12-2008

Stochastic Modeling and Estimation of Wireless Channels with Application to Ultra Wide Band Systems

Yanyan Li
University of Tennessee - Knoxville

To the Graduate Council:

I am submitting herewith a dissertation written by Yanyan Li entitled "Stochastic Modeling and Estimation of Wireless Channels with Application to Ultra Wide Band Systems." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Seddik Djouadi, Major Professor

We have read this dissertation and recommend its acceptance:

Paul Crilly, Michael Roberts, Jie Xiong

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Yanyan Li entitled "Stochastic Modeling and Estimation of Wireless Channels with Application to Ultra Wide Band Systems." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Seddik Djouadi, Major Professor

We have read this dissertation
and recommend its acceptance:

Paul Crilly

Michael Roberts

Jie Xiong

Accepted for the Council:

Carolyn R Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Stochastic Modeling and Estimation of Wireless Channels with Application to Ultra Wide Band Systems

A Dissertation
Presented for the
Doctor of Philosophy Degree
Electrical Engineering and Computer Science Department
The University of Tennessee, Knoxville

Yanyan Li
December 2008

DEDICATION

This dissertation is dedicated to my parents, my friends and my advisor. They are always standing beside me, believing in me, encouraging me and helping me to get rid of all the difficulties in my life.

Acknowledgements

I wish to thank all those who helped me complete my Doctor's degree in Electrical Engineering. I would like to thank my advisor Dr.Djouadi for his guidance and his patience in leading me to my research field. I would like to thank Dr. Crilly for making my understanding the digital communication deep. I would like to thank Dr. Robert for teaching me important concepts in digital signal processing. I would like to thank Dr. Xiong to expand my background in probability and statistics.

Finally, I would like to thank my family and friends, whose suggestions and encouragement are the source of my momentum.

Abstract

This thesis is concerned with modeling of both space and time variations of Ultra Wide Band (UWB) indoor channels. The most common empirically determined amplitude distribution in many UWB environments is Nakagami distribution. The latter is generalized to stochastic diffusion processes which capture the dynamics of UWB channels. In contrast with the traditional models, the statistics of the proposed models are shown to be time varying, but converge in steady state to their static counterparts.

System identification algorithms are used to extract various channel parameters using received signal measurement data, which are usually available at the receiver. The expectation maximization (EM) algorithm and the Kalman filter (KF) are employed in estimating channel parameters as well as the inphase and quadrature components, respectively. The proposed algorithms are recursive and therefore can be implemented in real time. Further, sufficient conditions for the convergence of the EM algorithm are provided. Comparison with recursive Least-square (LS) algorithms is carried out using experimental measurements.

Distributed stochastic power control algorithms based on the fixed point theorem and stochastic approximations are used to solve for the optimal transmit power problem and numerical results are also presented.

A framework which can capture the statistics of the overall received signal and a methodology to estimate parameters of the counting process based on the received signal is developed. Furthermore, second moment statistics and characteristic functions are computed explicitly and considered as an extension of Rice's shot noise analysis.

Another two important components, input design and model selection are also considered. Gel'fand n-widths and Time n-widths are used to represent the inherent error introduced by input design. Kolmogorov n-width is used to characterize the representation error introduced by model selection. In particular, it is shown that the optimal model for reducing the representation error is a finite impulse response (FIR) model and the optimal input is an impulse at the start of the observation interval.

Table of Contents

List of Figures

Introduction

1.1 UWB Radio Channel Models

Ultra-wideband (UWB) communication systems have recently attracted significant interest from both the research community and industry since the Federal Communications Commission (FCC) allowed limited unlicensed operation of UWB devices in the USA [82]. For example, in the digital home of now or in the not far away future, people will be sharing photos, music, video, data and voice among networked consumer electronics, PCs and mobile devices throughout the home and even remotely. What's more, users will be able to stream video content from a PC or consumer electronics device. [81].

Wireless technology which is designed for short range, high resolution and personal area network is a potential candidate for enabling this capability. In recent years, UWB is making the transition from laboratories to standardization. This is a key step toward the development of real world products [81].

Recently researchers have made great industry achievements. They built the formation of industry working groups that will define the UWB physical layer (PHY) and MAC layer and applications that will run on top of the radio platform. In the U.S., the FCC has mandated that UWB radio transmission can legally operate in the range from 3.1 GHz to 10.6 GHz, at a transmit power of 1 dBm/MHz. Japanese regulators have issued the first UQWB experimental license allowing the operation of a UWB transmitter in Japan [81].

UWB communications systems are commonly defined as systems that have either more than 20% relative bandwidth or more than 500 MHz absolute bandwidth. One of the most promising applications for UWB is sensor networks, where a large number of sensor nodes communicate among each other and with central nodes with high reliability. UWB systems have a number of advantages that make them attractive for consumer communications applications [1]. In particular, UWB systems

- have potentially low complexity and low cost;
- have noise-like signal;

- are resistant to severe multi-path and jamming;
- Have very good time domain resolution slowing for location and tracking applications [1].

The low complexity and low cost of UWB systems arises from the essentially baseband nature of the signal transmission. Unlike conventional radio systems, the UWB transmitter produces a very short time domain pulse, which is able to propagate without the need for an additional RF (radio frequency) mixing stage [1, 51].

This thesis is focused on the UWB channels which are very similar to a wideband channel as may be experienced in spread spectrum or CDMA systems [27, 35]. The main distinguishing feature of an 'ultra' wideband channel model is the extremely multi-path-rich channel profile [1, 2, 5 and 6].

A number of UWB channel models have been proposed in the literature. In [2] an ultra-wide bandwidth signal propagation experiment is performed in a typical modern office building in order to characterize the UWB signal propagation channels. This experiment setup is also used outdoors to find out UWB's suitability to narrowband loss models. The results in [3] show that the narrowband loss models can be applied to the UWB case. In [4] the author suggests a model for the frequency range below 1 GHz. Further, a comprehensive statistical model that is valid for a frequency range from 3-10 GHz is proposed in [5]. This model is accepted by the IEEE 802.15.4a task group as standard model for evaluation of UWB system proposals.

The real valued model in [6] is based on empirical measurements originally carried out in indoor environments in 1987. Due to the clustering phenomena observed at measured UWB indoor channels, the model proposed in IEEE 802.15 is derived from the Saleh and Valenzuela (SV) model using a log-normal distribution rather than Rayleigh distribution for the multi-path gain magnitude. An independent fading mechanism is assumed for each ray within the cluster. In the SV model, both the cluster and ray arrival times are modeled independently by a Poisson process [55, 56]. The phase of the channel impulse response can be either 0 or π. Therefore the model contains no imaginary component.

The main goals of the IEEE 802.15 model are the modeling of attenuation and delay dispersion. The former subsumes both shadowing and average path loss, while the latter describes the power delay profile and the small-scale fading statistics; from this, other parameters such as root mean square (rms) delay spread, number of multipath components are derived in [5].

The model in [6] starts with the physical realization that rays arrive in clusters. The cluster arrival times are modeled as Poisson arrival processes with some fixed rate Λ. Within each cluster, subsequent rays also arrive according to Poisson processes with another fixed rate λ. Typically, each cluster consists of many rays.

The classical SV model also uses a Poisson process for the ray arrival time. Due to the discrepancy in the fitting for the indoor residential and indoor and outdoor office environments, [8] proposes to model ray arrival times with mixtures of two Poisson processes.

For some environments, most notably the industrial environment, a dense arrival of Multi-path Component (MPCs) was observed, i.e., each resolvable delay bin contains significant energy. In that case, the concept of ray arrival rates loses its meaning and a realization of the impulse response based on a tapped delay line model with regular tap spacing is to be used [57].

Different from what previous people have done, this thesis introduces a fresh approach in modeling the wireless UWB indoor channels which describes the characteristics of the channel more accurately. A more accurate representation of the channel will lead to a more appropriate design of the transmit function, power control of the transmitter, choice of the modulation schemes, etc. A better understanding of the channel characteristics can help the receiver avoid interferences by other users because it can subtract the jammer from the received signals based on the channel information. So a better modeling scheme can help the communication system to minimize the interferences.

A major specialty of the modeling scheme is treating the attenuation coefficient as a diffusion process; while current models usually consider it as a random variable. The diffusion process is described by a stochastic differential equation, which can further be written in a stochastic state space form. Since the number of unknown parameters of state space is finite, it is easier to use, for instance, in channel estimation.

Besides having a channel model, a complete statistical analysis is necessary. The statistical analysis helps fully understand the whole system so as to have better system design. In previous works, the statistics of the received signal were computed by making some important simplifications. First, the channel parameters such as attenuation coefficient depend not only on time but also the space, during the computation of various statistics. Second, although for deterministic or a fixed sample path, the computation of the statistical properties of received signal is not affected by this omission, this is not the case when the ensemble statistics are analyzed. Ensemble statistics using a counting process as simple as the inhomogeneous Poisson process reveal an additional smoothing property associated with each propagation environment, which is expressed in terms of the rate of the Poisson process and the attenuations [76].

Although in some works pretty good results with these simplifications are obtained, it is because they apply those to some specific wireless environments. In order to fully solve the statistical analysis problem, we introduce shot-noise analysis of random noise, brought by Rice [36]. In [36], Rice brought in a powerful tool (Campbell's theorem [36]) to fully analyze the statistical properties of wireless fading channels. The statistical analysis includes not only the first and second moments of the received signal but also the distribution of the number of paths, which has received little attention in the literature.

In this thesis, state space models [52] are used to model the indoor wireless channels with a finite number of parameters to identify. A worst case estimation error [75] which depends on input and model set is used as the objecting function. We provide the optimal input and model set which minimize the worst case estimation error. The problem is divided in two parts, one part is input design, and the other is model selection. Both input design and model selection are viewed as an optimization problem individually. The identification error, optimized over all bounded inputs and all n-

parameter models, is related to the n-width in the sense of Gel'fand and Kolmogorov [72] respectively, two standard notions in metric complexity theory.

1.2 Power Control

Power control (PC) is important to improve performance of wireless communication systems. The goal of power control is to minimize the power of transmitters in the system such that the signal to interference ratio (SIR) of each receiver reaches certain threshold. The threshold can be different for different receivers, depending on their propagation environments. Power minimization not only increases battery life, but also increases overall network capacity [41, 42, 43 and 44]. Users only need to expand sufficient power for acceptable reception as determined by their quality of service (QoS) specifications that is usually characterized by the SIR [9]. The downlink Power Control is straightforward since each link has the same channel gains, while the uplink PC is more complex and is the problem we solve in this thesis.

Power Control algorithms (PCA) can be classified as centralized and distributed. Centralized PCAs require the information from every node of the network, while the distributed PCAs only require the base station to know its own information, such as, its Signal to Noise ratio (SNR), which is easily obtained from local measurements. These power allocation problems have been treated as an eigenvalue problem of a nonnegative matrix [58]. The power is updated iteratively based on the signal to interference ratio (SINR) and certain thresholds [26].

Stochastic PCAs (SPCA) that use noisy interference estimates have been introduced in [18], where conventional matched filter receivers are used. It is shown in [18] that the iterative SPCA, which uses stochastic approximations, converges to the optimal power vector under certain assumptions on the step-size sequence. These results were later extended to the cases when a nonlinear receiver or a decision feedback receiver is used [19]. Much of this previous work deals with discrete time PCAs and static time-invariant channel models.

A fixed point of a function is a point mapped to itself by the function. Fixed point theorem [50] gives the conditions under which case the functions have solutions. In this thesis, the distributed PCA based power optimization problem is treated as a fixed point problem. The power is updated recursively until the power converges. It is also shown that the optimal transmit power is given by the fixed point of a certain function.

In this work, The PCA only requires the received SIR at its intended receiver, while the received matched filter output (received SIR) at its intended receiver and the channel gain between the transmitter and its intended receiver are required in the SPCA presented in [18, 26].

1.3 Contributions

The contributions of the research undertaken here are as follows

1. Application of stochastic differential equations to model UWB wireless indoor channels.

Unlike narrow band short term fading [57], the distribution of each path is characterized by Nakagami distribution [5, 8]. Impulse response of the indoor channel can be approximated and factorized. Therefore, it can be modeled by stochastic differential equations (SDEs) [52].

2. Estimation of UWB channel parameters using Expectation Maximization (EM) algorithm.

Several groups of indoor multipath data were provided by UT the Antenna Laboratory. Different paths are separated by their statistical properties. The Kalman filter is used to estimate inphase, quadrature, envelope of the received signal. The EM algorithm is used to identify the model parameters. Moreover, recursive Least-Square (LS) algorithm is also discussed and compared to the EM algorithm using experiment data.

3. Development of a PCA based on the Fixed Point theorem.

Centralized [58] and Distributed PCAs are developed based on the TV channel models. Further, a stochastic PCA is based on fixed point [65] and stochastic approximation is discussed. Numerical results are provided and the convergence results are compared.

4. *Statistical analysis of multipath channels*.

Develop a framework which can capture the statistics of the overall received signal. We assume the analysis with the most general case where the amplitude, phase, arrival times and the input signal are random processes; the delays are time-varying, while the counting process can be either homogeneous or inhomogeneous Poisson process [55]. Statistics such as first and second moments, characteristic function are calculated explicitly. Furthermore, we extend the Rice's shot noise analysis result to our framework.

5. *Input design and model selection*

Formulate the problem of modeling the wireless channels using two standard optimization problems, input design and model selection, with the criterion of minimizing the worst case estimation error. These problems are related to n-widths in the sense of Gel'fand and Kolmogorov in metric complexity theory. Further, the optimal estimation error is calculated based on the Gel'fand and Kolmogorov n-widths.

1.4 Thesis Organization

The thesis is organized as follows

Chapter 2 presents the modeling of UWB indoor wireless channels. A necessary and sufficient condition for representing any impulse response in stochastic state space from is that it is factorizable into the product of two separate functions of time and space. Howerer, in general this is not the case for impulse response of wireless channels. We show that the impulse response of wireless channels can be approximated in the mean square sense as closely as desired by factorizable impulse responses that can be realized by SDEs in state space form. New dynamical spatial and temporal models for

the power loss and attenuation coefficient are presented. Furthermore, another TV model Autoregressive Moving Average (ARMA) is provided and compared with the proposed SED models using experimental data.

Chapter 3 describes the EM algorithm together with the Kalman Filter to estimate system parameters in state space form. They are used to estimate the inphase, quadrature, envelop and they are suited to Nakagami fading. Simulation and experimental results are presented to verify the validity of the model; comparison is shown between the EM algorithm and recursive LS algorithm. Also, the sufficient conditions for the EM algorithm to converge are also discussed.

Chapter 4 presents the PCAs based on the dynamical models described in Chapter 2. Distributed PCAs are developed based on the TV channel models. Further, a stochastic PCA is based on fixed point and stochastic approximation is discussed to minimize the total transmitted power.

In Chapter 5 the statistical analysis of the overall received signal is established. This is accomplished by extending and adapting accordingly Rice's shot-noise analysis to complex signals using homogeneous or inhomogeneous Poisson processes for describing the number of paths. A parameter estimation method and Cramer-Rao Lower Bound are also discussed. Campbell's theorem (established by Rice) is extended to our framework in order to calculate the statistics such as the first and second moments of the received signals.

In chapter 6 we focus on channel identification problem in discrete time. The objective is to minimize the worst case identification error. The worst case identification error can be separated to two terms, i.e., inherent error and representation error. The inherent error is generated during the information acquisition, while the representation is brought up from the information processing. Minimizing the inherent and representation error is related to the n-width in the sense of Gel'fand and Kolmogorov.

Chapter 2

UWB Indoor Wireless Channel Modeling

Small-scale propagation models characterize the rapid fluctuations of the received signal. This occurs when the transmitter-receiver separation distance is small. The main propagation mechanism for this type of environment is scattering. Signal is scattered at the vicinity of the receiver. Many copies of the transmitted signal arrive at the receiver via a number of different paths at slightly different times. This phenomenon is called multi-path in indoor communication [35]. At the receiver, plane waves add vectorially giving rise to rapid and severe signal fluctuations. This gives rise to the short-term fading channel model, where the term fading refers to the rapid and severe received signal fluctuations due to scattering in this type of environment.

2.1 Review of Previous UWB Channel Model

A number of UWB channel models [5, 8] have been proposed in the literature. For example, the measurement data collected in a series of multipath propagation studies can also be used to model UWB path loss. The impact of the link distance on the received signal energy can be determined by propagation loss calculations which define the fraction of the transmitted power that can be received at a distance d. General propagation physics approaches are valid in the case of UWB transmission, which means that the longer the distance is, the lower the received signal's energy is. The following equation [1] can be used to fit the measurement data.

$$\mathrm{P}(\mathrm{d}) = n10\log_{10}(d) + a \tag{1}$$

where n corresponds to the path loss factor, d is the distance and a is a power scaling constant.

Further, the IEEE 802.15.3 group for wireless personal area networks and especially its channel modeling subcommittee decided to use the so called modified Saleh-Valenzuela (SV) model as a reference for UWB channel model [5, 8].

The impulse response (in complex baseband) of the SV model is given in general as

$$h_{discr}(t) = \sum_{l=0}^{L}\sum_{k=0}^{K} a_{k,l}\exp(j\phi_{k,l})\delta(t - T_l - \tau_{k,l}) \tag{2}$$

where $a_{k,l}$ is the tap weight of the k^{th} component in the l^{th} cluster , T_l is the delay of the l^{th} cluster, $\tau_{k,l}$ is the delay of the k^{th} MPC relative to the l^{th} cluster arrival T_l. The phase $\phi_{k,l}$ is uniformly distributed, i.e., for a band pass system, the phase is taken as a uniformly distributed random variable from the range $[0, 2\pi]$. Following [7], the number of cluster L is an important parameter of the model. It is assumed to be Poisson-distributed [55]

$$pdf_L(L) = \frac{(\bar{L})^L \exp(-\bar{L})}{L!} \tag{3}$$

So the mean $\bar{L}$ completely characterizes the distribution.

In [8], $\tau_{0,l}$ is taken to be zero. The distribution of the cluster arrival times is given by a Poisson process

$$p(T_l \mid T_{l-1}) = \Lambda_l \exp[-\Lambda_l(T_l - T_{l-1})], l > 0 \tag{4}$$

where Λ_l is the cluster arrival rate (assumed to be independent of l). The classical SV model also uses a Poisson process for the ray arrival time. Due to the discrepancy in the fitting for the indoor residential and indoor and outdoor office environments, they propose to model ray arrival times with mixtures of two Poisson processes as follows [5,8]

$$p(\tau_{k,l} \mid \tau_{(k-1),l}) = \beta\lambda_1 \exp[-\lambda_1(\tau_{k,l} - \tau_{(k-1),l})] + (\beta - 1)\lambda_2 \exp[-\lambda_2(\tau_{k,l} - \tau_{(k-1),l})], k > 0 \tag{5}$$

where β is the mixture probability, while λ_1 and λ_2 are the ray arrival rates.

Although this model characterizes the UWB environment in certain cases, it is not general enough to capture both space and time variation of UWB indoor channels. All the channel parameters in this model are random variables. However, in practice, the channels are varying, i.e., they are random processes with time-varying statistics. Therefore, a model which captures both time and spatial character should be considered.

2.2 General Time-varying Model

The general time-varying (TV) model of wireless fading channels is typically represented by the following multipath TV low-pass equivalent impulse response [30]:

$$H_l(t;\tau) = \sum_{n=1}^{N(t)} r_n(t,\tau) e^{j\Phi_n(t,\tau)} \delta\left(\tau - \tau_n(t)\right) = \sum_{n=1}^{N(t)} \left(I_n(t,\tau) + jQ_n(t,\tau)\right) \delta\left(\tau - \tau_n(t)\right) \tag{6}$$

where $H_l(t;\tau)$ is the response of the channel at time t, due to an impulse applied at time $t-\tau$, $N(t)$ is the random number of multi-path components impinging on the receiver, while the set $\left\{r_n(t,\tau), \Phi_n(t,\tau), \tau_n(t)\right\}_{n=1}^{N(t)}$ describes the random TV attenuation, overall phase shift and arrival time of the different paths, respectively. $\left\{I_n(t,\tau), Q_n(t,\tau)\right\}_{n=1}^{N(t)} \triangleq \left\{r_n(t,\tau)\cos\Phi_n(t,\tau), r_n(t,\tau)\sin\Phi_n(t,\tau)\right\}_{n=1}^{N(t)}$ are defined as the inphase and quadrature components of each path. Let $s_l(t)$ be the low pass equivalent representation of the transmitted signal, then the low pass equivalent representation of the received signal is given by [30]

$$y_l(t) = \int_{-\infty}^{\infty} C_l(t;\tau) s_l(t-\tau)\, d\tau = \sum_{n=1}^{N(t)} r_n\left(t,\tau_n(t)\right) e^{j\Phi_n\left(t,\tau_n(t)\right)} s_l\left(t-\tau_n(t)\right) \tag{7}$$

and the multi-path TV band pass impulse response is given by [30]

$$\begin{aligned} H(t;\tau) &= \mathrm{Re}\left\{\sum_{n=1}^{N(t)} \left[r_n(t,\tau) e^{j\Phi_n(t,\tau)} \right] e^{j\omega_c t} \right\} \delta\left(\tau - \tau_n(t)\right) \\ &= \sum_{n=1}^{N(t)} \left(I_n(t,\tau)\cos\omega_c t - Q_n(t,\tau)\sin\omega_c t \right) \delta\left(\tau - \tau_n(t)\right) \end{aligned} \tag{8}$$

where ω_c is the carrier frequency and the band pass representation of the received signal is given by

$$y(t)=\sum_{n=1}^{N(t)}\left(I_n(t,\tau)\cos\omega_c t-Q_n(t,\tau)\sin\omega_c t\right)s_l\left(t-\tau_n(t)\right) \tag{9}$$

It can be seen that all the parameters such as inphase and quadrature component in this time-varying impulse response model are random processes and depend on both t and τ. The inphase and quadrature component are general functions of t and τ. As a result, they cannot be specified by a finite number of parameters. In next section we show that the impulse response can be approximated in a mean square sense by stochastic state space. The latter is completely characterized by a finite number of parameters.

2.3 Impulse Response Factorization

Now we want to represent the time-varying model in (9) with a stochastic state-space form. The following theorem states a necessary and sufficient condition about the realization of the impulse response.

Theorem 2.1 [52]: An impulse response $h(t;\tau)$ of a system has a state space realization if and only if it can be factorized, that is, if there exist two functions $g(\cdot)$ and $f(\cdot)$ such that for all t and τ we have

$$h(t;\tau)=g(t)\,f(\tau) \tag{10}$$

It is readily seen from the expression of the impulse response $H_l(t;\tau)$ of the channel that in general it can not be factorized in the form (10). However, we show that if $H_l(t;\tau)$ has finite energy, i.e., $H_l(t;\tau)\in L^2\left([0,\infty)\times[0,\infty)\right)$ (see definition in Appendix A) can be approximated as closely as desired by a factorizable impulse response function in the following theorem.

Theorem 2.2: If the impulse response $H_l(t;\tau)$ of the wireless channel in (6) has finite energy, it can be approximated as closely as desired by a factorizable impulse response in the mean square sense.

Proof: See Appendix A.

According to the realization theory in [52], $H_l(t;\tau)$ can be realized by the following linear state equation

$$\begin{aligned} \dot{x}(t) &= f(t)u(t) \\ y(t) &= g(t)x(t) \end{aligned} \tag{11}$$

However, the realization is not unique; we choose the following form of realization as our framework.

Let us denote $\{X_I(t), X_Q(t)\}$ as the state vectors for the inphase and quadrature components, respectively, and $\{I(t), Q(t)\}$ are the inphase and quadrature components of the channel, then the path can be written in compact form as

$$\begin{aligned} \begin{bmatrix} dX_I(t) \\ dX_Q(t) \end{bmatrix} &= \begin{bmatrix} A_I & 0 \\ 0 & A_Q \end{bmatrix} \begin{bmatrix} X_I(t) \\ X_Q(t) \end{bmatrix} dt + \begin{bmatrix} b_I & 0 \\ 0 & b_Q \end{bmatrix} \begin{bmatrix} dW_I(t) \\ dW_Q(t) \end{bmatrix}, \begin{bmatrix} X_I(0) \\ X_Q(0) \end{bmatrix} \\ \begin{bmatrix} I(t) \\ Q(t) \end{bmatrix} &= \begin{bmatrix} C_I & 0 \\ 0 & C_Q \end{bmatrix} \begin{bmatrix} X_I(t) \\ X_Q(t) \end{bmatrix} + \begin{bmatrix} D_I & 0 \\ 0 & D_Q \end{bmatrix} \begin{bmatrix} dW_I(t) \\ dW_Q(t) \end{bmatrix} + \begin{bmatrix} f_I(t) \\ f_Q(t) \end{bmatrix} \end{aligned} \tag{12}$$

$\{W_I(t), W_Q(t)\}_{t\geq 0}$ are independent standard Brownian motions which are independent of the initial random variables $X_I(0)$ and $X_Q(0)$ and $\{f_I(s), f_Q(s);\ 0 \leq s \leq t\}$ are random processes representing the LOS components, respectively. The band pass representation of the received signal corresponding to the j^{th} path is expressed as

$$y(t) = \left[\left(C_I\, X_I(t) + f_I(t)\right)\cos\omega_c t - \left(C_Q\, X_Q(t) + f_Q(t)\right)\sin\omega_c t\right] s_l\left(t-\tau_j\right) \qquad (13)$$

2.4 UWB Indoor Wireless Channel Model

For narrow-band systems, complex Gaussian fading is conventionally used to describe the small-scale fading. More precisely, the equivalent complex baseband representation consists of Rayleigh-distributed amplitude and uniformly distributed phase. This can be related theoretically to the fact that a large number of multi-path components fall into each resolvable delay bin, so that the central limit theorem is valid [27]. In UWB systems, this is not true anymore and a number of alternative amplitude distributions have been proposed in the literature. The most common empirically determined amplitude distribution in many UWB environments is Nakagami distribution, which is observed in [28] and [29] and considered in the IEEE 802.15.4a standard. Its probability density function is given by:

$$f(x) = \frac{2}{\Gamma(m)}\left(\frac{m}{\Omega}\right)^m x^{2m-1}\exp\left(-\frac{m}{\Omega}x^2\right) \qquad (14)$$

where $m \geq 0.5$ is the shape parameter, $\Gamma(m)$ is the gamma function and Ω controls the spread of the distribution. The m-parameter is often modeled as a random variable [6]. For integer values of m, the distribution describes m orthogonal independent Rayleigh distributed random variables X_i, the probability density function of random variable Y, is defined as $Y = \sqrt{\sum_{i=1}^{N} X_i^2}$ and is given by a Nakagami distribution with parameter $m = N$.

The general wireless fading channel can be represented in state space form in terms of the inphase and quadrature components for the i^{th} path as [12], [13]

$$\begin{aligned} dX_i^I(t) &= A_I X_i^I(t)\,dt + B_I dW_i^I(t) \\ I_i(t) &= C_I X_i^I(t) + f_i^I(t) \\ dX_i^Q(t) &= A_Q X_i^Q(t)\,dt + B_Q dW_i^Q(t) \\ Q_i(t) &= C_Q X_i^Q(t) + f_i^Q(t) \end{aligned} \qquad (15)$$

where $I_i(t)$ and $Q_i(t)$ are respectively the inphase and quadrature components for the i^{th} path, $X_i^I(t)$ and $X_i^Q(t)$ are respectively the state vectors of the inphase and quadrature components, $\left\{W_i^I(t)\right\}_{t\geq 0}$ and $\left\{W_i^Q(t)\right\}_{t\geq 0}$ are two independent standard Brownian motions which correspond to the inphase and quadrature components respectively, $f_i^I(t)$ and $f_i^Q(t)$ are arbitrary functions representing the line-of-sight (LOS) of the inphase and quadrature components respectively, and

$$A_I = A_Q = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_0 & a_1 & a_2 & & a_{n-1} \end{bmatrix}, \; B_I = B_Q = \begin{bmatrix} b_{n-1} \\ \vdots \\ \vdots \\ b_1 \\ b_0 \end{bmatrix}, \tag{16}$$

$$C_I = C_Q = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix}.$$

It is shown in [13] that a second order state space model (n = 2) is sufficient to capture the dynamics of wireless channel inphase and quadrature components.

