Channel Prediction Based on Sinusoidal Modeling

MING CHEN

Department of Signals and Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2005

Abstract

Long range channel prediction is considered as one of the most important enabling technologies to future wireless communication systems. The prediction of Rayleigh fading channels is studied in the frame of sinusoidal modeling in this thesis.

A stochastic sinusoidal model to represent Rayleigh fading channel is proposed. The average of the conditional power spectrum of this model is shown to be the well known Jake's model. Given Doppler frequencies to be deterministic, the Cramér-Rao Lower Bound (CRLB) for the frequency estimates is derived. An algorithm to calculate the compressed CRLB is also proposed.

Using measurement data, the Jake's model is confirmed by the Normalized Mean Doppler Spectrum (NMDS) in both urban and suburban environments. The analysis of the time varying property of the model parameters shows that the model parameters are more consistent in suburban than in urban environment. A strong dominant sinusoid was observed in most suburban measurements, which might be due to the direct path in Line-Of-Sight (LOS).

Based on the statistical sinusoidal modeling, three different predictors are proposed. These methods outperform the standard LP in Monte Carlo simulations, but underperform with real measurement data. A subjective study of the LMMSE prediction methods to nearby tones are performed by simulations. The Unconditioned LMMSE predictor is found to be more suitable for the prediction of closely separated sinusoids.

Later, a Joint Moving Average and Sinusoidal (JMAS) model is proposed for channel prediction, which predict the channel by LP and sinusoidal predictor jointly, together with a simple SVD based the i^{th} biggest gradient model selection method. This method is termed Joint LMMSE predictor. It outperforms all the other predictors in suburban environment, but still performs slightly worse than the standard LP in urban environments.

Keywords: Wireless communications, channel prediction, channel estimation, system identification, parameter estimation, channel modeling, CRLB, MIMO.

Contents

Acknowledgements

I became curious and anxious to become a P.D student when I was with Ericsson Research in Kista several years ago. This idea got support from my manager, Dr. Sören Andersson. I tried to take graduate courses and at the same time worked as a full time research engineer. But it turned out to be tough for me to perform well on both. To make the dream come true, I decided to come back to Chalmers as a full time Ph.D student in the early spring of 2003. Two and a half years later, when I finish writing this thesis, I would like to say that the study at Chalmers is so fruitful and enjoyable. To make the thesis as what it is now, some people are deserved to be acknowledged.

First, I appreciate my supervisor, Prof. Mats Viberg, for offering me the opportunity to study in the signal processing group. He is such a good teacher who can always make the abstract and complex concepts and theories easy to understand. The discussions with him are always heuristic, inspiring and fruitful. Without his help, the thesis will be far from finish.

I am also grateful to Dr. Torbjörn Ekman at UNIK in Oslo, for the measurement data and his valuable inputs and comments on our coauthored conference paper. I would like to sincerely thank Prof. Mikael Sternad in Uppsala University for his encouragements and all the other professors and Ph.D students in the Wireless IP project for the intriguing discussions during the project meetings.

I also own many thanks to the previous and current colleagues in the signal processing group for their helps at different time and aspects during my study. It is also these open minded, smart and knowledgeable persons, together with Prof. Mats Viberg, who make the working environment so friendly and enjoyable. Some names that I would like to mention are LicEng Joakim Gunnarsson, MSc Andy Backhouse, LicEng Mikael Coldrey, Dr. Lennart Svensson, LicEng Ingemar Andersson, and LicEng Maria Lanne.

I would like to express my deepest appreciation to my parents and my sister in China for their abundant love and support.

Finally and most importantly, my thanks go to Hong. Without her sincere love, understanding and support, my life is hard to be enjoyable.

Ming Chen
Göteborg, August 2005

Abbreviations and acronyms

ACF	Autocorrelation Function
ANMSE	Adjusted Normalized Mean Square Error
AWGN	Additive White Gaussian Noise
BF	Beam Forming
BS	Base Station
CDF	Cumulative Density Function
CRLB	Cramer-Rao Lower Bound
DFT	Discrete Fourier Transform
DOA	Direction-Of-Arrival
DOD	Direction-Of-Departure
DSR	Down Sampling Ratio
ESPRIT	Estimation of Signal Parameters via Rotation Invariance Techniques
EV	EigenVector
EVD	EigenValue Decomposition
FIR	Finite Impulse Response
IID	Independent Identical Distribution
IVM	Instrumental Variable Method
JMAS	Joint Moving Average and Sinusoidal
LMMSE	Linear Minimum Mean Square Error
LMS	Least Mean Square
LOS	Line-Of-Sight
LP	Linear Prediction
LS	Least Square
MA	Moving Average
MPS	MUSIC Pseudo Spectrum
MIMO	Multiple-In-Multiple-Out
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
MUSIC	MUltiple SIgnal Classification
NLOS	Non-Line-Of-Sight
NMDS	Normalized Mean Doppler Spectrum
NDS	Normalized Doppler Spectrum
NMSE	Normalized Mean Square Error

NSE	Normalized Square Error
OFDM	Orthogonal Frequency Division Multiplexing
OSF	Over Sampling Factor
PDF	Probability Density Function
RI	Rotation Invariant
RLS	Recursive Least Squares
RMSE	Root Mean Square Error
SIMO	Single-In-Multiple-Out
SISO	Single-In-Single-Out
SNR	Signal Noise Ratio
SVD	Singular Value Decomposition
TOA	Time-Of-Arrival

Notations

In this thesis, matrices and vectors are denoted on boldface. The upper-case letters are assigned to matrices, and lower-case is assigned to vectors. If there is no explicitly statement, the meaning of these notations are

$\mathbf{A}^T$	Transpose operation.
$\mathbf{A}^H$	Hermitian transpose.
$\bar{\mathbf{A}}$	Complex conjugation without transposition.
$\mathbf{A}^{1/2}$	Hermitian (*i.e.* $(\mathbf{A}^{1/2})^H = \mathbf{A}^{1/2}$) square root factor, *i.e.* $\mathbf{A} = \mathbf{A}^{1/2}\mathbf{A}^{1/2}$.
$\lvert\mathbf{A}\rvert$	Matrix determinant.
$[\mathbf{A}]_{ij}$ or a_{ij}	The $(i,j)^{\text{th}}$ element of the matrix $\mathbf{A}$.
$\lVert\mathbf{a}\rVert$	Euclidian norm of vector $\mathbf{a}$.
$\mathbf{I}_N$	The $N \times N$ identity matrix.
$\mathbf{X}(m:n, p:q)$	The submatrix of $\mathbf{X}$ contains the elements from m^{th} to n^{th} row and from p^{th} to q^{th} column.
$\mathbf{X}^\dagger$	Pseudo inverse of $\mathbf{X}$, *i.e.* $(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H$.
$\lvert x\rvert$	The modulus of the (possibly complex) scalar x.
x^*	The complex conjugate of x.
$\hat{x}$	An estimate of x.
$\mathbf{x} \sim \mathcal{CN}(\mu, \mathbf{C})$	The random vector $\mathbf{x}$ is distributed as complex normal random vector with mean μ and covariance matrix $\mathbf{C}$.
$f(\mathbf{x}\lvert\mathbf{y})$	The conditioned PDF of $\mathbf{x}$ given $\mathbf{y}$.
$E[\mathbf{x}\lvert\mathbf{y}]$	The conditioned expectation of $\mathbf{x}$ given $\mathbf{y}$.
j	The imaginary unit; $j^2 = -1$.
$\delta_{k,l}$	Kronecker delta, *i.e.* $\delta_{k,l} = 1$, if $k = l$ and $\delta_{k,l} = 0$ otherwise.
$\delta(t)$	The Dirac's delta function.
$\arg\min_x f(x)$	The minimizing argument of the function $f(x)$.
$E[\cdot]$	The expectation of a random variable or vector.
$Tr[A]$	Trace operation.
$\mathbf{A} \odot \mathbf{B}$	The Hadamard product.
$\mathbf{A} \otimes \mathbf{B}$	The Kronecker product.
$diag(\mathbf{A})$	The column vector formed from the diagonal elements of the square matrix $\mathbf{A}$.

$vec(\mathbf{A})$	The column vector obtained by stacking the columns of $\mathbf{A}$.
$\mathbf{Re}[\cdot]$	The real part of a variable, vector or matrix.
$\mathcal{F}\{\cdot\}$	The DFT operation.

Chapter 1

Introduction

This thesis deals with the problem of channel prediction in wireless communications in the frame of sinusoidal modeling.

1.1 Background

Radio signal propagation in wireless communications is much more complicated than in wired systems. This is mainly due to the multi-path propagation environment and the relative movement between the mobile terminal and the Base Station (BS) as in Fig. 1.1.

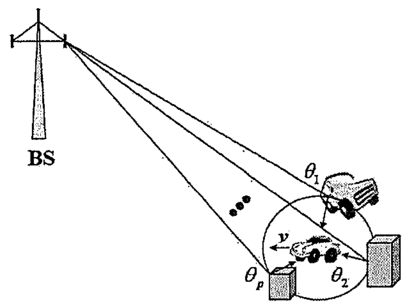

Figure 1.1: Example of multi-path propagation environment

Usually the angle between the impinging path and the velocity of the mobile terminal is termed Direction-Of-Arrival (DOA). The received signals via different paths suffer independent or correlated attenuations, phase shifts, frequency drifts and delays. Depending on the relative length of *excessive* delay and the symbol duration, the channels are divided into *narrow* band channel or *flat* fading channel, when the excessive delay is shorter than symbol duration, and *wide* band channel, or *frequency selective* channel otherwise. In practice, the number of paths could be large, so the complex channel coefficient can be modelled as a complex Gaussian random variable according to the *Central Limited Theorem*. Its envelope behaves as a random variable with Rayleigh distributed Probability Density Function (PDF). Such a channel is called a Rayleigh fading channel.

In a Rayleigh fading channel, there are deep minima over every a few portions of the wavelengths in space. This is due to the constructive and destructive addition of the impinging rays. Historically, the deep fading in Rayleigh fading channel is considered as one of the major difficulties for signal transmission in wireless communications. A number of techniques, such as diversity, equalization, and power control, were developed to mitigate its negative influence on the system performance. However, in recent research and development activities for future wireless communication systems, the technique of multi-user diversity combined with adaptive modulation attracts many research interests [Fal02, KH95]. These techniques exploit the instantaneous high SNR among the uncorrelated Rayleigh fading channels of different users in service. To enables these techniques, a good channel prediction is necessary.

In this thesis the problem of the prediction of a (narrow band) Rayleigh fading channel is studied in a frame of sinusoidal modeling. In general, we assume the availability of a vector $\mathbf{y} = [y(t), y(t-1), \cdots, y(t-N+1)]^T$ containing the N channel observations

$$\mathbf{y} = \mathbf{x} + \mathbf{e}, \tag{1.1}$$

where $\mathbf{x} = [x(t), x(t-1), \cdots, x(t-N+1)]^T$ is the *true* channels, and $\mathbf{e} = [e(t), e(t-1), \cdots, e(t-N+1)]^T$ is the additive estimation errors, which are assumed to be IID white Gaussian random variables. The value $x(t+L)$ is to be predicted from observations $\mathbf{y}$, where L is the prediction horizon. In practice, a prediction horizon corresponding to a distance of a half wave length travelled by the mobile is considered challenging, where λ is the wave length of the radio signal [SEA01].

1.2 Overview of Previous Works

During the last several years, many efforts have been made on algorithm design and performance evaluation of channel prediction methods [AJHF99, DHHH00, HW98, SK03, ESA02]. Most published predictors can be divided into two categories. One is the standard Linear Prediction (LP), and the other is deterministic sinusoidal modeling based channel prediction. In the first category, a d^{th} order LP of $x(t+L)$ is

$$\hat{x}(t+L) = \sum_{k=0}^{d-1} \beta_k y(t-k) = \boldsymbol{\beta}^H \mathbf{y}_d, \tag{1.2}$$

where $\mathbf{y}_d = [y(t), y(t-1), \cdots, y(t-d+1)]^T$, and $\boldsymbol{\beta} = [\beta_0, \cdots, \beta_{d-1}]^H$ is the coefficient vector of the LP. A large number of algorithms can be find in the

literature to estimate $\boldsymbol{\beta}$ with different optimization criteria.

In the second category, the channel is modelled as superimposed deterministic complex sinusoids, which correspond to the Doppler frequencies as in the Jake's model [Jak74],

$$y(t) = \sum_{i=1}^{p} \rho_i e^{j\phi_i} e^{j\omega_i t} + e(t), \tag{1.3}$$

where p is the number of sinusoids, ρ_i, ϕ_i and ω_i are the real amplitude, phase and the Doppler frequency associated with the i^{th} path respectively, and $e(t)$ is the additive complex Gaussian estimation error, $e(t) \sim \mathcal{CN}(0, \sigma_e^2)$. Assume there is no moving reflection objects, the i^{th} Doppler frequency is

$$\omega_i = \frac{2\pi v}{\lambda} \cos\theta_i, \tag{1.4}$$

where v is the velocity of the mobile terminal, θ_i is the i^{th} DOA. In this model, there are $(3p+1)$ parameters to be estimated. They are

$$\boldsymbol{\psi} = [\boldsymbol{\rho}^T, \boldsymbol{\phi}^T, \boldsymbol{\omega}^T, \sigma_e^2]^T, \tag{1.5}$$

where

$$\begin{aligned} \boldsymbol{\rho} &= [\rho_1, \cdots, \rho_p]^T, & (1.6)\\ \boldsymbol{\phi} &= [\phi_1, \cdots, \phi_p]^T, & (1.7)\\ \boldsymbol{\omega} &= [\omega_1, \cdots, \omega_p]^T. & (1.8)\\ & & (1.9) \end{aligned}$$

These model parameters are assumed to be stationary over the observation interval. In these methods, the first step is to compute parameter estimates of $\boldsymbol{\rho}$, $\boldsymbol{\phi}$, and $\boldsymbol{\omega}$. Then the prediction of $x(t+L)$ is

$$\hat{x}(t+L) = \sum_{i=1}^{p} \hat{\rho}_i e^{j\hat{\phi}_i} e^{j\hat{\omega}_i(t+L)}, \tag{1.10}$$

Let $s_i = \rho_i e^{j\phi_i}$ to be the complex amplitude, (1.3) can then be written as

$$y(t) = \sum_{i=1}^{p} s_i e^{j\omega_i t} + e(t), \tag{1.11}$$

and the prediction of $x(t+L)$ is

$$\hat{x}(t+L) = \sum_{i=1}^{p} \hat{s}_i e^{j\hat{\omega}_i(t+L)}. \tag{1.12}$$

A thorough performance evaluation of all these methods were made in [SK03], where these methods were tested by simulations and real world data. According to the reported results, the LP outperforms the deterministic sinusoidal modeling based methods. All these studies were performed in Single-In-Single-Out (SISO) systems. Later similar results were reported in Multiple-In-Multiple-Out (MIMO) scenarios in [SWS03]. It was also reported that the calculation complexity of LP increases exponentially with the increase of the dimensions of the MIMO systems. It makes LP costly for high dimension MIMO channel prediction. This is one of the motivations to pursue channel prediction methods based on sinusoidal modeling, where a less number of model parameters is expected than the LP in MIMO scenarios.

1.3 Contributions and Thesis Outline

Two major contributions are made in this thesis. The first contribution is a statistical sinusoidal modeling to Rayleigh fading channel and the three relative Linear Minimum Mean Square Error (LMMSE) prediction methods, which are the Conditioned LMMSE predictor, the Unconditioned LMMSE predictor, and the Combined LMMSE and LP predictor. The second contribution is a Joint Moving Average and Sinusoidal (JMAS) modeling to channel prediction, which leads to a Joint LMMSE predictor. Some of the contents in this thesis are contained in the following publications.

- M. Chen, and M. Viberg. LMMSE channel prediction based on sinusoidal modeling. In *Proc. of 3rd IEEE Sensor Array and Multichannel Signal Processing Workshop*, Barcelona, Spain, Jul. 2004. [CV04]

- M. Chen, T. Ekman, and M. Viberg, Two new approaches for channel prediction based on sinusoidal modeling. In *Proc. of IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, Jul. 2005.[CEV05]

A brief introduction of the content each chapter is given below.

Chapter 2 *Signal Models*

In this chapter, an introduction of *wide* band and *narrow* band channel modeling is given. A statistical sinusoidal modeling of the Rayleigh fading channel is proposed for SISO scenarios. The average of the conditional power spectrum, given frequencies, of this model is shown to be the Jake's model. The extension of this model to Single-In-Multiple-Out (SIMO) scenarios and MIMO scenarios are given.

