frac{1}{\tau_1 - \tau_K + 2} \tag{C.7}$$

$$d_K = \frac{1}{\tau_7 - \tau_8} - \frac{1}{\tau_1 - \tau_K + 2} \tag{C.8}$$

$$d_i = \frac{1}{\tau_{i-1} - \tau_i} + \frac{1}{\tau_i - \tau_{i+1}} \quad 2 \leq i \leq K-1 \tag{C.9}$$

Then, trace($\boldsymbol{R_{11}^{-1}}$) is equal to $\left(-\frac{1}{2}\sum_{i=1}^{K} d_i\right)$, which can be calculated using the above equations:

$$\text{trace}(\boldsymbol{R_{11}^{-1}}) = \frac{1}{(2 + \tau_1 - \tau_K)} + \sum_{i=1}^{K-1} \frac{1}{\tau_{i+1} - \tau_i} \tag{C.10}$$

Therefore, Eq. (C.6) is true for $N = 1$.

Induction step: Suppose Eq. (C.6) is true for N. We need to show that it also holds for $N + 1$, i.e.,

$$\text{trace}((\boldsymbol{R^{(N+1)}})^{-1}) = \frac{(N)(N+2)}{3(1 + \tau_1 - \tau_K)} + \frac{2N+3}{3(N + 2 + \tau_1 - \tau_K)} + \frac{(N+1)(N+3)}{3} \sum_{i=1}^{K-1} \frac{1}{\tau_{i+1} - \tau_i}$$

Because matrix $\boldsymbol{R}$ follows a recursive structure, $\boldsymbol{R^{N+1}}$can be presented as follows:

$$\boldsymbol{R^{N+1}} = \begin{bmatrix} (\boldsymbol{R^N})_{NK \times NK} & (\boldsymbol{L})_{NK \times K} \\ (\boldsymbol{L}^T)_{K \times NK} & (\boldsymbol{R_{11}})_{K \times K} \end{bmatrix} \tag{C.11}$$

where $\boldsymbol{L}^T = [\boldsymbol{0}_{K\times K}, \ldots, \boldsymbol{0}_{K\times K}, (\boldsymbol{R_{21}})_{K\times K}]$. For calculating the inverse of $\boldsymbol{R^{N+1}}$, we use the following lemma for matrix inversion in block form.

Lemma C.1. *Let na* $(m+n) \times (m+n)$ *matrix* $\boldsymbol{T}$ *be partitioned into a block form:*

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}$$

where the $m \times m$ *matrix* $\boldsymbol{A}$ *and* $n \times n$ *matrix* $\boldsymbol{D}$ *are invertible. Then, we have:*

$$\boldsymbol{T}^{-1} = \begin{bmatrix} \boldsymbol{M}^{-1} & -\boldsymbol{M}^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{C}\boldsymbol{M}^{-1} & \boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{C}\boldsymbol{M}^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \end{bmatrix}$$

where $\boldsymbol{M} = \boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}$ *[7].*

Here, $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ are equal to $\boldsymbol{R^N}$, $\boldsymbol{L}$, $\boldsymbol{L}^T$ and $\boldsymbol{R_{11}}$, respectively. Therefore, $\boldsymbol{M}$ is equal to:

$$\boldsymbol{M} = \boldsymbol{R^N} - \boldsymbol{L}(\boldsymbol{R_{11}})^{-1}\boldsymbol{L}^T \tag{C.12}$$

Now, we need to find the inverse of $\boldsymbol{M}$. By defining $\boldsymbol{Z}$ as $(\boldsymbol{R^N})^{-1}$, the inverse of $\boldsymbol{M}$ can be presented as:

$$\boldsymbol{M}^{-1} = \begin{bmatrix} \boldsymbol{I}_K & \dots & \boldsymbol{0}_{K,K} & \boldsymbol{Z_{1N}}\boldsymbol{Q}(\boldsymbol{I}_K - \boldsymbol{Z_{NN}}\boldsymbol{Q})^{-1} \\ \boldsymbol{0}_{K\times K} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \boldsymbol{I}_K & \boldsymbol{Z_{(N-1)N}}\boldsymbol{Q}(\boldsymbol{I}_K - \boldsymbol{Z_{NN}}\boldsymbol{Q})^{-1} \\ \boldsymbol{0}_{K\times K} & \dots & \boldsymbol{0}_{K\times K} & (\boldsymbol{I}_K - \boldsymbol{Z_{NN}}\boldsymbol{Q})^{-1} \end{bmatrix} \boldsymbol{Z}$$

where $\boldsymbol{Q} = \boldsymbol{R_{12}}{\boldsymbol{R_{11}}}^{-1}\boldsymbol{R_{21}}$ and $\boldsymbol{Z_{ij}}$s are $K \times K$ partitioning blocks of $\boldsymbol{Z}$. Also, $\boldsymbol{I}_k$ and $\boldsymbol{0}_{i\times j}$ are a $k \times k$ identity matrix and a $i \times j$ all-zero matrix, respectively.