In indoor UWB channels, a number of multipath components fall into each resolvable delay bin and the resultant received signal amplitude distribution is Nakagami. Thus, the UWB state space model together with the received signal level, $y(t)$, is represented as:

$$\begin{aligned} dX_i(t) &= A_i X_i(t)\,dt + B_i dW_i(t) \\ dX(t) &= AX(t)\,dt + B\,dW(t) \\ y(t) &= \sqrt{\sum_{i=1}^{N}\left(C(t)\,X_i(t)\right)^2}\; s(t-\tau_i) + D(t)\,v(t) \end{aligned} \tag{17}$$

where

$$
\begin{aligned}
&X_i(t) = \begin{bmatrix} X_i^I(t) & X_i^Q(t) \end{bmatrix}^T, \\
&W_i(t) = \begin{bmatrix} W_i^I(t) & W_i^Q(t) \end{bmatrix}^T, \\
&A_i = \begin{bmatrix} 0 & 1 & 0 & 0 \\ a_0^i & a_1^i & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & a_0^i & a_1^i \end{bmatrix}, \; B_i = \begin{bmatrix} b_1^i & 0 \\ b_0^i & 0 \\ 0 & b_1^i \\ 0 & b_0^i \end{bmatrix}, \\
&X(t) = \begin{bmatrix} X_1(t) & X_2(t) & \cdots & X_N(t) \end{bmatrix}^T, \\
&W(t) = \begin{bmatrix} W_1(t) & W_2(t) & \cdots & W_N(t) \end{bmatrix}^T, \\
&A = \begin{bmatrix} A_1 & 0 \cdots & 0 \\ 0 & A_2 \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 \cdots & A_N \end{bmatrix}, \; B = \begin{bmatrix} B_1 & 0 \cdots & 0 \\ 0 & B_2 \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 \cdots & B_N \end{bmatrix}, \\
&C(t) = \begin{bmatrix} \cos(\omega_c t) & 0 & -\sin(\omega_c t) & 0 \end{bmatrix}, \\
&D(t) = \begin{bmatrix} \cos(\omega_c t) & -\sin(\omega_c t) \end{bmatrix}, \\
&v(t) = \begin{bmatrix} v_I(t) & v_Q(t) \end{bmatrix}^T,
\end{aligned}
\tag{18}
$$

ω_c is the carrier frequency, $v(t)$ is the measurement noise which is assumed to be Gaussian, N is the number of resolvable delay bins, T is vector or matrix transpose and $i = 1, ..., N$.

The model introduced above has many applications. One of them is indoor channel estimation which will be discussed in detail in chapter 3. Since the channel impulse response is infinite, it is better to use finite number of parameters to model the channel. Furthermore, it can be used for power control in chapter 4.

SDE is not the only way to model the TV processes. There is a large body of literature in modeling TV processes such as [83] and [84]. ARMA is one of the processes which are widely used. ARMA assume the channel state is completely observable, which in reality is not the case due to additive noise and require long observation intervals. Figure 1 compares the ARMA models with the proposed SDE models using experimental data and show that the SDE models are a better fit.

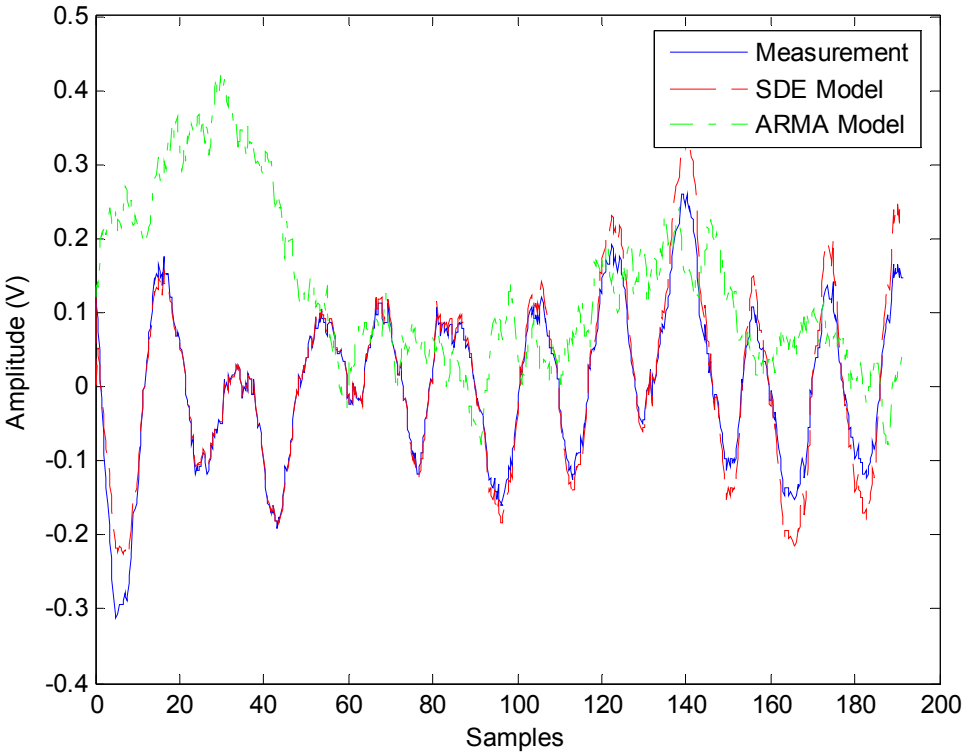

Figure.1 Comparison using received inphase component measurements between proposed SDE and ARMA models

Although the estimation process using SDE models may require more computational cost than ARMA models, the proposed SDE model shows excellent agreement with experimental data.

2.5 Ergodicity of the Channel

Ergodicity means averages computed from a sample of a stochastic process can be ultimately identified with corresponding ensemble averages [46]. It is an important property in classical statistical mechanics. An ergodic communication channel means the channel gain process is ergodic, i.e., the time average is equal to the ensemble average. In other words, the randomness of the channel gain can be averaged out over time. So

long-term constant bit rates can be supported, like the additive white Gaussian noise (AWGN) channel.

In some works such as [26], the channel is assumed to be ergodic; however, in general, this case is not true. There are several conditions that the channel should satisfy to be ergodic. We provide these conditions in this section.

Consider the continuous model in (12), the solution of the this state-space model is given by [52]

$$X_L(t) = \Phi_L(t,t_0)\left[X_L(t_0) + \int_{t_L}^{t} \Phi_L^{-1}(u,t_0) B_L(u) dW_L(u)\right] \tag{19}$$

where $L = I$ or Q, $\Phi_L(t,t_0)$ is the fundamental matrix, defined by the Peano-Baker series

$$\begin{aligned}\Phi(t,\tau) = I &+ \int_\tau^t A(\sigma_1)d\sigma_1 + \int_\tau^t A(\sigma_1)\int_\tau^{\sigma_1} A(\sigma_2)d\sigma_2 d\sigma_1 \\ &+ \int_\tau^t A(\sigma_1)\int_\tau^{\sigma_1} A(\sigma_2)\int_\tau^{\sigma_2} A(\sigma_3)d\sigma_3 d\sigma_2 d\sigma_1 + \cdots\end{aligned} \tag{20}$$

and $\Phi_L(t,t_0)$ satisfies [52]

$$\dot{\Phi}_L(t,t_0) = A_L(t)\Phi_L(t,t_0) \text{ with initial condition } \Phi_L(t_0,t_0) = I$$

where I is the identity matrix. Therefore, the mean of $X_L(t)$ is:

$$E\left[X_L(t)\right] = \Phi_L(t,t_0) E\left[X_L(t_0)\right] \tag{21}$$

and the covariance matrix of $X_L(t)$ is:

$$\sum\nolimits_L(t,s) = \Phi_L(t,t_0)\left[\begin{array}{l} Var\left[X_L(t_0)\right] + \\ \int_{t_0}^{t} \Phi^{-1}{}_L(u,t_0) B_L(u) du \int_{t_0}^{s} B_L^T(u)\left(\Phi_L^{-1}(u,t_0)\right)^T du \end{array}\right] \Phi_L^T(s,t_0) \tag{22}$$

According to [46], there are two conditions which are sufficient for the stochastic process $X_L(t)$ to be ergodic:

a) Suppose that the covariance function $\sum_L(t,s)$ is continuous for $t, s \geq 0$, and that

$$\frac{1}{T}\int_0^T \int_0^T \sum\nolimits_L(t,s) dt ds \to 0, \text{ as } T \to \infty$$

b) The covariance function $\sum_L (t,s)$ should satisfy the relation $\left|\sum_L (t,s)\right| < K\frac{t^\alpha + s^\alpha}{1+\left|t-s\right|^\beta}$,

where K, α and β are constants such that $K>0$, $0 \le 2\alpha < \beta < 1$

However, in practice, we may need to discretize the model. Consider the discrete version of the state space model

$$\begin{aligned} x_I(t+1) &= A_I(t)x_I(t) + B_I(t)u(t) \\ I(t) &= C_I(t)x_I(t) + D_I(t)u(t) \qquad t = 1,2,\ldots \end{aligned} \tag{23}$$

The solution of the discretized state space equation is [52]

$$x_I(t) = \Phi_I(t,t_0)x_0 + \sum_{j=t_0}^{t-1} \Phi_I(t,j+1)B(j)u(j), t \ge t_0 + 1 \tag{24}$$

Where $x_0 = x(t_0)$

$$\Phi_I(t,j) = A(t-1)A(t-2)\cdots A(j),\ t \ge j+1 \tag{25}$$

Therefore, the mean of $x_I(t)$ is

$$E[x_I(t)] = \Phi_L(t,t_0)x_0 \tag{26}$$

And the covariance matrix of $x_I(t)$ is:

$$\begin{aligned} K(s,t) = \Phi_I(t,t_0)[&Var(x_0) + \\ &\sum_{j=t_0}^{t} \Phi_I(t,j+1)B_I(j)u(j)\sum_{i=s_0}^{s} u^T(i)B_I^T(i)\Phi_I^T(s,i+1)]\Phi_I^T(s,s_0) \end{aligned} \tag{27}$$

From [54], for $t = 1,2,\ldots$ let $C(t)$ be the covariance between the t^{th} sample mean M_t and the observation $X(t)$

$$C(t) = Cov[X(t), M_t] = \frac{1}{t}\sum_{s=1}^{t} K(s,t) \tag{28}$$

The necessary and sufficient condition for the sample means of the stochastic process to be ergodic is

$$\lim_{t\to\infty} C(t) = 0 \tag{29}$$

From the conditions mentioned above, a sufficient condition for the channel to be ergodic is the state space realization should be uniform exponential stable, i.e., there exist finite positive constants γ, λ such that for any t_0 and x_0, the corresponding solution (19) should satisfy [52].

$$\left\|x(t)\right\| \le \gamma e^{-\lambda(t-t_0)} \left\|x_0\right\|, t \ge t_0 \tag{30}$$

Chapter 3

UWB Indoor Wireless Channel Estimation

System identification is a common problem in process modeling. It is also an important research area in wireless communications. In this chapter, Kalman Filtering together with Expectation Maximization (EM) algorithm is used to estimate channel parameters as well as the inphase and quadrature components from received signal measurements. The data is obtained from the UT Antenna Laboratory.

3.1 Kalman Filter Theories

In 1960, R.E.Kalman described a recursive solution to the discrete-data linear filtering problem. Currently the Kalman filter has been the subject of extensive research and application, particularly in the area of autonomous or assisted navigation. The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) solution of the least-squares method. The filter is very powerful in several aspects: it supports estimations of past, present and even future states. It can do so even when the precise nature of the modeled system is unknown [40, 59].

The general problem of estimating an unknown state $x(t)$ in engineering applications involves the cost function [59, 63]

$$J(t) = E\left[\left(x(t) - \hat{x}(t \mid t)\right)^T \left(x(t) - \hat{x}(t \mid t)\right)\right] \tag{31}$$

where $\hat{x}(t \mid t)$ denotes the estimated state of $x(t)$ and $E(\cdot)$ is the expectation operator.

The problem is to minimize the estimation error criterion $J(t)$ defined in (31). If the measurement noise is uncorrelated, zero-mean and white Gaussian, then the Kalman

filter is the best linear solution to the above problem. The Kalman filters give the best estimate when the objective function is linear. The optimal state estimate $\hat{x}(t \mid t)$ is expressed as

$$\hat{x}(t \mid t) = E\left[x(t) \mid Y_N(t)\right] \tag{32}$$

where $Y_N(t)$ is observation data.

- The discrete-time Kalman filter for Rayleigh fading channels [59,61,62]

Kalman filter equations provide solutions for estimating the least-square components of the received signal, such as inphase component, quadrature component, envelope component and phase component $\left\{I(t_n), Q(t_n), r^2(t_n), \phi(t_n)\right\}_{t_n \geq 0}$

The Kalman filter equations for the Rayleigh channels are given as follows [59, 61]

Initialization conditions

$$\hat{x}(0) = E\left[x(0)\right], \sigma^2 = Var(x(0)); \tag{33}$$

Error Covariance (Riccati Equation):

$$\begin{aligned} P(t_{n+1}) &= AP(t_n)A^T + BQB^T - AP(t_n)C^T\left[CP(t_n)C^T + DRD^T\right]^{-1} CP(t_n)A^T \\ P_0 &= \sigma^2 \end{aligned} \tag{34}$$

In this case (Rayleigh fading channel),

$R = \dfrac{E_0}{10^{\frac{(SNR)_{dB}}{10}}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, E_0 is a constant and SNR denotes Signal to Noise Ratio.

Filter Gain

$$K(t_n) = AP(t_n)C^T\left[CP(t_n)C^T + DRD^T\right]^{-1} \tag{35}$$

The state estimate

$$\hat{x}(t_{n+1} \mid t_{n+1}) = A\hat{x}(t_n \mid t_n) + K(t_n)\left[y(t_n) - C\hat{x}(t_n \mid t_n)\right] \tag{36}$$

Estimations of the inphase and quadrature components

$$\begin{aligned} \hat{I}(t_n) &= E\left[I(t_n) \mid Y_N(t)\right] = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}\hat{x}(t_n \mid t_n) \\ \hat{Q}(t_n) &= E\left[Q(t_n) \mid Y_N(t)\right] = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}\hat{x}(t_n \mid t_n) \end{aligned} \tag{37}$$

Estimation of the square-envelope component:

$$\hat{r}^2(t_n) = \hat{I}^2(t_n) + \hat{Q}^2(t_n) + e_I^2(t_n) + e_Q^2(t_n) \tag{38}$$

In this equation, $e_I^2(t_n)$ and $e_Q^2(t_n)$ denote the mean square errors (MSEs) and are defined as:

$$e_I^2(t_n) = E\left[(I(t_n) - \hat{I}(t_n))^2 \right] \; and \; e_Q^2(t_n) = E\left[(Q(t_n) - \hat{Q}(t_n))^2 \right] \tag{39}$$

Estimate of the phase component:

$$\hat{\phi}(t_n) = \tan^{-1}\left(\frac{\hat{Q}(t_n)}{\hat{I}(t_n)} \right) \tag{40}$$

- The discrete-time Kalman filter of the frequency-selective fading channels can be presented as [59, 61, 62]

Error Covariance (Riccati Equation):

$$\begin{aligned} &P(t_{n+1}) = AP(t_n)A^T + BQB^T - AP(t_n)C^T \left[CP(t_n)C^T + DRD^T \right]^{-1} CP(t_n)A^T \\ &P_0 = \sigma^2 \end{aligned} \tag{41}$$

In this case (frequency-selective fading channels),

$$R = \frac{E_0}{10^{\frac{(SNR)_{dB}}{10}}} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{N \times N}, Q = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{N \times N} \tag{42}$$

Filter Gain

$$K(t_n) = AP(t_n)C^T \left[CP(t_n)C^T + DRD^T \right]^{-1} \tag{43}$$

Estimation of the state

$$\hat{x}(t_{n+1} \mid t_{n+1}) = A\hat{x}(t_n \mid t_n) + K(t_n)\left[y(t_n) - C\hat{x}(t_n \mid t_n) - L(t_n) \right] \tag{44}$$

Estimations of the inphase and quadrature components on the i^{th} path:

$$\begin{aligned} &\hat{I}_i(t_n) = E\left[I_i(t_n) \mid Y_N(t) \right] = \left[0 \; 0 \; 0 \; 0 \cdots \overbrace{1 \; 0 \; 0 \; 0}^{i} \cdots 0 \; 0 \; 0 \; 0 \right] \hat{x}(t_n \mid t_n) \\ &\hat{Q}_i(t_n) = E\left[Q_i(t_n) \mid Y_N(t) \right] = \left[0 \; 0 \; 0 \; 0 \cdots \overbrace{0 \; 0 \; 1 \; 0}^{i} \cdots 0 \; 0 \; 0 \; 0 \right] \hat{x}(t_n \mid t_n) \end{aligned} \tag{45}$$

Estimate of the square-envelope component:

$$\hat{r}_i^2(t_n) = \hat{I}_i^{\,2}(t_n) + \hat{Q}_i^{\,2}(t_n) + e_{I_i}^2(t_n) + e_{Q_i}^2(t_n) \tag{46}$$

In this equation, $e_{I_i}^2(t_n)$ and $e_{Q_i}^2(t_n)$ denote the mean square errors (MSEs) on the i^{th} path and are defined as:

$$e_{I_i}^2(t_n) = E\left[(I_i(t_n) - \hat{I}_i(t_n))^2\right] \; and \; e_{Q_i}^2(t_n) = E\left[(Q_i(t_n) - \hat{Q}_i(t_n))^2\right] \tag{47}$$

Estimate of phase component:

$$\hat{\phi}_i(t_n) = \tan^{-1}\left(\frac{\hat{Q}_i(t_n)}{\hat{I}_i(t_n)}\right) \tag{48}$$

Consequently, Kalman filter is used to solve the discrete time Rayleigh fading channels and discrete time frequency fading channels, components like inphase, quadrature and phase can be estimated from the received signal. These methods will be extended to Nakagami fading channels in this chapter to estimate components for indoor multipath channels.

3.2 Introduction to System Identification

System identification is a process of constructing a mathematical model for a dynamic system from observations and prior knowledge (see Figure 2). Least-square or maximum likelihood criterion is usually used with some mathematical models and tools. Computer programs are often used to find general mathematical descriptions that give the best fit to a series of recorded input and output signals. In practical situations, the observation process is monitored at discrete times only, but the system identification could be determined in either continuous-time or discrete-time. The mathematical models via system identification are based on observations, so they are more accurate. Also, the process of system identification is fast, predictable and completely universal [59, 63].

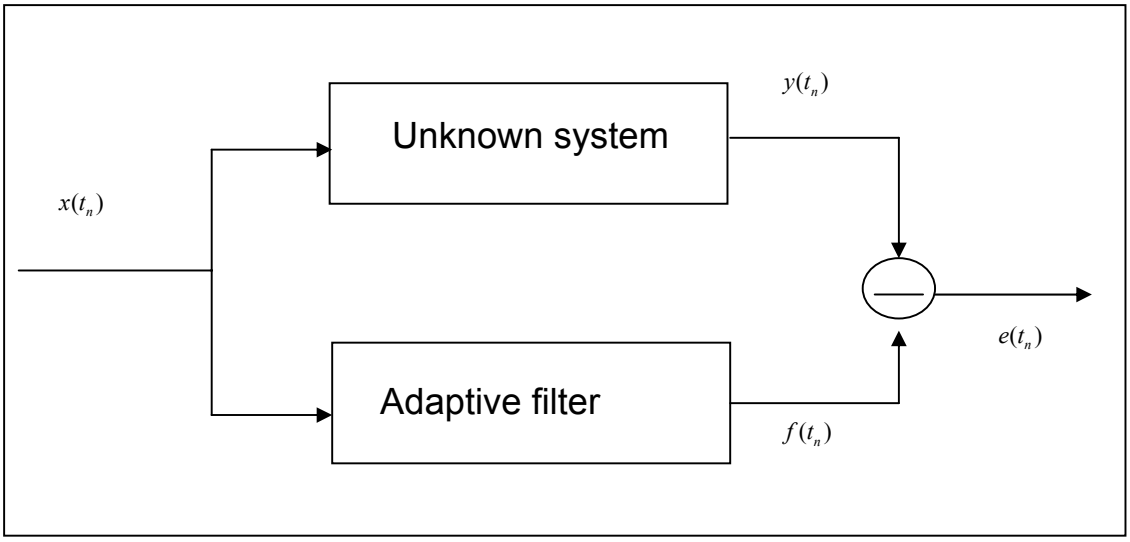

Figure.2 System identification structure

System identification is usually formulated as an optimization problem. It is assumed that the unknown system is needed to be identified and the adaptive filter is a known system. The variable $x(t_n)$ is usually a random variable [54]. When $x(t_n)$ is fed to both the unknown system and the adaptive filter, the difference between the system response $y(t_n)$ and the adaptive filter output $f(t_n)$ is termed as the error signal $\varepsilon(t_n)$. The adaptive filter coefficients are updated according to some mathematical algorithms in order to minimize $\varepsilon(t_n)$. If $\varepsilon(t_n)$ is less than a small given number, the adaptive filter response is said to give a good estimate of the unknown system [59, 63].

The algorithms of updating the adaptive filter coefficients are the main subject of system identification. Several algorithms with different convergence speed, computational complexity, performance and stability, have been considered. For example, the maximum likelihood estimation (MLE) is popular for deriving estimators [60, 63].

- Maximum likelihood estimation (MLE)

The basic idea of the MLE is to obtain the most likely values of system parameters for a given distribution that best describe the samples. The goal of the MLE is to determine the system parameters which maximize the probability of the samples [64].

Let $X_1,\ldots X_n$ be independent and identically distributed (i.i.d.) samples, $\Theta=\left(\theta_1,\theta_2,\cdots,\theta_k\right)$ denote the k unknown system parameters needed to be estimated. The probability mass function is denoted by $f\left(x;\Theta\right)$. Therefore, the likelihood function is expressed as the following equation [64]

$$L(x_1,\ldots x_n\mid\theta_1,\theta_2,\cdots\theta_k)=L=\prod_{i=1}^{n}f(x_i;\Theta);\quad i=1,2,\cdots,n \tag{49}$$

In the same way, the logarithmic likelihood function of L is given as

$$\Lambda=\ln L=\sum_{i=1}^{n}\ln f(x_i,\Theta) \tag{50}$$

A necessary condition to maximize Λ is given by the following equation

$$\frac{\partial(\Lambda)}{\partial\theta_i}=0;i=1,2,\cdots,k \tag{51}$$

The above equation is known as the maximum likelihood equation. The system parameters $\Theta=\left(\theta_1,\theta_2,\cdots,\theta_k\right)$ are obtained by maximizing L or Λ.

Intuitively, the MLE is a reasonable choice for an estimator, since the parameter calculated from the observed sequence is the most likely. In general, the MLE is a good point estimator. However, there is one inherent drawback associated with finding the maximum of a function and hence the maximum likelihood estimation. That is, how sensitive is the estimate to small change in the data, or how to deal with the problem of "missing data" problem [64]? Further, since both the channel states and channel parameters are unknown, how can we estimate them only with the measurement? Rather than detailing the procedure for solving for the MLE, we specify an algorithm that is guaranteed to converge to the MLE in the next section.

3.3 Expectation Maximization (EM) Algorithm

This section describes the procedure used to estimate the channel model parameters and states associated with the state space model using the EM algorithm [39] combined with Kalman filtering [40]. However, for simplicity we consider only the discrete-time version of (15) given by

$$\begin{aligned} x_{t+1} &= A_t x_t + B_t w_t \\ y_t &= C_t x_t + D_t v_t \end{aligned} \tag{52}$$

where $x_t \in \Re^n$ is a state vector, $y_t \in \Re^d$ is a measurement vector, $w_t \in \Re^m$ is state noise and $v_t \in \Re^d$ is measurement noise. The noise processes w_t and v_t are assumed to be independent zero mean and unit variance Gaussian processes. Further, the noises are independent of the initial state x_0, which is assumed to be Gaussian distributed.

The unknown system parameters $\theta_t = \{A_t, B_t, C_t, D_t\}$ as well as the system states x_t are estimated through the received signal measurement data $Y_N = \{y_1, y_2, ..., y_N\}$. The methodology used is recursive and based on the EM algorithm together with the Kalman filter. The Kalman filter is introduced next [57].

- Channel state estimation: the Kalman filter

The Kalman filter estimates the channel states x_t for given system parameter $\theta = \{A, B, C, D\}$ and measurements Y_N. It is described by the following equations [40]:

$$\begin{aligned} \hat{x}_{t|t} &= A\hat{x}_{t-1|t-1} + P_{t|t} C^T D^{-2} \left(y_t - CA\hat{x}_{t-1|t-1} \right) \\ \hat{x}_{t|t-1} &= A\hat{x}_{t-1|t-1} \\ \hat{x}_{0|0} &= m_0 \end{aligned} \tag{53}$$

where $t = 0, 1, 2, ..., N$ and $P_{t|t}$ is given by

$$\begin{aligned}
\overline{P}_{t|t}^{-1} &= P_{t-1|t-1}^{-1} + A^T B^{-2} A \\
P_{t|t}^{-1} &= C^T D^{-2} C + B^{-2} - B^{-2} \overline{P}_{t|t} A^T B^{-2} \\
P_{t|t-1} &= A P_{t-1|t-1} A^T + B^2
\end{aligned} \tag{54}$$

The channel parameters $\theta = \{A, B, C, D\}$ are estimated using the EM algorithm which is introduced next.

- Channel parameter estimation: the EM algorithm

The EM algorithm uses a bank of Kalman filters to yield a maximum likelihood (ML) parameter estimate of the Gaussian state space model.

Let $\theta = \{A, B, C, D\}$ denote the system parameters, P_0 denotes a fixed probability measure; and $\{P_\theta;\ \theta \in \Theta\}$ denotes a family of probability measures induced by the system parameters θ. If the original model includes a white noise sequence, then $\{P_\theta;\ \theta \in \Theta\}$ is absolutely continuous with respect to P_0 [39]. Moreover, it is shown that under P_0 we have

$$P_0: \quad \begin{cases} x_{t+1} = w_t \\ y_t = v_t \end{cases} \tag{55}$$

The EM algorithm is an iterative scheme for computing the ML estimate of the system parameters θ, given the data Y_N. The likelihood function for computing the estimate of the parameters θ based on the observation in Y_N is [57]

$$L_N(\theta) = E_0\left[\frac{dP_\theta}{dP_0} | Y_N\right] \tag{56}$$

And the maximum likelihood estimate is

$$\hat{\theta} \in \arg\max_{\theta \in \Theta} L_N(\theta) \tag{57}$$

Specifically, each iteration of the EM algorithm consists of two steps: The expectation step and the maximization step [57, 64].

The expectation step evaluates the conditional expectation of the log-likelihood function given the complete data, which is described by [57, 64]

$$\Lambda(\theta,\widehat{\theta}_l)=E_{\widehat{\theta}_l}\left\{\log\frac{dP_\theta}{dP_{\widehat{\theta}_l}}\mid Y_N\right\} \tag{58}$$

where θ_l denote the estimated system parameters at the l^{th} iteration. The maximization step finds:

$$\widehat{\theta}_{l+1}\in\arg\max_{\theta\in\Theta}\ \Lambda\left(\theta,\,\theta_l\right) \tag{59}$$

The expectation and maximization steps are repeated until the sequence of model parameters converge by testing $\left\|\widehat{\theta}_l-\widehat{\theta}_{l+1}\right\|$ to be less than the required accuracy.