Chapter 3 *Parameter Estimation Techniques*

Most parameter estimation techniques involved in the predictor design in this thesis are reviewed in this chapter. These techniques include:

- The Minimum Mean Square Error (MMSE) and the Least Square (LS) estimates of the coefficients of the standard LP;

- The Linear MMSE;

- Frequency estimation using the Unitary ESPRIT [PRK85, RPK86] and Instrumental Variable Method (IVM) [SVO94].

The Cramér-Rao Lower Bound (CRLB) of the frequency estimates of the statistical sinusoidal modeling is derived. But the bound does not exist when the frequencies are random. An algorithm to compute the compressed CRLB is proposed when the frequencies are close in some realizations.

Chapter 4 *Measurements and Data Analysis*

The measurement data is introduced. The Jake's model is confirmed by the Normalized Mean Doppler Spectrum (NMDS) of the measurements in both urban and suburban environments. An analysis of the stationary properties of the model parameters are made by using the MUSIC pseudo spectrum and parameter tracking techniques.

Chapter 5 *Model-Based LMMSE Predictors*

In this chapter, three LMMSE predictors based on sinusoidal modeling are proposed. They are the Conditioned LMMSE predictor, the Unconditioned LMMSE predictor, and the Combined LMMSE and LP predictor.

Chapter 6 *Performance Evaluation of Model-Based LMMSE Predictors*

In this chapter, two measures for the performance evaluation are defined, which are the Normalized Mean Square Error (NMSE) and the Adjusted NMSE (ANMSE). The performance of the model-based LMMSE predictors are evaluated by simulations and measurement data. A simple Singular Value Decomposition (SVD) based the i^{th} biggest gradient model selection method is proposed to the processing of the real data.

Chapter 7 *Joint LMMSE Predictors*

A Joint Moving Average and Sinusoidal (JMAS) approach to the prediction of Rayleigh fading channel is proposed. This new model leads to an LMMSE predictor, but computed as model-free, which is termed Joint LMMSE predictor. Its performance is evaluated using measurement data also.

Chapter 8 *Conclusions and Future Works*

This chapter contains the conclusions of the thesis. The comments and discussions on future works are given.

Chapter 2

Signal Models

This chapter starts with an introduction to channel modeling. A statistical sinusoidal model for the Rayleigh fading channel is proposed.

2.1 Physical Channel Models

Multi-path propagation is a common phenomenon in wireless communications. In such a scenario, the received signal can be modelled as the superimposition of a number of duplicates of the original transmitted signals. These signals suffer independent or correlated power attenuations, phase rotations, frequency shifts and delays. Numerous publications on channel modeling and channel characterizations can be found. Some overview on this topic can be found in [ECS+98, YO02]. The published channel models can be divided into *physical* and *non-physical* models, according to if detailed propagation parameters, such as Direction-Of-Arrival (DOA), Direction-Of-Departure (DOD), and Time-Of-Arrival (TOA) etc., are characterized. These parameters are necessary for many other topics, such as Beam Forming (BF), Transmit Diversity, and Multiuser Diversity etc. The disadvantage of this model is, in general, that it is difficult to fully describe the channel using a small number of parameters. The statistical properties might be less accurate. This category of channel models are mainly obtained by performing extensive measurements in different environments. The results are hard to validate. For instance, an effective or $false$ DOA might be observed in the data, while a clear reflection point cannot be found in reality using ray tracing methods [CA01].

The physical wave propagation over a local area can be described by an ellipsoidal model given in Fig. 2.1. In this figure, the transmitter and receiver stay on the two separated foci of a series of common-foci ellipsoids. Assume that the received signals through different paths experience no more than one reflection or diffraction. The shortest path is the direct path in Line-Of-Sight (LOS), which also has the smallest path loss. The paths from scatters on one ellipsoid has the same path length and delay, but different reflection attenuation. In general, the path has larger path loss than the LOS path. The inter-distance (resolution) of the ellipsoids is determined by the sampling frequency at the receiver. The path loss increases exponentially with the increase of the path length. So the outer ellipsoids could be ignored when the propagation paths are too long.

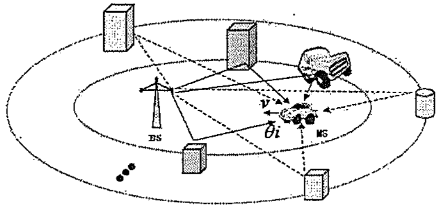

Figure 2.1: Physical ellipsoidal channel models.

2.2 Wide Band and Narrow Band Channel Models

Mathematically the channel impulse response can be modelled as an FIR filter, where each tap is corresponding to one reflection ellipsoid as shown in Fig. 2.1. Let $x(t,\tau)$ denote the channel impulse response at time t

$$x(t,\tau) \quad = \quad \sum_{n=0}^{k-1} x(t,\tau_n)\delta(\tau-\tau_n), \tag{2.1}$$

where τ is the excessive delay, k is the number of tap, $x(t,\tau_n)$ is the n^{th} taps of the FIR filter, which can be further modelled as

$$x(t,\tau_n) \quad = \quad \sum_{i=1}^{p_n} s_{n,i}e^{j\omega_{n,i}(t+\tau_n)}. \tag{2.2}$$

where p_n is the number of paths falling into the n^{th} tap, $s_{n,i}$ and $\omega_{n,i}$ are the amplitude and Doppler frequency associated with the i^{th} path in this tap. When $k>1$, (2.1) is called a wide band channel, or frequency selective fading channel. When $k=1$, this model can be simplified as

$$x(t) \quad = \quad \sum_{i=1}^{p} s_i e^{j\omega_i t}, \tag{2.3}$$

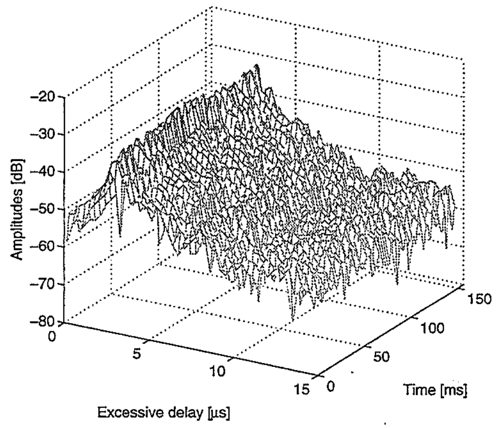

Figure 2.2: Example of a measured wide band

which is usually referred to as a narrow band channel, or flat fading channel. An example of a measured wide band channel is given in Fig. 2.2.

Due to the constructive and destructive addition of the Doppler frequencies, the envelope of the channel suffers deep minima over every a few portions of the wavelength. As mentioned before, when p is large, $x(t)$ can be modelled as a complex Gaussian random variable with independent real and imaginary parts. Its envelope $|x(t)|$ becomes a random variable with Rayleigh distribution as

$$f_{|x|}(|x|) = \frac{|x|}{\eta^2}e^{-|x|^2/2\eta^2}, \quad \eta > 0. \tag{2.4}$$

An example of a narrow band channel can be obtained from a single tap of the wide band channel. For instance, one power delay profile of the wide band channel in Fig. 2.2 is given in subplot (a) in Fig. 2.3. The envelope of the maximum tap is plotted as a function of time in subplot (b) in the same figure. The subplot (c) gives the Doppler spectrum of this tap, and the subplot (d) is the PDF of the channel envelopes (solid curve). A theoretical Rayleigh distribution PDF with $\eta = 5.5 \times 10^{-3}$ (dashed curve) is given also.

In general such a multi-path propagation structure is time-variant, but its variation is much slower than the variation of the envelope of the channel. This makes it possible to predict the channel using sinusoidal modeling.

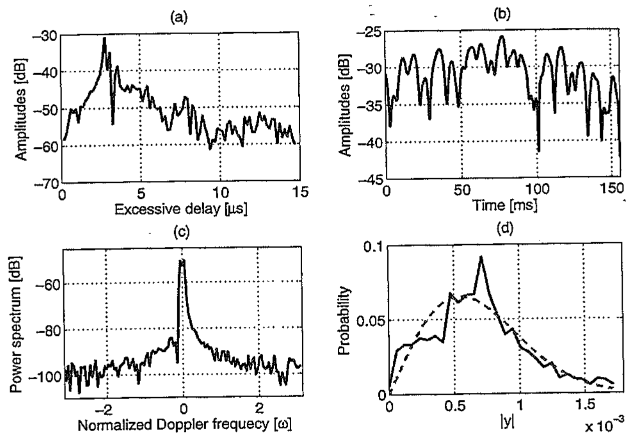

Figure 2.3: Example of a flat Rayleigh fading channel

2.3 The Jake's Model

The Jake's model is one of the most widely used models for a flat Rayleigh fading channel [Jak74]. It approaches the second order property of $x(t)$ by defining the *tub* shape power spectrum as in (2.5) [Rap96],

$$p(f) = \frac{1}{\pi f_m \sqrt{1-(\frac{f-f_c}{f_m})^2}}, \tag{2.5}$$

where f_m is the maximum Doppler frequency and f_c is the carrier frequency. To complete this definition, define $p(f) = 0$, when $|f - f_c| > f_m$. There are two assumptions made in this model:

1. A large number of uniformly distributed scattering objects surround the mobile terminal (in the far field);

2. The reflection signals from different scattering objects have equal power.

To show the relation between the Jake's model in (2.5) and the narrow band channel model in (2.3) explicitly, assume, without loss of generality, there is just a single path, with given DOA θ_0, in (2.3). The associated Doppler frequency ω_0 is

$$\omega_0 = \frac{2\pi v}{\lambda}\cos\theta_0. \tag{2.6}$$

In this case, the channel is a pure sinusoidal signal,

$$x_0(t) = s_0 e^{j\omega_0 t}, \tag{2.7}$$

where s_0 is the complex amplitude.

Then let θ_0 to be a random variable with uniform distribution, $\mathcal{U}[-\pi,\pi)$. It is easy to derive the PDF of the normalized Doppler frequency, which is

$$f(\omega_0) = \frac{1}{\pi\sqrt{1-\omega_0^2}}, -1 < \omega_0 < 1. \tag{2.8}$$

This PDF is identical (except a scaling factor) to the expression of the power spectrum given by the Jake's model (2.5). A plot of the PDF of ω_0 is given in Fig. 2.4, where the frequency is normalized.

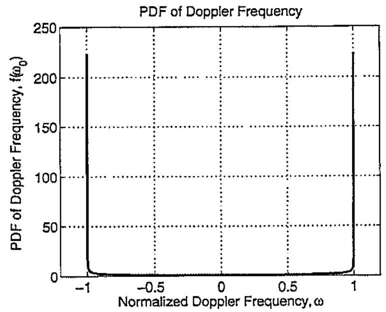

Figure 2.4: PDF of the normalized Doppler frequency

Assuming $s_0 = 1$, the conditional power spectrum of $x_0(t)$ is

$$p_{x_0}(\omega|\omega_0) = \delta(\omega - \omega_0), \tag{2.9}$$

where the scaling factor 2π is neglected. In this conditional spectrum, there is only one unit spectral line at $\omega = \omega_0$. This process is not ergodic. But averaged over many realizations, the unconditional spectrum gets the *tub* shape, which is

$$\begin{aligned} p_x(\omega) &= \int_{\omega_0} p_{x_0}(\omega|\omega_0) f(\omega_0) d\omega_0 \\ &= \int_{\omega_0} \delta(\omega - \omega_0) f(\omega_0) d\omega_0 \\ &= f(\omega_0), \end{aligned} \tag{2.10}$$

This discussion gives the following lemma.

Lemma 2.1

The average of the conditional *spectrum, given ω_0, is the Jake's spectrum.*

This lemma can be easily extended to multiple impinging rays with uniform DOA's and equal mean power. An averaged power spectrum of a single impinging ray with uniform DOA's from Monte Carlo simulations is given in Fig. 2.5.

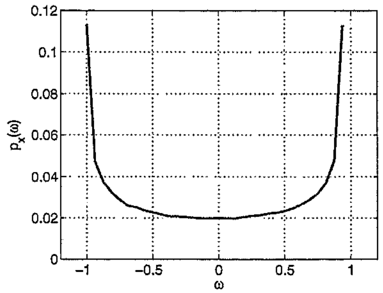

Figure 2.5: Average of the conditional power spectrum of a single impinging ray with uniformly distributed DOA's

2.4 Statistical Sinusoidal Modeling

Following the previous discussions, a flat Rayleigh fading channel can be modelled as

$$\begin{aligned} \mathbf{y} &= \mathbf{x}+\mathbf{e} \\ &= \mathbf{A}\mathbf{s}+\mathbf{e} \end{aligned} \tag{2.11}$$

where

$$\begin{aligned} \mathbf{y} &= [y(t), y(t-1), \cdots, y(t-N+1)]^T, & (2.12)\\ \mathbf{x} &= [x(t), x(t-1), \cdots, x(t-N+1)]^T, & (2.13)\\ \mathbf{e} &= [e(t), e(t-1), \cdots, e(t-N+1)]^T, & (2.14)\\ \mathbf{A} &= [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p], & (2.15)\\ \mathbf{a}_i &= [e^{j\omega_i t}, e^{j\omega_i(t-1)}, \cdots, e^{j\omega_i(t-N+1)}]^T, & (2.16)\\ \mathbf{s} &= [s_1, s_2, \cdots, s_p]^T, & (2.17) \end{aligned}$$

which are the channel observation vector $\mathbf{y}$, the true channel or the signal vector $\mathbf{x}$, and the estimation error or noise vector $\mathbf{e}$ respectively. The $\mathbf{A}$ matrix is a *Vandermonde* matrix and $\mathbf{a}_i$ is the DFT vector associated to the i^{th} frequency; $\mathbf{s}$ is the complex amplitude vector. The errors are assumed to be white Gaussian, $\mathcal{CN}(0, \delta_e^2 \mathbf{I}_N)$. The amplitudes are assumed to be independent and have equal mean power, $\mathbf{S} = E[\mathbf{s}\mathbf{s}^H] = \delta_s^2 \mathbf{I}_p$.

2.5 Extension to SIMO Scenarios

The signal model in (2.11) can be easily extended into a SIMO scenario with one transmit antenna and m receive antennas.

Assume that the DOA's $\boldsymbol{\theta}$ are identical at different antenna elements, but with independent amplitudes. In practice, inaccurate calibration of the array might give rise to such uncorrelated amplitudes. Let the amplitude matrix be

$$\mathbf{S}_{SIMO} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \cdots, \mathbf{s}_p^T]^T, \tag{2.18}$$

where $\mathbf{S}_{SIMO}$ is a $p \times m$ matrix, the m-vector $\mathbf{s}_i$ is one realization of the complex random amplitude vector associated to the i^{th} path. Assume these sinusoids have equal mean power σ_s^2, the power matrix could be written as $\mathbf{P} = E[\mathbf{S}_{SIMO}\mathbf{S}_{SIMO}^H] = m\sigma_s^2 \mathbf{I}_p$.

Then the statistical sinusoidal signal model for the SIMO system is

$$\mathbf{Y} = \mathbf{A}\mathbf{S}_{SIMO} + \mathbf{E}, \tag{2.19}$$

where

$$\begin{aligned} \mathbf{Y} &= [\mathbf{y}^T(t), \mathbf{y}^T(t-1), \cdots, \mathbf{y}^T(t-N+1)]^T, & (2.20)\\ \mathbf{E} &= [\mathbf{e}^T(t), \mathbf{e}^T(t-1), \cdots, \mathbf{e}^T(t-N+1)]^T, & (2.21) \end{aligned}$$

where the m-vector $\mathbf{y}(t)$ is the channel observations at the sensor array at time t, the additive noise vector $\mathbf{e}(t)$ is $1 \times m$ with covariance matrix $E\left[\mathbf{e}^H(k)\mathbf{e}(l)\right] = \sigma_e^2 \mathbf{I}_m \delta_{k,l}$. The $\mathbf{A}$ matrix is the same as in (2.15).