To show the correctness of Eq. (C.13), we need to take the following steps:

Step 1: By some calculations, it can be shown that $\boldsymbol{L}(\boldsymbol{R_{11}})^{-1}\boldsymbol{L}^T$ is equal to $\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Q} \end{bmatrix}$. As a result, we have:

$$\boldsymbol{M} = \boldsymbol{R^N} - \begin{bmatrix} \mathbf{0}_{(N-1)K\times(N-1)K} & \mathbf{0}_{(N-1)K\times K} \\ \mathbf{0}_{K\times(N-1)K} & \boldsymbol{Q} \end{bmatrix} \tag{C.13}$$

Step 2: If we multiply both sides by $\boldsymbol{Z}$, we will have:

$$\boldsymbol{ZM} = \boldsymbol{I}_{NK} - \begin{bmatrix} \mathbf{0}_{K\times K} & \dots & \mathbf{0}_{K\times K} & \boldsymbol{Z_{1N}Q} \\ \mathbf{0}_{K\times K} & \dots & \mathbf{0}_{K\times K} & \boldsymbol{Z_{2N}Q} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{K\times K} & \dots & \mathbf{0}_{K\times K} & \boldsymbol{Z_{NN}Q} \end{bmatrix} \tag{C.14}$$

Step 3: We denote the right hand side of Eq. (C.14) by $\boldsymbol{X}$, then, we can conclude that the inverse of $\boldsymbol{M}$ is equal to:

$$\boldsymbol{M}^{-1} = \boldsymbol{X}^{-1}\boldsymbol{Z} \tag{C.15}$$

Step 4: $\boldsymbol{X}^{-1}$ can be calculated as follows:

$$\boldsymbol{X}^{-1} = \begin{bmatrix} \boldsymbol{I}_K & \dots & \mathbf{0}_{K,K} & \boldsymbol{Z_{1N}Q}(\boldsymbol{I}_K - \boldsymbol{Z_{NN}Q})^{-1} \\ \mathbf{0}_{K\times K} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \boldsymbol{I}_K & \boldsymbol{Z_{(N-1)N}Q}(\boldsymbol{I}_K - \boldsymbol{Z_{NN}Q})^{-1} \\ \mathbf{0}_{K\times K} & \dots & \mathbf{0}_{K\times K} & (\boldsymbol{I}_K - \boldsymbol{Z_{NN}Q})^{-1} \end{bmatrix} \tag{C.16}$$

Step 5: Finally, if we plug $\boldsymbol{X}^{-1}$ in Eq. (C.15), we will reach Eq. (C.13).

If we denote $K \times K$ diagonal blocks of $\boldsymbol{M}^{-1}$ as $[\boldsymbol{M}^{-1}]_{i,i}$ $1 \leq i \leq N$, then, by use of Lemma C.1, trace$((\boldsymbol{R}^{\boldsymbol{N+1}})^{-1})$ can be written as:

$$\text{trace}((\boldsymbol{R}^{\boldsymbol{N+1}})^{-1}) = \sum_{i=1}^{N} \text{trace}([\boldsymbol{M}^{-1}]_{i,i}) + \text{trace}(\boldsymbol{R_{11}}^{-1} + \boldsymbol{R_{11}}^{-1}\boldsymbol{R_{21}}[\boldsymbol{M}^{-1}]_{N,N}\boldsymbol{R_{12}}\boldsymbol{R_{11}}^{-1}) \tag{C.17}$$

By simplifying Eq. (C.13), diagonal blocks of $\boldsymbol{M}^{-1}$ can be presented as follows:

$$1 \leq i \leq N-1 :$$

$$[\boldsymbol{M}^{-1}]_{i,i} = \boldsymbol{Z_{ii}} + \boldsymbol{Z_{iN}}\boldsymbol{Q}(\boldsymbol{I} - \boldsymbol{Z_{NN}}\boldsymbol{Q})^{-1}\boldsymbol{Z_{Ni}} \tag{C.18}$$

$$i = N :$$

$$[\boldsymbol{M}^{-1}]_{i,i} = (\boldsymbol{I} - \boldsymbol{Z_{NN}}\boldsymbol{Q})^{-1}\boldsymbol{Z_{NN}} \tag{C.19}$$