The EM algorithm is described by the following equations [39]

$$\begin{aligned}
\hat{A}&=E\left(\sum_{t=1}^{N}x_tx_{t-1}^T\mid Y_N\right)\times\left[E\left(\sum_{t=1}^{N}x_tx_t^T\mid Y_N\right)\right]^{-1}\\
\hat{B}^2&=\frac{1}{N}E\left(\sum_{t=1}^{N}\left(\left(x_t-Ax_{t-1}\right)\left(x_t-Ax_{t-1}\right)^T\right)\mid Y_N\right)\\
&=\frac{1}{N}E\left(\sum_{t=1}^{N}\left(\left(x_tx_t^T\right)-A\left(x_tx_{t-1}^T\right)^T-\left(x_tx_{t-1}^T\right)A^T+A\left(x_{t-1}x_{t-1}^T\right)A^T\right)\mid Y_N\right)\\
\hat{C}&=E\left(\sum_{t=1}^{N}y_tx_t^T\mid Y_N\right)\times\left[E\left(\sum_{t=1}^{N}x_tx_t^T\mid Y_N\right)\right]^{-1}\\
\hat{D}^2&=\frac{1}{N}E\left(\sum_{t=1}^{N}\left(\left(y_t-Cx_t\right)\left(y_t-Cx_t\right)^T\right)\mid Y_N\right)\\
&=\frac{1}{N}E\left(\sum_{t=1}^{N}\left(\left(y_ty_t^T\right)-A\left(y_tx_t^T\right)C^T-C\left(y_tx_t^T\right)^T+C\left(x_tx_t^T\right)C^T\right)\mid Y_N\right)
\end{aligned} \tag{60}$$

where $B^2=BB^T$, $D^2=DD^T$ $E(\cdot)$ denotes the expectation operator. These system parameters $\left\{\widehat{A},\widehat{B}^2,\hat{C},\hat{D}^2\right\}$ are computed from the following conditional expectations [39]

$$
\begin{aligned}
L_N^{(1)} &= E\left\{\sum_{t=1}^{N} x_t^T Q x_t \mid Y_N\right\} \\
L_N^{(2)} &= E\left\{\sum_{t=1}^{N} x_{t-1}^T Q x_{t-1} \mid Y_N\right\} \\
L_N^{(3)} &= E\left\{\sum_{t=1}^{N}\left[x_t^T R x_{t-1} + x_{t-1}^T R^T x_t\right] \mid Y_N\right\} \\
L_N^{(4)} &= E\left\{\sum_{t=1}^{N}\left[x_t^T S y_t + y_t^T S^T x_t\right] \mid Y_N\right\}
\end{aligned} \tag{61}
$$

where Q, R and S are given by

$$
\begin{aligned}
Q &= \left\{\frac{e_i e_j^T + e_j e_i^T}{2};\ i, j = 1, 2, \ldots n\right\} \\
R &= \left\{\frac{e_i e_j^T}{2};\ i, j = 1, 2, \ldots n\right\} \\
S &= \left\{\frac{e_i e_n^T}{2};\ i = 1, 2, \ldots m;\ n = 1, 2, \ldots d\right\}
\end{aligned} \tag{62}
$$

In which e_i is the unit vector in the Euclidean space; that is $e_i = 1$ in the i^{th} position and 0 elsewhere. For instance, consider the case $n = m = 2$, then $E\left(\sum_{t=1}^{N} x_t x_{t-1}^T \mid Y_N\right)$ can be computed as [57]

$$
E\left(\sum_{t=1}^{N} x_t x_{t-1}^T \mid Y_N\right) = \begin{bmatrix} L_N^{(3)}(R_{11}) & L_N^{(3)}(R_{12}) \\ L_N^{(3)}(R_{21}) & L_N^{(3)}(R_{22}) \end{bmatrix} \tag{63}
$$

where $R_{ij} = \left\{\frac{e_i e_j^T}{2};\ i, j = 1, 2\right\}$. The other terms in (62) can be computed similarly.

The conditional expectations $\left\{L_N^{(1)}, L_N^{(2)}, L_N^{(3)}, L_N^{(4)}\right\}$ can be estimated from measurements Y_N as follows

Filter estimate of $L_N^{(1)}$ [57]

$$
\begin{aligned}
L_N^{(1)} &= E\left\{\sum_{t=1}^{N} x_t^T Q x_t \mid Y_N\right\} \\
&= -\frac{1}{2} Tr\left(N_N^{(1)} P_{N|N}\right) - \frac{1}{2}\sum_{t=1}^{N} Tr\left(N_{t-1}^{(1)} \overline{P}_{t|t}\right) \\
&\quad -\frac{1}{2}\sum_{t=1}^{N}\begin{pmatrix} -2x_{t|t}^T P_{t|t}^{-1} r_t^{(1)} + 2x_{t|t-1}^T P_{t|t-1}^{-1} r_{t|t-1}^{(1)} - x_{t|t}^T N_t^{(1)} x_{t|t} \\ +x_{t|t-1}^T B^{-2} A\overline{P}_{t|t} N_{t-1}^{(1)} \overline{P}_{t|t} A^T B^{-2} x_{t|t-1} \end{pmatrix}
\end{aligned} \tag{64}
$$

where $Tr(\cdot)$ denotes the matrix trace. In (64), $r_t^{(1)}$ and $N_t^{(1)}$ satisfy the following recursions:

$$
\begin{cases}
r_t^{(1)} = \left(A - P_{t|t} C^T D^{-2} C A\right) r_{t-1}^{(1)} + 2P_{t|t} Q x_{t|t-1} \\
\qquad\qquad - P_{t|t} N_t^{(1)} P_{t|t} C^T D^{-2}\left(y_t - C x_{t|t-1}\right) \\
r_{t|t-1}^{(1)} = A r_t^{(1)} \\
r_0^{(1)} = 0_{m\times 1}
\end{cases}
$$

$$
\begin{cases}
N_t^{(1)} = B^{-2} A\overline{P}_{t|t} N_{t-1}^{(1)} \overline{P}_{t|t} A^T B^{-2} - 2Q \\
N_0^{(1)} = 0_{m\times m}
\end{cases} \tag{65}
$$

Filter estimate of $L_N^{(2)}$ [57]

$$
\begin{aligned}
L_N^{(2)} &= E\left\{\sum_{t=1}^{N} x_{t-1}^T Q x_{t-1} \mid Y_N\right\} \\
&= E_\theta\left\{x_0^T Q x_0 \mid Y_N\right\} + E_\theta\left\{\sum_{t=2}^{N} x_{t-1}^T Q x_{t-1} \mid Y_N\right\} \\
&= E_\theta\left\{x_0^T Q x_0 \mid Y_N\right\} + E_\theta\left\{\sum_{t=1}^{N} x_t^T Q x_t \mid Y_N\right\} - E_\theta\left\{x_N^T Q x_N \mid Y_N\right\}
\end{aligned} \tag{66}
$$

Therefore, $L_N^{(2)}$ can be obtained from $L_N^{(1)}$.

Filter estimate of $L_N^{(3)}$ [57]

$$
\begin{aligned}
L_N^{(3)} &= E\left\{\sum_{t=1}^{N}\left(x_t^T R x_{t-1} + x_{t-1}^T R^T x_t\right) \mid Y_N\right\} \\
&= -\frac{1}{2} Tr\left(N_N^{(3)} P_{N|N}\right) - \frac{1}{2}\sum_{t=1}^{N} Tr\left(N_{t-1}^{(3)} \overline{P}_{t|t}\right) \\
&\quad -\frac{1}{2}\sum_{t=1}^{N}\begin{pmatrix} -2x_{t|t}^T P_{t|t}^{-1} r_t^{(3)} + 2x_{t|t-1}^T P_{t|t-1}^{-1} r_{t|t-1}^{(3)} - x_{t|t}^T N_t^{(3)} x_{t|t} \\ +x_{t|t-1}^T B^{-2} A\overline{P}_{t|t} N_{t-1}^{(3)} \overline{P}_{t|t} A^T B^{-2} x_{t|t-1} \end{pmatrix}
\end{aligned} \tag{67}
$$

In this case, $r_t^{(3)}$ and $N_t^{(3)}$ satisfy the following recursions

$$\begin{cases} r_t^{(3)} = \left(A - P_{t|t}C^T D^{-2}CA\right) r_{t-1}^{(3)} - P_{t|t}N_t^{(3)}P_{t|t}C^T D^{-2}\left(y_t - Cx_{t|t-1}\right) \\ \qquad\quad + \left(2P_{t|t}R + 2P_{t|t}B^{-2}A\overline{P}_{t|t}R^T A\right) x_{t-1|t-1} \\ r_{t|t-1}^{(3)} = Ar_t^{(3)} \\ r_0^{(3)} = 0_{m\times 1} \end{cases} \tag{68}$$

$$\begin{cases} N_t^{(3)} = B^{-2}A\overline{P}_{t|t}N_{t-1}^{(3)}\overline{P}_{t|t}A^T B^{-2} - 2R\overline{P}_{t|t}A^T B^{-2} \\ \qquad - 2B^{-2}A\overline{P}_{t|t}R^T \\ N_0^{(3)} = 0_{m\times m} \end{cases}$$

Filter estimate of $L_N^{(4)}$ [57]

$$\begin{aligned} L_N^{(4)} &= E\left\{\sum_{t=1}^{N}\left(x_t^T Sy_t + y_t^T S^T x_t\right) \mid Y_N\right\} \\ &= \sum_{t=1}^{N}\left(x_{t|t}^T P_{t|t}^{-1} r_t^{(4)} - x_{t|t-1}^T P_{t|t-1}^{-1} r_{t|t-1}^{(4)}\right) \end{aligned} \tag{69}$$

Where $r_t^{(4)}$ satisfy the following recursions

$$\begin{cases} r_t^{(4)} = (A - P_{t|t}C^T D^{-2}CA) r_{t-1}^{(4)} + 2P_{t|t}Sy_t \\ r_{t|t-1}^{(4)} = Ar_t^{(4)} \\ r_0^{(4)} = 0_{m\times 1} \end{cases} \tag{70}$$

Using the filters for $L_N^{(i)}$ $(i = 1,2,3,4)$ and the Kalman filter described earlier, the system parameters $\theta_t = \{A_t, B_t, C_t, D_t\}$ can be estimated through the EM algorithm described in (61).

Since the original purpose of EM algorithm was to provide the iterative computation of maximum likelihood estimates, it is very important to know whether this computation converge or not, what is the condition for it to converge, even if it converge, does it converge to a global maximum, a local maximum, or a stationary value?

It is mentioned in [70] that using Jensen's inequality it can be shown that the sequence of model estimates $\{\theta_l\}, l \in Z^+$ from the EM algorithm are such that the sequence of likelihood $\Lambda(\theta, \theta_l), l \in Z^+$ is monotonically increasing with equality if and only if $\theta_{l+1} = \theta_l$.

The sufficient conditions for EM algorithm to converge are summarized to the following [70, 71]:

i. The parameter space Θ is a subset of some finite dimensional Euclidean space R^r

ii. $\Omega_{\theta_0} = \left\{\theta \in \Theta : L_N(\theta) \geq L_N\left(\hat{\theta}_0\right)\right\}$ is compact for any $L_N\left(\hat{\theta}_0\right) > -\infty$

iii. L_N is continuous in Θ and differentiable in the interior of Θ

iv. The function $\Lambda(\theta, \hat{\theta}_l)$ is continuous both in θ and $\hat{\theta}_l$

Then by [71], the limit of the sequence of EM estimates $\left\{\hat{\theta}_l\right\}$ is a stationary point $\bar{\theta}$ of L_N.

Also, $L_N\left(\hat{\theta}_l\right)$ converges monotonically to $\bar{L}_N = L_T\left(\bar{\theta}\right)$ for some stationary point $\bar{\theta}$.

3.4 Expectation Maximization (EM) Algorithm based on Experiment Data

This section introduces mathematical models for UWB indoor wireless channels based on the data measured in the UT's Antenna Laboratory. Each model is used to identify system parameters using EM algorithm together with the Kalman filter.

A simple Gaussian pulse with clean pulse shape and narrow pulse width has been chosen as UWB source signal. The experiment setup of the developed system is comprised of a 300psec Gaussian pulse (Figure 3) generated by a step-recovery diode pulse generator [79]. For pico-second pulses, step-recovery diodes are the best choice for the UWB pulse generation because they are not only simple to design, inexpensive, but also having low power consumption with relatively high output voltage swings [79].

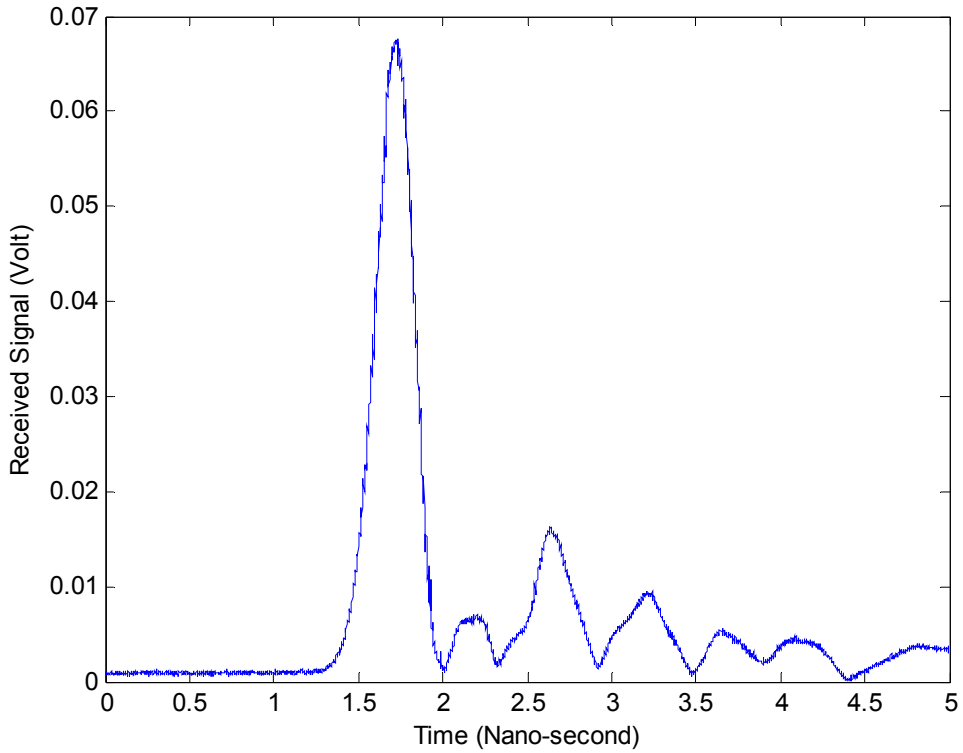

Figure.3 300 psec pulse shape

A step by step model of output pulse formation is given in detail to demonstrate the source of the distortion. The step-recovery diode has low impedance in the ON state and much higher impedance in the OFF state. After fast transition to OFF state, a step signal is generated and would propagate towards both the positive x-axis (step 'A') and in the direction of the short-circuited line arrives at the end and its completely reflected back out of phase and is step 'B'. Finally, step signal 'A' and 'B' combine to produce a Gaussian pulse with a width corresponding to the round trip delay along the short-circuited stub [79].

The modulation scheme used is on-off keying. The pulse modulates a carrier signal centered at 8 GHz, which is then transmitted through an Omni-directional UWB antenna. The modulated pulse signal has a 10-dB bandwidth of approximately 6 GHz, exceeding the 500-MHz minimum bandwidth required under the FCC rules governing UWB. Its

average output power spectral density satisfies the FCC indoor limit by a margin of more than 3 dB for a majority of the unstable bandwidth. The 8-GHz carrier signal leaks through the mixer and is shown as a peak at -16 dBm. This leakage could be suppressed by adding a band-notched filter or utilizing a band-notched monopole [78].

Multiple directional Vivaldi sub-array receiving antennas are located at distinct position in an indoor environment to receive to modulated pulse signal. Each received modulated Gaussian pulse is amplified through a low noise amplifier (LNA) and then demodulated to the I/Q components and stored in a multi-channel Tektronix TDS8200 sampling oscilloscope [78].

The Gaussian pulse with the signal-to-noise ratio of approximately 20 dB was detected at a range of up to 3.8m. The sampler has a sensitivity of -45 dBm. The two amplification stages increase the signal by 25 dB, while the I/Q down-converter has 8 dB conversion losses. This makes the overall sensitivity of each base station around -62 dBm. The dynamic range of the system is limited by the sampler and is over 50 dB [78].

A. Resolvable multipath experiment

Initially, the transmitter was put next to the metal wall. The receiving antenna was placed 10cm away from the transmitter and then moved along the same direction away from the transmitter for 20cm, 50cm, 1m and 2m. Then the transmitter was moved along the metal wall 70 cm away from the original place, the receiver repeated the same process. The 2D scenario is shown in Figure 4.

Under this situation, the transmitted signal is reflected by the metal wall, the ceiling, the floor, or scattered by the corner of the metal table and then come back to the receiver. Since there are no dense reflecting sources around the antenna, the coming multipath signal can be distinguished very clearly. See Figure 5

From Figure 5 it is very clear that received signal has three paths. They are treated as Nakagami fading separately. The measurement data from three paths are amplified by 1000 and fit to the Nakagami distribution respectively (see Figure 6 to Figure 8).

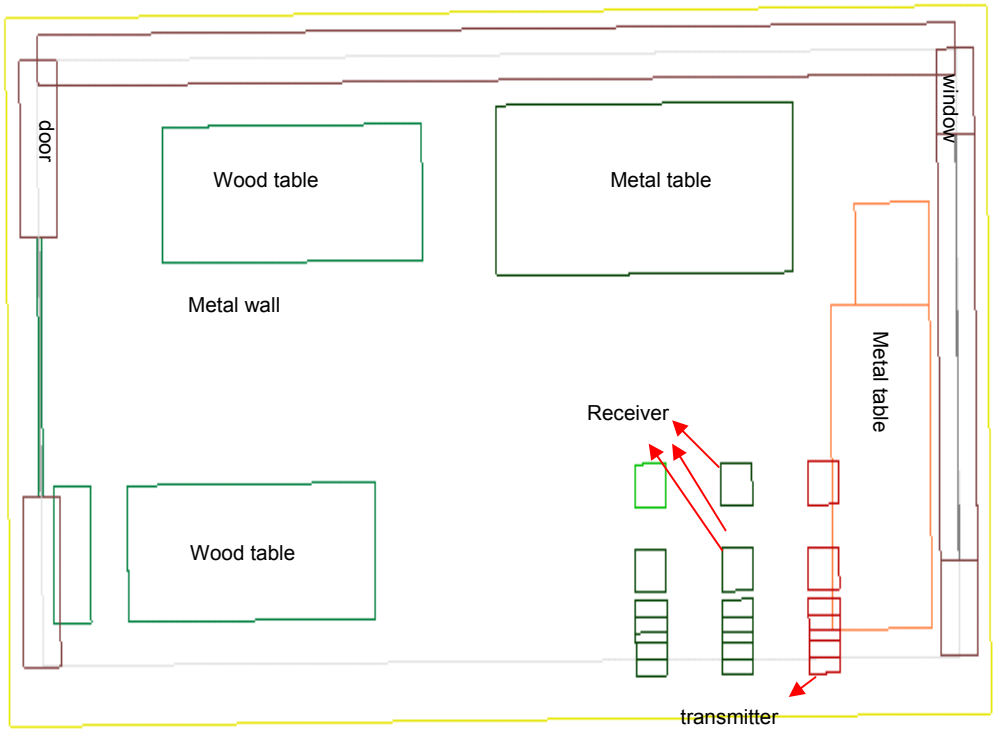

Figure.4 2D scenario of indoor measurement

The parameters which fit the 1^{st} path are $m = 2.5, \sigma = 0.28$, while the parameters fit the 2^{nd} path are $m = 4.7, \sigma = 0.4$, the 3^{rd} path are $m = 2.5, \sigma = 0.18$. The estimated system parameters are computed as follows

$$\hat{A} = \begin{bmatrix} \hat{A}_1 & 0 & 0 \\ 0 & \hat{A}_2 & 0 \\ 0 & 0 & \hat{A}_3 \end{bmatrix}, \quad B\hat{B}^T = \begin{bmatrix} B\hat{B}_1^T & 0 & 0 \\ 0 & B\hat{B}_2^T & 0 \\ 0 & 0 & B\hat{B}_3^T \end{bmatrix}$$
$$\hat{C} = \begin{bmatrix} \hat{C}_1 & \hat{C}_2 & \hat{C}_3 \end{bmatrix} \tag{71}$$

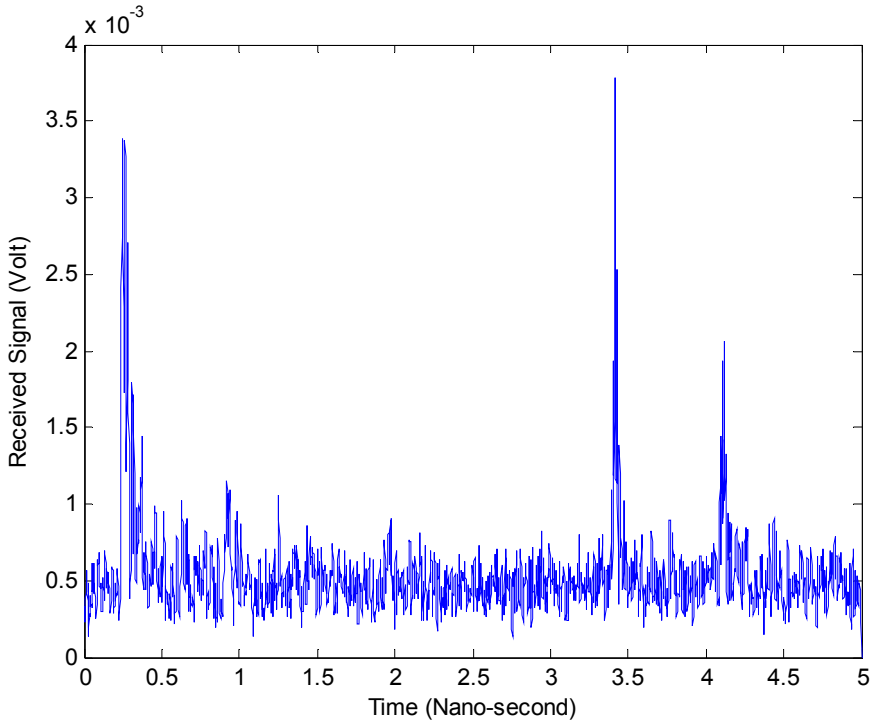

Figure.5 Resolvable multipath received signal

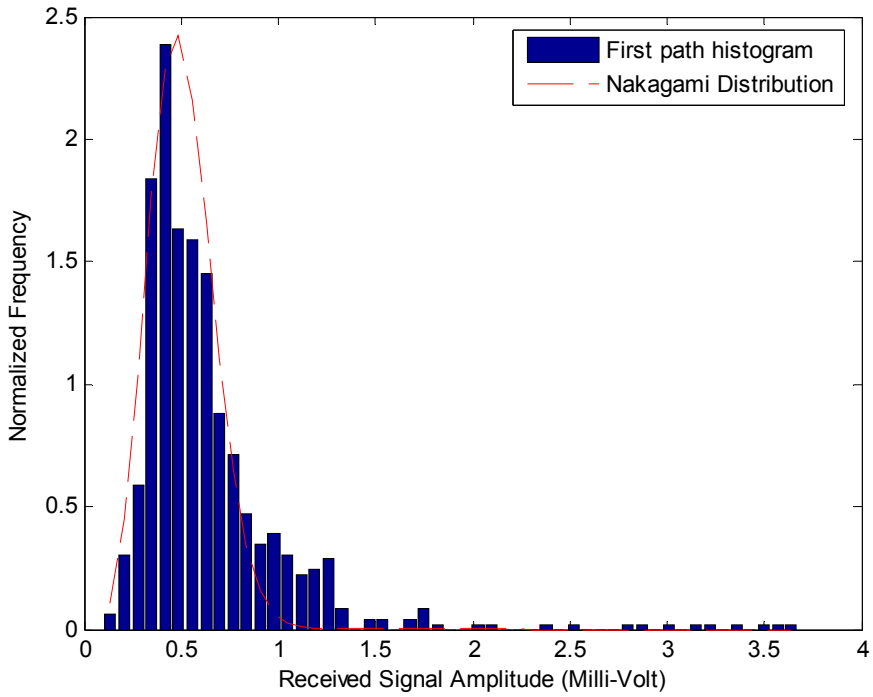

Figure.6 Nakagami distribution fitting the data (1^{st} path)

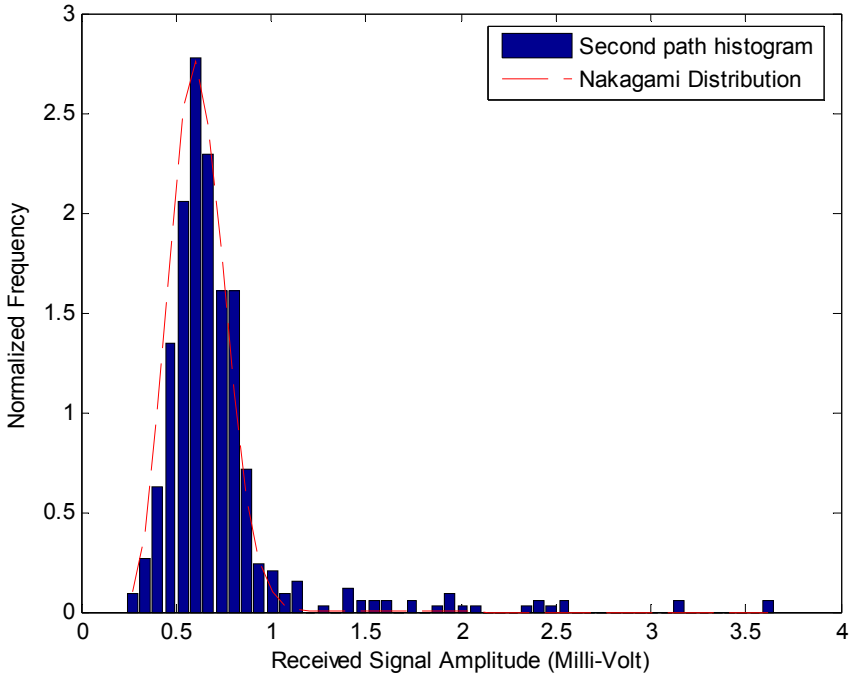

Figure.7 Nakagami distribution fitting the data (2^{nd} path)

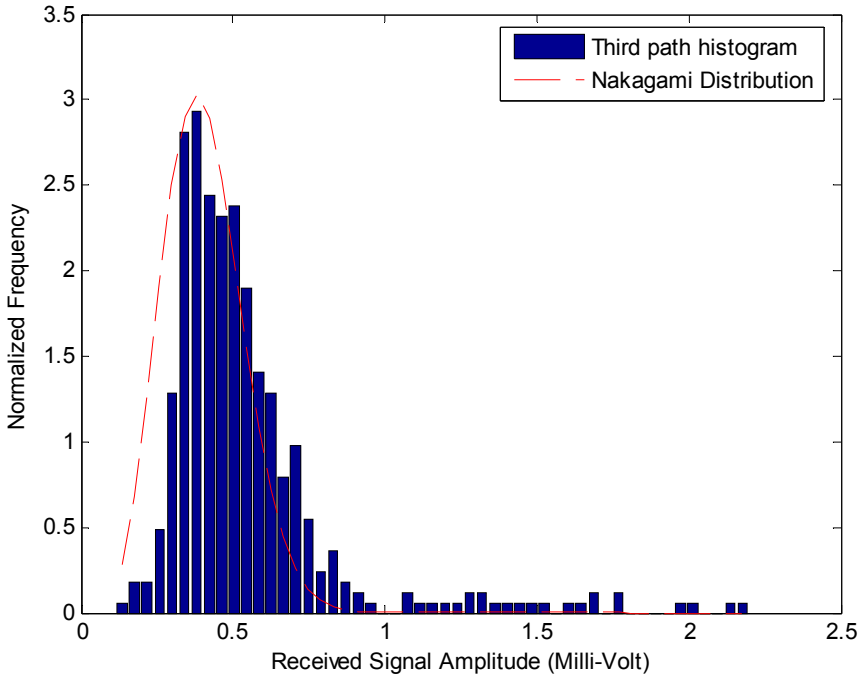

Figure.8 Nakagami distribution fitting the data (3^{rd} path)

The estimated results are the follows

$$\hat{A}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.1928 & 0.0556 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0.1484 & -0.0966 \end{bmatrix},$$

$$\hat{A}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.2917 & 0.0556 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0.3563 & -0.0682 \end{bmatrix}$$

$$\hat{A}_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.2014 & -0.0586 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0.1881 & -0.2669 \end{bmatrix} \quad (72)$$

$$B\hat{B}_1^{\,T} = \begin{bmatrix} 0.1699 & 0.2451 & 0.0122 & 0.0009 \\ 0.0004 & 0.0961 & 0.0287 & -0.0004 \\ 0.0122 & 0.0087 & 0.1628 & 0.0213 \\ 0.0077 & 0.0254 & 0.2959 & -0.0066 \end{bmatrix}$$

$$B\hat{B}_2^{\,T} = \begin{bmatrix} 0.0006 & 0.0182 & 0.1429 & 0.0077 \\ 0.0043 & -0.0256 & 0.0076 & -0.2597 \\ 0.0795 & 0.0056 & 0.0043 & 0.0218 \\ -0.0177 & 0.0192 & 0.1541 & -0.1632 \end{bmatrix}$$

$$B\hat{B}_3^{\,T} = \begin{bmatrix} 0.2561 & 0.0013 & 0.3310 & 0.0069 \\ 0.0129 & 0.0213 & -0.0107 & 0.0921 \\ 0.0044 & 0.0015 & 0.0824 & 0.0076 \\ 0.1541 & 0.1429 & 0.1632 & -0.3334 \end{bmatrix}$$

$$\hat{C}_1 = \hat{C}_2 = \hat{C}_3 = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}$$

Figures 9 to Figure 11 show the compared results between measurement data and estimated data for three different paths.