2.6 Extension to MIMO Scenarios

In this section, we extend the sinusoidal signal model to a MIMO system with n transmit antennas and m receive antennas. First let the SIMO channel, associated to the k^{th} transmit antenna, be

$$\mathbf{X}_k = \mathbf{A}\mathbf{S}_k, \tag{2.22}$$

where $\mathbf{A}$ and $\mathbf{S}_k$ are the same as in (2.19), except the subscripts. Vectorizing $\mathbf{X}_k$, we have

$$\begin{aligned} \mathbf{x}_k &= vec(\mathbf{X}_k) = vec(\mathbf{A}\mathbf{S}_k) \\ &= (\mathbf{I}_m \otimes \mathbf{A}) vec(\mathbf{S}_k), \\ &= (\mathbf{I}_m \otimes \mathbf{A}) \mathbf{s}_k, \end{aligned} \tag{2.23}$$

where $\otimes$ is the Kronecker product and $vec(\cdot)$ is the vectorization operation [Lüt96]. Stacking $\mathbf{x}_k$, the sinusoidal signal model for the MIMO channel is

$$\begin{aligned} \mathcal{X} &= [\mathrm{x}_1, \mathrm{x}_2, \cdots, \mathrm{x}_n], \\ &= [\mathrm{I}_m \otimes \mathrm{A}][s_1, s_2, \cdots, s_n], \\ &= \mathcal{AS} \end{aligned} \tag{2.24}$$

where $\mathcal{X}$ is $Nm \times n$ matrix and $\mathcal{S}$ is $pm \times n$ matrix. The power matrix is $\mathcal{P} = E\left[\mathcal{S}\mathcal{S}^H\right] = n\sigma_s^2\mathrm{I}_{pm}$, where we have assumed, again, that the amplitudes are independent with equal mean power σ_s^2. Assume the channel observations are white noise $\mathcal{E}$, the observed channel based on sinusoidal modeling for MIMO system becomes

$$\mathcal{Y} = \mathcal{AS} + \mathcal{E}, \tag{2.25}$$

where $\mathcal{E}$ is assumed to be Gaussian noise with PDF $\mathcal{CN}(\mathbf{0}, \sigma_e^2\mathrm{I}_n)$. Note that we have here assumed that all mn subchannels share the same Doppler frequencies. The extension to the general case is obvious, but introduces considerably more unknown parameters to be estimated.

Chapter 3

Parameter Estimation Techniques

Some parameter estimation techniques related to the predictor design in the thesis are discussed in this chapter. The MMSE and LS estimation techniques used in the LP are given in Section 3.1. The Linear MMSE estimation is introduced in Section 3.2. Frequency estimation using Unitary-ESPRIT is given in Section 3.3. The CRLB of the frequency estimates of the stochastic sinusoidal model for given frequencies is derived. An algorithm to calculate the compressed CRLB is proposed when frequencies are closely separated in some realizations.

3.1 MMSE and LS Estimation

Assume $y(t)$ is the observation of the signal $x(t)$ buried in white noise $e(t)$, which is independent of the signal,

$$y(t) = x(t) + e(t). \tag{3.1}$$

An L-step linear predictor of $x(t+L)$ with order d is

$$\begin{aligned} \hat{x}(t+L) &= \sum_{k=0}^{d-1} \beta_k y(t-k) \\ &= \boldsymbol{\beta}^H \mathbf{y}_d, \end{aligned} \tag{3.2}$$

where $\boldsymbol{\beta} = [\beta_0, \beta_1, \cdots, \beta_{d-1}]^H$ is the predictor coefficient vector, and $\mathbf{y}_d = [y(t), y(t-1), \cdots, y(t-d+1)]^T$.

3.1.1 MMSE Estimates

Define the Autocorrelation Functions (ACF) of $x(t)$, $e(t)$ and $y(t)$ as

$$r_{xx}(\tau) = E[x(t)x^*(t-\tau)], \tag{3.3}$$
$$r_{ee}(\tau) = E[e(t)e^*(t-\tau)] = \sigma_e^2\delta(\tau), \tag{3.4}$$
$$r_{yy}(\tau) = E[y(t)y^*(t-\tau)] = r_{xx}(\tau) + r_{ee}(\tau), \tag{3.5}$$

where τ is the lag index, $\delta(\tau)$ is the Dirac's delta function. These ACF's are assumed to be known. The LMMSE estimate of $x(t+L)$ is $\hat{\beta}^H_{MMSE}\mathbf{y}_d$, where

$$\hat{\beta}_{MMSE} = \arg\min_{\beta} E[\|x(t+L) - \beta^H \mathbf{y}_d\|^2], \tag{3.6}$$

where $E[\|x(t+L) - \beta^H \mathbf{y}_d\|^2]$ is the Mean Square Error (MSE). Setting the derivative of $E[\|x(t+L) - \beta^H \mathbf{y}_d\|^2]$ respect to β to be zero,

$$\begin{aligned}\frac{\partial E[\|x(t+L) - \beta^H \mathbf{y}_d\|^2]}{\partial \beta} &= -2E[\mathbf{y}_d(x^*(t+L) - \mathbf{y}_d^H \beta)] \\ &= -2E[\mathbf{y}_d x^*(t+L)] + 2E[\mathbf{y}_d \mathbf{y}_d^H]\beta \\ &= 0.\end{aligned}$$

We have

$$E[\mathbf{y}_d \mathbf{y}_d^H]\beta = E[\mathbf{y}_d x^*(t+L)],. \tag{3.7}$$

Since

$$\begin{aligned}E[\mathbf{y}_d y^*(t+L)] &= E[\mathbf{y}_d x^*(t+L)] + E[\mathbf{y}_d e^*(t+L)] \quad (3.8)\\ &= E[\mathbf{y}_d x^*(t+L)], \quad (3.9)\end{aligned}$$

where $E[\mathbf{y}_d e^*(t+L)] = 0$ for $L > 0$. Substitute (3.9) into (3.7),

$$\begin{aligned}E[\mathbf{y}_d \mathbf{y}_d^H]\beta &= E[y(t+L)\mathbf{y}_d^H], \\ \mathbf{R}_d \beta &= \mathbf{r}_d.\end{aligned} \tag{3.10}$$

where

$$\mathbf{R}_d = \begin{bmatrix} r_{yy}(0) & r_{yy}(1) & \cdots & r_{yy}(d-1) \\ r_{yy}(-1) & r_{yy}(0) & \cdots & r_{yy}(d-2) \\ \vdots & \vdots & \vdots & \ddots \\ r_{yy}(-d+1) & r_{yy}(-d+2) & \cdots & r_{yy}(0) \end{bmatrix}, \tag{3.11}$$

(3.12)

and

$$\mathbf{r}_d = \begin{bmatrix} r_{yy}(L) \\ r_{yy}(L+1) \\ \vdots \\ r_{yy}(L+d-1) \end{bmatrix}. \tag{3.13}$$

Then the MMSE estimate of $x(t+L)$ is obtained by

$$\hat{x}_{MMSE}(t+L) = \hat{\beta}^H_{MMSE}\mathbf{y}_d, \tag{3.14}$$

where

$$\hat{\beta}_{MMSE} = \mathbf{R}_d^{-1} r_d. \tag{3.15}$$

Since $\mathbf{R}_d$ is a positive definite matrix, the solution is unique and stable [SM97]. Note that when $L = 1$, (3.10) is the well known Yule-Walker equation [Yul86, Wal31].

In practice, $\mathbf{R}_d$ might not given directly. It can be estimated from the observations as

$$\hat{\mathbf{R}}_d = \frac{1}{N}\mathbf{Y}_d \mathbf{Y}_d^H, \tag{3.16}$$

where $\mathbf{Y}_d$ is a $d \times (N-d+1)$ Hankel matrix as

$$\mathbf{Y}_d = \begin{bmatrix} y(t) & y(t-1) & \cdots & y(t-N+d) \\ y(t-1) & y(t-2) & \cdots & y(t-N+d+1) \\ \vdots & \vdots & \ddots & \vdots \\ y(t-d+1) & y(t-d) & \cdots & y(t-N+1) \end{bmatrix}. \tag{3.17}$$

When N goes to infinite, it can be shown that $\hat{\mathbf{R}}_d \to \mathbf{R}_d$ in probability, *i.e.* $\hat{\mathbf{R}}_d$ is a consistent estimator.

3.1.2 LS Estimate

Instead of assuming that $y(t)$ is a stationary random variable with known ACF, the LS method deals with the observations as a deterministic sequence. By assuming the data to be zero outside the observation window, the linear dependence of the data in (3.2) can be written as

$$\left[\begin{array}{cccc} 0 & 0 & 0 & y(t) \\ \vdots & \vdots & \vdots & \vdots \\ y(t) & y(t-1) & \cdots & y(t-d+1) \\ \hline \vdots & \vdots & \vdots & \vdots \\ y(t-L) & y(t-L-1) & \cdots & y(t-L-d+1) \\ \vdots & \vdots & \vdots & \vdots \\ y(t-N+d) & y(t-N+d-1) & \cdots & y(t-N+1) \\ \hline \vdots & \vdots & \cdots & \vdots \\ y(t-N+k) & y(t-N+k-1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ y(t-N+1) & 0 & \cdots & 0 \end{array}\right] \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_{d-1} \end{bmatrix} = \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hline \vdots \\ y(t) \\ \vdots \\ y(t-N+L+d) \\ \hline \vdots \\ y(t-N+L+k) \\ \vdots \\ y(t-N+L+1) \end{array}\right]. \tag{3.18}$$

Let $\mathbf{Y}_z$ and $\mathbf{y}_z$ be the corresponding matrix and vector on the left and right side in (3.18), where the subscript z indicates zero padding. The above equation can be written as

$$\mathbf{Y}_z\beta = \mathbf{y}_z, \tag{3.19}$$

The LS estimate of β is

$$\hat{\beta}_{LS} = \arg\min_{\beta} \ \|\mathbf{y}_z - \mathbf{Y}_z\beta\|^2, \tag{3.20}$$

Similarly to the derivation of MMSE estimate, the LS estimate of β can be obtained by setting the derivative of $\|\mathbf{y}_z - \mathbf{Y}_z\beta\|^2$ with respect to β equal to zero. Solving the resulting linear equations, the LS estimate of β is given by

$$\begin{aligned} \hat{\beta}_{LS} &= (\mathbf{Y}_z^H\mathbf{Y}_z)^{-1}\mathbf{Y}_z^H\mathbf{y}_z \quad (3.21)\\ &= \mathbf{Y}_z^{\dagger}\mathbf{y}_z, \quad (3.22) \end{aligned}$$

where $\mathbf{Y}_z^{\dagger}$ is the pseudo inverse of $\mathbf{Y}_z$. In textbooks, this method is usually referred as the *autocorrelation* method or *Yule-Walker* method. While, if only the middle subset of the linear equations are used, it is called the *covariance* method [SM97]. It was also claimed that the covariance method had been found to be more accurate than the autocorrelation method in the sense that the estimated parameters of the former are on average closer to the true values than those of the latter [Mar87, Kay88]. One possible reason is that there is no windowing involved in the former method. But a drawback of the former method is the lack of guaranteed stability [SM97].

In this thesis, the covariance method is used to the LP. Let $\tilde{\mathbf{Y}}_z$ and $\tilde{\mathbf{y}}_z$ be the corresponding sub matrix and sub vectors of $\mathbf{Y}_z$ and $\mathbf{y}_z$ between the two horizontal lines as in (3.18). We define

$$\mathbf{f}_l^H = \beta^H = \tilde{\mathbf{Y}}_z^{\dagger}\tilde{\mathbf{y}}_z, \tag{3.23}$$

to be the prediction filter. The LP of $x(t+L)$ is

$$\hat{x}(t+L) = \mathbf{f}_l^H\mathbf{y}_d. \tag{3.24}$$

3.1.3 A Brief Survey of Computational Complexity

In practice, solving (3.15) or (3.22) is not a trivial task for large d due to the generalized inverse operation. Much effort have been made by mathematicians to find stable and cost efficient algorithms [Lev47, Dur60, BRS66, ABI74, Fle69, MC91]. These reported methods are divided into two categories, *recursive* algorithms and *nonrecursive* algorithms. The former is usually referred to as *fast algorithms*, like Levinson-Durbin or Schur algorithms. In practice, one can of course also track the LP recursively in time, with updates for each new data sample. This can be done using some classical adaptive methods, such as Least Mean Square (LMS), Recursive Least Square (RLS) or the Kalman filter. A detailed study of these

methods is outside the scope of this thesis. However, a quick conceptual survey of the computational complexity is within our interest.

As claimed that the major cost to solve these linear equations is due to the computation of the generalized inverse, *i.e.* $(\mathbf{Y}^H\mathbf{Y})^{-1}$. For a matrix $\mathbf{Y}_{k\times l}$, the calculation complexity of the generalized inverse by using *recursive* algorithms are between $O(3kl)$ and $O(4kl)$ measured by the number of float point operations. The corresponding orders for *nonrecursive* algorithms are between $O(5kl^2 + l^3/3)$ and $O(6kl^2 + l^3/2)$. In [SK03], the *modified covariance method* was found to have the best performance for channel prediction [Hay96], and its computational complexity is lower bounded by $O(3kl)$.

In $m \times n$ MIMO systems, the number of channels is mn. The calculation complexity for the LP predictors is then lower bounded by $O(3mnkl)$. It was found that the calculation complexity using LP predictors in a 4×4 MIMO system becomes unaffordable in [SWS03].

3.2 The Linear Minimum Mean Square Error Estimator

If the observations $\mathbf{y}$ can be modelled as

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{e}, \tag{3.25}$$

where $\mathbf{y}$ is an $N\times 1$ observations, $\mathbf{A}$ is a known $N\times p$ matrix, $\mathbf{s}$ is one realization of a $p\times 1$ random vector with zero mean and covariance $\mathbf{S} = E[\mathbf{s}\mathbf{s}^H] = \sigma_s^2\mathbf{I}_p$, $\mathbf{e}$ is an $N\times 1$ noise vector with PDF $\mathcal{CN}(\mathbf{0}, \sigma_e^2\mathbf{I}_N)$, and is uncorrelated with $\mathbf{s}$. Such a model is termed the *Bayesian linear model* [Kay93].

The LMMSE estimate of $\mathbf{s}$ can be derived element-wise to minimize the *Bayesian* MSE

$$\hat{s}_i = \arg\min_{\mathbf{f}_i} \ E[\|s_i - \mathbf{f}_i^H\mathbf{y}\|^2], \tag{3.26}$$

for $i \in [1,p]$, where the expectation is taken over s_i and $\mathbf{e}$. The LMMSE estimate of $\mathbf{f}_i$ is

$$\mathbf{f}_i = \mathbf{R}_{yy}^{-1}\mathbf{R}_{ys_i} \tag{3.27}$$

where $\mathbf{R}_{ys_i} = \mathbf{R}_{s_iy}^H = E[\mathbf{y}s_i^*]$. The LMMSE estimate of s_i is

$$\begin{aligned} s_i &= \mathbf{f}_i^H\mathbf{y}, \\ &= \mathbf{R}_{s_iy}\mathbf{R}_{yy}^{-1}\mathbf{y}, \quad (3.28) \end{aligned}$$

Finally the LMMSE estimation of $\mathbf{s}$ is *i.e.* [Kay93]

$$\begin{aligned} \hat{\mathbf{s}} &= \mathbf{F}^H\mathbf{y}, \\ &= \mathbf{R}_{sy}\mathbf{R}_{yy}^{-1}\mathbf{y}, \quad (3.29) \end{aligned}$$

where $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_p]$, $\mathbf{R}_{sy} = E[\mathbf{s}\mathbf{y}^H] = \sigma_s^2 \mathbf{A}^H$ and $\mathbf{R}_{yy} = E[\mathbf{y}\mathbf{y}^H] = \sigma_s^2 \mathbf{A}\mathbf{A}^H + \sigma_e^2 \mathbf{I}_N$.

3.3 Frequency Estimation Using ESPRIT Algorithms

The abbreviation ESPRIT stands for *Estimation of Signal Parameters via Rotation Invariance Techniques* [PRK85, RPK86]. It belongs to the eigenvector (EV) approaches to frequency estimates. This method exploits the Rotation Invariance (RI) property of the sinusoidal signal subspace.

3.3.1 Rotation Invariance Property of Sinusoidal Signal Subspace

In terms of a state space model, the sinusoidal model in (2.11) can be written as

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{\Phi}\mathbf{x}(t), && (3.30)\\ y(t) &= \mathbf{C}\mathbf{x}(t) + e(t),. && (3.31) \end{aligned}$$

where

$$\begin{aligned} \mathbf{x}(t) &= [s_1 e^{j\omega_1 t}, \cdots, s_p e^{j\omega_p t}]^T, && (3.32)\\ \mathbf{\Phi} &= \begin{bmatrix} e^{j\omega_1} & & & \\ & e^{j\omega_2} & & \\ & & \ddots & \\ & & & e^{j\omega_p} \end{bmatrix}, && (3.33)\\ \mathbf{C} &= [1, 1, \cdots, 1]. && (3.34) \end{aligned}$$

The vector $\mathbf{x}(t)$ is called the state vector, and (3.30) is called *state* equation. It describes the *dynamics* of the system. The matrix $\mathbf{\Phi}$ is the rotation operator, which is a diagonal unitary matrix in this case. The adjacent state vectors are related by the rotation operation matrix. The second equation is called *observation* equation. It provides the relation between observation $y(t)$ and the state vector $\mathbf{x}(t)$.