In Eq. (C.17), we set the diagonal blocks of $\boldsymbol{M}^{-1}$ as Eqs. (C.18) and (C.19). Then, by some manipulations, trace$((\boldsymbol{R}^{\boldsymbol{N+1}})^{-1})$ can be presented as:

$$\begin{aligned}\text{trace}((\boldsymbol{R}^{\boldsymbol{N+1}})^{-1}) = {} & \text{trace}((\boldsymbol{R}^{\boldsymbol{N}})^{-1}) + \text{trace}(\boldsymbol{R_{11}}^{-1}) \\ & + \sum_{i=1}^{N-1} \text{trace}(\boldsymbol{Z_{iN}}\boldsymbol{Q}(\boldsymbol{I} - \boldsymbol{Z_{NN}}\boldsymbol{Q})^{-1}\boldsymbol{Z_{Ni}}) \qquad \text{(C.20)} \\ & + \text{trace}((\boldsymbol{I} - \boldsymbol{Z_{NN}}\boldsymbol{Q})^{-1}\boldsymbol{Z_{NN}}) - \text{trace}(\boldsymbol{Z_{NN}}) \\ & + \text{trace}(\boldsymbol{R_{11}}^{-1}\boldsymbol{R_{21}}(\boldsymbol{I} - \boldsymbol{Z_{NN}}\boldsymbol{Q})^{-1}\boldsymbol{Z_{NN}}\boldsymbol{R_{12}}\boldsymbol{R_{11}}^{-1}) \qquad \text{(C.21)}\end{aligned}$$

The first and second terms in Eq. (C.21) can be calculated by induction hypothesis and induction base, respectively. Calculating other terms in Eq. (C.21) is tedious but similar for different values of K. Therefore, we only calculate it for $K = 2$ and skip

the rest. For $K = 2$, $\boldsymbol{Q}$ is equal to:

$$\boldsymbol{Q} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1-\tau}{1+\tau} \end{bmatrix} \tag{C.22}$$

where $\tau = \tau_2 - \tau_1$.

If we plug $\boldsymbol{Q} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1-\tau}{1+\tau} \end{bmatrix}$ in Eq. (C.21), after some calculations we will have:

$$\begin{aligned} \text{trace}((\boldsymbol{R^{N+1}})^{-1}) = \text{trace}((\boldsymbol{R^N})^{-1}) &+ \frac{2}{1-(1-\tau)^2} \\ &+ \frac{1-\tau}{(1+\tau)-(1-\tau)r(2N,2N)} \sum_{i=1}^{2N} (r(2N,i))^2 \\ &+ \frac{(1+\tau)(1+(\tau-1)^2)}{(2-\tau)^2[(1+\tau)-(1-\tau)r(2N,2N)]} r(2N,2N) \end{aligned} \tag{C.23}$$

where $r(i,j)$ is the (i,j)th element of matrix $(\boldsymbol{R^N})^{-1}$. By induction hypothesis, the first term in Eq. (C.23) is equal to $\frac{(N-1)(N+1)}{3(1-\tau_2+\tau_1)} + \frac{N(N+2)}{3(\tau_2-\tau_1)} + \frac{2N+1}{3(N+1-\tau_2+\tau_1)}$. For calculating Eq. (C.23), we also need values of $r(2N,i)$, $1 \le i \le 2N$, which are elements of the last row of $(\boldsymbol{R^N})^{-1}$. Due to the special structure of matrix $\boldsymbol{R}$, values of $r(2N,i)$ can be calculated as follows:

$$\begin{cases} r(2N,2i-1) = \frac{\tau-i}{\tau(N+1-\tau)} \\ r(2N,2i) = \frac{i}{\tau(N+1-\tau)} \end{cases} 1 \le i \le N \tag{C.24}$$

To verify Eq. (C.24), we can multiply the last row of $(\boldsymbol{R^N})^{-1}$, i.e.,

$[r(2N,1), r(2N,2), \ldots, r(2N,2N)]$, by different columns of $\boldsymbol{R^N}$ as follows:

$$\text{1st column: } \frac{(\tau-1)1}{\tau(N+1-\tau)} + \frac{(1)(1-\tau)}{\tau(N+1-\tau)} = 0$$

$$(2i)_{th} \text{ column: } \quad 1 \leq i \leq N-1$$

$$\frac{(\tau-i)(1-\tau)}{\tau(N+1-\tau)} + \frac{(i)1}{\tau(N+1-\tau)} + \frac{(\tau-(i+1))\tau}{\tau(N+1-\tau)} = 0$$

$$(2i-1)_{th} \text{ column: } \quad 2 \leq i \leq N$$

$$\frac{((i-1))(\tau)}{\tau(N+1-\tau)} + \frac{(\tau-i)1}{\tau(N+1-\tau)} + \frac{(i)(1-\tau)}{\tau(N+1-\tau)} = 0$$

$$2N\text{th column: } \frac{(\tau-N)(1-\tau)}{\tau(N+1-\tau)} + \frac{(N)1}{\tau(N+1-\tau)} = 1$$

These results verify that the last row of $(\boldsymbol{R^N})^{-1}$ follows the pattern in Eq. (C.24). The last step is to plug Eq. (C.24) into Eq. (C.23). As a result, trace$((\boldsymbol{R^{N+1}})^{-1})$ is equal to $\frac{(N)(N+2)}{3(1-\tau_2+\tau_1)} + \frac{(N+1)(N+3)}{3(\tau_2-\tau_1)} + \frac{2N+3}{3(N+2-\tau_2+\tau_1)}$, which verifies the induction step and completes the proof.