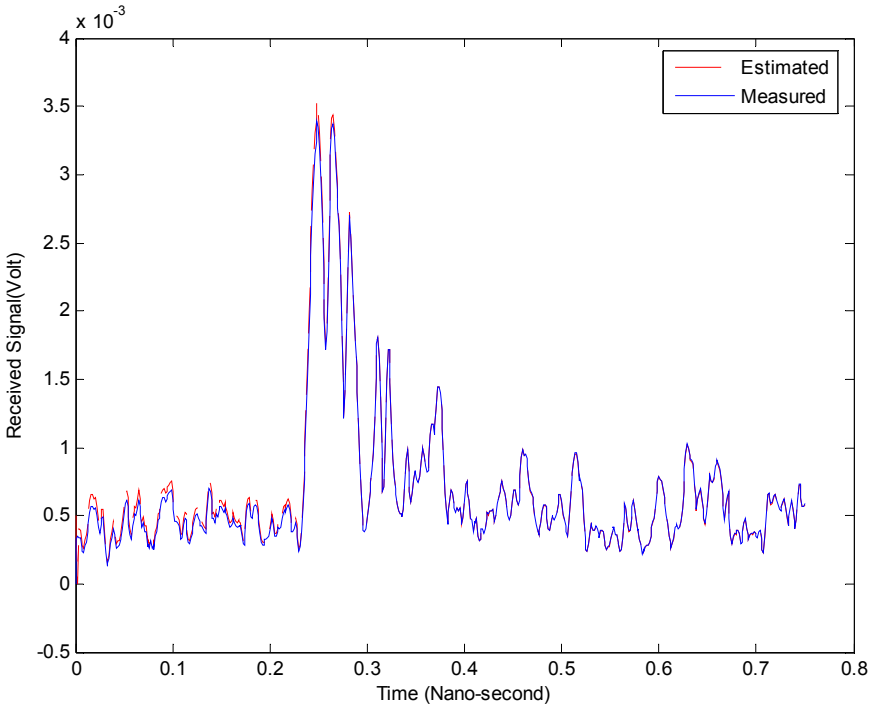

Figure.9 EM algorithm measurement vs. estimation for the 1^{st} path

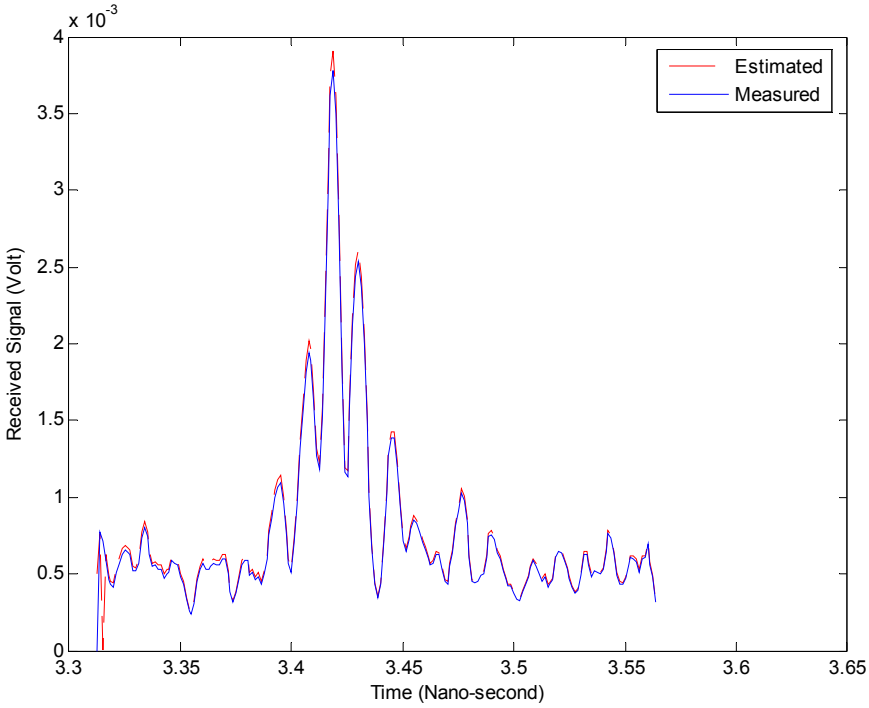

Figure.10 EM algorithm measurement vs. estimation for the 2^{nd} path

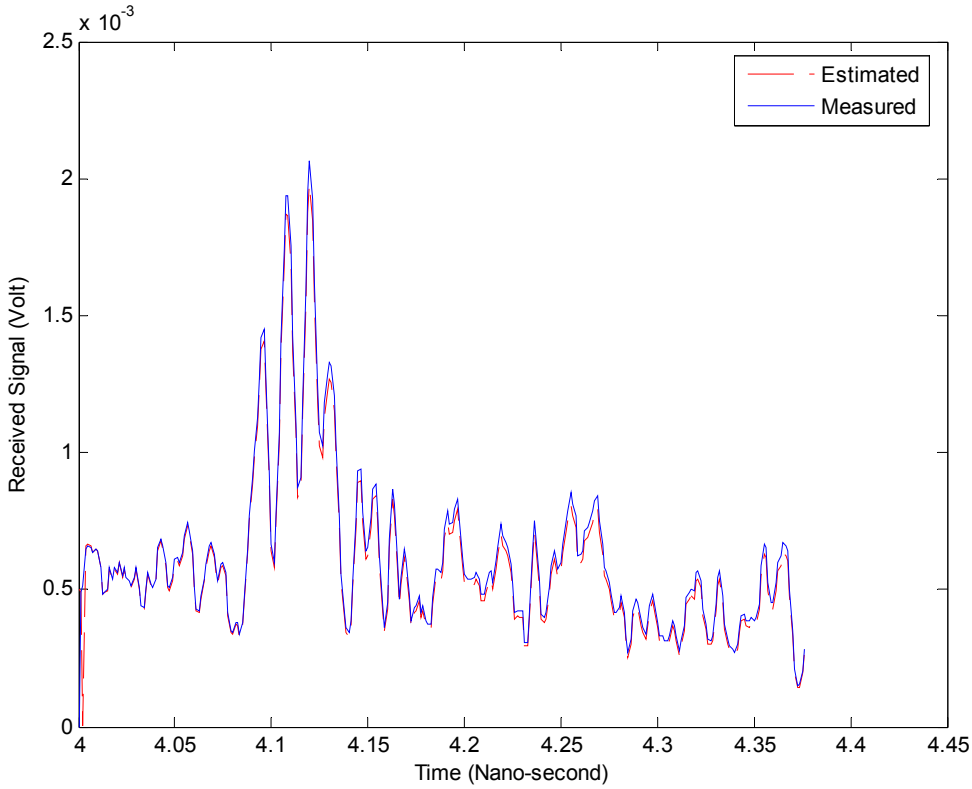

Figure.11 EM algorithm measurement vs. estimation for the 3^{rd} path

B. Non-resolvable multipath experiment

These experiments were performed by placing the transmitter and receiver in a fixed position, while progressively adding metal objects near the transmitter with the purpose of creating a harsh multipath environment [79].

Under this situation, the received signal is too dense to distinguish the paths, since the paths are close to each other and some paths are combined to one.

The Nakagami distribution given by m, Ω can be calculated by its mean and variance. The parameters for this measurement are $m = 1.37,\ \Omega = 0.0016 \times 10^{-3}$ .The estimated result is Figure 12

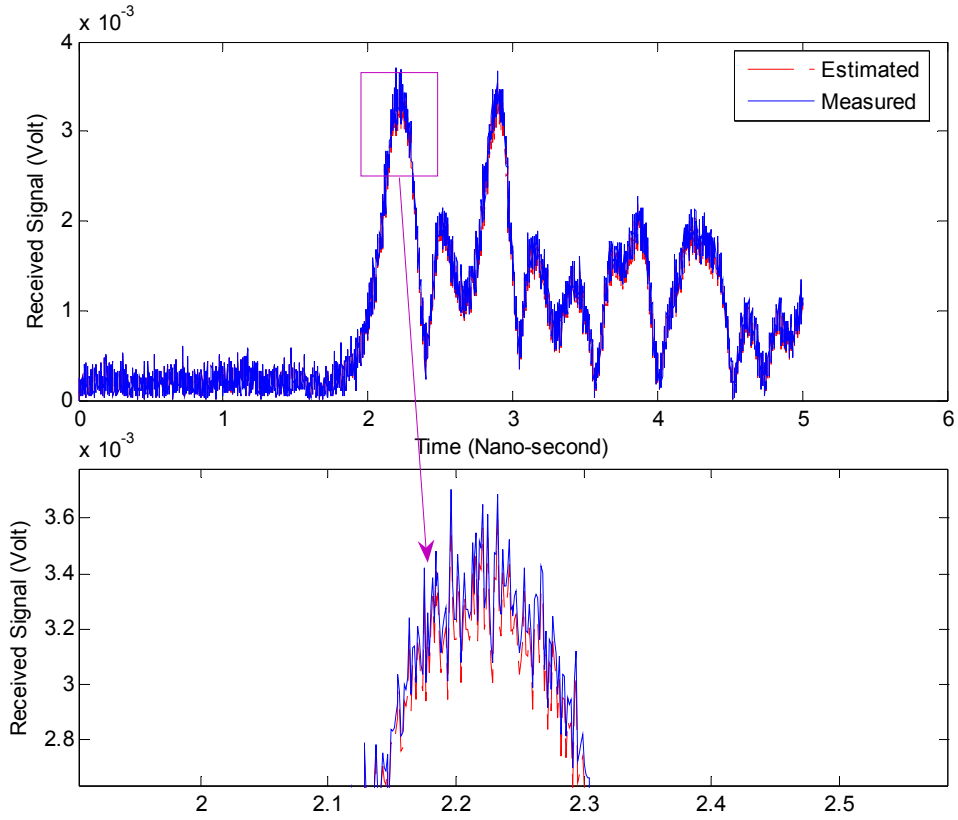

Figure.12 EM algorithm measurement vs. estimation for non-resolvable received signal

Although the non-resolvable received signal consists of probably more than one path, when it is applied by the EM algorithm, it is treated as one path in order to keep the data continuous. Therefore, the estimation error is slightly bigger than the resolvable signal (see Figure.13). From figure 13 we notice that resolvable UWB signals has been estimated with very high accuracy. Further, from the zoom in figure, we find that the estimation algorithm takes little time to converge, from 6×10^{-4} to around 1×10^{-4}.

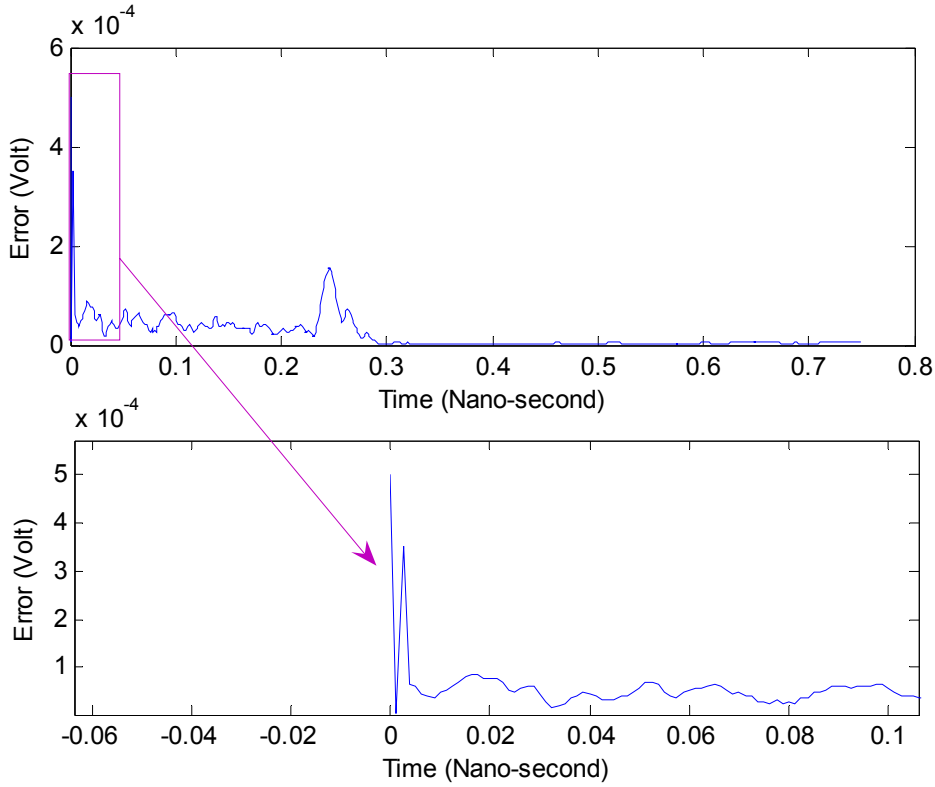

Figure.13 EM algorithm's error for the 1^{st} path

3.5 The Recursive Least-Square Algorithm

In last few sections the Kalman filter and maximum likelihood estimation are implemented to estimate the channel parameters recursively. In this section a recursive least-square algorithm [60] is introduced.

Consider the following state space model [60]:

$$\begin{aligned} x(t+1) &= A(\theta)x(t) + B(\theta)u(t) + \omega(t) \\ y(t) &= C(\theta)x(t) + v(t) \end{aligned} \tag{73}$$

where θ are unknown parameters needed to be estimated, $v(t)$ is measurement noise while the $\omega(t)$ is process noise acting on the states.

The Kalman filter estimation is given by [60]

$$\begin{aligned}
&\hat{x}(t+1,\theta) = A(\theta)\hat{x}(t,\theta) + B(\theta)u(t) \\
&\qquad + K(t,\theta)\left[y(t) - C(\theta)\hat{x}(t,\theta)\right] \\
&\hat{y}(t,\theta) = C(\theta)\hat{x}(t,\theta),\ \hat{x}(0,\theta) = x_0(\theta) \\
&K(t,\theta) = \left[A(\theta)P(t,\theta)C^T(\theta) + R_{12}(\theta)\right] \\
&\qquad \times\left[C(\theta)P(t,\theta)C^T(\theta) + R_2(\theta)\right]^{-1} \\
&P(t+1,\theta) = A(\theta)P(t,\theta)A^T(\theta) + R_1(\theta) \\
&\qquad - K(t,\theta)\times\left[C(\theta)P(t,\theta)C^T(\theta) + R_2(\theta)\right]K^T(t,\theta) \\
&P(0,\theta) = \Pi_0(\theta)
\end{aligned} \tag{74}$$

where $R_1(\theta) = E\omega(t)\omega^T(t), R_2(\theta) = Ev(t)v^T(t), R_{12}(\theta) = E\omega(t)v^T(t)$, $K(t,\theta)$ is Kalman gain, $P(t,\theta)$ is error covariance, $\hat{x}(t,\theta)$ is the estimate of the state, $\hat{y}(t,\theta)$ is the estimate of received signal.

The linear regression model is

$$\hat{y}(t\,|\,\theta) = \varphi^T(t)\theta \tag{75}$$

Therefore consider the following case

$$\begin{aligned}
&\theta(t+1) = \theta(t) \\
&y(t) = \hat{y}(t\,|\,\theta) + v(t) = \varphi^T(t)\theta + v(t)
\end{aligned} \tag{76}$$

So the prediction error becomes

$$\varepsilon(t,\theta) = y(t) - \varphi^T(t)\theta \tag{77}$$

The least-square criterion for the linear regression is [60]

$$V_N\left(\theta, Z^N\right) = \frac{1}{N}\sum_{t=1}^{N}\frac{1}{2}\left[y(t) - \varphi^T(t)\theta\right]^2 \tag{78}$$

Since (78) is the quadratic function of θ, it can be minimized analytically to get the least-square estimate [60]

$$\hat{\theta}^{LS} = \arg\min V_N\left(\theta, Z^N\right) = \left[\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)\right]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)y(t) \tag{79}$$

If the noise $v(t)$ is white and Gaussian, we apply the Kalman filter to (76), then the detailed updating steps are [60]

$$\begin{aligned}
&\hat{\theta}(t)=\hat{\theta}(t-1)+L(t)\left[y(t)-\varphi^T(t)\hat{\theta}(t-1)\right]\\
&L(t)=P(t-1)\varphi(t)\left[\lambda(t)\Lambda_t+\varphi^T(t)P(t-1)\varphi(t)\right]^{-1}\\
&P(t)=\frac{P(t-1)-P(t-1)\varphi(t)\left[\lambda(t)\Lambda_t+\varphi^T(t)P(t-1)\varphi(t)\right]^{-1}\varphi^T(t)P(t-1)}{\lambda(t)}
\end{aligned} \tag{80}$$

where $\lambda(t)\equiv 1$ and $\Lambda_t=R_2(t)$

However, sometimes the "true" parameter is considered to be random [60]

$$\theta(t+1)=\theta(t)+\omega(t) \tag{81}$$

The Kalman filter then gives the estimation as [60]:

$$\begin{aligned}
&\hat{\theta}(t)=\hat{\theta}(t-1)+L(t)\left[y(t)-\varphi^T(t)\hat{\theta}(t-1)\right]\\
&L(t)=\frac{P(t-1)\varphi(t)}{R_2(t)+\varphi^T(t)P(t-1)\varphi(t)}\\
&P(t)=P(t-1)-\frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{R_2(t)+\varphi^T(t)P(t-1)\varphi(t)}+R_1(t)
\end{aligned} \tag{82}$$

The simulation results based on the experimental data from the Antenna lab are shown in Figure 14 to Figure 16:

$$\begin{aligned}
A_1&=\begin{bmatrix}0.9905 & 0.12853\\ -0.0094795 & 0.36861\end{bmatrix}\\
C_1&=\begin{bmatrix}0.029535 & 0.0012354\end{bmatrix}\\
A_2&=\begin{bmatrix}0.97591 & -0.20306\\ 0.15454 & 0.41173\end{bmatrix}\\
C_2&=\begin{bmatrix}0.027824 & 0.0020553\end{bmatrix}\\
A_3&=\begin{bmatrix}0.98259 & -0.14872\\ 0.08759 & 0.10956\end{bmatrix}\\
C_3&=\begin{bmatrix}0.029638 & 0.0012522\end{bmatrix}
\end{aligned} \tag{83}$$

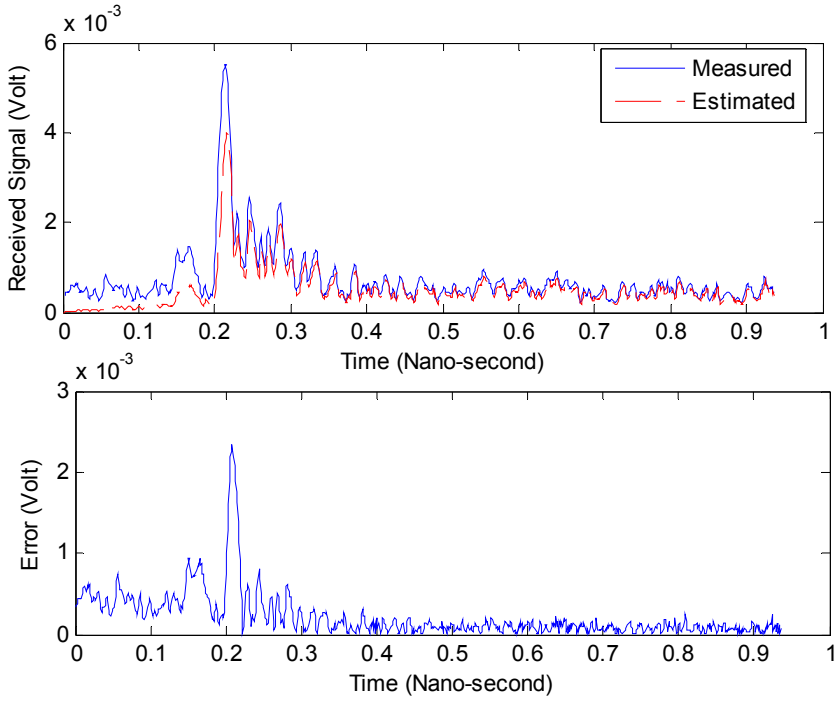

Figure.14 RLS measurement vs. estimation for the 1^{st} path

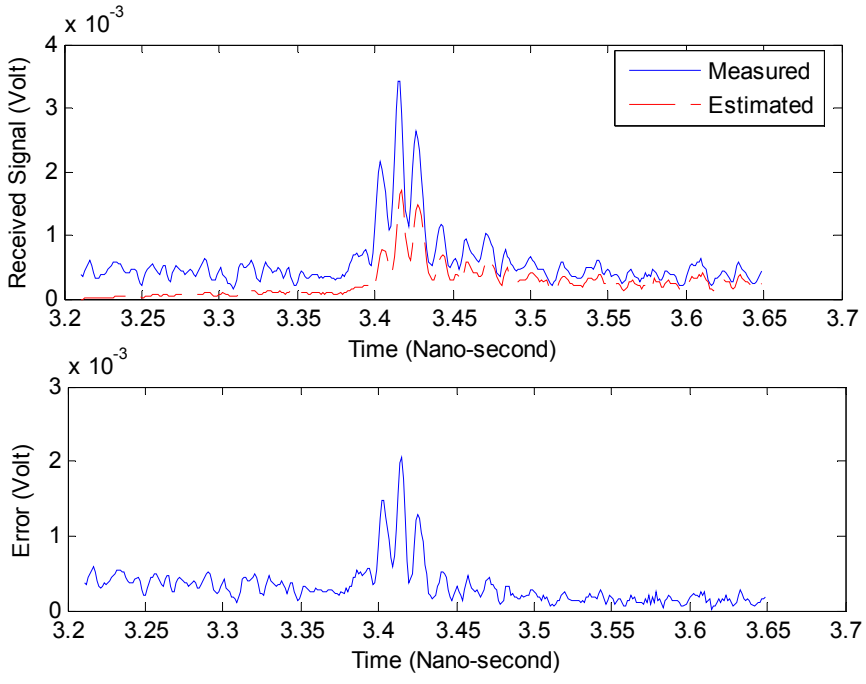

Figure.15 RLS measurement vs. estimation for the 2^{nd} path

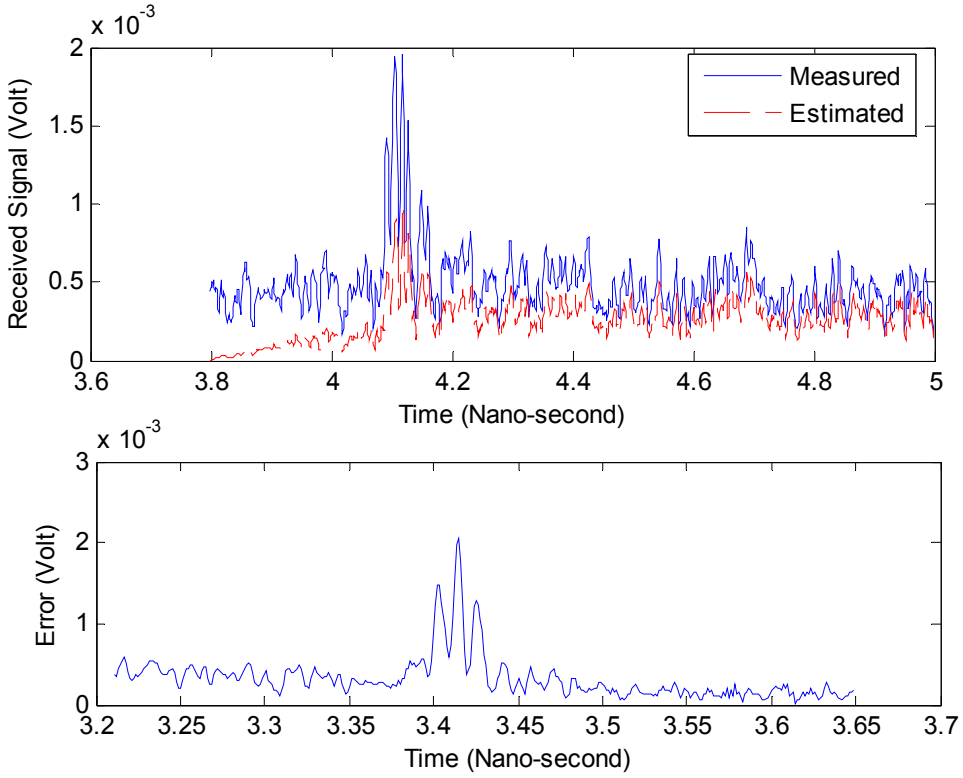

Figure.16 RLS measurement vs. estimation for the 3^{rd} path

From Figure 17, we can clearly see that although the recursive least-square algorithm can roughly capture measurement data, the error is still too big compared to the estimation result obtained from the EM algorithm. The error level of EM algorithm is around 10^{-4} while the error lever of RLS algorithm is around 10^{-3} and the error is extremely big at the peak of the signal. One of the possible reasons is the EM algorithm takes the expectation outside the maximum likelihood function, which reduce the error in some sense.

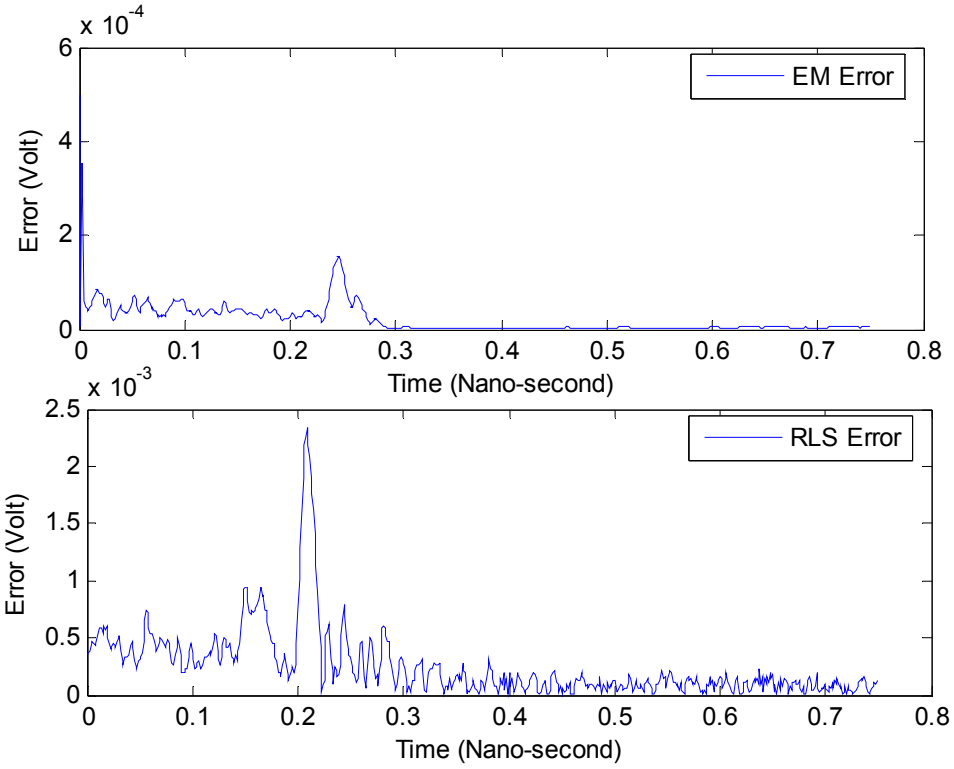

Figure.17 RLS Error vs. EM Error for the 1^{st} path

Chapter 4

Stochastic Power Control Algorithms for Time Varying Wireless Channel Networks

A general framework for continuous time power control algorithm (PCA) under time varying (TV) long term fading (LTF) wireless channels is developed.

PCAs can be classified as centralized and distributed. Centralized PCAs require the information from every node of the network, while the distributed PCAs only require the base station to know its own information, such as, its SNR, which is easily obtained from local measurements. These power allocation problems have been treated as an eigenvalue problem of a nonnegative matrix [41, 58]. The power is updated iteratively based on the SINR and certain thresholds [26].

Stochastic PCAs that use noisy interference estimates have been introduced in [18], where conventional matched filter receivers are used. It is shown in [18] that the iterative stochastic PCA, which uses stochastic approximations, converges to the optimal power vector under certain assumptions on the step-size sequence [44, 65].