Let us form the state matrix $\mathbf{X}$ by stacking N consecutive state vectors as

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}(t)^T \\ \mathbf{x}(t-1)^T \\ \vdots \\ \mathbf{x}(t-N+1)^T \end{bmatrix}. \qquad (3.35)$$

where $\mathbf{X}$ is $N \times p$ matrix. For convenience, let $\mathbf{X} = [\mathbf{v}_1, \cdots, \mathbf{v}_p]$, where $\mathbf{v}_i$ is the ith column vector. These columns span the p dimension signal space. Then we define two observation matrices, $\mathcal{O}_1$ and $\mathcal{O}_2$ as

$$\begin{aligned} \mathcal{O}_1 &= \mathbf{X}(1:N-1,:), && (3.36)\\ \mathcal{O}_2 &= \mathbf{X}(2:N,:), && (3.37) \end{aligned}$$

which contain the first $N-1$ and the last $N-1$ rows of $\mathbf{X}$ respectively. Note that $\mathcal{O}_1$ and $\mathcal{O}_2$ are two temporally displaced subsets of the basis functions of $\mathbf{X}$. They are related by the rotation matrix $\mathbf{\Phi}$ as in (3.38). This fact is interpreted as the Rotation Invariance (RI) property of the sinusoidal signal subspace:

$$\mathcal{O}_1 = \mathcal{O}_2 \mathbf{\Phi}. \qquad (3.38)$$

3.3.2 Frequency Estimation Using RI Property

Based on the RI property of the signal subspace in (3.38), the frequency estimate can be obtained by the following steps:

1. Find a set of basis functions which span the signal subspace, *i.e.* a similar matrix to $\mathcal{O}$;

2. Forming $\mathcal{O}_1$ and $\mathcal{O}_2$ as in (3.36) and (3.37);

3. The LS estimate the rotation matrix $\mathbf{\Phi}$ is

$$\begin{aligned} \hat{\mathbf{\Phi}} &= (\mathcal{O}_2^H \mathcal{O}_2)^{-1} \mathcal{O}_2^H \mathcal{O}_1 \\ &= \mathcal{O}_2^{\dagger} \mathcal{O}_1; && (3.39) \end{aligned}$$

4. The frequency estimate $\hat{\omega}$ can be obtained from the phases of the eigenvalues of $\hat{\mathbf{\Phi}}$.

However, there is still one question need to answer, which is

Does the RI property hold for any set of basis functions spanning the sinusoidal signal subspace?

The answer is *Yes* and the proof is given below.

Assume a matrix $\mathbf{B}_{N\times p} = [\mathbf{b}_1, \cdots, \mathbf{b}_p]$ contains another set of basis function spanning the same signal subspace of $\mathbf{X}$. Let

$$\begin{aligned} \mathbf{b}_i &= t_{1i}\mathbf{v}_1 + t_{2i}\mathbf{v}_2 + \cdots + t_{pi}\mathbf{v}_p, && (3.40)\\ &= \mathbf{X}\mathbf{t}_i, && (3.41) \end{aligned}$$

where $\mathbf{t}_i = [t_{1i}, \cdots, t_{pi}]^T$ is the coordinates of $\mathbf{b}_i$ with respect to the basis in $\mathbf{X}$. So there exists a $p \times p$ transform matrix

$$\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_p], \tag{3.42}$$

which satisfies

$$\mathbf{B} = \mathbf{X}\mathbf{T}, \tag{3.43}$$

where $\mathbf{T}$ is an invertible matrix [HJ85]. Define two sub matrixes

$$\begin{aligned} \mathcal{B}_1 &= \mathbf{B}(1:N-1,:), & (3.44)\\ \mathcal{B}_2 &= \mathbf{B}(2:N,:), & (3.45) \end{aligned}$$

which contain the first $N-1$ and the last $N-1$ rows of $\mathbf{B}$ respectively. So

$$\begin{aligned} \mathcal{B}_1 &= \mathcal{O}_1\mathbf{T} = \mathcal{O}_2\mathbf{\Phi}\mathbf{T} = (\mathcal{O}_2\mathbf{T})(\mathbf{T}^{-1}\mathbf{\Phi}\mathbf{T}) \\ &= \mathcal{B}_2\mathbf{\Psi}, \end{aligned} \tag{3.46}$$

where $\mathbf{\Psi} = \mathbf{T}^{-1}\mathbf{\Phi}\mathbf{T}$. The matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are similar, which implies that they have the same eigenvalues. So the frequency estimate can be obtained from the phases of the eigenvalues of $\mathbf{\Psi}$ also. This proof leads to the following lemma.

Lemma 3.1

The Rotation Invariant property of sinusoidal signal subspace can be interpreted as in (3.38). A generalized form of this property is given by (3.46). The matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are similar, and thus share the same eigenvalues. The transform matrix $\mathbf{T}$ between these two sets of basis functions in $\mathbf{X}$ and $\mathbf{B}$ can be obtained as in (3.43).

3.3.3 Estimating the Signal Subspace

In practice, the observed sinusoidal signal is buried in noise $e(t)$, which is independent of the signals. The signal subspace can be estimated by making an Eigenvalue Decomposition (EVD) of its covariance matrix, or equivalently by computing the Singular Value Decomposition (SVD) of particularly formed Hankel matrix.

Eigenvalue Decomposition Methods

The covariance matrix of the data sequence $\mathbf{y}$ is

$$\begin{aligned} \mathbf{R}_{yy} &= E[\mathbf{y}\mathbf{y}^H], \\ &= E[\mathbf{x}\mathbf{x}^H] + E[\mathbf{e}\mathbf{e}^H], \\ &= \mathbf{R}_{xx} + \mathbf{R}_{ee}, \end{aligned} \tag{3.47}$$

where

$$\begin{aligned} \mathbf{R}_{xx} &= E[\mathbf{A}\mathbf{s}\mathbf{s}^H\mathbf{A}^H] = \mathbf{A}\mathbf{S}\mathbf{A}^H, & (3.48)\\ \mathbf{R}_{ee} &= \sigma_e^2\mathbf{I}_N. & (3.49) \end{aligned}$$

Taking the eigenvalue decomposition of $\mathbf{R}_{yy}$,

$$\begin{aligned} \mathbf{R}_{yy} &= [\mathbf{U}_s \ \mathbf{U}_n]\begin{bmatrix} \Lambda_s & \\ & \Lambda_n \end{bmatrix}\begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix}, & (3.50)\\ &= \mathbf{U}_s\Lambda_s\mathbf{U}_s^H + \mathbf{U}_n\Lambda_n\mathbf{U}_n^H, & (3.51) \end{aligned}$$

where the signal subspace of $\mathbf{R}_{yy}$ is spanned by the columns of $\mathbf{U}_s$. In this procedure a rank estimation algorithm is necessary, and the eigenvalues are ordered in non-increasing order.

Singular Value Decomposition Method

This method is also termed *Kung's algorithm* [Kun78]. This method is based on the SVD of the Hankel matrix $\mathbf{H}$ formed by the data sequence as

$$\mathbf{H} = \begin{bmatrix} y(t) & y(t-1) & \ldots & y(t-N+M) \\ y(t-1) & y(t-2) & \ldots & y(t-N+M-1] \\ \vdots & \vdots & \ddots & \vdots \\ y(t-M+1) & y(t-M) & \ldots & y(t-N+1) \end{bmatrix}, \tag{3.52}$$

where M is chosen as an integer close to $N/2$ as a rule of thumb. The SVD of $\mathbf{H}$ is

$$\mathbf{H} = [\mathbf{U}_s \ \mathbf{U}_n]\begin{bmatrix} \Sigma_s & \\ & \Sigma_n \end{bmatrix}\begin{bmatrix} \mathbf{V}_s^H \\ \mathbf{V}_n^H \end{bmatrix}. \tag{3.53}$$

where the singular values are ordered in non-increasing order. The *observability* matrix $\mathcal{O}$ is formed as

$$\mathcal{O} = \mathbf{U}_s \cdot \Sigma_s^{1/2}. \tag{3.54}$$

and $\mathcal{O}_1$ and $\mathcal{O}_2$ can be obtained as in (3.36) and (3.37). Finally, $\hat{\mathbf{\Phi}}$ is obtained as in (3.39).

3.3.4 Unitary ESPRIT

It was observed that the eigenvalues of $\mathbf{\Phi}$ have unit norm. To exploit this property, the Unitary ESPRIT algorithms was proposed in [HN95].

The Unitary ESPRIT algorithms begin with forming the data matrix $\mathbf{H}_u$,

$$\mathbf{H}_u \stackrel{\text{def}}{=} [\mathbf{H} \ \ \Pi_M\bar{\mathbf{H}}], \tag{3.55}$$

where $\mathbf{H}$ is the same as in (3.52). The over bar denotes complex conjugation without transposition. The Π_M is the $M \times M$ exchange matrix with ones on its anti-diagonal and zeros elsewhere as given in (3.56).

$$\Pi_M = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \cdot & & \\ 1 & & & \end{bmatrix}. \tag{3.56}$$

Then the SVD is performed upon $\mathbf{H}_u$ instead of on $\mathbf{H}$ as in Kung's algorithm. The remaining steps are identical to Kung's algorithm. It was found that Unitary ESPRIT has better performance than ESPRIT [HN95].

In practice, the noise might be colored in (2.11). An Instrumental Variable Method (IVM) to frequency estimation is proposed in [SVO94]. An introduction of this method is given in Appendix A.

3.4 CRLB of the Frequency Estimation

Assume that $\boldsymbol{\omega}$ is deterministic. The CRLB of the frequency estimate in the given model in (2.11) is derived in this section. Since the mean value of $\mathbf{y}$ is zero, only $(p+2)$ parameters are involved in the frequency estimate. They are $\boldsymbol{\xi} = [\boldsymbol{\omega}^T \, \sigma_s^2 \, \sigma_e^2]^T$. The CRLB of the estimate of $\boldsymbol{\xi}$ based on complex data with zero mean complex Gaussian PDF is given by $E\left[(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})^T \leq \mathbf{J}^{-1}(\boldsymbol{\xi})\right]$, where $\mathbf{J}(\boldsymbol{\xi})$ is the Fisher information matrix given by (3.57), for $i, j = 1, 2, \cdots, p+2$, i.e. [Kay93],

$$[\mathbf{J}(\xi)]_{ij} = Tr\left[\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial \xi_i}\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial \xi_j}\right], \tag{3.57}$$

where the subscription of $\mathbf{R}_{yy}$ is dropped for notation simplification.

In block matrix form, the Fisher information matrix is found as

$$\mathbf{J}(\xi) = \begin{bmatrix} \mathbf{J}_{\omega\omega} & \mathbf{J}_{\omega p} & \mathbf{J}_{\omega\sigma} \\ \mathbf{J}_{p\omega} & \mathbf{J}_{pp} & \mathbf{J}_{p\sigma} \\ \mathbf{J}_{\sigma\omega} & \mathbf{J}_{\sigma p} & \mathbf{J}_{\sigma\sigma} \end{bmatrix}, \tag{3.58}$$

where $\mathbf{J}_{\omega\omega}$ is $p \times p$, $\mathbf{J}_{pp}$ and $\mathbf{J}_{\sigma\sigma}$ are scalars. The other block matrices have comfortable dimensions. Explicit expressions of these block matrixes are given below.

$$\begin{aligned}
\mathbf{J}_{\omega\omega} &= 2(\sigma_s^2)^2 \mathrm{Re}[\mathbf{D}^H\mathbf{R}^{-1}\mathbf{D} \odot (\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A})^T, \\
&\quad + \mathbf{A}^H\mathbf{R}^{-1}\mathbf{D} \odot (\mathbf{A}^H\mathbf{R}^{-1}\mathbf{D})^T], \\
\mathbf{J}_{\omega p} &= 2(\sigma_s^2)\mathrm{Re}[diag(\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A}\mathbf{A}^H\mathbf{R}^{-1}\mathbf{D})], \\
\mathbf{J}_{\omega\sigma} &= 2(\sigma_s^2)\mathrm{Re}[diag(\mathbf{A}^H\mathbf{R}^{-2}D)], \\
\mathbf{J}_{pp} &= Tr[(\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A})^2], \\
\mathbf{J}_{p\sigma} &= Tr[\mathbf{A}^H\mathbf{R}^{-2}\mathbf{A}], \\
\mathbf{J}_{\sigma\sigma} &= Tr[\mathbf{R}^{-2}].
\end{aligned}$$

where $\mathbf{D} = [\mathbf{d}_1, \cdots, \mathbf{d}_p]$ and is $N \times p$, and $\mathbf{d}_i = \partial \mathbf{a}_i / \partial \omega_i$. The $diag(\mathbf{M})$ represents the column vector formed from the diagonal elements of the matrix $\mathbf{M}$. The others are $\mathbf{J}_{p\omega} = \mathbf{J}_{\omega p}^T$, $\mathbf{J}_{\sigma\omega} = \mathbf{J}_{\omega\sigma}^T$ and $\mathbf{J}_{\sigma p} = \mathbf{J}_{p\sigma}^T$. Note that the CRLB given in (3.58) is different from that given in e.g. [TV01], where the complex amplitudes are treated as deterministic. It should also be mentioned that the CRLB does not exist when the frequencies are random, at least not with the *tub* distribution in (2.8). Note that since they are random, they are close in some realizations and well separated in others.

However, in practice, we might be interested in the best possible estimated frequencies, even if some of them are too closely separated, in which case the CRLB is impractically large. A method to calculate a compressed CRLB is proposed, which is given in Appendix B.

Chapter 4

Measurements and Data Analysis

The measurement data is introduced in this chapter. The Jake's model is confirmed by the Normalized Mean Doppler Spectrum (NMDS) of the measurements in both urban and suburban environments. An analysis of the stationarity properties of the model parameters is made by using the MUSIC pseudo spectrum and parameter tracking techniques.

4.1 Measurement Data

A measurement campaign was performed in the urban area of Stockholm, and the suburban area, Kista, which is a few kilometers north to central Stockholm. The measurements were performed at 1880 MHz and the band width is 5MHz. In total, 45 and 31 effective measurements were performed at different locations in urban and suburban respectively. The velocities of the mobile terminal were between 30 to 90 Km/h during the measurements, except a few stand still cases. Each measurement records channel sounding data over 156.4 ms and contains 1430 repetitions of channel impulse responses. This gives rise to the channel update rate of 9.1 KHz. Each channel impulse response (or power delay profile) is described by 120 taps with a sampling frequency of 6.4 MHz. The time interval between two neighboring taps is $0.156\mu s$. An example of the measured channel is given in Fig. 2.2. It is represented in Fig. 4.1 for convenience.

Noise reduction by smoothing was performed over the raw data [Ekm02]. An example of the effect of noise reduction is given in Fig. 4.2, where the smoothing process is made over 20 neighboring samples. In this figure, the left and right subplots are from a channel tap with low SNR before and after smoothing respectively. The smoothing effect is significant. But in case of high SNR the effect of smoothing is negligible. In this thesis, the channel taps with high SNR are used in the performance evaluations of the proposed channel prediction methods, which can be considered as a narrow band Rayleigh fading channel or a single frequency bin in an Orthogonal Frequency Division Multiplexing (OFDM) system.

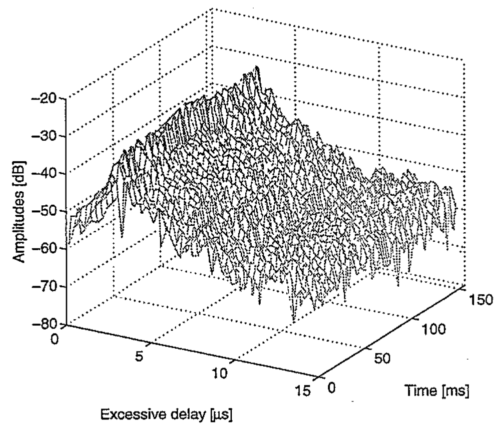

Figure 4.1: Example of a measured wide band channel. Each power delay profile contains 120 taps and the sampling frequency is 6.4 MHz. The duration of the measurement is 156 ms.

4.2 Normalized Mean Doppler Spectrum

To verify the Jake's model, the Normalized Mean Doppler Spectrum (NMDS) is introduced. Assume $\mathbf{y}$ is the observed channel, the NMDS is obtained as follow:

1. Estimate the $d \times d$ covariance matrix of $\mathbf{y}$, $\mathbf{R}_d$, using (3.16);

2. The Doppler spectrum of $\mathbf{y}$ is

$$p_y(\omega) = \mathcal{F}\{\mathbf{R}_d(:,1)\}, \tag{4.1}$$

   where $\mathcal{F}\{\cdot\}$ is the DFT operation, and $\mathbf{R}_d(:,1)$ is the first column of $\mathbf{R}_d$;

3. Normalize $p_y(\omega)$ by its maximum value;

4. Stretch the normalized $p_y(\omega)$ within $[-f_{max}, f_{max}]$ to the normalization frequency grid, $[-1,1]$, where the maximum Doppler frequency f_{max} is estimated beforehand;

5. The average of the normalized and stretched $p_y(\omega)$ over all measurements is defined as the NMDS.

In Fig. 4.3, an example of the normalized Doppler power spectrum (left) and the stretched and normalized Doppler spectrum (right) from one measurement is

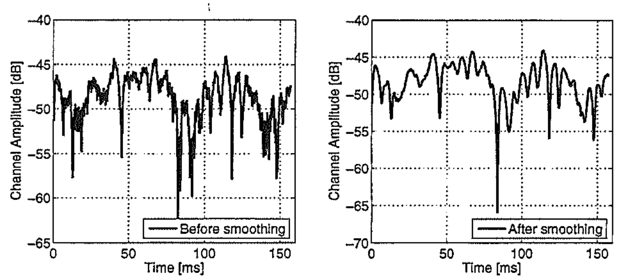

Figure 4.2: Example of the effect of noise reduction. The left and right subplots are from a channel tap with low SNR before and after smoothing respectively. The smoothing effect is significant. The smoothing process was made over 20 neighboring samples.

given. The parts between the two vertical dash lines are the counterparts in these two plots.

In Fig. 4.4 and Fig. 4.5, the resulting NMDS in urban and suburban area are given respectively. Both have *tub* shape spectrum, which is described by the Jake's model.

4.3 Stationarity of Model Parameters

For sinusoidal modeling based channel prediction methods, the stationarity property of model parameters plays an important role on prediction accuracy. In a stationary scenario, the channel can be predicted as long as the model parameters remain unchanged. On the other hand, an increased prediction error is inevitable due to the model errors as the prediction horizon is increased. In this section, an intuitive study of the stationarity of the model parameters are investigated by means of MUSIC Pseudo Spectrum and parameter tracking.

4.3.1 MUSIC Pseudo Spectrum

The MUSIC Pseudo Spectrum (MPS) is defined as [KV96]

$$p_{mu}(\omega) = \frac{\mathbf{a}(\omega)^H \mathbf{a}(\omega)}{\mathbf{a}(\omega)^H \mathbf{\Pi}^{\perp} \mathbf{a}(\omega)}, \qquad \omega \in [-f_{max}, f_{max}] \tag{4.2}$$

where $\mathbf{\Pi}^{\perp} = \mathbf{U}_n \mathbf{U}_n^H$, and $\mathbf{U}_n$ is given in (3.51). The $\mathbf{a}(\omega)$ is the DFT vector associated with frequency ω, and f_{max} is the maximum Doppler frequency. The MPS is calculated over each segmented data block, where a sliding window is applied for data extraction. Examples of typical time varying MPS's in urban and suburban channels are presented in Fig. 4.6 and Fig. 4.7 respectively. In these two figures, the sampling frequency is 910 Hz and the number of sinusoids are assumed

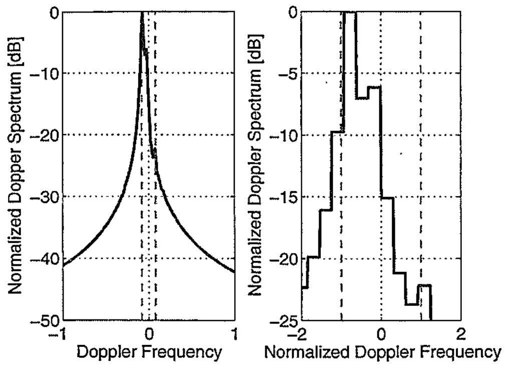

Figure 4.3: Example of stretching of normalized Doppler spectrum. The left plot is the normalized Doppler spectrum of a measured Rayleigh fading channel; the right plot is the corresponding stretched normalized Doppler spectrum, where the frequency is normalized by the maximum Doppler frequency.

to be 8 and 7 respectively. In these figures, the color of the contours indicates the powers in the frequency bins.

In Fig. 4.6, the MPS from an urban measurement is given, where a number of frequency bins with deeper color appear in the spectrum. A number of relatively high power frequency bins can be observed. They distribute on both the positive and negative sides of the spectrum. But the relative power between these bins are low. Some of them are consistent over the measurements, the others are not. These bins have blurry edges, which implies that these bins are contributed by a clutter of close reflectors. Such a spectrum agrees with the typical Doppler spectrum in a rich scattering urban environment. These non-consistent frequency bins might make the sinusoidal modeling based channel prediction painful.

The MPS in Fig. 4.7 is obtained from a measurement in a suburban area. It has much fewer high power frequency bins compared to what is observed in urban measurement as given in Fig. 4.6. The distribution of these frequency bins in the spectrum are biased to the negative side in this example. This is probably due to that the mobile terminal was moving away from the base station or the reflection object when the measurement was made. The edges of these high power frequency bins are much sharper than those in Fig. 4.6. This indicates that each bin might contain only one pure sinusoid. This is also coincident with a typical Doppler spectrum in suburban or rural area in LOS. The power in the strongest frequency bin is much higher than the others and is consistent over the whole measurement interval.

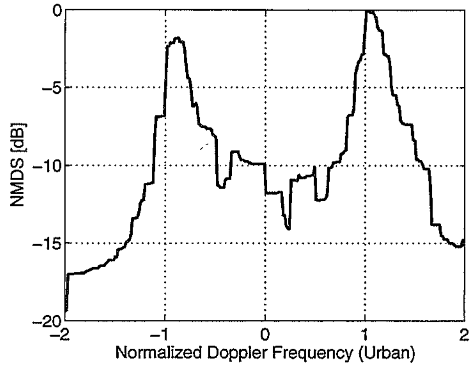

Figure 4.4: NMDS in urban area. The average of the stretched normalized Doppler spectrum is made over 45 measurements in urban area. A *tub* spectrum is obtained.

4.3.2 Model Parameter Tracking

Similar to the study of the MPS's, estimates of the model parameters (frequencies and amplitudes) are computed for each data block using the Unitary-ESPRIT and LS methods. The estimates of the frequencies and amplitudes using the urban and suburban example measurement are plotted as a function of number of data blocks in Fig. 4.8 and Fig. 4.9 respectively. The number of sinusoids are assumed to be 8 and 7 for urban and suburban example measurements respectively. In these figures, the channel amplitudes and phases are given in the subplot (a) and (b). The estimated frequencies and the associated amplitudes are given in subplot (c) and (d), where different markers are used for frequency and amplitude estimates associate with each sinusoid.

In Fig. 4.8, which is the urban case, the dynamic range of the channel amplitudes is larger than the one in the suburban measurement, as given in subplot (a) in Fig. 4.9. The channel phases are linear in just a part of the measurement, which indicates that there might be just one dominant sinusoid in the linear part of the channel, but more comparable dominant sinusoids in the rest part. In subplot (d), one amplitude is much higher than others, which is marked by an asterisk. Its frequency is close to -0.1 as seen in subplot (c). The amplitude of this ray is consistent over the observation interval. Besides this sinusoid, a number of sinusoids with well separated frequencies are found in subplot (c), but their amplitudes are time varying and comparable. The variations of these amplitudes are much larger than the first frequency. The presence of these sinusoids might explain the large variation of the channel amplitudes and nonlinear phase in this measurement.

In Fig. 4.9, the parameter tracking using the suburban example measurement is given. Both the average and the variation of the channel amplitude are smaller

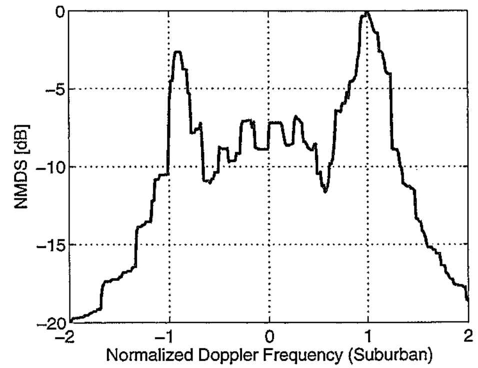

Figure 4.5: NMDS in suburban area. The average of the stretched normalized Doppler spectrum is made over 31 measurements in suburban area. A *tub* spectrum is obtained.

than in the urban measurement, as seen in subplot (a) in Fig. 4.8. The phase of the channel is quite linear over the whole measurement interval. This implies that there might be just one dominant sinusoidal signal in this channel. But, in subplot (d), three sinusoids are found that have higher and comparable amplitudes than others. Two of these amplitudes, marked by a triangle and a pentagon respectively, are much more consistent than the other marked by a circle. But it was found that the frequencies associated to the sinusoids marked by a triangle and a circle are very close in subplot (c). The appearance of these two strong and closely separated frequencies might be due to the over estimated number of sinusoids in the Unitary-ESPRIT method. A guess is that they might belong to the same impinging ray.

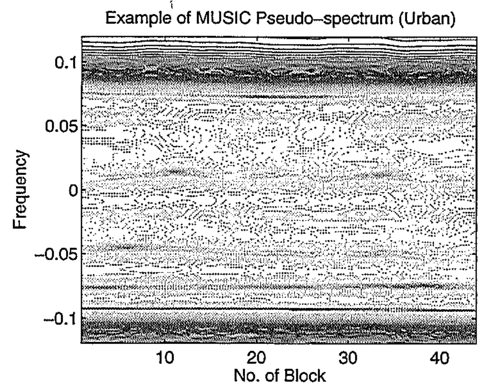

Figure 4.6: Example of a MUSIC Pseudo Spectrum of a measurement in urban area. A number of high power frequency bins can be observed. They are distributed on both the positive and negative sides of the spectrum. But the relative power between these bins and the remaining frequency bins is low. Some of them are consistent over the measurements, the others are not. These bins have blurry edges.

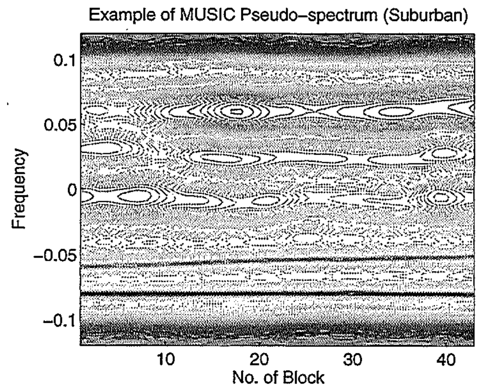

Figure 4.7: Example of a MUSIC Pseudo Spectrum of a measurement in suburban area. A less number of high power frequency bins are observed compared to Fig. 4.6. The locations of these frequency bins are biased to the negative side in the spectrum. The edges of these high power frequency bins are much sharper than those in Fig. 4.6. The power in the strongest frequency bin is much higher than others and is consistent over the whole measurement interval.

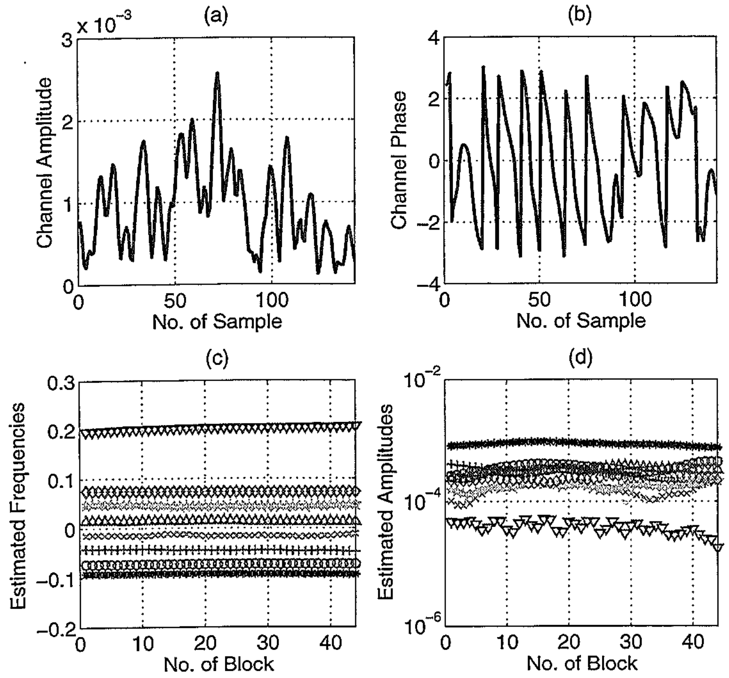

Figure 4.8: Example of the time variation of model parameters in urban area. (a) is the channel amplitudes, whose dynamic range is large in this example; (b) is the phases of the channel, which is linear in a part of this measurement; (c) is the estimated frequencies of the sinusoids in each block, the number of sinusoids is assumed to be 8. (d) The amplitudes of the sinusoids associated to the frequencies in (c).

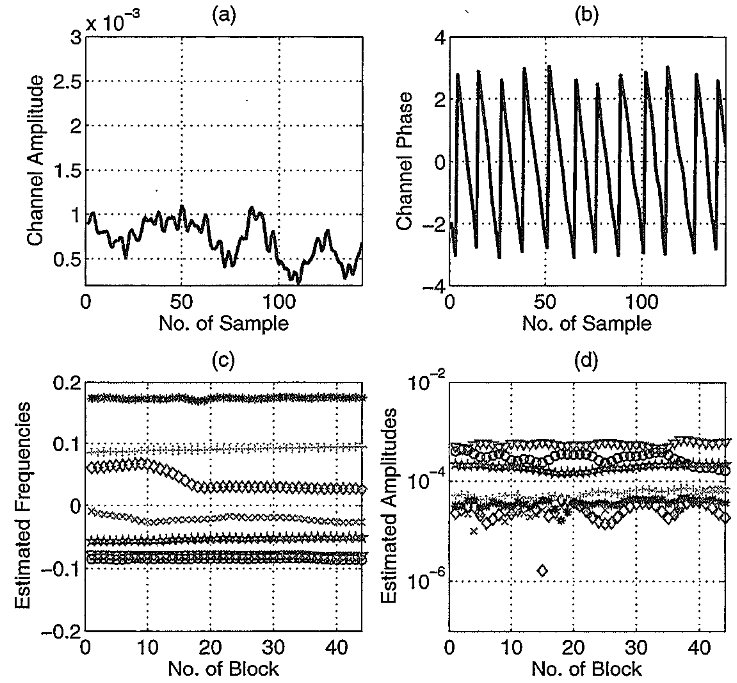

Figure 4.9: Example of the time variation of model parameters in suburban area. (a) is the channel amplitudes, whose dynamic range is smaller than in Fig. 4.8; (b) is the phases of the channel, which is linear in the whole measurement interval; (c) is the estimated frequencies of the sinusoids in each block, the number of sinusoids is assumed to be 7. (d) The amplitudes of the sinusoids associated to the frequencies in (c).

Chapter 5

Model-Based LMMSE Predictors

Based on the sinusoidal modeling of the Rayleigh fading channel given in (2.11), a series model-based LMMSE prediction methods are proposed in this chapter.

5.1 Conditioned LMMSE Predictors

Assuming the frequency estimates $\hat{\omega} = [\hat{\omega}_1, \cdots, \hat{\omega}_p]^T$ are given, the LS estimate of the complex amplitudes $\mathbf{s}$ is

$$\hat{\mathbf{s}}_{LS} = \arg\min_{\mathbf{s}} \|\mathbf{y} - \hat{\mathbf{A}}\mathbf{s}\|^2, \tag{5.1}$$

where $\hat{\mathbf{A}} = \mathbf{A}(\hat{\omega})$, and $\|\mathbf{y} - \hat{\mathbf{A}}\mathbf{s}\|^2$ is the LS cost function, denoted as ε. The solution is

$$\begin{aligned} \hat{\mathbf{s}}_{LS} &= (\hat{\mathbf{A}}^H\hat{\mathbf{A}})^{-1}\hat{\mathbf{A}}^H\mathbf{y} \\ &= \hat{\mathbf{A}}^\dagger\mathbf{y}. \end{aligned} \tag{5.2}$$

Substituting $\mathbf{y} = \hat{\mathbf{A}}\mathbf{s} + \mathbf{e}$ into (5.2), the LS estimate of $\mathbf{s}$ can be written as

$$\begin{aligned} \hat{\mathbf{s}}_{LS} &= (\hat{\mathbf{A}}^H\hat{\mathbf{A}})^{-1}\hat{\mathbf{A}}^H(\hat{\mathbf{A}}\mathbf{s} + \mathbf{e}) \\ &= \mathbf{s} + (\hat{\mathbf{A}}^H\hat{\mathbf{A}})^{-1}\hat{\mathbf{A}}^H\mathbf{e} \\ &= \mathbf{s} + \hat{\mathbf{A}}^\dagger\mathbf{e}, \end{aligned} \tag{5.3}$$

where the last term is the estimation error of $\mathbf{s}$, which is denoted as $\Delta\mathbf{s}$ with covariance $\sigma_e^2(\hat{\mathbf{A}}^H\hat{\mathbf{A}})^{-1}$.