In time-invariant models, channel parameters are random but do not depend on time and remain constant throughout the observation and estimation phase. This contrasts with TV models, where the channel dynamics become TV stochastic processes [21]. These models take into account relative motion between transmitters and receivers and temporal variations of the propagating environment such as moving scatters. They exhibit more realistic behavior of wireless networks. In this chapter, we consider dynamical TV LTF channel modeling. The dynamics of LTF wireless channels are captured by SDEs. The SDE model proposed allows viewing the wireless channel as a dynamical system, which shows how the channel evolves in time and space. In addition, it allows well-developed tools of adaptive and non-adaptive estimation and identification

techniques (to estimate the model parameters) to be applied to this class of problems [44, 65].

The correct usage of any PCA and thereby the power optimization of the channel models, require the use of TV channel models that capture both temporal and spatial variations of the wireless channel. Since few temporal or even spatio-temporal dynamical models have so far been investigated with the application of any PCA, the suggested dynamical model and PCAs will thus provide a far more realistic and efficient optimal control for wireless channels [41-44, 65].

The materials of this chapter have been published in [42-44, 65].

4.1 Time Varying Lognormal Fading Channel Model

Wireless radio channels experience both long-term fading (LTF) and short-term fading (STF). LTF is modeled by lognormal distribution and STF are modeled by Rayleigh or Rican distribution [35]. In general, LTF and STF are considered as superimposed and may be treated separately [35]. In this chapter, we consider dynamical modeling and power control for LTF channels which are predominate in suburban areas. The STF case has been considered in [45].

The time-invariant power loss (PL) in dB for a given path is given by [35]

$$PL(d)[dB] = \overline{PL}(d_0)[dB] + 10\alpha \log\left(\frac{d}{d_0}\right) + \tilde{Z} \tag{84}$$

where $d \geq d_0$, $\overline{PL}(d_0)$ is the average PL in dB at a reference distance d_0 from the transmitter, α is the path loss exponent which depends on the propagation medium and $\tilde{Z}$ is a zero-mean Gaussian distributed random variable, which represents the variability of the PL due to numerous reflections occurring along the path and possibly any other uncertainty of the propagation environment from one observation instant to the next. In TV LTF models, the PL becomes a random process denoted by $\{X(t,\tau)\}_{t\geq 0,\tau\geq\tau_0}$, which is a function of both time t and location represented by τ, where $\tau = d/c$, d is the path

length, c is the speed of light, $\tau_0 = d_0 / c$ and d_0 is the reference distance. The process $\{X(t,\tau)\}_{t\geq 0,\tau\geq\tau_0}$ represents how much power the signal looses at a particular distance as a function of time. The signal attenuation is defined by $S(t,\tau)\triangleq e^{kX(t,\tau)}$, where $k = -\ln(10)/20$ [35].

The process $X(t,\tau)$ is generated by a mean-reverting version of a general linear SDE given by [42]

$$\begin{aligned} dX(t,\tau) &= \beta(t,\tau)\big((\gamma(t,\tau) - X(t,\tau)\big)dt + \delta(t,\tau)dW(t), \\ X(t_0,\tau) &= N\left(\overline{PL}(d)[dB];\sigma_{t_0}^2\right) \end{aligned} \tag{85}$$

where $\{W(t)\}_{t\geq 0}$ is a standard Brownian motion (zero drift, unit variance) which is assumed to be independent of $X(t_0,\tau)$, $N(\mu;\kappa)$ denotes a Gaussian random variable with mean μ and variance κ and $\overline{PL}(d)[dB]$ is the average path loss in dB. The parameter $\gamma(t,\tau)$ models the average TV PL at distance d from transmitter, which corresponds to $\overline{PL}(d)[dB]$ at d indexed by t. This model tracks and converges to this value as time progresses. The instantaneous drift $\beta(t,\tau)\big(\gamma(t,\tau) - X(t,\tau)\big)$ represents the effect of pulling the process towards $\gamma(t,\tau)$, while $\beta(t,\tau)$ represents the speed of adjustment towards this value. Finally, $\delta(t,\tau)$ controls the instantaneous variance or volatility of the process for the instantaneous drift. Define $\{\theta(t,\tau)\}_{t\geq 0}\triangleq\{\beta(t,\tau),\gamma(t,\tau),\delta(t,\tau)\}_{t\geq 0}$. If the random processes in $\{\theta(t,\tau)\}_{t\geq 0}$ are measurable and bounded, then (77) has a unique solution for every $X(t_0,\tau)$ given by [42]

$$X(t,\tau) = e^{-\beta([t,t_0],\tau)}\cdot\left\{X(t_0,\tau) + \int_{t_0}^{t} e^{\beta([u,t_0],\tau)}\left(\beta(u,\tau)\gamma(u,\tau)du + \delta(u,\tau)dW(u)\right)\right\} \tag{86}$$

where $\beta([t,t_0],\tau) \triangleq \int_{t_0}^{t} \beta(u,\tau)du$. This model captures the temporal and spatial variations of the propagation environment as the random parameters $\{\theta(t,\tau)\}_{t\geq 0}$ which can be used to model the time and space varying characteristics of the channel.

At every instant of time $X(t,\tau)$ is Gaussian with mean and variance given by [65]

$$E[X(t,\tau)] = e^{-\beta([t,t_0],\tau)} \cdot \left(X_0 + \int_{t_0}^{t} e^{\beta([u,t_0],\tau)} \beta(u,\tau)\gamma(u,\tau)\,du \right)$$

$$Var[X(t,\tau)] = e^{-2\beta([t,t_0],\tau)} \cdot \left(\int_{t_0}^{t} e^{2\beta([u,t_0],\tau)} \delta^2(u,\tau)\,du + \sigma_{t_0}^2 \right) \tag{87}$$

Moreover, the distribution of $S(t,\tau) = e^{kX(t,\tau)}$ is lognormal with mean and variance given by

$$\begin{aligned} E[S(t,\tau)] &= \exp\left(\frac{2kE[X(t,\tau)] + k^2 Var[X(t,\tau)]}{2} \right) \\ Var[S(t,\tau)] &= \exp\left(2kE[X(t,\tau)] + 2k^2 Var[X(t,\tau)]\right) \\ &\quad - \exp\left(2kE[X(t,\tau)] + k^2 Var[X(t,\tau)]\right) \end{aligned} \tag{88}$$

The mean and variance in (87) and (88) show that the statistics of the communication channels vary as a function of both time t and space τ.

In this section, we consider the uplink channel of a cellular network and we assume that users are already assigned to their base stations. Let M be the number of mobiles (users) and N be the number of base stations. The received signal of the i^{th} mobile at its assigned base station at time t is given by [65]

$$y_i(t) = \sum_{j=1}^{M} \sqrt{p_j(t)}\, s_j(t)\, S_{ij}(t) + n_i(t) \tag{89}$$

where $p_j(t)$ is the transmitted power of mobile j at time t, which acts as a scaling on the information signal $s_j(t)$, $n_i(t)$ is the channel disturbance or noise at the base station of mobile i and $S_{ij}(t)$ is the signal attenuation coefficient between mobile j and the base station assigned to mobile i. Therefore, in a cellular network the spatio-temporal model described in (85) for M mobiles and N base stations can be described as

$$\begin{aligned} dX_{ij}(t,\tau) &= \beta_{ij}(t,\tau)\left(\gamma_{ij}(t,\tau) - X_{ij}(t,\tau)\right)dt + \delta_{ij}(t,\tau)\,dW_{ij}(t), \\ X_{ij}(t_0,\tau) &= N\left(\overline{PL}(d)[dB]_{ij};\sigma^2_{t_0}\right),\ 1 \le i,\ j \le M \end{aligned} \tag{90}$$

and the signal attenuation coefficients $S_{ij}(t,\tau)$ are generated using the relation $S_{ij}(t,\tau) = e^{kX_{ij}(t,\tau)}$, where $k = -\ln(10)/20$. Moreover, correlation between the channels in a multi-user/multi-antenna model can be induced by letting the different Brownian motions W_{ij} to be correlated, i.e., $E\left[\mathbf{W}(t)\mathbf{W}(t)^T\right] = \mathbf{Q}(\tau)\cdot t$, where $\mathbf{W}(t) \triangleq \left(W_{ij}(t)\right)$ and $\mathbf{Q}(\tau)$ is some (not necessarily diagonal) matrix that is a function of τ and dies out as τ becomes large [65].

The TV LTF channel models in (90) are used to generate the link gains of wireless networks for the PCA proposed in the next section.

4.2 Stochastic PCA in TV Wireless Networks

The aim of the PCAs described here is to minimize the total transmitted power of all users while maintaining acceptable QoS for each user. The measure of QoS can be defined by the SIR for each link to be larger than a target SIR.

Consider the cellular network described in the previous section, the centralized PC problem for time-invariant channels can be stated as follows [45]

$$\min_{(p_1\geq 0,\ldots p_M\geq 0)} \sum_{i=1}^{M} p_i, \quad \text{subject to}$$
$$\frac{p_i g_{ii}}{\sum_{j\neq i}^{M} p_j g_{ij} + \eta_i} \geq \varepsilon_i, \quad 1 \leq i \leq M \tag{91}$$

where p_i is the power of mobile i, $g_{ij} > 0$ is the time-invariant channel gain between mobile j and the base station assigned to mobile i, $\varepsilon_i > 0$ is the target SIR of mobile i and $\eta_i > 0$ is the noise power level at the base station of mobile i. The generalization to (91) for the TV LTF channel models in (85), described using path-wise QoS of each user over a time interval $[0, T]$, is given by [26]

$$\min_{(p_1(t)\geq 0,\ldots p_M(t)\geq 0)} \left\{ \sum_{i=1}^{M} p_i(t) \right\}, \quad \text{subject to}$$
$$\frac{E\left\{ p_i(t) s_i^2(t) S_{ii}^2(t) \right\}}{E\left\{ \sum_{k\neq i}^{M} p_k(t) s_k^2(t) S_{ik}^2(t) + n_i^2(t) \right\}} \geq \varepsilon_i(t) \tag{92}$$

where $E\{\cdot\}$ is the expectation operator, $t \in [0, T]$ and $i = 1, \ldots, M$. Define

$$\mathbf{p}(t) \triangleq \begin{pmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_M(t) \end{pmatrix}, \ \mathbf{u}(t) \triangleq \begin{pmatrix} \dfrac{\varepsilon_1(t)\eta_1(t)}{E\left[S_{11}^2(t) \right]} \\ \dfrac{\varepsilon_2(t)\eta_2(t)}{E\left[S_{22}^2(t) \right]} \\ \vdots \\ \dfrac{\varepsilon_M(t)\eta_M(t)}{E\left[S_{MM}^2(t) \right]} \end{pmatrix},$$
$$\mathbf{F}(t) = \begin{cases} 0, & i = j \\ \dfrac{E\left[S_{ij}^2(t) \right] \varepsilon_i(t)}{E\left[S_{ii}^2(t) \right]}, & i \neq j \end{cases} \tag{93}$$

where $i, j = 1, \ldots, M$, note that $E\left[S_{ij}^2(t) \right]$ can be calculated from (87) and (88).

4.2.1 Fixed Point Problem

If the power control problem in (92) is feasible, then the optimal power satisfies [65]

$$\mathbf{p}^*(t) = \mathbf{F}(t)\mathbf{p}^*(t) + \mathbf{u}(t) \tag{94}$$

where $\mathbf{p}^*(t)$ is the optimal power vector. Expression (94) shows that the optimal power is the fixed point of the following function

$$\mathbf{\Phi}(\mathbf{p})(t) \triangleq \mathbf{F}(t)\mathbf{p}(t) + \mathbf{u}(t) \tag{95}$$

The power vector $\mathbf{p}(t)$ is assumed to be a continuous and bounded function from $[0,T]$ to $\mathbf{R}^{+M}$, that is, it belongs to the Banach space [77] (see Appendix B) of continuous and bounded functions defined on $[0,T]$, which is denoted by $\mathbf{C}([0,T];\mathbf{R}^M)$, under the supremum norm given by

$$\left\|\mathbf{p}(t)\right\|_\infty = \sup_{0\le t\le T}\left(\sum_{i=1}^{M}\left|p_i(t)\right|^2\right)^{1/2} \tag{96}$$

Assuming $\mathbf{u}(t) \in \mathbf{C}([0,T];\mathbf{R}^M)$, then the map $\mathbf{\Phi}(\cdot)$ is defined as

$$\mathbf{\Phi}(\cdot):\ \mathbf{C}\left([0,T];\mathbf{R}^M\right) \ \rightarrow\ \mathbf{C}\left([0,T];\mathbf{R}^M\right) \tag{97}$$

The existence of a fixed point for $\mathbf{\Phi}(\cdot)$ is guaranteed by the contraction mapping theorem or Banach's fixed-point theorem [50], which states that if

$$\left\|\mathbf{\Phi}(\mathbf{p}_1)(t) - \mathbf{\Phi}(\mathbf{p}_2)(t)\right\|_\infty \le k\left\|\mathbf{p}_1(t) - \mathbf{p}_2(t)\right\|_\infty \tag{98}$$

For some $k<1$ and all $\mathbf{p}_1(t), \mathbf{p}_2(t) \in \mathbf{C}\left([0,T];\mathbf{R}^M\right)$, then $\mathbf{\Phi}(\cdot)$ has a uniform fixed point. Expression (98) can be rewritten as

$$\left\|\mathbf{F}(t)\mathbf{p}_1(t) - \mathbf{F}(t)\mathbf{p}_2(t)\right\|_\infty \le k\left\|\mathbf{p}_1(t) - \mathbf{p}_2(t)\right\|_\infty,\ \forall\, \mathbf{p}_1(t), \mathbf{p}_2(t) \in \mathbf{C}\left([0,T];\mathbf{R}^M\right) \tag{99}$$

which is equivalent to

$$\frac{\left\|\mathbf{F}(t)\left(\mathbf{p}_1(t) - \mathbf{p}_2(t)\right)\right\|_\infty}{\left\|\mathbf{p}_1(t) - \mathbf{p}_2(t)\right\|_\infty} \le k < 1, \forall\, \mathbf{p}_1(t), \mathbf{p}_2(t) \in \mathbf{C}\left([0,T];\mathbf{R}^M\right) \tag{100}$$

Expression (100) holds if and only if the following holds

$$\sup_{\substack{\mathbf{p}_1(t)\neq\mathbf{p}_2(t)\\ \mathbf{p}_1(t),\mathbf{p}_2(t)\in\mathbf{C}\left([0,T];\mathbf{R}^M\right)}} \frac{\left\|\mathbf{F}(t)\left(\mathbf{p}_1(t)-\mathbf{p}_2(t)\right)\right\|_\infty}{\left\|\mathbf{p}_1(t)-\mathbf{p}_2(t)\right\|_\infty} \leq k < 1 \tag{101}$$

The LHS in (101) is equal to the induced norm of $\mathbf{F}(t)$ viewed as a multiplication acting from $\mathbf{C}([0,T];\mathbf{R}^M)$ into $\mathbf{C}([0,T];\mathbf{R}^M)$, i.e.,

$$\left\|\mathbf{F}(t)\right\| \leq k < 1 \tag{102}$$

where

$$\left\|\mathbf{F}(t)\right\| = \sup_{\substack{\mathbf{p}(t)\in\mathbf{C}\left([0,T];\mathbf{R}^M\right)\\ \|\mathbf{p}(t)\|_\infty \leq 1}} \left\|\mathbf{F}(t)\mathbf{p}(t)\right\|_\infty \tag{103}$$

$\left\|\mathbf{F}(t)\right\|$ is equal to supremum with respect to t of the largest singular value of $\mathbf{F}(t)$, that is [50]

$$\left\|\mathbf{F}(t)\right\| = \sup_{t\in[0,T]} \bar{\sigma}\left(\mathbf{F}(t)\right) < 1 \tag{104}$$

where $\bar{\sigma}(\cdot)$ denotes the largest singular value of $\mathbf{F}(t)$. Expression (104) gives a sufficient condition on the channels' attenuation coefficients for the existence of an optimal power vector. Expression (104) is satisfied if and only if

$$\bar{\sigma}\left(\mathbf{F}(t)\right) < 1, \quad \forall\, t \in [0,T] \tag{105}$$

Thus, if (105) is satisfied, the following continuous time PCA will converge to the minimal power

$$\mathbf{p}_{k+1}(t) = \mathbf{F}(t)\mathbf{p}_k(t) + \mathbf{u}(t) \tag{106}$$

Note that in (106) index k is the iteration on the continuous time power vector and therefore does not represent the time variable as in most PCAs in the literature.

4.2.2 Stochastic Approximation Problem

Consider a set of discrete time strategies $\left\{p_i(t_k)\right\}_{i=1}^M$, $0 = t_0 < t_1 < \ldots < t_k < t_{k+1} < \ldots \leq T$. Equation (92) is equivalent to [43, 44]

$$\min_{\mathbf{p}(t_{k+1})>0} \sum_{i=1}^{M} p_i(t_{k+1}) \quad \text{subject to}$$
$$\mathbf{p}(t_{k+1}) \geq \boldsymbol{\Gamma}\mathbf{G}_I^{-1}(t_k, t_{k+1}) \times \left(\mathbf{G}(t_k, t_{k+1})\mathbf{p}(t_{k+1}) + \boldsymbol{\eta}(t_{k+1})\right) \tag{107}$$

where

$$g_{ij}(t_k, t_{k+1}) \coloneqq \int_{t_k}^{t_{k+1}} s_j^2(t) S_{ij}^2(t)\, dt,\ 1 \leq i,\ j \leq M \quad \text{for LTF},$$
$$\coloneqq \int_{t_k}^{t_{k+1}} s_j^2(t) \left[H(t) X_{ij}(t)\right]^2 dt,\ 1 \leq i,\ j \leq M \quad \text{for STF},$$
$$\mathbf{p}(t_{k+1}) \coloneqq \left(p_1(t_{k+1}), \cdots, p_M(t_{k+1})\right)^T,\ \eta_i(t_k, t_{k+1}) \coloneqq \int_{t_k}^{t_{k+1}} n_i^2(t)\, dt, \tag{108}$$
$$\mathbf{G}_I(t_k, t_{k+1}) \coloneqq diag\left(g_{11}(t_k, t_{k+1}), \cdots, g_{MM}(t_k, t_{k+1})\right),$$
$$\mathbf{G}(t_k, t_{k+1}) \coloneqq \begin{Bmatrix} 0 & , \text{if } i = j \\ g_{ij}(t_k, t_{k+1}) & , \text{if } i \neq j \end{Bmatrix},\ 1 \leq i,\ j \leq M,$$
$$\boldsymbol{\eta}(t_k, t_{k+1}) \coloneqq \left(\eta_1(t_k, t_{k+1}), \cdots, \eta_M(t_k, t_{k+1})\right)^T,\ \boldsymbol{\Gamma} \coloneqq diag(\varepsilon_1, \cdots, \varepsilon_M),$$

and $diag(\cdot)$ denotes a diagonal matrix with its argument as diagonal entries. Here $[t_k,\ t_{k+1}]$ is a time varying interval such that the channel model does not change significantly.

The distributed version of (107) can be written as [43, 44]

$$\left(\mathbf{I} - \boldsymbol{\Gamma}\mathbf{G}_I^{-1}(t_k, t_{k+1})\mathbf{G}(t_k, t_{k+1})\right)\mathbf{p}(t_{k+1}) \geq \boldsymbol{\Gamma}\mathbf{G}_I^{-1}(t_k, t_{k+1})\boldsymbol{\eta}(t_{k+1}) \tag{109}$$

Define $\mathbf{F}(t_k, t_{k+1}) \triangleq \boldsymbol{\Gamma}\mathbf{G}_I^{-1}(t_k, t_{k+1})\mathbf{G}(t_k, t_{k+1})$ and $\mathbf{u}(t_k, t_{k+1}) \triangleq \boldsymbol{\Gamma}\ \mathbf{G}_I^{-1}(t_k, t_{k+1})\ \boldsymbol{\eta}(t_{k+1})$, (109) can be written as [43, 44]

$$\left(\mathbf{I} - \mathbf{F}(t_k, t_{k+1})\right)\mathbf{p}(t_{k+1}) \geq \mathbf{u}(t_k, t_{k+1}) \tag{110}$$

We briefly introduce one basic result on stochastic approximation that will be needed later in this section. Consider an unknown measurable function $\mathbf{h}(\mathbf{x})$. A zero point of $\mathbf{h}$, $\overline{\mathbf{x}}$ is defined by $\mathbf{h}(\overline{\mathbf{x}}) = 0$ and can be calculated by various rapidly convergent methods such as Newton's method.

Assume now that the observation function is provided by $\mathbf{h}(\cdot)$ but subject to an additive measurement noise. The k^{th} measurement is given by [43, 44]

$$\mathbf{y}(k) = \mathbf{h}\left(\mathbf{x}(k)\right) + \boldsymbol{\xi}(k) \tag{111}$$

where $\mathbf{y}(k)$ is the observation at the k^{th} time and $\boldsymbol{\xi}(k)$ is the zero mean measurement error at the k^{th} time and may be dependent on $\mathbf{x}(k)$. In 1951, Robbins and Monro [48] suggested a method for solution of this and a more general problem, which they called the method of stochastic approximation. The purpose of using stochastic approximation is to find a zero point $\bar{\mathbf{x}}$ based on the noisy observation $\mathbf{y}(k)$. Given an arbitrary initial point $\mathbf{x}(0)$ and an arbitrary sequence of positive numbers $a(k)$ such that [43, 44]:

$$\sum_{k=0}^{\infty} a(k) = \infty, \ \sum_{k=0}^{\infty} a(k)^2 < \infty \tag{112}$$

Then, it is shown in [67] that the following approximation sequence,

$$\mathbf{x}(k+1) = \mathbf{x}(k) - a(k)\mathbf{y}(k) \tag{113}$$

converges to the zero point $\bar{\mathbf{x}}$ of $\mathbf{h}(\cdot)$ with probability one.

Now, since the matrix $\mathbf{F}(t_k, t_{k+1})$ has non-negative elements and if the SIR targets are feasible, then it has been shown that the power vector minimizes the sum of the transmitted power, i.e., the power vector $\mathbf{p}(t_{k+1})$ that satisfies

$$\left(\mathbf{I} - \mathbf{F}(t_k, t_{k+1})\right)\mathbf{p}(t_{k+1}) - \mathbf{u}(t_k, t_{k+1}) = 0 \tag{114}$$

is the minimum power vector. Since the link gains are random, given $\mathbf{F}(t_k, t_{k+1})$ and $\mathbf{u}(t_k, t_{k+1})$ in practice obtained through measurement and the estimation error in these is represented by $\boldsymbol{\xi}(k)$, an additive zero mean random noise. Therefore, applying the stochastic approximation algorithm in (113) to (114) we get

$$\mathbf{p}(t_{k+1}) = \mathbf{p}(t_k) - a(t_k)\left[\left(\mathbf{I} - \mathbf{F}(t_k, t_{k+1})\right)\mathbf{p}(t_k) - \mathbf{u}(t_k, t_{k+1})\right] \quad (115)$$

which can be written as

$$\mathbf{p}(t_{k+1}) = \left(\mathbf{I} - a(t_k)\right)\mathbf{p}(t_k) + a(t_k)\left[\mathbf{F}(t_k, t_{k+1})\mathbf{p}(t_k) + \mathbf{u}(t_k, t_{k+1})\right] \quad (116)$$

The distributed version of (116) is:

$$p_i(t_{k+1}) = \left(1 - a(t_k)\right) p_i(t_k) + a(t_k)\frac{\varepsilon_i(t_k)}{R_i(t_k)} p_i(t_k) \quad (117)$$

If the PC problem is feasible, the distributed SPCA in (117) converges to the optimal power vector when the step-size sequence satisfies certain conditions. Two different types of convergence results are shown in [49] under different choices of the step-size sequence. If the step-size sequence satisfies $\sum_{k=0}^{\infty} a(t_k) = \infty$ and $\sum_{k=0}^{\infty} a(t_k)^2 < \infty$, then the SPCA in (117) converges to the optimal power vector with probability one. However, due to the requirement for the SPCA to track TV environments, the iteration step-size sequence is not allowed to decrease to zero. So we consider the case where the condition $\sum_{k=0}^{\infty} a(t_k)^2 < \infty$ is violated. This includes the situation where the step-size sequence decreases slowly to zero and the situation when the step-size is fixed at a small constant. In the first case when $a(t_k) \to 0$ slowly, the SPCA in (117) converges to the optimal power vector in probability. While in the second case the power vector clusters around the optimal power [49]. In fact, the error between the power vector and the optimal value does not vanish for non-vanishing step-size sequence; this is the price paid in order to make the algorithm in (117) able to track TV environments [43, 44].

This algorithm is fully distributed in the sense that each user iteratively updates its power level by estimating the received SIR of its own channel. It does not require any knowledge of the link gains and state information of other users. It is worth mentioning that the proposed distributed SPCA in (117) is different from the algorithm proposed in [26] where two parameters, namely, the received SIRs $R_i(t_k)$ and the channel gains

$g_{ii}(t_k, t_{k+1})$, are required to be known, while only $R_i(t_k)$ are required in (117).

The selection of an appropriate $[t_k, t_{k+1}]$ will have a significant impact on the system performance. For small $[t_k, t_{k+1}]$, the power control updates will be more frequent and thus convergence will be faster. However, frequent transmission of the feedback on the downlink channel will effectively decrease the capacity of the system since more system resources will have to be used for power control [43, 44].

4.3 Numerical Results

Consider a network of three users, the channels between different users are ergodic, i.e., the expected value of the gain matrix is constant. We assume that $\varepsilon_i = 5$ and $\eta_i = 1$ for each transmitter. The time and iteration is seperated in the simulation, i.e., the optimized power value obtained from previous time is used as the initial value of power for the current time. Simulation results are presented in Figure 18 to Figure 19.

From figure (18), it is clear that the transmit power converges under the fixed point algorithm. But the power looks more “noisy” because of the existence of noise in the data. The stochastic approximation algorithm works more stable than the fixed point algorithm since the former algorithm works better at the noisy environment. When the step size is smaller, the SA algorithm converges faster. Figure 19 expresses the relation between the power and iteration number for the fixed point algorithm. It can be seen that the power converges to the optimal power less than five iterations.

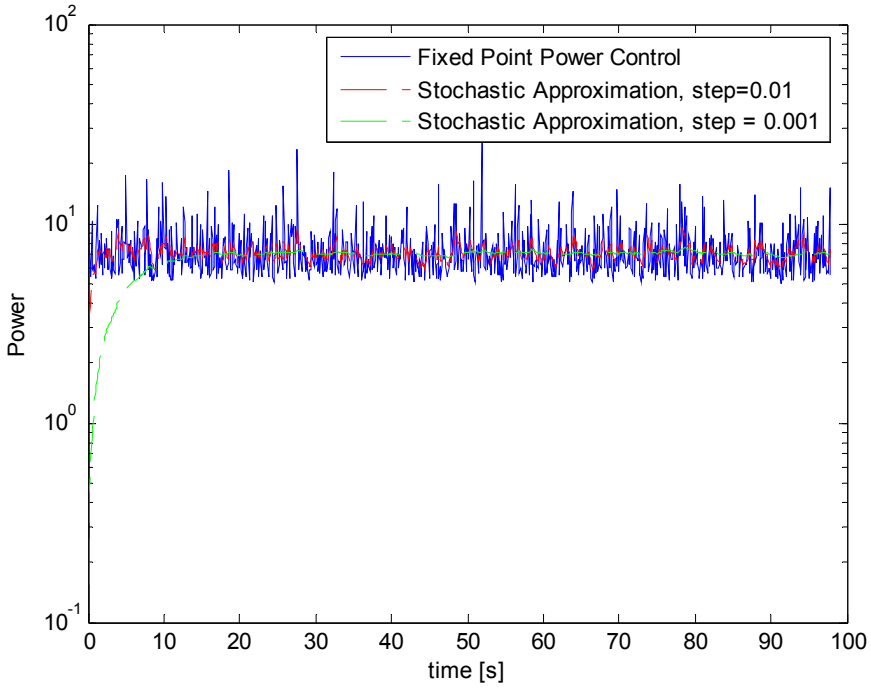

Figure.18 Transmit power vs. time by the Fixed Point algorithm

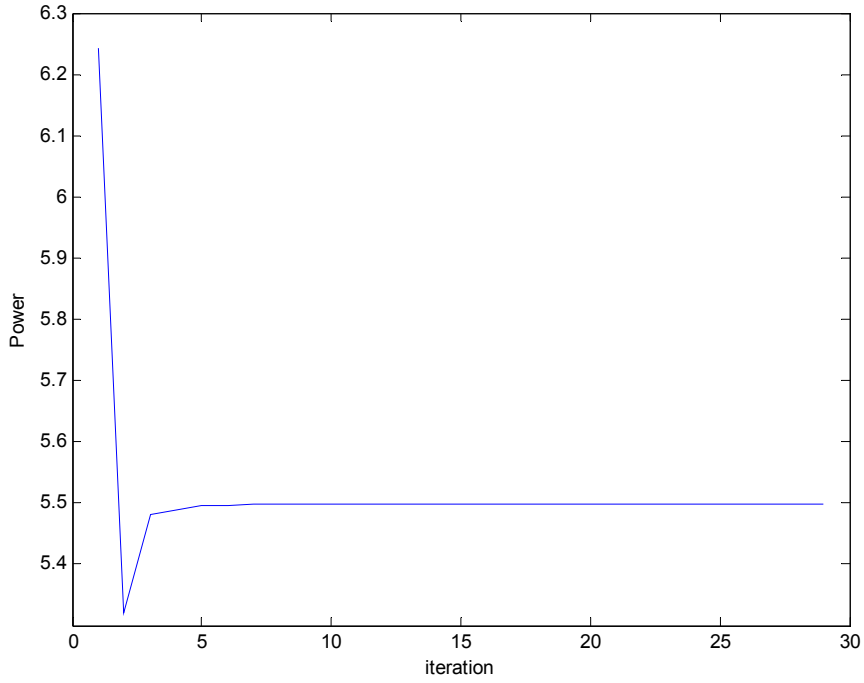

Figure 19 Transmit power vs. iteration

Chapter 5

Statistical Analysis of Multipath Channels

In chapter two UWB indoor channel modeling is discussed. However, in various design stages one is also interested in the statistics of the overall received signal. For example, one major problem in [8] is how to identify clusters since the position and size of the clusters will highly depend on the environment and physical structures. They select the cluster region by visual inspection because of the difficulty in developing a robust algorithm to identify the cluster region. Also, in UWB localization problem, it is interesting to find out which path forms the first peak. [6] claims that the distribution of the cluster arrival times satisfies Poisson processes, while in [8], they modified the model from one Poisson process to two Poisson processes in order to fit for the data of indoor residential environment.