Obviously, if $(\hat{\mathbf{A}}^H\hat{\mathbf{A}})$ is poorly conditioned due to closely spaced frequencies, the noise will be amplified by the pseudo inverse matrix $\hat{\mathbf{A}}^\dagger$. In fact, the probability distribution function of the Doppler frequency, given in Fig. 2.4, shows that the probability of the Doppler frequency in the region close to ± 1 in the normalized spectrum is much higher than the other part. Therefore, we have good reasons to expect the *true* $\mathbf{A}$ matrix to be ill-conditioned.

5.1.1 Regularized LS Estimate of Amplitudes

Since the ill-conditioned $\hat{\mathbf{A}}$ matrix leads to large $\|\mathbf{s}\|$, a penalty term could be added into the LS cost function ε and the new cost function becomes

$$\varepsilon_\alpha = \|\mathbf{y} - \hat{\mathbf{A}}\mathbf{s}\|^2 + \alpha\|\mathbf{s}\|^2, \tag{5.4}$$

and the amplitude estimate is

$$\hat{\mathbf{s}}_{Reg} = \arg\min_{\mathbf{s}} \{\|\mathbf{y} - \hat{\mathbf{A}}\mathbf{s}\|^2 + \alpha\|\mathbf{s}\|^2\}, \tag{5.5}$$

where $\alpha > 0$ is called the *regularization* parameter. Taking the derivative of ε_α with respect to s

$$\frac{\partial \varepsilon_\alpha}{\partial \mathbf{s}} = -2\hat{\mathbf{A}}^H(\mathbf{y} - \hat{\mathbf{A}}\mathbf{s}) + 2\alpha\mathbf{s}, \tag{5.6}$$

and set it to be zero, the regularized LS estimate of s is

$$\begin{aligned} \hat{\mathbf{s}}_{Reg} &= (\hat{\mathbf{A}}^H\hat{\mathbf{A}} + \alpha\mathbf{I}_p)^{-1}\hat{\mathbf{A}}^H\mathbf{y} && (5.7)\\ &= \mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y}, && (5.8) \end{aligned}$$

where $\mathbf{R}_\alpha$ is the regularized covariance matrix. Its eigenvalues are displaced from those of $(\hat{\mathbf{A}}^H\hat{\mathbf{A}})$ by the regularization parameter α. So the ill-conditioned problem (zero eigenvalues) in (5.3) is alleviated.

5.1.2 Conditioned LMMSE Estimate of Complex Amplitudes

The LMMSE estimate of s given $\hat{\omega}$ is $\hat{\mathbf{F}}^H\mathbf{y}$, where

$$\hat{\mathbf{F}} = \arg\min_{\mathbf{F}} E\left[\|\mathbf{s} - \mathbf{F}^H\mathbf{y}\|^2\right]. \tag{5.9}$$

From (3.29), the LMMSE estimate is

$$\begin{aligned} \hat{\mathbf{s}} &= \hat{\mathbf{F}}^H\mathbf{y} = \mathbf{R}_{sy}\mathbf{R}_{yy}^{-1}\mathbf{y} && (5.10)\\ &= \hat{\mathbf{A}}^H\left(\hat{\mathbf{A}}\hat{\mathbf{A}}^H + \frac{\sigma_e^2}{\sigma_s^2}\mathbf{I}_N\right)^{-1}\mathbf{y}, && (5.11) \end{aligned}$$

where

$$\begin{aligned} \mathbf{R}_{sy} &= E[\mathbf{s}\mathbf{y}^H] = \sigma_s^2\hat{\mathbf{A}}^H, && (5.12)\\ \mathbf{R}_{yy} &= E[\mathbf{y}\mathbf{y}^H] = \sigma_s^2\hat{\mathbf{A}}\hat{\mathbf{A}}^H + \sigma_e^2\mathbf{I}_N. && (5.13) \end{aligned}$$

Let $\alpha = \sigma_e^2/\sigma_s^2$, the inverse of Signal Noise Ratio (SNR). Using the matrix inversion lemma[1], we have

$$\begin{aligned} (\hat{\mathbf{A}}\hat{\mathbf{A}}^H + \alpha\mathbf{I}_N)^{-1} &= \alpha^{-1}\mathbf{I}_N - \alpha^{-1}\hat{\mathbf{A}}(\mathbf{I}_N + \alpha^{-1}\hat{\mathbf{A}}^H\hat{\mathbf{A}})^{-1}\hat{\mathbf{A}}^H\alpha^{-1}\\ &= \alpha^{-1}\mathbf{I}_N - \alpha^{-1}\hat{\mathbf{A}}(\hat{\mathbf{A}}^H\hat{\mathbf{A}} + \alpha\mathbf{I}_N)^{-1}\hat{\mathbf{A}}^H. && (5.14) \end{aligned}$$

[1] $(A+BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B+C^{-1})^{-1}DA^{-1}$.

Inserting (5.14) into (5.11) shows that the LMMSE estimate of s is obtained as

$$\begin{aligned} \hat{\mathbf{s}} &= (\hat{\mathbf{A}}^H\hat{\mathbf{A}} + \alpha\mathbf{I}_N)^{-1}\hat{\mathbf{A}}^H\mathbf{y}\\ &= \mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y}. && (5.15) \end{aligned}$$

This result also answered the question of the optimal value of the regularization parameter in (5.8).

5.1.3 Conditioned LMMSE Channel Prediction

Given $\hat{\omega}$, the Conditioned LMMSE predictor is given by the Wiener filter [Kay93],

$$\hat{x}(t+L) = \mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{y}, \tag{5.16}$$

where

$$E\left[\begin{bmatrix} x(t+L)\\ \mathbf{y}\end{bmatrix}\begin{bmatrix} x(t+L)^H & \mathbf{y}^H\end{bmatrix}\right] = \begin{bmatrix} \sigma_x^2 & \mathbf{R}_{xy}\\ \mathbf{R}_{yx} & \mathbf{R}_{yy}\end{bmatrix}. \tag{5.17}$$

where the expectation is taken over amplitudes and noise. Define the p-row vector

$$\hat{\mathbf{a}}(L)^H = [e^{j\hat{\omega}_1(t+L)}, \cdots, e^{j\hat{\omega}_p(t+L)}]. \tag{5.18}$$

The matrix $\mathbf{R}_{xy}$ can then be written as

$$\begin{aligned} \mathbf{R}_{xy} &= E[\hat{\mathbf{a}}(L)^H\mathbf{s}\mathbf{y}^H],\\ &= \hat{\mathbf{a}}(L)^H E[\mathbf{s}\mathbf{y}^H],\\ &= \hat{\mathbf{a}}(L)^H\mathbf{R}_{sy}. && (5.19) \end{aligned}$$

Finally the Conditioned LMMSE predictor of $x(t+L)$ is

$$\begin{aligned} \hat{x}(t+L) &= \hat{\mathbf{a}}(L)^H\mathbf{R}_{sy}\mathbf{R}_{yy}^{-1}\mathbf{y}, && (5.20)\\ &= \hat{\mathbf{a}}(L)^H\hat{\mathbf{s}}, && (5.21) \end{aligned}$$

where $\hat{\mathbf{s}}$ is given in (5.15). We could further define

$$\mathbf{f}_c^H = \hat{\mathbf{a}}(L)^H\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H. \tag{5.22}$$

Then the Conditioned LMMSE predictor (given $\hat{\omega}$) is

$$\hat{x}(t+L) = \mathbf{f}_c^H\mathbf{y}. \tag{5.23}$$

For Gaussian signals, the LMMSE predictor is also an MMSE predictor. The proof is given in Appendix C.

5.2 Unconditioned LMMSE Predictors

In previous studies, the frequency estimates $\hat{\omega}$ were regarded as exact. But $\hat{\omega}$ is estimated by some method, *e.g.* Unitary ESPRIT [HN95] and subject to errors in practice. These frequency estimation errors bring large prediction error for long range channel prediction methods. An LMMSE predictor is proposed by modeling the frequency estimate error as Gaussian random variables, which is called Unconditioned LMMSE predictor.

We can model ω_k as a random variable with mean $\hat{\omega}_k$ and variance $E[\Delta\omega_k^2] = \sigma^2_{\Delta\omega_k}$, where

$$\omega_k = \hat{\omega}_k + \Delta\omega_k. \tag{5.24}$$

The error variance, in theory, is determined by the SNR and number of samples [Kay93]. Such a problem was investigated in [SS91].

The Unconditioned LMMSE predictor has the same form as in (5.16)

$$\hat{x}(t+L) = \mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{y}. \tag{5.25}$$

But the expectation is taken over frequencies, amplitudes and noise.

The covariance matrix for N observations is

$$\mathbf{R}_{yy} = E[\mathbf{y}\mathbf{y}^H] = \mathbf{R}_{xx} + \mathbf{R}_{ee}. \tag{5.26}$$

Under the assumption that the amplitudes and frequency estimation errors are independent stochastic variables, we can write the $(m,n)^{\text{th}}$ element of $\mathbf{R}_{xx}$ as

$$\begin{aligned}[\mathbf{R}_{xx}]_{mn} &= E\left[\mathbf{A}\mathbf{s}\mathbf{s}^H\mathbf{A}^H\right]_{mn} \\ &= \sum_{k=1}^{p} E\left[|s_k|^2\right] E\left[e^{j(\hat{\omega}_k+\Delta\omega_k)(n-m)}\right].\end{aligned} \tag{5.27}$$

The expectation over the frequency estimation error can be expressed as

$$\begin{aligned}E\left[e^{j(\hat{\omega}_k+\Delta\omega_k)(n-m)}\right] &= e^{j\hat{\omega}_k(n-m)}E\left[e^{j\Delta\omega_k(n-m)}\right] \\ &= e^{j\hat{\omega}_k(n-m)}\Phi_{\Delta\omega_k}(n-m)\end{aligned} \tag{5.28}$$

where $\Phi_{\Delta\omega_k}(\tau)$ is the characteristic function of the frequency estimation error $\Delta\omega_k$. If we assume the frequency errors to be Gaussian distributed, then

$$\Phi_{\Delta\omega_k}(n-m) = e^{-\frac{\sigma^2_{\Delta\omega_k}}{2}(n-m)^2}. \tag{5.29}$$

The expectation over the ensemble of amplitudes is just the variance

$$E\left[|s_k|^2\right] = \sigma_s^2. \tag{5.30}$$

The $(m,n)^{\text{th}}$ element of the covariance matrix is thus obtained as

$$[\mathbf{R}_{xx}]_{mn} = \sum_{k=1}^{p} \sigma_s^2 e^{j\hat{\omega}_k(n-m)} e^{-\frac{\sigma^2_{\Delta\omega_k}}{2}(n-m)^2}. \tag{5.31}$$

For simplicity, let all the frequency errors be IID with $\sigma^2_{\Delta\omega_k} = \sigma^2_{\Delta\omega}$. We then have

$$[\mathbf{R}_{xx}]_{mn} = \sigma_s^2\left[\sum_{k=1}^{p} e^{j\hat{\omega}_k(n-m)}\right] e^{-\frac{\sigma^2_{\Delta\omega}}{2}(n-m)^2}. \tag{5.32}$$

Define the damping matrix $\mathbf{\Gamma}$ by

$$[\mathbf{\Gamma}]_{mn} = e^{-\frac{\sigma^2_{\Delta\omega}}{2}(n-m)^2}, \tag{5.33}$$

The covariance matrix can then be expressed as

$$\mathbf{R}_{xx} = \sigma_s^2\hat{\mathbf{A}}\hat{\mathbf{A}}^H \odot \mathbf{\Gamma} \tag{5.34}$$

where $\odot$ denotes the Hadamard product.
Similarly, the cross covariance $\mathbf{R}_{xy}$ is given by

$$\begin{aligned}\mathbf{R}_{xy} &= E\left[\hat{\mathbf{a}}(L)^H\mathbf{s}\mathbf{s}^H\hat{\mathbf{A}}^H\right] \\ &= \sum_{k=1}^{p} E\left[|s_k|^2\right] E\left[\left[e^{j(\hat{\omega}_k+\Delta\omega_k)L}, \cdots, e^{j(\hat{\omega}_k+\Delta\omega_k)(L+N-1)}\right]\right].\end{aligned} \tag{5.35}$$

Define the damping vector

$$\gamma = \left[e^{\frac{-\sigma^2_{\Delta\omega}}{2}L^2}, \cdots, e^{\frac{-\sigma^2_{\Delta\omega}}{2}(L+N-1)^2}\right]. \tag{5.36}$$

The cross covariance is then obtained as

$$\begin{aligned}\mathbf{R}_{xy} &= \sigma_s^2\sum_{k=1}^{p}\left\{\left[e^{j\hat{\omega}_kL}, \cdots, e^{j\hat{\omega}_k(L+N-1)}\right]\right\} \odot \gamma \\ &= \sigma_s^2\hat{\mathbf{a}}(L)^H\hat{\mathbf{A}}^H \odot \gamma,\end{aligned} \tag{5.37}$$

where $\hat{\mathbf{a}}(L)$ is the same as in (5.18). Finally, the Unconditioned LMMSE predictor is

$$\begin{aligned}\hat{x}(t+L) &= \mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{y}, \\ &= (\sigma_s^2\hat{\mathbf{a}}(L)^H\hat{\mathbf{A}}^H \odot \gamma)(\sigma_s^2\hat{\mathbf{A}}\hat{\mathbf{A}}^H \odot \mathbf{\Gamma} + \sigma_e^2\mathbf{I}_N)^{-1}\mathbf{y}, \\ &= (\hat{\mathbf{a}}(L)^H\hat{\mathbf{A}}^H \odot \gamma)(\hat{\mathbf{A}}\hat{\mathbf{A}}^H \odot \mathbf{\Gamma} + \alpha\mathbf{I}_N)^{-1}\mathbf{y}.\end{aligned} \tag{5.38}$$

Define the unconditioned LMMSE prediction filter as

$$\mathbf{f}_u^H = (\hat{\mathbf{a}}(L)^H\hat{\mathbf{A}}^H \odot \gamma)(\hat{\mathbf{A}}\hat{\mathbf{A}}^H \odot \mathbf{\Gamma} + \alpha\mathbf{I}_N)^{-1}, \tag{5.39}$$

(5.38) can be expressed as

$$\hat{x}(t+L) = \mathbf{f}_u^H\mathbf{y}. \tag{5.40}$$

Clearly, the influence of old observations is reduced by the damping matrix $\mathbf{\Gamma}$ and the damping vector γ. The damping vector γ is also dependent on the prediction range. The further ahead we are looking, the smaller is the gain of the predictor. The way the frequency error is taken into account can be interpreted as a convolution of the line spectrum of the signal with the error distribution. The filter design is thus done for distributed sources.

The errors in the frequency estimates force the predictor to miss-trust older data, as even a small frequency error can cause a large phase error if one waits long enough. As the special case, when $\sigma^2_{\Delta\omega} = 0$, (5.38) degenerates to (5.23). When $\sigma^2_{\Delta\omega} > 0$, (5.38), although being LMMSE, is not the MMSE due to the nonlinear dependency on the frequency estimates.

5.3 Combined LMMSE and LP Predictor

In the study of the LMMSE Prediction methods, it was found that the prediction residue signal is colored. It could be predicted by a low order LP predictor. This method is referred to as Combined LMMSE and LP predictor.

Define the *residue* signal $\epsilon = [\epsilon(t), \cdots, \epsilon(t-N+1)]^T$ as the part of the channel which can not be accurately reconstructed by the estimated sinusoidal model:

$$\epsilon = \mathbf{y} - \hat{\mathbf{x}},$$

where $\hat{\mathbf{x}} = \hat{\mathbf{A}}\hat{\mathbf{s}}$. In this definition, $\epsilon(t)$ contains both the channel estimation error $e(t)$ and the prediction error $x(t) - \hat{x}(t)$. In the imperfect modeling case, a low order LP can be used to predict the colored residues $\epsilon(t+L)$. So the Combined LMMSE and LP is

$$\hat{x}_{comb}(t+L) = \hat{x}(t+L) + \hat{\epsilon}(t+L). \tag{5.41}$$

The predictor $\hat{\epsilon}(t+L)$ is computed based on past residues as

$$\begin{aligned} \hat{\epsilon}(t+L) &= \sum_{k=0}^{d-1} \beta_k \epsilon(t-k), \\ &= \beta_d^H \epsilon_d. \end{aligned} \tag{5.42}$$

where $\epsilon_d = [\epsilon(t), \cdots, \epsilon(t-d+1)]^T$. The coefficients are computed by solving an LS problem. In principle, the color of the residuals should be taken into account also when estimating the frequencies [KN94]. However, at the relatively high SNR considered here, this will not substantially change the estimates.

Chapter 6

Performance Evaluation of Model-Based LMMSE Predictors

Two measures used for the performance evaluation are defined, which are Normalized Mean Square Error (NMSE), and Adjusted NMSE (ANMSE). The evaluations are made by Monte Carlo simulations and using measured data.