Thus, establishing a framework of the statistics of the overall received signal is necessary. After knowing the distribution information of the received signal such as mean, variance, correlation, joint distribution etc., we can understand the channel much better and improve receiver design.

However, in most of the work found in the literature [5, 8] authors compute the statistics of the channel making some important simplifications [66]. First, they often omit the explicit dependence of r_i on τ and thus $\tau_i(t)$, during the computation of various statistical properties of the received signal. Although for deterministic, or a fixed sample path of $\{N(s); 0 \le s \le t\}$, the computation of the statistical properties of $H_l(t;\tau)$ is not affected by this omission, this is not the case when the ensemble statistics are analyzed. Second, they compute the various statistics based solely on a fixed number of paths, $N(t) = N$ due to the difficulty in assuming that $\{N(s); 0 \le s \le T_s\}$ is a random process.

Ensemble statistics are then computed by using sample averages, i.e.，averaging over a number of different realizations [66].

The above omissions are due either to the fact that the models used up to date provided a satisfactory performance for the first and second generation wireless communication systems or perhaps to the fact that it had not yet been recognized that the shot-noise analysis of random noise, brought forward by Rice [36], constitutes a powerful tool in putting forward a general framework for investigating the overall statistical properties of multipath dispersive fading channels [66].

A shot noise effect [36, 66], refers to the output of a linear random dynamic system whose input is a train of impulses arriving at random times distributed according to a counting process, which may be as simple as a non-homogeneous Poisson process. When the responses associated with each occurrence are described by i.i.d. random variables or random processes, then the closed form expressions for the statistics of the resulting process can be computed [36].

5.1 Poisson Counting Process

The Poisson process [55] is the simplest process associated with counting random number of points. The Poisson counting process is important in at least three aspects [55]. First, Poisson process has been proved to be an accurate modeling method in many applications. Second, it is a basis for the more complicated counting process.

Definition 5.2.1 [55]: Let $\{N(t); t \geq t_0\}$ be a Poisson counting process, for $t_0 \leq s < t$, the increment $N_{s,t} = N(t) - N(s)$ is Poisson distributed with parameter $\Lambda_t - \Lambda_s$, so

$$\Pr\left[N_{s,t} = n\right] = (n!)^{-1} (\Lambda_t - \Lambda_s)^n \exp\left[-(\Lambda_t - \Lambda_s)\right], \text{ for } n = 0,1,2,\ldots \tag{118}$$

where Λ_t is a nonnegative, non-decreasing function of t.

If Λ_t is an absolutely continuous function of t, it can be expressed as $\Lambda_t \triangleq \int_{t_0}^{t} \lambda_\sigma d\sigma$, for all $t \geq t_0$, where λ_t is a nonnegative function of t for $t \geq t_0$. We call λ_t the intensity function of the process. At any time $t \geq t_0$, λ_t is the instantaneous average rate of the points occurring.

Under the definition, (118) can be rewritten as [55]

$$\Pr\left[N_{s,t} = n\right] = (n!)^{-1}\left(\int_s^t \lambda_\sigma d\sigma\right)^n \exp\left[-\left(\int_s^t \lambda_\sigma d\sigma\right)\right], \text{ for } n = 0,1,2,\ldots \qquad (119)$$

when the intensity λ_t is a constant independent of time, the corresponding Poisson counting process is called Homogeneous Poisson process (HPP) [55]. In this case, $\int_s^t \lambda_\sigma d\sigma$ is proportional to $t-s$, which implies that the counting statistics on the interval $[s,t)$ are the same as those on $[s+\tau, t+\tau)$ for all τ such that $t_0 \leq s+\tau$. Whenever λ_t is not a constant, the corresponding Poisson process is called inhomogeneous Poisson process [55].

It follows from (119) that, for $t_0 \leq t < s$,

$$\Pr\left[N(s) - N(t) > 0\right] = 1 - \exp\left[-\int_t^s \lambda_\sigma d\sigma\right] \qquad (120)$$

It is clear that the Poisson process rate acts as a filter which models the filtering properties of each propagation environment [55]. The number of paths is related to the number of Poisson points in that time interval, i.e., if the Poisson process rate is high; the number of paths is high.

After discussing the interval characteristics which relates to the number of points occurring in arbitrary intervals of time, another important term is the location characteristics which relate to the point locations and interpoint spacings. In particular, the statistics of the location is an important tool in calculating the corresponding statistics of the whole received signals. Several new concepts will be introduced here in the following.

Consider Figure 20 [55], the sequence $\{w_n\}$ is the occurrence time sequence; the sequence $\{t_n\}$ is the interarrival time sequence. t_n is the n^{th} interarrival time, which is random.

Consider $p_w^{(n)}(W)$ as the joint probability density for the joint probability density for the first n occurrence times $w=(w_1,w_2,\dots w_n)$. For an inhomogeneous Poisson counting process with intensity λ_t on $[t_0,t)$, we have [55]

$$p_w^{(n)}(W)=\begin{cases}\left(\prod_{i=1}^{n}\lambda_{W_i}\right)\exp\left(-\int_{t_0}^{W_n}\lambda_\sigma d\sigma\right), & t_0\le W_1\le W_2\le\cdots\le W_n\\ 0, & otherwise\end{cases} \tag{121}$$

The joint occurrence density for a *homogeneous Poisson process* with constant intensity λ is given by [55]

$$p_w^{(n)}(W)=\begin{cases}\left(\lambda^n\right)\exp\left(-\lambda\left(W_n-t_0\right)\right), & t_0\le W_1\le W_2\le\cdots\le W_n\\ 0, & otherwise\end{cases} \tag{122}$$

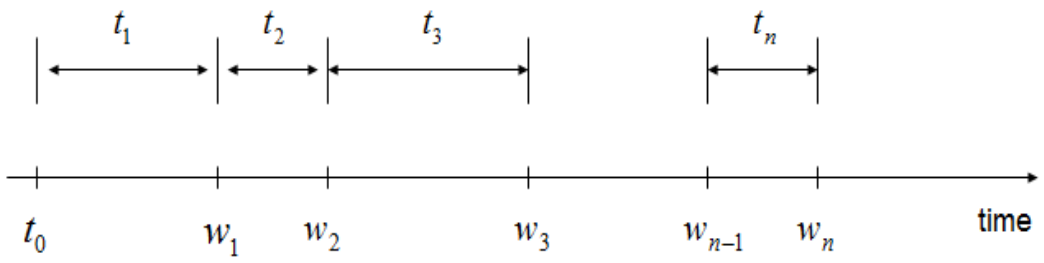

Figure.20 Waiting and interarrival time sequences [55]

The occurrence time sequence $w_1, w_2, \ldots w_n$ for an *inhomogeneous Poisson process* and also for the homogeneous Poisson process, is a Markov sequence with the transition density [55]

$$p_{w_n|w_{n-1}}\left(W_n \mid W_{n-1}\right) = \lambda_{W_n} \exp\left[-\left(\Lambda_{W_n} - \Lambda_{W_{n-1}}\right)\right] \tag{123}$$

where $\Lambda_t \triangleq \int_{t_0}^{t} \lambda_\sigma d\sigma$, using the definition of conditional probability densities, we have [55]

$$p_{w_n|w_{n-1},\ldots w_n}\left(W_n \mid W_{n-1}, \ldots W_1\right) = \frac{p_w^{(n)}(W)}{p_w^{(n-1)}(W)} = \lambda_{W_n} \exp\left(-\int_{W_{n-1}}^{W_n} \lambda_\sigma d\sigma\right) \tag{124}$$

Since the conditional density of w_n given that $w_i = W_i$, for $i = 1, 2, \ldots n-1$ is not a function of W_i, for $i = 1, 2, \ldots n-2$, so the occurrence time sequence for inhomogeneous Poisson process is a Markov sequence with the transition density given by (124)

The next important quantity is the forward-occurrence density, which is the conditional probability density for the n^{th} interarrival time, t_n given the $n-1$ occurrence times $w_1, w_2, \ldots w_n$, for $n = 2, 3, \ldots$. Since the inhomogeneous Poisson process is a Markov sequence, it implies [55]

$$\begin{aligned} p_{t_n|w_{n-1},\ldots,w_1}\left(T \mid W_{n-1,\ldots,} W_1\right) &= p_{w_n|w_{n-1,\ldots,}w_1}\left(W_{n-1} + T \mid W_{n-1,\ldots,} W_1\right) \\ &= p_{w_n|w_{n-1}}\left(W_{n-1} + T \mid W_{n-1}\right) \end{aligned} \tag{125}$$

It follows that the forward-occurrence density for a Poisson process is given by [55]

$$p_{t_n|w_{n-1,\ldots,}w_1}\left(T \mid W_{n-1,\ldots,} W_1\right) = \lambda_{W_{n-1}+T} \exp\left[-\left(\Lambda_{W_{n-1}+T} - \Lambda_{W_{n-1}}\right)\right] \tag{126}$$

Sample function density also plays an important role in problems of statistical inference for observed Poisson processes. It is defined as [55]

$$p\left[N(\sigma); t_0 \le \sigma < t\right] \triangleq \begin{cases} \Pr\left[N(t) = 0\right], & N(t) = 0 \\ p_w\left(W, N(t) = n\right), & N(t) = n \ge 1 \end{cases} \tag{127}$$

where $p_w\left(W, N(t) = n\right) \triangleq \Pr\left[N(t) = n \mid w_1 = W_1, w_2 = W_2, \ldots w_n = W_n\right] p_w^{(n)}(W)$

The sample function defines the probability of obtaining a particular realization of the point process on $[t_0, t)$ with $N(t) = n$ points located, for $n \geq 1$, at times $w_1 = W_1, w_2 = W_2, \ldots, w_n = W_n$.

Follow in (121), (127) and the relation

$$\Pr\left[N(t) = n \mid w_1 = W_1, w_2 = W_2, \ldots w_n = W_n \right] = \Pr\left[N(t) - N\left(W_n^+\right) = 0 \right] = \exp\left(-\int_{W_n}^{t} \lambda_\sigma d\sigma \right) \tag{128}$$

the sample-function density of inhomogeneous for N defined in (127) has the form [55]:

$$p\left[\left\{ N(\sigma); t_0 \leq \sigma < t \right\} \right] = \begin{cases} \exp\left(-\int_{t_0}^{t} \lambda_\sigma d\sigma \right), & N(t) = 0 \\ \left(\prod_{i=1}^{n} \lambda_{W_i} \right) \exp\left(-\int_{t_0}^{t} \lambda_\sigma d\sigma \right), & N(t) = n \geq 1 \end{cases} \tag{129}$$

One of the applications of the sample function density is to derive a useful additional property of the occurrence times. According to (129) and Poisson distribution for $N(T)$, we can obtain [55]

$$p_w\left(W \mid N(T) = n \right) = \frac{p_w\left(W, N(T) = n \right)}{\Pr\left[N(T) = n \right]} = (n!) \prod_{i=1}^{n} \left[\lambda_{W_i} \left(\int_{t_0}^{T} \lambda_\sigma d\sigma \right)^{-1} \right] \tag{130}$$

for $t_0 \leq W_1 \leq \cdots \leq W_n \leq T$

Under the condition $N(T) = n$, where $n \geq 1$, the n occurrence times $w_1, w_2, \ldots w_n$ for inhomogeneous with intensity λ_t have the same distribution as the order statistics [64] of n independent, identically distributed random variables $\tau_1, \tau_2, \ldots \tau_n$ with the probability density being [55]

$$p_{\tau_i}(t) = \begin{cases} \lambda_t \left(\int_{t_0}^{T} \lambda_\sigma d\sigma \right)^{-1}, & t_0 \leq t \leq T \\ 0, & otherwise \end{cases} \tag{131}$$

In summary, we have discussed several important statistical properties of the occurrence and interarrival times for a Poisson counting process. First, the joint occurrence time density is given. Second, the occurrence time sequence is Markov. Third, given the number of points in an interval n, the n occurrence times have the same distributions as the order statistics for n independent and identically distributed random variables with distribution given by (131) [55].

5.2 Poisson Process Parameter Estimation

In the last section, we use the inhomogeneous Poisson process to capture the distribution of the number of path in wireless channels. Now we are interested in solving the following problem: given the data influenced by the parameters, how to design an analysis to perform the estimation and predict the accuracy of the result.

The general problem of estimating the parameters for inhomogeneous Poisson process can be stated as follows [55]: Let $\{N(t); t \geq t_0\}$ be an inhomogeneous Poisson process with intensity $\lambda_t(X)$, where λ a known function of is t and X, where X is a vector of unknown parameter in the space χ of possible parameter values. Let $D_{t_0,T}$ be the observation data on the interval $[t_0, T)$, or the entire path $\{N(\sigma); t_0 \leq \sigma < T\}$. Let $P(D_{t_0,T} \mid X)$ denote the statistics of $D_{t_0,T}$. The problem is restated as follows: knowing $P(D_{t_0,T} \mid X)$ for every $X \in \chi$ and a particular set of data, estimate the parameter X. In other words, we are trying to determine $\hat{X}$, which is an estimate of X.

If the form of the intensity is known, the maximum likelihood estimation can be used. The maximum likelihood estimate $\hat{X}_{ML} = \hat{X}_{ML}(n_1, n_2, \ldots, n_k)$ of X in terms of the observed subinterval counts $n_1, n_2, \ldots, n_k$ is by definition the value of X that maximizes the probability of having observed $n_1, n_2, \ldots, n_k$. The probability of observing $n_1, n_2, \ldots, n_k$ for a given parameter vector X is given by [55, 64]

$$\Pr\left[N_{t_{i-1},t_i}=n_i;i=1,\ldots,k\mid X\right]=\prod_{i=1}^{k}\left\{\left[(n_i)!\right]^{-1}\left(\int_{t_{i-1}}^{t_i}\lambda_\sigma(X)\,d\sigma\right)^{n_i}\exp\left[-\int_{t_{i-1}}^{t_i}\lambda_\sigma(X)\,d\sigma\right]\right\} \tag{132}$$

The log-likelihood function is

$$l(X)=-\int_{t_0}^{T}\lambda_\sigma(X)\,d\sigma+\sum_{i=1}^{k}n_i\ln\left(\int_{t_{i-1}}^{t_i}\lambda_\sigma(X)\,d\sigma\right) \tag{133}$$

In some practical applications, there is no analytic solution for the value of X maximizing the likelihood function. Thus, the estimate of X has to be obtained numerically [55].

Theoretically, the maximum likelihood estimate $\widehat{X}_{ML}$ maximizes the probability of the observation, however, in practice, $\widehat{X}_{ML}$ will not be equal to the actual and unobservable value of X. A nonzero error, defined by $e=\widehat{X}_{ML}-X$, exists in any experiment. We are interested in finding a statistical variable which describes this error. The characterization will be on the ensemble of errors that result with repetitions of the experiment to collect data and generate $\widehat{X}_{ML}$, i.e., treating e as a random variable. The most complete characterization of e is its probability distribution. However, the distribution of e is difficult to determine analytically and we look for a weaker characterization such as the moments of e. The first moment is given by the following [55, 64]

$$b(X)\triangleq E\left[\widehat{X}_{ML}-X\mid X\right] \tag{134}$$

It is named as the bias of the maximum likelihood estimate, the second moment, defined as [55, 64]

$$\sum(X)\triangleq E\left[\left(\widehat{X}_{ML}-X\right)\left(\widehat{X}_{ML}-X\right)'\mid X\right] \tag{135}$$

The second moment is also called the mean square-error matrix. If the bias is zero, the error covariance matrix is the same as the mean square error matrix.

Sometimes, even the first and second moments are not easy to determine and it is convenient to consider a lower bound on the mean square error matrix for an arbitrary

estimate in terms of available data. Then we compare mean square error matrix with this bound. Under favorable circumstances the maximum likelihood estimates can achieve the lower bound so that its mean square error matrix is the smallest possible using any other conceivable estimate with the same data [55, 64].

The lower bound is called Cramer-Rao Lower Bound, introduced in [68] and stated as follows.

Theorem 5.3.1 [55]: Let $\{N(t); t \geq t_0\}$ be an inhomogeneous Poisson process with a parameterized intensity function $\lambda_t(X)$ for $t \geq t_0$. Let X^* denote an arbitrary estimate of X based on observations of N on the interval $[t_0, T)$. Then expression (135) satisfies

$$\sum(X) \geq b(X)b'(X) + \left[I + \frac{\partial b(X)}{\partial X}\right] F^{-1}(X) \left[I + \frac{\partial b(X)}{\partial X}\right]' \tag{136}$$

where $b(X)$ is the bias given in (134), $\frac{\partial b(X)}{\partial X}$ is the Jacobin matrix of the bias, I denote the identity matrix, $F(X)$ is the Fisher information matrix defined by

$$F(X) = \sum_{i=1}^{k} \left[\int_{t_{i-1}}^{t_i} \lambda_\sigma(X)\, d\sigma\right]^{-1} \left[\int_{t_{i-1}}^{t_i} \frac{\partial \lambda_\sigma(X)}{\partial X} d\sigma\right] \left[\int_{t_{i-1}}^{t_i} \frac{\partial \lambda_\sigma(X)}{\partial X} d\sigma\right]' \tag{137}$$

Equality holds if and only if estimate X^* satisfies

$$X^* = X + b(X) + \left[I + \frac{\partial b(X)}{\partial X}\right] F^{-1}(X) \left[\frac{\partial l(X)}{\partial X}\right] \text{ for all } X \tag{138}$$

Some important remark can be obtained from the previous theorem [55]

- An estimate having a mean square error matrix satisfying (136) with equality is said to be efficient. Suppose there exists and efficient estimate X^*. Then, in terms of this estimate, it follows that the maximum likelihood estimate $\widehat{X}_{ML}$ will be efficient if and only if the bias is zero. Further, if there exists an unbiased, efficient estimate, then the maximum likelihood estimate is both unbiased and efficient.

- Suppose X^* is an unbiased estimate of X, then the mean square error matrix satisfies $\sum(X) \geq F^{-1}(X)$. Consider the experiment of collecting data to estimate X is repeated M times. Denote the estimate of X based on M trials by X_M^*. Then the mean square error matrix for X_M^* satisfies $\sum_M(X) \geq (1/M)F^{-1}(X)$. The estimated X_M^* is said to be asymptotically efficient if $\left[M\sum_M(X) - F^{-1}(X)\right]$ tends to zero as M tends to infinity. It is said to be asymptotically unbiased if $b(X)$ tends to zero as M tends to infinity. The maximum likelihood estimate is both asymptotically efficient and asymptotically unbiased under relatively weak conditions [69].

5.3 Extension of Campbell's Theorem

Campbell's theorem [36] gives the information about the average value and the standard deviation of the probability distribution of the received signals generated by shot noise. In this section we are going to discuss the distribution of the received signals which are defined in (7) and (9). In order to simplify the calculation, the received signals can be rewritten as

$$y_l(t) = \sum_{i=1}^{N(T_s)} r_i(\tau_i) e^{j\phi_i} s_l(t-\tau_i) = \sum_{i=1}^{N(T_s)} h_l\left(t, \tau_i, m_i(\tau_i)\right),\ m_i(\tau_i) = \left(r_i(\tau_i), \phi_i\right) \tag{139}$$

$$y(t) = \sum_{i=1}^{N(T_s)} h\left(t, \tau_i, m_i(\tau_i)\right) \tag{140}$$

Assumption 5.4.1 [76]: let $\left\{\lambda_T(s) \triangleq c \times f(s),\ 0 \leq s \leq t\right\}$ denotes the nonnegative rate of the counting process $\left\{N(s),\ 0 \leq s \leq t\right\}$, where c is constant and non-random and $f(t)$ is a time-varying non-random function. For fixed τ_i, the random processes $h_l\left(t, \tau_i, m_i(\tau_i)\right),\ i = 1,2,3,...$ are i.i.d. having the same distribution and independent of $\left\{N(s),\ 0 \leq s \leq t\right\}$.

The mean of the received signal $y_l(t)$, is defined by [76]

$$\overline{y}_l(t) \triangleq E\left[\sum_{i=1}^{N(T_s)} h_l\left(t,\tau_i,m_i(\tau_i)\right)\right] \tag{141}$$

The variance is defined in [76]

$$Var\left(y_l(t)\right) \triangleq E\left[y_l^*(t)\,y_l(t)\right] - \overline{y}_l^*(t)\,\overline{y}_l(t) \tag{142}$$

Under the previous assumption, the delay times $\{\tau_i\}$ are independent identically distributed with density given by $f(t) = \dfrac{\lambda_T(t)}{\int_0^{T_s}\lambda_T(t)\,dt}$, $0 \le t \le T_s$ (see (131)).

Hence, $\overline{y}_{l,k}(t)$ is calculated as follows

$$\overline{y}_{l,k}(t) = E\left[\sum_{i=1}^{N(T_s)} h_l\left(t,\tau_i,m_i(\tau_i)\right)\middle| N(T_s)=k\right] = \sum_{i=1}^{k}\int_0^{T_s} f(\tau)\,E\left[h_l\left(t,\tau,m_i(\tau)\right)\right]d\tau \tag{143}$$

The ensemble average of the received signal is obtained from

$$E\left[y_l(t)\right] = \sum_{k=1}^{\infty}\overline{y}_{l,k}(t)\,\mathrm{P}\left\{N(T_s)=k\right\} \tag{144}$$

Similarly, the second moment can be calculated as [76]

$$E\left[\left(y_l(t)\right)^2\right] = \sum_{k=1}^{\infty}\overline{y}_{l,k}^2(t)\,\mathrm{P}\left\{N(T_s)=k\right\} \tag{145}$$

where

$$\begin{aligned}\overline{y}_{l,k}^2(t) = E\left[y_l^2(t)\middle|N(T_s)=k\right] &= \sum_{i=1}^{k}\int_0^{T_s} f(\tau)\,E\left[h_l\left(t,\tau,m_i(\tau)\right)\right]^2 d\tau \\ &+ \sum_{\substack{i,j=1\\ i\neq j}}^{k}\int_0^{T_s} f(\tau_i)\,d\tau_i\int_0^{T_s} f(\tau_j)\,d\tau_j\,E\left[h_l^*\left(t,\tau_i,m_i(\tau_i)\right)h_l\left(t,\tau_i,m_i(\tau_i)\right)\right]\end{aligned} \tag{146}$$

$$\begin{aligned}Var\left(y_l(t)\right) &= \sum_{k=1}^{\infty}\overline{y}_{l,k}^2(t)\,P\left(N(T_s)=k\right) - \left(\overline{y}_l(t)\right)^2 \\ &= \sum_{k=1}^{\infty}P\left(N(T_s)=k\right)\sum_{i=1}^{k}\left\{\begin{array}{l}\int_0^{T_s} f(\tau)\,E\left[r_i^2(\tau)\right]\left(s_l(t-\tau)^2\right)d\tau - \\ \left(\int_0^{T_s} f(\tau)\,E\left[r_i(\tau)e^{j\phi_i}s_l(t-\tau)\right]d\tau\right)^2\end{array}\right\}\end{aligned} \tag{147}$$

If the previous assumption is satisfied, then (144) and (147) can be calculated explicitly

by the following form which is a generalization of the shot-noise effect discussed by Rice [36]

$$E\left[y_l(t)\right]=\int_0^{T_s}\lambda_T(\tau)E\left[h_l\left(t,\tau,m(\tau)\right)\right]d\tau=\int_0^{T_s}\lambda_T(\tau)E\left[r(\tau)e^{j\phi}\right]s_l(t-\tau)\,d\tau \tag{148}$$

$$\begin{aligned}Var\left(y_l(t)\right)&=\int_0^{T_s}\lambda_T(\tau)E\left[\left(h_l\left(t,\tau,m(\tau)\right)\right)^2\right]d\tau\\&=\int_0^{T_s}\lambda_T(\tau)E\left[r^2(\tau)\right]\left(s_l(t-\tau)\right)^2d\tau\end{aligned} \tag{149}$$

Follow in the same idea, the correlation of $y_l(t_1)$ and $y_l(t_2)$ can be defined as [76]

$$\begin{aligned}R_{yl}(t_1,t_2)&\triangleq E\left[y_l^*(t_1)\,y_l(t_2)\right]\\&=E\left[\sum_{i=1}^{N(T_s)}h_l^*\left(t_1,\tau_i,m_i(\tau_i)\right)\sum_{j=1}^{N(T_s)}h_l\left(t_2,\tau_j,m_j(\tau_j)\right)\right]\\&=\sum_{k=1}^{\infty}R_{yl,k}(t_1,t_2)\,P\left(N(T_s)=k\right)\end{aligned} \tag{150}$$

where

$$\begin{aligned}R_{yl,k}(t_1,t_2)&\triangleq E\left[y_{l,k}^*(t_1)\,y_{l,k}(t_2)\right]=E\left[\sum_{i=1}^{k}h_l^*\left(t_1,\tau_i,m_i(\tau_i)\right)h_l\left(t_2,\tau_j,m_j(\tau_j)\right)\right]\\&\quad+E\left[\sum_{\substack{i,j=1\\i\neq j}}^{k}h_l^*\left(t_1,\tau_i,m_i(\tau_i)\right)h_l\left(t_2,\tau_j,m_j(\tau_j)\right)\right]\\&=\sum_{i=1}^{k}\frac{1}{\int_0^{T_s}\lambda_T(t)\,dt}\int_0^{T_s}\lambda_T(t)E\left[r_i^2(\tau)\right]s_l^*(t_1-\tau)s_l(t_2-\tau)\,d\tau\\&\quad+\sum_{\substack{i,j=1\\i\neq j}}^{k}\frac{1}{\int_0^{T_s}\lambda_T(t)\,dt}E\left[\begin{array}{l}e^{j\phi_i}\int_0^{T_s}\lambda_T(t)r_i(\tau)s_l^*(t_1-\tau)\,d\tau\\\times\dfrac{1}{\int_0^{T_s}\lambda_T(t)\,dt}\int_0^{T_s}\lambda_T(t)r_j(\tau)s_l^*(t_2-\tau)\,d\tau\end{array}\right]\end{aligned} \tag{151}$$

If assumption 5.4.1 was satisfied, the correlation can be calculated explicitly as [76]

$$
\begin{aligned}
R_{yl}(t_1,t_2) &= c\int_0^{T_s} f(\tau)E\left[h_l^*(t_1,\tau,m(\tau))h_l(t_2,\tau,m(\tau))\right]d\tau \\
&\quad + c\int_0^{T_s} f(\tau)E\left[h_l^*(t_1,\tau,m(\tau))\right]d\tau \times c\int_0^{T_s} f(\tau)E\left[h_l(t_2,\tau,m(\tau))\right]d\tau \\
&= c\int_0^{T_s} f(\tau)E\left[r^2(\tau)\right]s_l^*(t_1-\tau)s_l(t_2-\tau)d\tau \\
&\quad + c\int_0^{T_s} f(\tau)E\left[r(\tau)e^{-j\phi}\right]s_l^*(t_1-\tau)d\tau \times c\int_0^{T_s} f(\tau)E\left[r(\tau)e^{-j\phi}\right]s_l(t_2-\tau)d\tau
\end{aligned}
\tag{152}
$$

5.4 Distribution and Characteristic function

Distribution and characteristic functions are important statistics for describing the received signals. We discuss both of them in this section.

Let I denotes the indicator function. The probability density function is defined by [76]

$$
f_{y_l}(x_l,t)dx_l = E\left[I_{\{y_l(t)\in dx_l\}}\right] \tag{153}
$$

The characteristic function is defined by [76]

$$
\Phi_{y_l}(u) \triangleq E\left[e^{ju\times y_l(t)}\right] \tag{154}
$$

For fixed $N(T_s)=k$, the density of $y_l(t)$ is [76]

$$
f_{y_{l,k}}(x_l,t)dx_l = E\left[I_{\{y_l(t)\in dx_l\}}\middle|N(T_s)=k\right] = P\left\{\sum_{i=1}^{k} h_l(t,\tau_i,m_i(\tau_i))\in dx\right\} \tag{155}
$$

For fixed $N(T_s)=k$, the characteristic function is

$$
\begin{aligned}
\Phi_{y_{l,k}}(t) &= E\left[\exp\left\{ju\sum_{i=1}^{N(T_s)} h_l(t,\tau_i,m_i(\tau_i))\right\}\middle|N(T_s)=k\right] \\
&= \frac{1}{T_s^k}\int_0^{T_s} d\tau_1\int_0^{T_s} d\tau_2 \ldots \int_0^{T_s} d\tau_k E\left[\prod_{i=1}^{k} e^{juh_l(t,\tau_i,m_i(\tau_i))}\right] \\
&= \prod_{i=1}^{k}\frac{1}{T_s}\int_0^{T_s} d\tau E\left[e^{juh_l(t,\tau_i,m_i(\tau_i))}\right]
\end{aligned}
\tag{156}
$$

If the assumption 5.4.1 is satisfied, the characteristic function of $y_l(t)$ can be calculated

as [76]

$$\Phi_{y_l}(t) = E\left[e^{ju\times y_l(t)}\right] = \exp\left\{c\int_0^{T_s} f(\tau) E\left[e^{ju\times h_l(t,\tau,m(\tau))} - 1\right] d\tau\right\} \tag{157}$$

and its density function is given by [76]

$$f_{y_l}(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} du e^{-jux} \exp\left\{c\int_0^{T_s} f(\tau) E\left[e^{ju\times h_l(t,\tau,m(\tau))} - 1\right] d\tau\right\} \tag{158}$$

The joint characteristic function of $y_l(t_1),\ldots y_l(t_n)$ and their cumulates are obtained from [76]

$$\begin{aligned}\Phi_{y_l}(u_1,t_1;\ldots u_n,t_n) &= E\left\{\exp\left(\sum_{i=1}^{n} ju_i y_l(t_i)\right)\right\} \\ &= \exp\left\{c\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} ju_i \times h_l(t_i,\tau,m(\tau)) - 1\right] d\tau\right\}\end{aligned} \tag{159}$$

Further, we are going to illustrate the generalization of the Gaussianity of shot-noise described in [36].

Define the centered random variables [76]

$$y_{l,c}(t_i) = \frac{y_l(t_i) - \overline{y}(t_i)}{\sigma_{y_l}(t_i)},\ \sigma_{y_l}(t_i) = \sqrt{Var\left(y_l(t_i)\right)} \tag{160}$$

The joint-characteristic function of the centered random variables [76]

$$\begin{aligned}&\Phi_{y_{l,c}}(ju_1,t_1;\ldots ju_n,t_n) = E\left[\exp\left\{\sum_{i=1}^{n} ju_i y_{l,c}(t_i)\right\}\right] \\ &= \exp\left\{-j\sum_{i=1}^{n} u_i \frac{\overline{y}_l(t_i)}{\sigma_{y_l}(t_i)}\right\} \times \exp\left\{c\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} \exp\left(\sum_{i=1}^{n} \frac{ju_i}{\sigma_{y_l}(t_i)} h_l(t_i,\tau,m(\tau)) - 1\right)\right] d\tau\right\}\end{aligned} \tag{161}$$

(157) can be expanded in power series as

$$\begin{aligned}&c\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} \exp\left(\sum_{i=1}^{n} \frac{ju_i}{\sigma_{y_l}(t_i)} h_l(t_i,\tau,m(\tau)) - 1\right)\right] d\tau \\ &= c\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} \frac{ju_i}{\sigma_{y_l}(t_i)} h_l(t_i,\tau,m(\tau))\right] d\tau + \frac{1}{2} c\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} \frac{ju_i}{\sigma_{y_l}(t_i)} h_l(t_i,\tau,m(\tau))\right]^2 d\tau + \ldots\end{aligned} \tag{162}$$

Since the order of $\sigma_{y_l}(t_i)$ is proportional to $c^{\frac{1}{2}}$, the order of first item in (162) is proportional to $c^{\frac{1}{2}}$, the order of the second item is proportional to 1, the third item is proportional to $c^{-\frac{1}{2}}$. If we ignore the terms of high order, we can get the following approximation [76]

$$\begin{aligned} & c\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} \exp\left(\sum_{i=1}^{n} \frac{ju_i}{\sigma_{y_l}(t_i)} h_l(t_i,\tau,m(\tau)) - 1\right)\right] d\tau \\ & = c\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} \frac{ju_i}{\sigma_{y_l}(t_i)} h_l(t_i,\tau,m(\tau))\right] d\tau + \frac{1}{2} c\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} \frac{ju_i}{\sigma_{y_l}(t_i)} h_l(t_i,\tau,m(\tau))\right]^2 d\tau \end{aligned} \tag{163}$$

Then we substitute (161) with (163),

$$\Phi_{y_{l,c}}(ju_1,t_1;\ldots ju_n,t_n) \approx \exp\left\{-\frac{\lambda}{2}\int_0^{T_s} f(\tau) E\left[\sum_{i=1}^{n} \frac{ju_i}{\sigma_{y_l}(t_i)} h_l(t_i,\tau,m(\tau))\right]^2 d\tau\right\} \tag{164}$$

From (164), we can see that joint characteristics function of $y_{l,c}$ is quadratic. Further, $y_{l,c}(t_1),\ldots y_{l,c}(t_n)$ is approximately Gaussian, with zero mean and covariance matrix identified. When $\lambda \to \infty$, $y_{l,c}(t_j) \sim N(0,1), 1 \le j \le n$.

5.5 Numerical Result

Based on the theoretical results in previous sections, some simulation results about inhomogeneous Poisson process are provided in this section. The approach is based on the observations which have the rate function $\lambda(t)$, the increment $N_t - N_s, t_0 \le s < t$ is distributed as a Poisson random process with intensity $\Lambda_t = \int_{t_0}^{t} \lambda_\sigma d\sigma$. Let the cumulative distribution function be $F(t)$, following from (120),

$$F(t) = 1 - \exp\left(-\int_0^t \lambda(s+v)\, dv\right) \tag{165}$$

Next the inverse F^{-1} is found and then a random variable T with the distribution F is generated.

The simulation is accomplished as follows

Step 1: Compute the realization of the inhomogeneous Poisson process, which is equivalent to compute the occurrence times. The detailed step is the following: first, set $T_0 = 0$, second, for $i = 1, 2, \ldots n$, generate a random variable Y uniformly distributed on $(0,1)$, update T by $T_i = T_{i-1} + F^{-1}(Y)$.

Step 2: Compute the attenuation coefficient for each path.

The simulation is shown in Figure 21.

The first plot in this Figure shows the intensity function $\lambda(t)$, the red lines indicate the occurrence times calculated from step 1. The second plot is the signal magnitude based on the occurrence times.

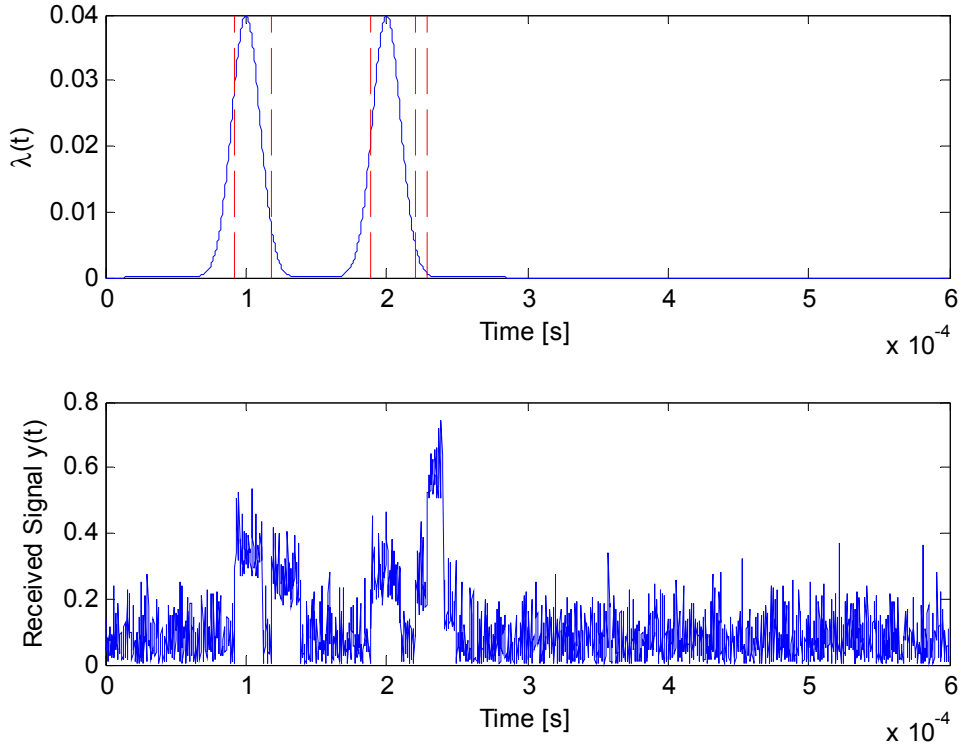

Figure 21 Simulation of inhomogeneous Poisson process

Chapter 6

Input Design and Model Selection

This thesis uses state space model to model the UWB indoor wireless channel. In the literature [80] Finite Impulse Response (FIR) models are popular in particular in communication systems, such as, in Orthogonal Frequency Division Multiplexing (OFDM) systems. FIR model can be realized in state space form. In this chapter we show that FIR models are optimized in some sense. Further, after the model is selected, certain input can be chosen to minimize the worst identification error. We also show that the model selection and input design can be done independently. The model selection is not affected by the experiment conditions. Also, the input design is independent of the model set.

6.1 Introduction to Fast Identification Problem

In general, identification can be divided to two steps [75], one is information acquisition, the other is information processing. During the information acquisition, the experiment data is collected. Then information processing gives us a representation of the available information. In order to achieve fast identification, we should acquire information quickly and process it promptly.

Let's consider the wireless channel as a discrete time stable linear time invariant system which is represented by a convolution operator of the following form [75]

$$y(t) \triangleq \sum_{\tau=0}^{\infty} h(\tau)u(t-\tau), t \in Z \tag{166}$$

Assume U, Y represent inputs and outputs spaces respectively. We assume that the set U is contained in $L^2[0,\infty)$. For fixed $u \in U$, define the map $\Phi_u(h) \triangleq Hu = y$ is a linear map from impulse response to outputs.

In practice, the observations are corrupted by noise, i.e.,

$$y(t)=\sum_{\tau=0}^{\infty}h(\tau)u(t-\tau)+v(t),t\in Z \tag{167}$$

(167) can be rewritten in a more compact form as

$$y=\Phi_u(h)+v \tag{168}$$

We assume that the a priori knowledge of the channels defined by the set: $S_a \triangleq \{h:h\in L^2[0,\infty)\}$ ($L^2[0,\infty)$ is defined in Appendix A). Since the structure of the real channel is unknown, the accuracy of the estimate will be measured by the distance between the real channel and estimates from the observation in a finite length interval.

Based on the observations $\{y(t_0),y(t_0+1),\ldots,y(t_0+T-1)\}$, the true kernel is well defined in the following set [75]

$$S(y):=\left\{h\in S_a:\sum_{\tau=0}^{\infty}h(\tau)u(t-\tau)=y(t)+v(t),\forall t\in[t_0,t_0+T)\right\} \tag{169}$$

$S(y)$ is called the posterior information set about the real channel. It can be written in a more compact form as [75]

$$S(y):=\left\{h\in S_a:P_{[t_0,t_0+T)}\left(\Phi_u(h)-y\right)=P_{[t_0,t_0+T)}(v)\right\} \tag{170}$$

Now we are going to define the worst case error which is to measure the difference [75]:

$$e_{y,h_{est}}(u):=\sup_{h\in S(y)}\|h-h_{est}\| \tag{171}$$

The optimized estimate can be computed by minimizing the worst case error

$$e_y(u):=\inf_{h_{est}\in L^2[0,\infty)}\sup_{h\in S(y)}\|h-h_{est}\| \tag{172}$$

However, if we let the estimate h_{est} be in $L^2[0,\infty)$, an infinite number of parameters will have to be identified to determine the channel, which means it is computationally not feasible. In order to achieve fast identification, we minimize the identification error by choosing a finite dimensional model set, denoted X_n. The model set can be an ARMA model, a state space model, etc.

If a parameter set X_n is chosen, the minimum identification error becomes [75]

$$e_y^T\left(u, X_n\right) = \inf_{h_{est} \in X_n} \sup_{h \in S(y)} \left\| h - h_{est} \right\| \tag{173}$$

Therefore, the worst case identification error becomes [75]

$$e^T\left(u, X_n\right) = \sup_{h_{true} \in S_a} \sup_{v_{true} \in V} \inf_{h_{est} \in X_n} \sup_{h \in S(y)} \left\| h - h_{est} \right\| \tag{174}$$

The main task in this chapter is to minimize the worst case identification error. Some new concepts will be introduced to solve this problem.

6.2 Partition of the Worst Case Identification Error

The worst case identification error can be separated into two parts, the inherent error and representation error, which depend on the input and the model set, respectively.

We introduce the notion of inherent error as follows [75]

$$\delta^T\left(u\right) \triangleq \sup\left\{ \left\| h \right\| : h \in S_a, \Phi_u\left(h\right)\left(t\right) = v\left(t\right), \forall t \in \left[t_0, t_0 + T\right] \right\} \tag{175}$$

The second part is the representation error defined as [75]

$$dist\left(S_a, X_n\right) := \sup_{h \in S_a} \inf_{g \in X_n} \left\| h - g \right\| \tag{176}$$

The inherent error is due to the information collection such as lack of data and incomplete measurement. This error is irreducible no matter what model is chosen and what identification algorithm is used. The representation error is due to inaccurately representation of the dataset, which represents the loss of information in the data processing stage. The worst case identification error is minimized when both of inherent and representation errors are minimized [75].

The following theorem gives the lower and upper bounds of the worst case identification error $e^T\left(u, X_n\right)$ in terms of the inherent and representation errors.

Theorem 6.2.1: If S_a and V are convex symmetric subsets of $L^2[0,\infty)$ and Y respectively, then $e^T(u, X_n)$ satisfies

$$\max\left\{\delta^T(u), dist(S_a, X_n)\right\} \le e^T(u, X_n) \le 3\max\left\{\delta^T(u), dist(S_a, X_n)\right\} \qquad (177)$$

Proof: see [75]

The minimum representation and inherent errors are shown in next two sections, from which the optimal model and optimal input will be obtained.

6.3 Representation Error and Kolmogorov n-Width

By the definition of representation error, the minimum representation error of a n-dimensional subspace is [72]

$$d_n\left(S_a, L^2[0,\infty)\right) := \inf_{X_n \subset L^2[0,\infty)} dist(S_a, X_n) \qquad (178)$$

Definition 6.3.1 [72]: Consider the Hilbert space $L^2[0,\infty)$ and S_a a subset of $L^2[0,\infty)$. The n-Width, in the sense of Kolmogorov, of S_a in $L^2[0,\infty)$ is given by

$$d_n\left(S_a, L^2[0,\infty)\right) \triangleq \inf_{X_n} \sup_{h \in S_a} \inf_{g \in X_n} \|h - g\| \qquad (179)$$

where the infimum is taken over all n-dimensional subspace X_n of $L^2[0,\infty)$. If

$$d_n\left(S_a, L^2[0,\infty)\right) = \sup_{h \in S_a} \inf_{g \in X_n} \|h - g\|_{L^2[0,\infty)} \qquad (180)$$

If the infimum in (180) is achieved for some subspace X_n of dimension at most n, then X_n is said to be optimal subspace for $d_n\left(S_a, L^2[0,\infty)\right)$.

Since the expression of the received signal has a convolution in time domain, it is easier to take the z-transform for both sides of (166), i.e., switch the problem to frequency domain.

Without losing generality, we choose the priori set to be $A=\left\{h:h\in H^2,\left\|h\right\|_{H^2}\le 1\right\}$, where H^2 is the class of analytic functions in the unit disk with norm given by [72]

$$\left\|h\right\|_{H^2}=\left(\sum_{j=0}^{\infty}\left|a_j\right|^2\right)^{1/2} \tag{181}$$

where $h(z)=\sum_{j=0}^{\infty}a_j z^j$. Let μ be a positive measure on a compact set K , $K\subseteq\left\{\left|z\right|\le r,r<1\right\}$ we wish to calculate the distance $d_n\left(A,L^2\left(K,d\mu\right)\right)$ and determine the optimal subspaces for d_n.

Consider Cauchy's integral formula [72],

$$f(z)=\frac{1}{2\pi}\int_0^{2\pi}\frac{f\left(e^{i\theta}\right)}{1-ze^{-i\theta}}d\theta \tag{182}$$

For $z\in K$, set $T\left(f(z)\right)=f(z)$. T is regarded as a map from H^2 to $L^2\left(K,d\mu\right)$, then [72]

$$d_n\left(A,L^2\left(K,d\mu\right)\right)=d_n\left(T\left(H^2\right),L^2\left(K,d\mu\right)\right) \tag{183}$$

According to [72], the d_n is equal to the square root of the eigenvalue which is arranged in decreasing order of magnitude to their multiplicity of TT', where T' is the adjoint of T and the n eigenfunctions of TT' corresponding to the n largest eigenvalues span an optimal subspace for d_n.

Remark: Before calculating d_n, we have to determine T'. T' maps $L^2\left(K,d\mu\right)$ to H^2 and satisfies

$$\left(f,Tg\right)_{L^2(K,d\mu)}=\left(T'f,g\right)_{H^2} \tag{184}$$

For all $f\in L^2\left(K,d\mu\right)$ and $g\in H^2$. Further,

$$\left(T'f\right)\left(e^{i\theta}\right)=\int_K\frac{f(w)}{1-\bar{w}e^{i\theta}}d\mu(w) \tag{185}$$

and

$$\left(TT'f\right)(z)=\int_K \frac{f(w)}{1-\bar{w}z}d\mu(w) \tag{186}$$

So the eigenvalue eigenfunction problem becomes

$$\lambda f(z)=\int_K \frac{f(w)}{1-\bar{w}z}d\mu(w) \tag{187}$$

Based on the Theorem 2.2 and Proposition 2.4 in [72], the $span\left\{1,re^{i\theta},\ldots,r^{n-1}e^{i(n-1)\theta}\right\}$ is optimal for $d_n\left(A,L^2\left(K,d\mu\right)\right)$. Since $z=re^{i\theta}$, the $span\left\{1,re^{i\theta},\ldots,r^{n-1}e^{i(n-1)\theta}\right\}$ is optimal and corresponds to the FIR models.

6.4 Inherent Error and Gel'fand n-Width

Definition 6.4.1 [72, 75]: Let S_a be a subset of $L^2\left[0,\infty\right)$. The Gel'fand n-width of S_a in $L^2\left[0,\infty\right)$ is given by

$$d^n\left(S_a,L^2\left[0,\infty\right)\right)\triangleq\inf_{L^n}\sup_{h\in S_a\cap L^n}\left\|h\right\|_{L^2[0,\infty)} \tag{188}$$

where the infimum is taken over all subspace L^n of $L^2\left[0,\infty\right)$ of codimension n. A subspace is said to be of codimension n independent bounded linear functions $f_1,\ldots,f_n$ such that $L^n=\left\{h\in L^2\left[0,\infty\right):f_i\left(h\right)=0,i=1,\ldots,n\right\}$. If L^n is a subspace of codimension at most n for which $d^n\left(S_a,L^2\left[0,\infty\right)\right)=\sup\left\{\left\|h\right\|_{L^2[0,\infty)}:h\in S_a\cap L^n\right\}$, then L^n is called an optimal subspace for the Gel'fand n-width $d^n\left(S_a,L^2\left[0,\infty\right)\right)$.

Definition 6.4.2 [75]: For any $n\in\mathbb{Z}_+$ and arbitrary $t_0\in\mathbb{Z}$, the time n-width is defined as

$$\theta^n\left(S_a,L^2\left[0,\infty\right)\right)\triangleq\begin{cases}\inf_{u\in U}\delta_0^n\left(u\right) & n>0\\ \left\|S_a\right\|_{L^2[0,\infty)}, & n=0\end{cases} \tag{189}$$

$$\delta_0^T(u) \triangleq \sup\left\{\|h\|_{L^2[0,\infty)} : h \in S_a, \Phi_u(h)(t) = 0, \forall t \in [t_0, t_0+T)\right\}$$
where
$$= \sup\left\{\|h\|_{L^2[0,\infty)} : h \in S_a \cap Null\left(P_{[t_0,t_0+T)}\Phi_u\right)\right\} \quad (190)$$

and $Null(\cdot)$ represents the null space. The optimal input is the one for which the infimum is reached in (189).

The time n-width is the best achieved inherent error with n consecutive output observations. Furthermore, the time n-width and Gel'fand n-width are related to each other under some condition. Time n-width is bounded below by Gel'fand n-width in certain sense (see [75] Proposition 3.2).

Before introducing the theorem to estimate the Gel'fand n-width, we need the following definitions.

Definition 6.4.3 [75]: For any subset S of $L^2[0,\infty)$, $S\big|_{[t_1,t_2)}$ denotes the subset of functions of S in the interval $[t_1,t_2)$ of $\mathbb{Z}$, i.e., $S\big|_{[t_1,t_2)} \triangleq S \cap P_{[t_1,t_2)}S$. For a set of the form $S\big|_{[0,m)}$, a p-section of $S\big|_{[0,m)}$ is the intersection with any p-dimensional subspace of $L^2[0,\infty)\big|_{[0,m)}$.

Definition 6.4.4 [75]: S_a will be called q-monotone decreasing, $1 \le q \le \infty$ if every p-section X_p in $S_a\big|_{[0,m)}$ satisfies

$$\|X_p\|_{L^2[0,\infty)} \ge \left\|S_a\big|_{[m-p,m)}\right\|_{L^2[0,\infty)} \quad (191)$$

where $1 \le p < m < \infty,\ p \le q \le \infty$

Next we introduce a theorem which can be used to estimate the Gel'fand n-width.

Theorem 6.4.5 [75]: if the priori set S_a is q-monotone decreasing, $(1 \le q \le \infty)$, then the

Gel'fand n-width d^n has bounds

$$\left\| S_a \Big|_{[n,n+p]} \right\|_{L^2[0,\infty)} \le d^n\left(S_a, L^2[0,\infty)\right) \le \left\| S_a \Big|_{[n,\infty)} \right\|_{L^2[0,\infty)}, \left(p \le q, p < \infty\right) \tag{192}$$

Moreover, if

$$\lim_{p \to q} \left\| S_a \Big|_{[n,n+p)} \right\|_{L^2[0,\infty)} = \left\| S_a \Big|_{[n,\infty)} \right\|_{L^2[0,\infty)} \tag{193}$$

The $d^n\left(S_a, L^2[0,\infty)\right) = \left\| S_a \Big|_{[n,\infty)} \right\|_{L^2[0,\infty)}$ and the subspace $L_{opt}^n = \left\{h : h(i) = 0, i = 0, \ldots, n-1\right\}$ is optimal for d^n.

Applying theorem 6.4.5, we can get the estimate of time n-width of a q-monotone decreasing data set.

Theorem 6.4.6 [75]: If the priori set S_a is q-monotone decreasing (see [75], page 35), then the n-width θ^n has bounds

$$\left\| S_a \Big|_{[n,n+p]} \right\|_{L^2[0,\infty)} \le \theta^n\left(S_a, L^2[0,\infty)\right) \le \left\| S_a \Big|_{[n,\infty)} \right\|_{L^2[0,\infty)}, \left(p \le q, p < \infty\right) \tag{194}$$

Moreover, if

$$\lim_{p \to q} \left\| S_a \Big|_{[n,n+p)} \right\|_{L^2[0,\infty)} = \left\| S_a \Big|_{[n,\infty)} \right\|_{L^2[0,\infty)} \tag{195}$$

Then $\theta^n\left(S_a, L^2[0,\infty)\right) = d^n\left(S_a, L^2[0,\infty)\right) = \left\| S_a \Big|_{[n,\infty)} \right\|_{L^2[0,\infty)}$ and an impulse at the start of the observation interval is optimal for θ^n.

From this theorem, we obtain the principle based on the property of monotone decrease of the data sets. For such data set, the optimal input is an impulse at the start of the observation interval [75].

Chapter 7

Conclusion and Future Work

7.1 Conclusion

This dissertation proposes a new time-varying wireless channel model which captures both space and time variation of UWB indoor channels. The distribution of attenuation coefficient of received signal of the indoor channels satisfies the Nakagami distribution. The dynamics of UWB indoor channels are captured by stochastic diffusion processes which are represented by stochastic differential equations. Furthermore, ergodicity property of the channel is discussed and certain conditions for it are derived.

System identification is a process of constructing a mathematical model for a dynamic system from observations and prior knowledge. MLE is studied and its limitation is discussed. The EM algorithm and the Kalman filter are used in estimating channel parameters as well as the inphase and quadrature components, respectively. The proposed algorithms are recursive and therefore can be implemented in real time.

An optimal Distributed SPCA based on the developed models is proposed. The optimal DPCA is shown to reduce to a fixed point problem and solved iteratively using stochastic approximation. Numerical results show that there are potentially large gains to be achieved by using TV stochastic models and the distributed SPCA provides better power stability and consumption than the distributed SPCA.

A framework of the statistics of the overall received signal is established. A connection between Rice's shot-noise analysis and wireless channels is built. Some statistics of the receiver signals such as second moment functions are calculated, explicit expressions are also provided under certain conditions. Moreover, the number of path is also treated as a random process which described by homogeneous or inhomogeneous Poisson

process. A parameter estimation method of the Poisson process is provided and the Cramer-Rao Lower Bound is also calculated.

The worst case identification error of the wireless channel is divided into two parts, the input design and model selection errors. These errors are related to n-widths in the sense of Gel'fand and Kolmogorov in metric complexity theory. The results are derived for time invariant channels models. It is shown that the optimal model to achieve the minimum worst case representation error is the FIR model, while the optimal input to minimize the inherent error is an impulse at the start of the observation interval.

7.2 Direction for Future Work

This thesis contains several aspects which are worth further study and are listed as follows

1. The UWB channel model is valid for Nakagami distribution and needs to be extended to more general indoor channel models.

2. In the power control problem defined in expression (82); to simplify the problem we took expectations of the numerator and denominator separately. However, in general expectation of the whole expression should be evaluated, which makes the problems more complicated to solve.

3. Input design and model selection based on metric complexity need to be extended to more general wireless channel modeling, in particular, to time varying impulse responses.

4. More general models such as the doubly inhomogeneous Poisson process needs to be investigated to derive more realistic models for the number of paths in MPCs.

LIST OF REFERENCES

LIST OF REFERENCES

[1] UWB Theory and Applications, Ian Oppermann

[2] Win, M.Z.; Ramirez-Mireles, F.; Scholtz, R.A.; Barnes, M.A., 'Ultra-wide Bandwidth (UWB) signal propagation for outdoor wireless communications', Vehicular Technology Conference, 1997 IEEE 47th, Volume 1, 4-7 May 1997, Page(s):251 - 255 vol.1

[3] Win, M.Z. and R.A. Scholtz, 'Comparisons of analogue and digital impulse radio for Wireless multiple access communications', Proceedings of the 1997 IEEE International Conference on Communications, Montreal, Canada, pp.91-95

[4] R. C. Qiu, "Propagation effects," in UWB Communication Systems: A Comprehensive Overview, M. G. Di Benedetto, Ed. Lausanne, Switzerland: EURASIP, 2005

[5] Andreas F. Molisch, etc. "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," IEEE Transactions on Antennas and Propagation, vol.54, No.11, pp 3151-3166, November 2006

[6] Saleh, A.; Valenzuela, R., "A Statistical Model for Indoor Multipath Propagation", IEEE Journal on Selected Areas in Communications, Volumn5, Issue 2, Feb 1987 Page(s):128-137

[7] M. Steinbauer and A. F. Molisch, Directional Channel Modelling, Wiley, 2001.