6.1 Measures of Prediction Performance

The performance of the predictors are evaluated by Monte Carlo simulations and measurements. Normalized Mean Square Error (NMSE) is adopted in this thesis to evaluate the performances of channel predictors. First we define the Normalized Square Error (NSE) in a single realization as

$$e^2_{NSE}(t+L) = \frac{N\|x(t+L) - \hat{x}(t+L)\|^2}{\mathbf{x}^H\mathbf{x}}. \tag{6.1}$$

Then the NMSE is the mean value of $e^2_{NSE}(t+L)$ taken over the Monte Carlo simulations. However, an Adjusted NMSE (ANMSE) is also introduced to get rid of the influence of a small amount of outliers which might ruin the whole averaged performance. So the outliers are dropped when the NMSE of their power prediction is larger than 0.04, or in other words, when the prediction error of the complex amplitude is larger than 20% of the Root Mean Square (RMS) channel amplitudes.

Un dollat begrepp

The concept *level of confidence* is also used in the performance evaluation. For example, when we say the NMSE of a channel predictor is -10 dB with the level of confidence of 95%, we mean that the average of the NSE's of the best 95% of the channel predictions ~~from~~ simulations or measurements ~~is less than~~ -10 dB, while the worst 5% cases are excluded.

It is also worth to note that when the measurement data is used, the prediction error is normalized by the mean power of $\mathbf{y}$, instead of $\mathbf{x}$ as given in (6.1).

6.2 Performance Evaluation of Model-Based LMMSE Predictors with Nearby Tones

As discussed in Chapter 2, the probability of the Doppler frequency appearing at the maximum Doppler shift is high according to the Jake's model. Of particular interest, the prediction capability for closely separated sinusoids of the LMMSE predictors is investigated by simulations in this section.

The simulation settings are given below:

- Number of sinusoids = 2 ($p = 2$, which is assumed to be known);
- Number of samples = 100 ($N = 100$);
- The first frequency $\omega_1 = 0.2\pi$ and is fixed, and the second frequency $\omega_2 = \omega_1 - \Delta f$, where Δf is the frequency separation, which is adjusted from $\frac{\pi}{20N}$ to $\frac{\pi}{2N}$;
- The complex amplitudes are fixed at $s_1 = s_2^* = e^{j\frac{\pi}{4}}$;
- Spatial channel sampling interval = 0.1 λ ($\Delta l = 0.1$);
- Prediction horizon = 0.5 λ ($L = 5$);
- The SNR is 10 dB.
- Number of Monte Carlo simulations = 200;
- The frequency error variance is used as a design parameter and is taken to be $\sigma^2_{\Delta\omega} = 1/N^3$. This is somewhat arbitrary, but is motivated by the CRLB [Kay93];
- The Unitary-ESPRIT method is used for frequency estimation in the LMMSE prediction.
- Order of LP is 2;

In Fig. 6.1, the NMSE's of the Conditioned LMMSE predictor, the Unconditioned LMMSE predictor and the standard LP are presented. It is observed that the Conditioned LMMSE predictor outperforms the Unconditioned LMMSE predictor when Δf is large, while the opposite result is obtained when Δf is small. These results imply that Unconditioned LMMSE predictor is more suitable for the prediction of closely separated sinusoidal signals than the Conditioned LMMSE predictor.

6.3 Performance Evaluation of Model-Based LMMSE Predictors for a Rayleigh Fading Channel

The performances of the proposed LMMSE channel predictors are evaluated by Monte Carlo simulations. The simulation setup is as follows:

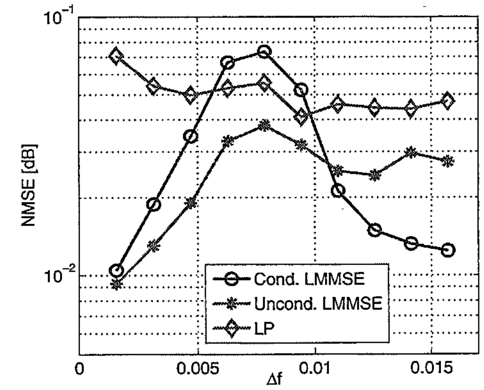

Figure 6.1: LMMSE Prediction methods of nearby tones

- SISO scenario;
- Number of sinusoids = 8 ($p = 8$, which is assumed to be known);
- Number of Monte Carlo simulations = 1000;
- Order of LP is two times the number of sinusoids;
- In the combined method, the sinusoids with $|\hat{s}_k| > 2\sigma_e$ are predicted by the LMMSE method, the residue is predicted using an LP with order 2 ($d = 2$);
- Other settings are the same as in Section 6.2.

In Fig. 6.2, the Cumulative Density Function (CDF) of the NMSE are presented. All sinusoidal model based predictors have light tails, but the LP has a heavy tail.

The ANMSE of different predictors are given in Fig. 6.3. It can be seen that the LMMSE with known frequency and the LP has the best and the worst performance, respectively. The LMMSE predictors have much better performance than the LP in the investigated scenarios. The Unconditioned LMMSE is slightly better than the Conditioned LMMSE predictor. Finally the combined method has extremely good performance, which approaches the performance of the LMMSE with known frequency with the increase of SNR.

Note that, in these simulations, for a given carrier frequency, f, and velocity, v, the spatial sampling interval can be easily converted into time sampling interval as $\Delta t = \frac{\Delta l \cdot c}{f \cdot v}$, where c is the speed of light.

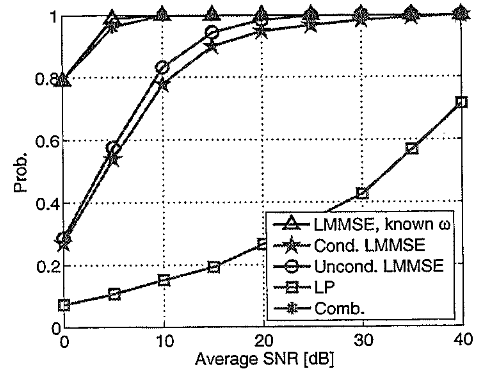

Figure 6.2: Probability of channel prediction error less than 20% ($L = 0.5\lambda$, $N = 100$, 8 Rays)

6.4 Performance Evaluation of Model-Based LMMSE Predictors Using Measurement Data

The measurement data introduced in Chapter 4 is used in this section to evaluate the proposed LMMSE prediction methods.

6.4.1 Model Selection

The *model selection* in the design of LMMSE predictors refers to the selection of the strong and stationary sinusoidal signals estimated from the data block.

In practice, it is probably a wise idea to overestimate the rank, and estimate un-necessarily many sinusoids. But only a subset of these are used in the predictor. The tricky part is to know which ones to keep. With only one data set, one would probably go for the ones with the largest estimated amplitude. With several data sets, one could try to evaluate the stationarity of the various frequency estimates, but this requires a clever sorting (data association) to keep track of the tracks.

In this study, a simple SVD based model selection method is proposed. First the number of the consistent sinusoidal signals is estimated using the i^{th} biggest gradient method. It is simply choosing the index of the descending ordered singular value which gives the i^{th} biggest gradient in linear scale. For example, an m-vector

$$\boldsymbol{\sigma} = [\sigma_1, \cdots, \sigma_{r-1}, \sigma_r, \cdots, \sigma_m]^T \tag{6.2}$$

contains the singular values from the SVD of $\mathbf{Y}_d$, and they are arranged in the descending order. If $(\sigma_{r-1} - \sigma_r)$ gives the i^{th} biggest gradient between adjacent singular values, σ_r is called the *break* point, and the index r is then selected as the number of the consistent sinusoidal signals. In the second step, p frequencies

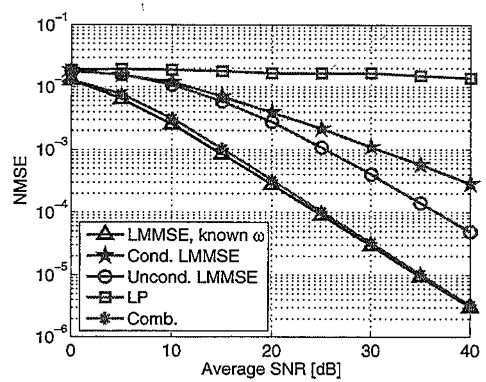

Figure 6.3: Adjusted NMSE of channel prediction ($L = 0.5\lambda$, $N = 100$, 8 Rays)

are estimated from the data block, where $p > r$. The frequencies that have the r biggest estimated amplitudes are then assume to be consistent.

The selection of i in this simple method is environment dependent. In a suburban area, most rays can be expected to be consistent. So we could choose a larger i. But in urban area, this method might not fit, since some sinusoids are strong, but not stationary. So only a few strong paths are assumed to be consistent. This means that a small i should be selected for urban measurements to be conservative.

The design parameter i is chosen as 1 and 3 for urban and suburban measurements, respectively, in the following studies. This is just a rule of thumb according to the measurement data. The searching for better model selection methods is kept for future work.

However, this method is tested by using the same example measurements in urban and suburban area. In Fig. 6.4 and Fig. 6.5, the circles on the curves are the singular values in descending order. The short vertical lines show the location of the break point giving the 1st and 3rd biggest gradient. The estimated number of consistent sinusoids are 2 and 4 respectively in these two example measurements.

The proposed model selection method is also compared with some fixed model order setting, *i.e.*, 2, 6 and 10. The model parameters estimated with these settings are used in the LMMSE prediction methods, respectively. The NMSE's using urban and suburban example measurements are given in Fig. 6.6 and Fig. 6.10, respectively. In these figures, the subplots (a), (b), and (c), correspond to fixed model orders of 2, 6, and 10, respectively. The subplot (d) is the NMSE's using the estimated model order. The data Down Sampling Ratio (DSR) is 10. The data length is 100. The prediction horizon L is 5 in these tests. The SNR's are estimated from the whole data sequences. The variance of $\Delta\omega$ is set to be $1/N^3$. The order of the LP is set to be same as the order of the sinusoidal models.

In Fig. 6.6, it can be seen that the higher the order of the standard LP, the

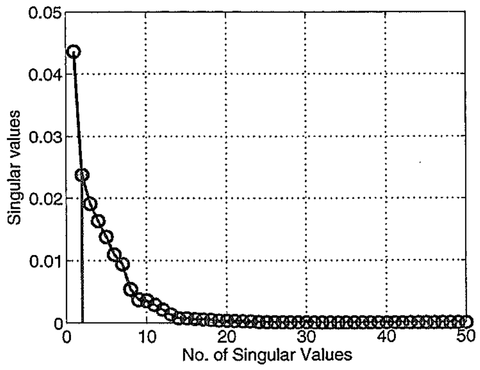

Figure 6.4: Example of model selection (urban measurement)

better the performance. But the difference is small when the order becomes too high.

Both of the best performances of the Conditioned LMMSE predictor and the Unconditioned LMMSE predictor are obtained with fixed order of 2 in the urban measurement. The performances of these two methods degrade *slightly* and *significantly* with the increasing of the fixed order. The performances of these methods with estimated model order, which is given in subplot (d), are close to the best fixed order. In this model selection, the 1^{st} biggest gradient is used. This implies that there are less strong and consistent model parameters in urban environment.

In Fig. 6.7, the frequency response of the Conditioned LMMSE predictors in (5.23) with fixed order 2, 6, and 10 are presented in the subplots (a), (b), and (c) respectively. The subplot (d) is the frequency response of the Conditioned LMMSE predictor with proposed model selection method. Compared with the MPS of this measurement in Fig. 4.6, it can be seen that only a few of the highest power frequency bins can pass through this prediction filter with fixed order 2 and using the proposed model order selection method. With fixed order 6 and 10, there are more high power frequency bins that can pass through the prediction filter, but these frequency bins might be non stationary, so it degrades the prediction performance to allow them pass through the model-based prediction filter as seen in Fig. 6.6.

In Fig. 6.8, the power spectrum of the data blocks in the urban measurement are given in subplot (a), together with the frequency responses of the Conditioned LMMSE predictor in (5.23), the Unconditioned LMMSE predictor in (5.39) and the LP in (3.24) in subplots (b), (c), and (d) respectively. It can be seen that both Conditioned and Unconditioned LMMSE predictors are conservative. Only a small number of the frequencies, which are assumed to be consistent, can pass through the predictor filters. The Unconditioned LMMSE predictor has slightly

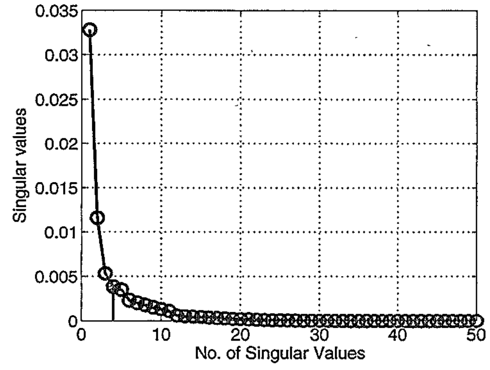

Figure 6.5: Example of model selection (suburban measurement)

wider passband bins. This might explain the improved performance as seen in Fig. 6.6, because the nearby frequencies to the strong and stationary frequency are probably also stationary, and the Unconditioned LMMSE predictor was found to be good at prediction of nearby sinusoids in Section 6.2. The frequency response of the LP in this measurement is a high pass filter, but the channel is a band limited low pass signal as seen in subplot (a). Even though the LP gives the best performance in this test. This is somewhat unexpected. But a close look of these frequency responses and the signal power spectrum gives the answer. In Fig. 6.9, the aligned signal power spectrum and frequency responses of channel predictors in the urban measurement is given. It can be seen that the signal is band limited. The strongest frequency component is at the negative maximum Doppler frequency, while there are many other smaller but significant frequency components distributed over the band limited spectrum. The model-based LMMSE predictors only allow the highest frequency bin at negative maximum Doppler frequency pass through the prediction filter, and all the other frequency components are heavily attenuated. The LP performs as a high pass, but *flat* filter, which not only allows the high power bins at plus and minus maximum Doppler frequency pass through the prediction filter, but also the low frequency components are filtered with limited attenuation. Actually this is exactly what the Jake's model tells.

In Fig. 6.10, the suburban example measurement is used. It can be seen, still, that the performance of LP with higher order is better than those with lower orders. But the difference is small when the orders are high.

The Conditioned LMMSE predictors with fixed order 10 outperforms the other two fixed orders. But it is just marginally better than the fixed order of 6. The performance of the Unconditioned LMMSE predictor with fixed order of 6 outperforms other fixed orders. By using estimated order, its performance is even better, which outperforms the LP in this measurement. In this order estimate,

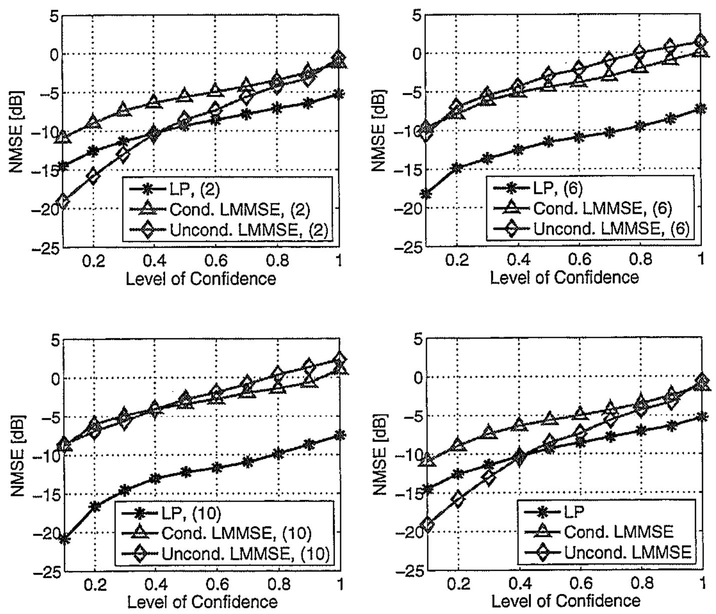

Figure 6.6: Example of the influence of model selection (urban)

the 3rd biggest gradient is used. This is because there are more consistent model parameters in suburban environments.

In Fig. 6.11, the frequency response of the Conditioned LMMSE predictors (5.23) with fixed order 2, 6, and 10 in this suburban measurement are given in the subplots (a), (b), and (c) respectively. The subplot (d) is the frequency response of the Conditioned LMMSE predictor with our proposed model selection method. Compared with the MPS of this measurement in Fig. 4.7, it can be seen that the two highest power frequency bins are very well captured by the prediction filter with order 2. These two highest power bins are also well followed by the predictor with the proposed model selection method as seen in (d), where a wider bin width are observed. Since the model parameters are consistent, to allow more frequencies to pass through the predictor filter can help to improve the prediction accuracy, which has been seen in Fig. 6.10. This also probably explains that the Conditioned LMMSE prediction is the best at $p = 10$, and the Unconditioned LMMSE predictor is the best at $p = 6$ as seen in Fig. 6.10.