[8] Andreas F. Molisch, etc. "IEEE 802.15.4a channel model-final report"

[9] M. Huang, P.E. Caines, C.D. Charalambous, and R. Malhame, "Power control in wireless systems: A stochastic control formulation", Proceedings of the American Control Conference, pp. 750-755, Arlington, VA, USA, 2001.

[10] J. Zander, "Performance of optimum transmitter power control in cellular radio systems", IEEE Trans. on Vehicular Tech., vol. 41, no.1, pp. 57-62, Feb. 1992.

[11] J. Aein, "Power balancing in systems employing frequency reuse", COMSAT Technical Review, vol. 3, 1973.

[12] N. Bambos and S. Kandukuri, "Power-controlled multiple access schemes for next-generation wireless packet networks", IEEE Wireless Communications, vol. 9, no. 3, pp. 58-64, June 2002

[13] G.J. Foschini and Z. Miljanic, “A simple distributed autonomous power control algorithm and its convergence”, IEEE Trans. on Vehicular Tech., vol. 42, no.4, pp. 641-646, Nov. 1993.

[14] S.A. Grandhi, J. Zander and R. Yates, “Constrained power control”, International Journal of Wireless Personal Communications, vol. 1, no. 4, April 1995.

[15] X. Li and Z. Gajic, “An improved SIR-based power control for CDMA systems using Steffensen iterations”, Proceeding of the Conference on Information Science and Systems, pp. 287-290, Princeton NJ, Mar. 2002.

[16] S. Kandukuri and S. Boyd, “Optimal power control in interference-limited fading wireless channels with outage-probability specifications”, IEEE Transactions on Wireless Communications, vol. 1, no. 1, pp. 46-55, 2002.

[17] A.M. Viterbi and A.J. Viterbi, “Erlang capacity of a power-controlled CDMA system”, IEEE J. Select. Areas Commun., vol. 11, pp. 892-900, Sep. 1993.

[18] S. Ulukus and R. Yates, “Stochastic power control for cellular radio systems”, IEEE Transactions on Communication, vol. 46, pp. 784-798, 1998.

[19] M.K. Varanasi and D. Das, “Fast stochastic power control algorithms for nonlinear multiuser receivers”, *IEEE Trans. Com.*, vol. 50, no. 11, pp. 1817-1827, Nov. 2002.

[20] N.B. Mandayam, P.C. Chen, and J.M. Holtzman, “Minimum duration outage for cellular systems: A level crossing analysis”, in Proc. 46th IEEE Conf. Vehicular Technol., pp. 879-883, Atlanta, GA, Apr. 1996.

[21] R. Buche and H.J. Kushner, “Control of mobile communications with time-varying channels in heavy traffic”, IEEE Trans. Automat. Contr., vol. 47, pp. 992-1003, June 2002.

[22] M. Huang, P.E. Caines, and R.P. Malhame, “Uplink power adjustment in wireless communication systems: A stochastic control analysis”, IEEE Trans. Automat. Contr., vol. 49, no. 10, pp. 1693-1708, Oct. 2004.

[23] J. F. Chamberland and V.V. Veeravalli, “Decentralized dynamic power control for cellular CDMA systems”, IEEE Trans. Wireless Commun., vol. 2, pp. 549-559, May 2003.

[24] L. Song, N.B. Mandayam, and Z. Gajic, “Analysis of an up/down power control algorithm for the CDMA reverse link under fading”, IEEE J. Select. Areas Commun., vol. 19, pp. 277-286, Feb. 2001.

[25] C.W. Sung and W.S. Wong, “Performance of a cooperative algorithm for power control in cellular systems with a time-varying link gain matrix”, Wireless Networks, vol. 6, no. 6, pp. 429-439, 2000.

[26] T. Holliday, A.J. Goldsmith, and P. Glynn, “Distributed power control for time-varying wireless networks: Optimality and convergence”, Proceedings of Allerton Conference on Communications, Control, and Computing, Oct. 2003.

[27] A. F. Molisch, Wireless Communications. New York: IEEE Press/ Wiley, 2005.

[28] D. Cassioli, M. Z. Win, and A. F. Molisch “The ultra-wide bandwidth indoor channel: From statistical model to simulations,” IEEE J. Sel. Areas Commun., vol. 20, pp. 1247-1257, Aug 2002.

[29] J. Kunisch and J. Pamp, “Measurement results and modeling aspects for the UWB radio channel,” in Proc .IEEE UWBST, pp.19-23, 2002.

[30] J.G. Proakis, Digital communications, Fourth Edition, McGraw Hill, 2000.

[31] E.Wong and B.Hajek. Stochastic Processes in Engineering Systems. Springer-Verlag, New York, 1985

[32] D.C.Youla, on the factorization of rational matrices. IRE Transactions on Information Theory, IT-7:172-189, July 1961.

[33] M.C.Davis, Factoring the spectral matrix. IEEE Transactions on AC, AC-8(4): 296-305, October 1963.

[34] T. Aulin, “A modified model for fading signal at a mobile radio channel,” IEEE Trans. Veh. Technol., vol. 28, no. 3, pp. 182-203, 1979.

[35] T.S.Rappaport, “Wireless communications: Principles and practice”, Prentice Hall, 2nd Edition, 2002.

[36] P.L.Rice, Mathematical analysis of random noise. Bell Systems Technical Journal, 24: 46-156, 1944.

[37] G.L.Turin, F.D.Clapp, T.L.Johnston, S.B.Fine, and D.Lavry. A statistical model of urban multipath propagation. IEEE Transactions on Vehicular Technology, VT-21(6): 1-9, 1972

[38] H.Suzuki, A statistical model for urban radio propagation. IEEE Transactions on Communications, 25:673-680, 1977

[39] C.D. Charalambous and A. Logothetis, “Maximum-likelihood parameter estimation from incomplete data via the sensitivity equations: The continuous-time case”, IEEE Transaction on Automatic Control, vol. 45, no. 5, pp. 928-934, May 2000.

[40] G. Bishop and G. Welch, An introduction to the Kalman filters, University of North Carolina, 2001.

[41] Mohammed M. Olama, Seddik M. Djouadi, and Charalambos D. Charalambous, "Stochastic power control for time-varying long-term fading wireless networks", EURASIP Journal on Applied Signal Processing, vol. 2006, Article ID 89864, 13 pages, 2006.

[42] M.M. Olama, S.M. Shajaat, S.M. Djouadi, and C.D. Charalambous, "Stochastic power control for time-varying long-term fading wireless channels", Proceedings of the American Control Conference, pp. 1817-1822, June 8-10, 2005.

[43] M.M. Olama, S.M. Djouadi, and C.D. Charalambous, "Stochastic power control for time-varying wireless fading communication networks", Proceedings of the 13th IEEE Mediterranean Conference on Control and Automation, Limassol, Cyprus, pp. 713-718, June 27-29, 2005.

[44] M.M. Olama, S.M. Shajaat, S.M. Djouadi, and C.D. Charalambous, "Stochastic Power Control for Time-Varying Short-Term Flat Fading Wireless Channels," Proceedings of the 16th IFAC World Congress held in Prague, July 4-8, 2005.

[45] C.D. Charalambous, S.M. Djouadi, and S.Z. Denic, "Stochastic power control for wireless networks via SDE's: Probabilistic QoS measures", IEEE Trans. on Information Theory, vol. 51, no. 2, pp. 4396-4401, Dec. 2005.

[46] H.Cramer and M.R.Leadbetter, Stationary and Related Stochastic Processes. New York: Wiley, 1967.

[47] A.I. Mees, Nonlinear Dynamics and Statistics, Birkhauser, Boston, 2001.

[48] H. F. Chen, Stochastic approximation and its applications, Kluwer Acad. Publishers, 2002.

[49] J. Luo, S. Ulukus and A. Ephremides, "Standard and quasi-standard stochastic power control algorithms", IEEE Trans. on Information Theory, vol. 51, no. 7, pp. 2612-2624, July 2005.

[50] B. Bolloba's, linear analysis (Cambridge Mathematical textbooks, 1992).

[51] M. Ghavami, L. Michael, R. Kohno, Ultra Wideband Signals and Systems in Communication Engineering, 2007

[52] W. J. Rugh, Linear System Theory, Prentice Hall 1996

[53] S.M. Djouadi, R. Chris Camphouse and James H. Myatt, “Optimal order reduction for the two-dimensional Burgers' Equation, IEEE Conference on Control and Decision, pp. 3507-3512, Dec. 2007

[54] E. Parzen, “Stochastic Process”, Society for Industrial Mathematics, 1962.

[55] Donald L. Snyder, “Random Point Processes”, John Wiley & Sons Inc, 1975.

[56] Harald Cramer, “Stationary and Related Stochastic Process”, Dover Publications, 2004.

[57] M.M. Olama, Y. Li, S.M. Djouadi, and C.D. Charalambous, “Time Varying Wireless Channel Modeling, Estimation, Identification, and Power Control from Measurements”, Proceedings of the American Control Conference, 2007.

[58] R. D. Yates, “A Framework for Uplink Power Control in Cellular Radio Systems”, IEEE Journal on selected areas in communications, vol.13, No.7, September 1995.

[59] J. Zhan, “Web-Based Radio Fading Channel Modeling, Estimation and Identification”, University of Ottawa, master’s thesis, 2003

[60] L. Ljung, “System Identification: Theory for User (2nd Edition)”, Prentice Hall PTR, 1999

[61] C. D. Charalambous, J. zhan, X. Li and R. J. C.Bultitude. “Radio Channel Modeling, Estimation and Identification from Measurement Data”, European Cooperation in the field of science and technical research, 2003

[62] E. Brookkner, “Tracking and Kalman Filtering Made Easy”, John Wiley & Sons, 1998

[63] M. H. A. Davis, R. B. Vinter, “Stochastic Modeling and Control”, Chapman and Hall, 1985

[64] G. Casella, R. L. Berger, “Statistical Inference”, Duxbury Press, 2 edition,2001

[65] M. M. Olama, S. M. Djouadi, C. D. Charalambous, “A General Framework for Continuous Time Power Control in Time Varying Long Term Fading Wireless Networks”, Proceedings of the Ninth Iasted International Conference, May 30-June1, 2007, Montreal, Quebec, Canada.

[66] N. Menemenlis, “Stochastic Models for Multipath Fading Channels”, McGill University, Doctoral thesis, 2002

[67] D. R. Cox, P. A. W. Lewis, “The Statistical Analysis of Series of Events”, Monographs on Applied Probability and Statistics, London: Chapman and Hall, 1966

[68] C. R. Rao, Linear Statistical Inference and Its Applications, Wiley, New York, 1965

[69] S. S. Wilks, Mathematical Statistics, Wiley, New York, 1962

[70] R. J. Elliott, V. Krishnamurthy, "New Finite-Dimensional Filters for Parameter Estimation of Discrete-Time Linear Gaussian Models", IEEE Transactions on Automatic Control, Vol. 44, No.5, May,1999

[71] C. F. Jeff Wu, "On the Convergence Properties of the EM Algorithm", the Annals of Statistics, Vol. 11, No. 1, Mar, 1983, pp. 95-103

[72] A. Pinkus, "n-Widths in Approximation Theory", Springer-Verlag, 1985.

[73] Schatten R. "Norm Ideals of Completely Continuous Operators", Springer-Verlag, Berlin, Gottingen, Heidelberg, 1960.

[74] F. Riesz and B.Sz.-Nagy, "Functional Analysis", Dover, 1990.

[75] L. Lin, "Fast identification for robust adaptive control: a metric complexity approach", McGill University, PhD thesis, 1993

[76] C. D. Charalambous, "Statistical Analysis of Multipath Fading Channels using Generalizations of Shot-Noise", preprint

[77] R. E. Megginson, "An Introduction to Banach Space Theory", Springer, 1998

[78] M. R. Mahfouz, C. Zhang, B. Merkl, M. Kuhn, A. E. Fathy, "Investigation of High-Accuracy Indoor 3-D Positioning Using UWB Technology", IEEE Transactions on Microwave Theory and Techniques, Volume 56, Issue 6, June 2008 Page(s): 1316-1330

[79] C. Zhang, A. E. Fathy, "Reconfigurable Pico-Pulse Generator for UWB Applications", IEEE MTT-S International Microwave Symposium Digest, June 2006

[80] M. K. Ozdemir, H. Arslan, "Channel Estimation for Wireless OFDM Systems", IEEE Communications Surveys, 2nd Quarter 2007, Volume 9, No.2

[81] R.Kolic, "An introducetion to Ultra Wideband (UWB) wirelesss", Intel White Paper, Feb 24, 2004.

[82] Federal Communications Commission, "First report 02-48," 2002.

[83] A. Stamoulis, S.N. Diggavi and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," IEEE Trans. on Signal Processing, vol. 50, No. 10, pp. 2451 – 2464, Oct. 2002.

[84] K. Baddour and N.C. Beaulieu, "Autoregressive Modeling for Fading Channel Simulation," IEEE Trans. On Wireless Communication, vol. 4, No. 4, pp. 1650-1662, July 2005.

APPENDIX

Appendix A

Proof of Theorem 2.2:

Let $L^2\left([0,\infty)\times[0,\infty)\right)$ be the Hilbert space of Lebesgue measurable and square integrable complex valued functions defined on $[0,\infty)\times[0,\infty)$ with the following mean square norm

$$\|f\|_2^2 := \iint_{[0,\infty)\times[0,\infty)} |f(t;\tau)|^2 d\tau dt < \infty, \quad f(t;\tau)\in L^2([0,\infty)\times[0,\infty)) \tag{A1}$$

Likewise define $L^2\left([0,\infty)\right)$ as the standard Hilbert space of Lebesgue measurable and square integrable complex valued functions defined on $[0,\infty)$ under the norm

$$\|x\|_2^2 := \int_0^\infty |x(t)|^2 dt, \quad x\in L^2\left([0,\infty)\right) \tag{A2}$$

The space $L^2\left([0,\infty)\right)$ contains all finite energy signals defined $\text{on}[0,\infty)$.

The impulse response $C_{IR}(t;\tau)$ of the channel has finite energy in time and space and therefore belongs to $L^2\left([0,\infty)\times[0,\infty)\right)$, that is,

$$\|C_{IR}\|_2^2 := \iint_{[0,\infty)[0,\infty)} |C_{IR}(t;\tau)|^2 d\tau dt < \infty \tag{A3}$$

For fixed n, define the shortest distance minimization in the $\|\bullet\|_2$-norm from the impulse response $C_{IR}(t;\tau)$ to the subspace S, by

$$\mu \triangleq \inf_{s\in S} \|C_{IR}(t;\tau) - s(t;\tau)\|_2 \tag{A4}$$

where the subspace S is defined as

$$S \triangleq \left\{ \sum_{i=1}^{n} \alpha_i(t)\varphi_i(\tau) : \alpha_i(t)\in L^2\left([0,\infty)\right),\ \varphi_i(t)\in L^2\left([0,\infty)\right);\ \forall n \text{ integer} \right\} \tag{A5}$$

Note that the distance minimization problem (A4) is posed in an infinite dimensional space.

Since the transmitted and received signals are finite energy signals, the impulse response can be viewed as an integral operator mapping transmitted signals in $L^2([0,\infty))$ into $L^2([0,\infty))$, i.e., if $s_l \in L^2([0,\infty))$ then

$$\begin{aligned} y_l(t) &= \int_0^\infty C_{IR}(t;\tau)\, s_l(t-\tau)\, d\tau \\ &= \sum_{j=1}^{J(t)} \left(I_j(t,\tau)\cos(\omega_c t) - Q_j(t,\tau)\sin(\omega_c t) \right) s_l\left(t-\tau_j(t)\right) \in L^2([0,\infty)) \end{aligned} \tag{A6}$$

It is known that such an operator is compact [72], that is, an operator which maps bounded sets into pre-compact sets. The operator T is said to be a Hilbert-Schmidt or a trace class 2 operators [73]. Let us denote the class of Hilbert-Schmidt operators acting from $L^2([0,\infty))$ into $L^2([0,\infty))$, by C_2 and the Hilbert-Schmidt norm $\|\bullet\|_{HS}$. For the operator T

$$\|T\|_{HS} = \sqrt{\iint_{[0,\infty)[0,\infty)} |C_{IR}(t;\tau)|^2 \, d\tau dt} \tag{A7}$$

Define the adjoint of T^* as the operator acting from $L^2([0,\infty))$ into $L^2([0,\infty))$ by

$$\begin{aligned} < Tf, g >_2 &\triangleq \int_0^\infty \int_0^\infty C_{IR}(t;\tau)\, f(\tau)\, d\tau\, g(t)\, dt \\ &= \int_0^\infty f(\tau) \int_0^\infty C_{IR}(t;\tau)\, g(t)\, dt\, d\tau \\ &\triangleq < f, T^* g >_1 \end{aligned} \tag{A8}$$

showing that

$$(T^* g)(\tau) = \int_0^\infty C_{IR}(t;\tau)\, g(t) dt \tag{A9}$$

Using the polar representation of compact operators [73] $T = U(T^*T)^{1/2}$, where U is a partial isometry and $(T^*T)^{1/2}$ is the square root of T, which is also a Hilbert-Schmidt operator, and admits a spectral factorization of the form [73]

$$(T^*T)^{1/2} = \sum_i \lambda_i \nu_i \otimes \nu_i \tag{A10}$$

where $\lambda_i > 0$ and $\lambda_i \downarrow 0$ as $i \uparrow \infty$ are the eigenvalues of $(T^*T)^{1/2}$, and ν_i form the corresponding orthonormal sequence of eigenvectors, i.e., $(T^*T)^{1/2}\nu_i = \lambda_i \nu_i$, $i = 1, 2, \cdots$. Putting $U\nu_i = \psi_i$, we can write

$$T = \sum_i \lambda_i \nu_i \otimes \psi_i \tag{A11}$$

Both $\{\nu_i\}$ and $\{\psi_i\}$ are orthonormal sequences in $L^2([0,\infty))$. The sum (A11) has either a finite or countable infinite number of terms. The above representation is unique.

Noting that the polar decomposition of $T^* = U^*(TT^*)^{1/2}$, a similar argument yields

$$(TT^*)^{1/2} = \sum_i \lambda_i \psi_i \otimes \psi_i \tag{A12}$$

and

$$T^* = \sum_i \lambda_i \psi_i \otimes \nu_i \tag{A13}$$

which shows that $\{\psi_i\}$ from an orthonormal sequence of eigenvectors of $(TT^*)^{1/2}$ corresponding to the eigenvalues λ_i. From (A10) and (A13) it follows that

$$T\psi_i = U(T^*T)^{1/2}\psi_i = \lambda_i \nu_i \tag{A14}$$

and

$$T^*\nu_i = U^*(TT^*)^{1/2}\nu_i = \lambda_i \psi_i \tag{A15}$$

We say that ψ_i and ν_i constitute a Schmidt pair [74]. In terms of integral operators expressions, identities (A14) and (A15) can be written, respectively, as

$$\nu_i(t) = \int_0^\infty C_{IR}(t;\tau)\psi_i(\tau)d\tau \tag{A16}$$

and

$$\psi_i(t) = \int_0^\infty C_{IR}(t;\tau)\nu_i(t)dt \tag{A17}$$

In terms of the eigenvalues λ_i's of T , the Hilbert-Schmidt norm $\|\bullet\|_{HS}$ is given by [73]

$$\|T\|_{HS} = \sqrt{\sum_i \lambda_i^2} = \sqrt{\iint_{[0,\infty)[0,\infty)} |C_{IR}(t;\tau)|^2 d\tau dt} \tag{A18}$$

Note that since the operator T is Hilbert-Schmidt the sum in (A18) is finite. The Hilbert-Schmidt norm is also induced by the operator inner product defined below.

By interpreting each elements of the subspace S defined in (A5) as a Hilbert-Schmidt operator as we did for $C_{IR}(t;\tau)$, we see that S is the subspace of Hilbert-Schmidt operators of rank n , i.e.,

$$S = \left\{ s = \sum_{i=1}^n \vartheta_j f_j(t) \otimes \chi_j(\tau) : f_j(t) \in L^2([0,\infty)), \chi_j(x) \in L^2([0,\infty)), \vartheta_j \in \mathbb{R} \right\} \tag{A19}$$

In addition, the distance minimization (A4) is then the minimal distance from T to Hilbert-Schmidt operators of rank n . In other terms, we have

$$\mu = \min_{s \in S} \| T - s \|_{HS} \tag{A20}$$

The space of Hilbert-Schmidt operators is in fact a Hilbert space with the inner product [73], denoted $(\cdot,\cdot)$, if A and B are two Hilbert-Schmidt operators defined on $L^2([0,\infty))$,

$$(A,B) \triangleq tr(B^* A) \tag{A21}$$

where tr denotes the trace, which in this case is given by the sum of the eigenvalues of the operator B^*A which is necessarily finite [73]. Note that the inner product (A21) induces the Hilbert-Schmidt norm $\| A \|_{HS} = \left(tr(A^* A) \right)^{\frac{1}{2}}$. In the case where A and B are integral operators with kernels $A(t,\tau)$ and $B(t,\tau)$, respectively, the inner product can be realized concretely by

$$(A,B) = \int_0^\infty \int_0^\infty A(t,\tau)\, B(t,\tau)\, dtd\tau \tag{A22}$$

The solution to the distance minimization is simply given by the orthogonal projection of T onto S. To compute the latter, note that the eigenvectors of $(TT^*)^{\frac{1}{2}}$ and $(T^*T)^{\frac{1}{2}}$ form orthonormal bases (by completing them if necessary) for $L^2\left([0,\infty)\right)$ and $L^2\left([0,\infty)\right)$, respectively. In terms of the eigenvectors ν_j and ψ_j the subspace S can be written as

$$S = Span\{\nu_j \otimes \psi_j, j = 1,2,\cdots,n\} \tag{A23}$$

Since the shortest distance minimization (A20) is posed in a Hilbert space, by the principle of orthogonally it is solved by the orthogonal projection P_S acting from C_2 onto S. The latter can be computed by first determining the orthogonal projection P_ν onto $Span\{\nu_j, j = 1,2,\cdots,n\}$, and the orthogonal projection P_ψ onto $Span\{\psi_j, j = 1,2,\cdots,n\}$. These projections have finite rank and since ν_j's and ψ_j's are orthogonal vectors in $L^2\left([0,\infty)\right)$ and $L^2\left([0,\infty)\right)$, respectively, it can be easily verified that P_ν and P_ψ are give by

$$(P_\nu f)(t) = \sum_{j=1}^{n} < f, \nu_j >_1 \nu_j(t) = \sum_{j=1}^{n} (\int_0^\infty f(t)\nu_j(t)dt)\nu_j(t) \tag{A24}$$

and

$$(P_\psi G)(\tau) = \sum_{j=1}^{n} < G, \psi_j >_2 \psi_j = \sum_{j=1}^{n} (\int_0^\infty G(\tau)\psi_j(\tau)d\tau)\psi_j(\tau) \tag{A25}$$

The overall orthogonal projection P_S can be computed as

$$P_S = P_\nu \otimes P_\psi \tag{A26}$$

That is, if $W \in C_2$ has spectral decomposition $\sum_{i=1} \eta_i u_i \otimes \upsilon_i$, where $u_i \in L^2([0,\infty))$, $\upsilon_i \in L^2([0,\infty))$, then

$$P_S W = \sum_{i=1} \eta_i P_S(u_i \otimes \upsilon_i) \tag{A27}$$

$$= \sum_{i=1} \eta_i \left(\left(\sum_{j=1}^{n} \int_0^\infty u_i(t)\nu_j(t)dt \right) \nu_j \otimes \left(\int_0^\infty \upsilon_i(\tau)\psi_j(\tau)d\tau \right) \psi_j \right) \tag{A28}$$

$$= \sum_{j=1}^{n} \theta_j \nu_j \otimes \psi_j, \ \exists \text{ scalars } \theta_j \tag{A29}$$

where the last finite sum is obtained thanks to orthogonality, i.e., only the u_i's and υ_i's that live in the span of ν_j's and ψ_j's, respectively, are retained. For the orthogonality property we only need verify that

$$x \otimes y - (P_\nu \otimes P_\psi)(x \otimes y) \perp u \otimes \upsilon \tag{A30}$$

$$x \in L^2([0,\infty)), \ y \in L^2([0,\infty)), \ u \otimes \upsilon \in S \tag{A31}$$

Computing the inner product, we get

$$< x - P_\nu x, u >_1 < y - P_\psi, \upsilon >_2 = 0 \tag{A32}$$

Because P_ν is the orthogonal projection of $L^2([0,\infty))$ onto $Span\{\nu_j, j=1,2,\cdots,n\}$, and P_ψ the orthogonal projection of $L^2([0,\infty))$ onto $Span\{\psi_j, j=1,2,\cdots,n\}$. The minimizing operator $s_o \in S$ in (A20) is then given by

$$s_o \triangleq P_S T = \sum_{i=1}^{n} \lambda_i \nu_i \otimes \psi_i \tag{A33}$$

and

$$\mu = \min_{s\in S} \| w(t,\tau) - s(t,\tau) \|_2 = \| T - P_S T \|_{HS} = (\sum_{i=n+1}^{n} \lambda_i^2)^{\frac{1}{2}} \tag{A34}$$

and as $n \uparrow \infty$, $\| T - P_S T \|_{HS} \searrow 0$. Therefore, the minimizing function $s_o(t,\tau)$ in (A4) corresponds to the kernel of s_o, which is given by

$$s_o(t,\tau) = \sum_{i=1}^{n} \lambda_i \nu_i(t) \psi_i(\tau) \tag{A37}$$

This implies that in the $\|\bullet\|_2$ -norm $C_{IR}(t;\tau)$ can be approximated to any desired accuracy by an impulse response of the form $\sum_{i=1}^{n} \lambda_i \nu_i(t) \psi_i(\tau)$ which is factorizable by putting

$$g(t) := [\lambda_1 \nu_1(t) \quad \lambda_2 \nu_2(t) \quad \cdots \quad \lambda_n \nu_n(t)] \text{ and } f(t) := [\psi_1(\tau) \quad \psi_2(\tau) \quad \cdots \quad \psi_n(\tau)]^T \tag{A38}$$

where T denotes the transpose .

Appendix B

Introduction of Banach Space [77]:

Definition B.1: A vector space or linear space over $\mathbb{F}$ (either $\mathbb{R}$ or $\mathbb{C}$) is a set X of objects called vectors along with an operation $+$ from $X \times X$ into X called addition of vectors and an operation from $\mathbb{F} \times X$ into X called multiplication of vectors by scalars satisfying these conditions:

(1) addition is commutative and associative ;
(2) there is a zero vector 0 in X, sometimes called the origin of X, such that $x + 0 = x$ for each vector x;
(3) for each vector x there is a vector $-x$ such that $x + (-x) = 0$;
(4) for all scalars α and β and all vectors x and y, $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$, $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$, and $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$
(5) for each vector x, $1 \cdot x = x$

Definition B.2: Let X be a vector space. A norm on X is a real-valued function $\|\cdot\|$ on X such that the following conditions are satisfied by all members x and y of X and each scalar α:

(1) $\|x\| \geq 0$, and $\|x\| = 0$ if and only if $x = 0$;
(2) $\|\alpha x\| = |\alpha| \|x\|$;
(3) $\|x + y\| \leq \|x\| + \|y\|$

The ordered pair $(X, \|\cdot\|)$ is called a normed space or normed vector space or normed linear space.

Definition B.3: A Banach norm or complete norm is a norm that induces a complete metric. A normed space is a Banach space or complete normed space if its norm is a Banach norm.

Example B.1: Let c_0 be the collection of all sequences of scalars that converge to 0, with the same vector space operations and norm as l_∞. Then c_0 is a Banach space since it is a closed subspace of l_∞.

Vita

Yanyan Li was born in China on September 27, 1984. She graduated from high school in 2001. From there, she went to Beijing University of Posts and Telecommunications and received her Bachelor degree in Electrical Engineering in 2005.

She is currently pursuing her doctorate in Electrical Engineering at the University of Tennessee, Knoxville.