In Fig. 6.12, the power spectrum of the data blocks of the suburban mea-

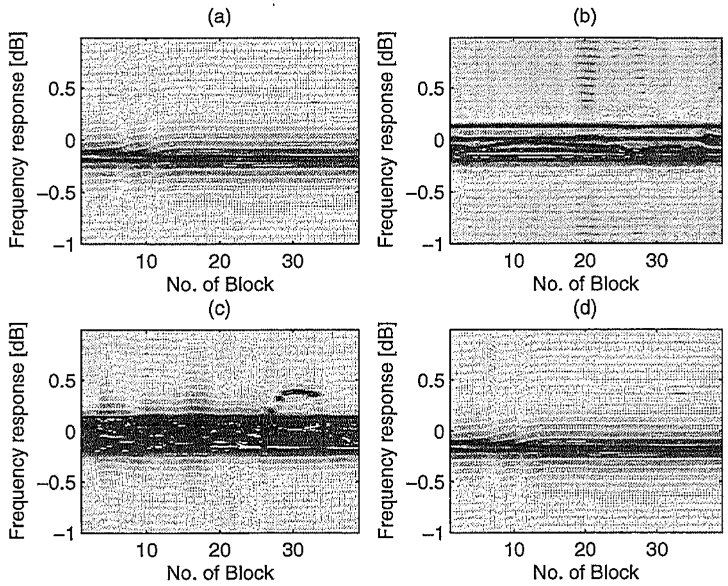

Figure 6.7: Example of frequency response of Conditioned LMMSE predictors with fixed order (urban); $p = 2$ (a), $p = 6$ (b), $p = 10$ (c). In (d), the estimated order is used.

surement are given in subplots (a), together with the frequency responses of the Conditioned LMMSE predictor in (5.23), the unconditioned LMMSE predictor in (5.39) and the LP in (3.24) in subplot (b), (c), and (d) respectively. It can be seen that both the Conditioned and Unconditioned LMMSE predictors capture the high power frequency bins very well. The Conditioned LMMSE predictor has the narrowest pass band width. The pass band width is wider of the LP. The Unconditioned LMMSE predictor has the widest pass band width. In this case, the frequencies are much consistent and its spectrum is more like discrete spectral lines. This indicates the sinusoidal model based method might be better as seen in subplot (d) in Fig. 6.10.

In Fig. 6.13, the aligned frequency responses of the channel predictors and the channel power spectrum is given. It can be seen that the high power frequency components are very well captured by the LMMSE predictors. The Unconditioned LMMSE predictor has less attenuation on the frequencies aside the frequency peaks than what the Conditioned LMMSE predictor does. The LP still has a high pass property.

These results show that the sinusoidal modeling based LMMSE prediction

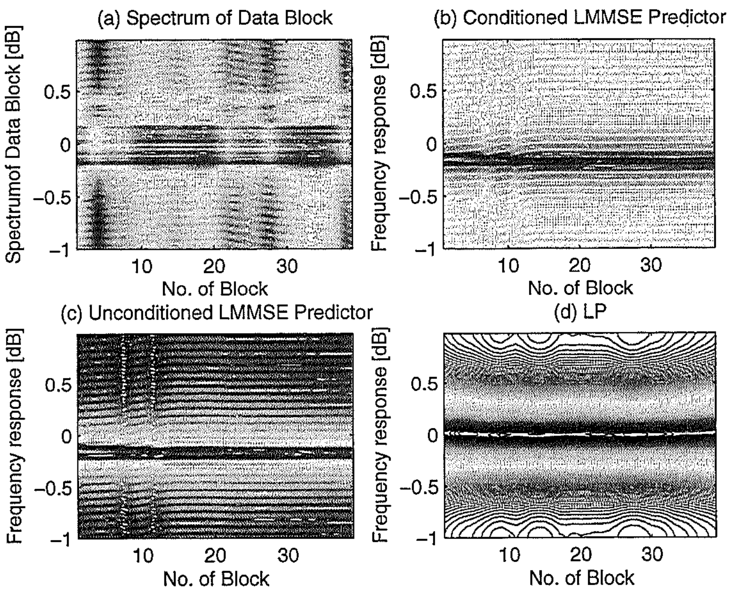

Figure 6.8: Example of frequency response of Model-based LMMSE predictors and LP in urban area

methods are *sensitive* to model selection. Both under estimate and over estimate the order of the *consistent* signal model can degrade the prediction accuracy. In general, the model parameters are more stationary in a suburban environment than in an urban environment.

6.4.2 Performance Evaluations of LMMSE Prediction Methods Using Measurement Data

The original data is down sampled by a DSR of 10 to reduce the calculation load and increase the prediction horizon. After the down sampling, the evaluation parameter setting are:

- The length of data block N is set to be 100, which is corresponding to a measurement interval of 0.92 m or 2.75 m in distance at the speed of 30 Km/h or 90 Km/h respectively;
- The prediction horizon L is 5, which is about $\frac{\lambda}{3}$ or 1λ at the same speeds and the frequency band of 1880 MHz;

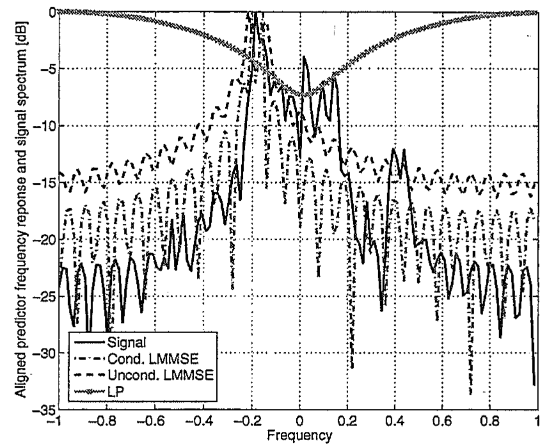

Figure 6.9: Example of aligned signal power spectrum and frequency responses of channel predictors in the urban measurement

- The SNR's used in the Conditioned LMMSE prediction methods are estimated from the relative power delay profiles;
- The order of the LP in the Combined LMMSE and LP methods are 2, which is fixed;
- Unitary ESPRIT is used for frequency estimation;
- The order of the standard LP is set to be the same as the signal order estimated for the data block;
- The 1^{st} and 3^{rd} biggest gradient is used for the model selection based on SVD for urban and suburban environments respectively.

With these settings there are 39 blocks of data in each measurement. In total, 1755 and 1209 blocks of data are obtained in urban and suburban area respectively.

In Fig. 6.14 the overall NMSE's of different prediction methods using all measurement data in urban area is presented vs. the level of confidence. It is meaningless to discuss the achievable NMSE's in this case due to the limited number of measurements. We put our attentions on the relative performance instead. In this figure, the LP has the best performance. The Conditioned LMMSE predictor has the worst performance. The Unconditioned LMMSE predictor outperforms the Conditioned LMMSE predictor. The Combined LMMSE and LP predictor with order 2 outperforms the Conditioned LMMSE predictor and the Unconditioned LMMSE predictor, and is slightly worse than LP.

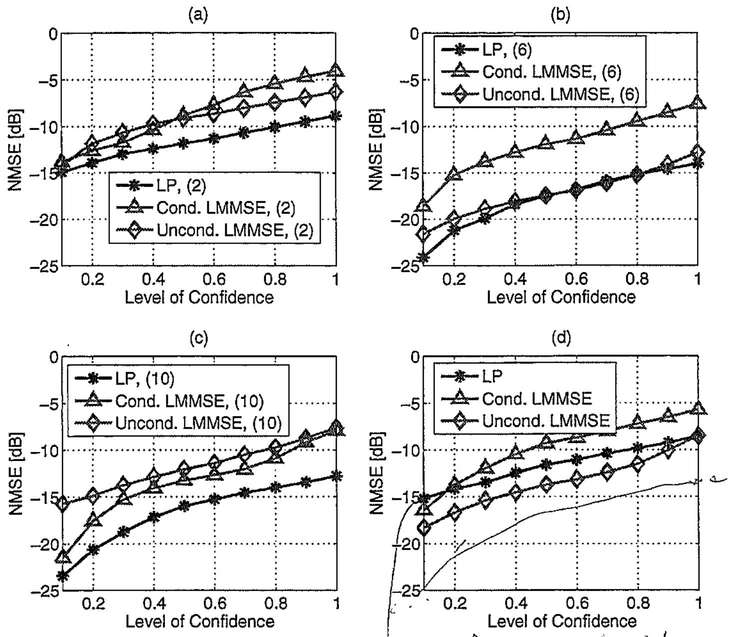

Figure 6.10: Example of the influence of model selection (suburban)

In Fig. 6.15, the results using suburban measurements are given. Similar results are observed. In these results, the performances of the Conditioned LMMSE and the Unconditioned LMMSE predictor are even closer compared to those in urban area. This is mainly due to that the frequency separation is larger in suburban than in urban environments. The Combined LMMSE and LP predictor outperforms the LMMSE prediction methods and underperforms the LP.

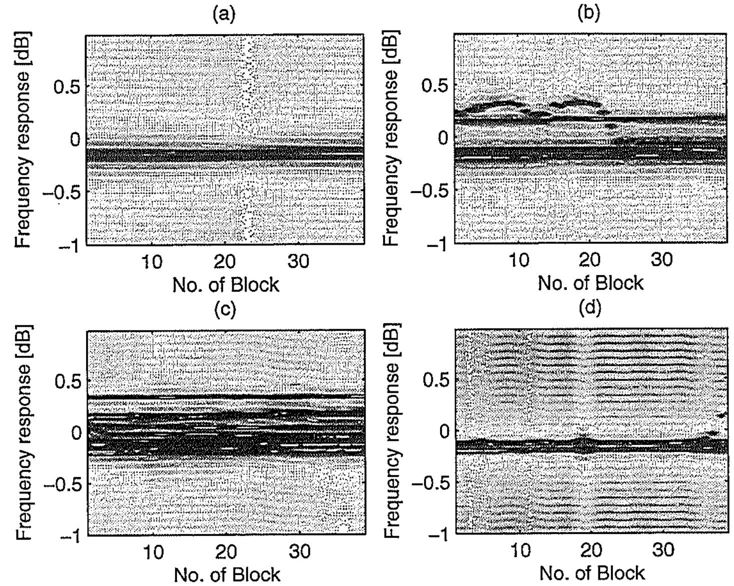

Figure 6.11: Example of frequency response of Conditioned LMMSE predictors with fixed order (suburban); $p = 2$ (a), $p = 6$ (b), $p = 10$ (c). In (d), the estimated order is used.

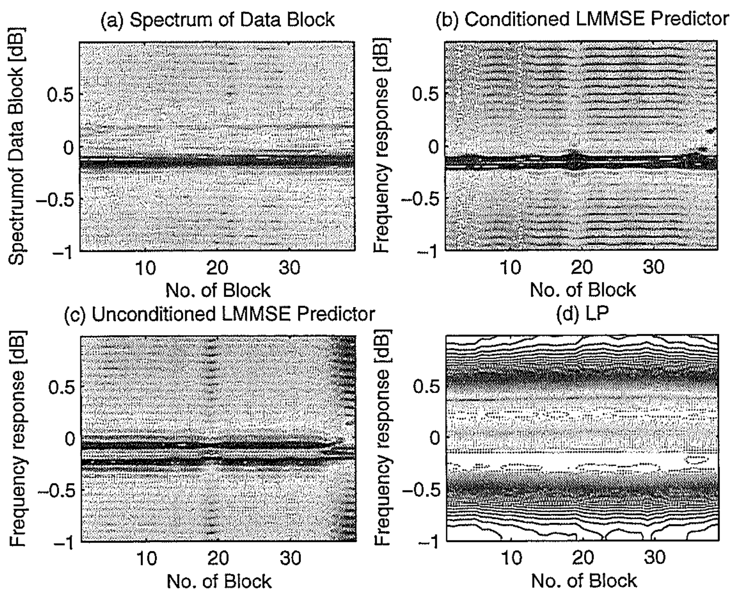

Figure 6.12: Example of frequency response of Model-based LMMSE predictors and LP in suburban area

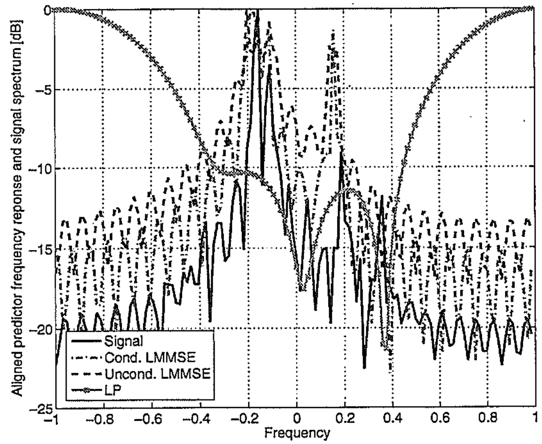

Figure 6.13: Example of aligned signal power spectrum and frequency responses of channel predictors in the suburban measurement

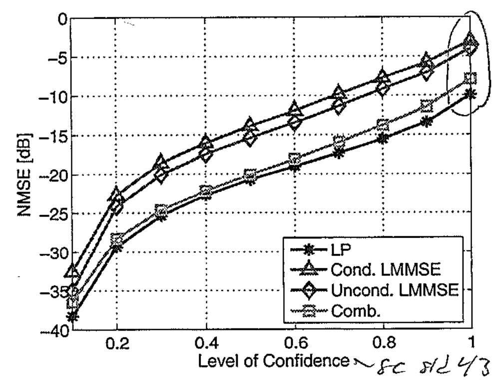

Figure 6.14: Performance evaluation of LMMSE prediction methods in urban area

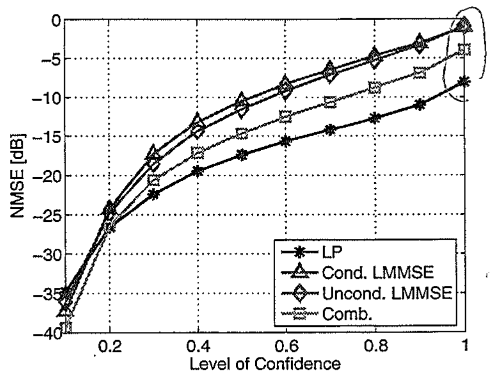

Figure 6.15: Performance evaluation of LMMSE prediction methods in suburban area

Chapter 7

Joint LMMSE Predictors

In previous chapters, it was found that the real channels cannot be correctly described by the sinusoidal modeling due to the non station of model parameters. This motivates us to model the channel prediction using a Joint Moving Average and Sinusoidal (JMAS) model. The associated channel predictor is termed Joint LMMSE predictor.

7.1 Joint Sinusoidal Modeling and Non-Parametric Predictor

Let

$$y(t) = x(t) + e(t), \tag{7.1}$$

where $y(t)$, $x(t)$, and $e(t)$ are the channel estimate, true channel, and estimation error respectively as before. Further, let the estimate of $x(t+L)$ to be

$$\hat{x}(t+L) = \mathbf{y}_d^T \boldsymbol{\beta}_d + \hat{\mathbf{a}}(L)^H \mathbf{s}, \tag{7.2}$$

where

$$\boldsymbol{\beta}_d = [\beta_t, \cdots, \beta_{t-d+1}]^T, \tag{7.3}$$
$$\mathbf{y}_d = [y(t), \cdots, y(t-d+1)]^T, \tag{7.4}$$

and $\hat{\mathbf{a}}(L)$ is defined as in (5.18), and assume $\hat{\omega}$ is given. With this model, the channel is divided into two parts. One part contains the consistent and strong sinusoidal signals, i.e. the direct rays in LOS scenarios. The second part captures all the remaining power in the data. This prediction model is called the Joint Moving Average and Sinusoidal (JMAS) model.

7.2 Joint LMMSE Predictors

There are different methods to estimate the model parameters in (7.2). One approach is given in the Combined LMMSE and LP predictor in Section 5.3. But it was observed that the sinusoidal model errors introduce extra system dynamics into the residue signal, $\epsilon(t)$, which has to be taken care by LP. An iterative estimate of the model parameters in the sinusoidal part and the MA part might help to improve the performance. But it increases the calculation complexity. However these model parameters can be estimated jointly due to the linearity of the model.

Given $\hat{\omega}$, (7.2) can be rewritten as

$$\hat{x}(t+L) = \left[\ \mathbf{y}_d^T \ \ \hat{\mathbf{a}}(L)^H \ \right] \begin{bmatrix} \beta \\ \mathbf{s} \end{bmatrix}. \tag{7.5}$$

In matrix form,

$$\begin{aligned} \hat{\mathbf{x}}_{N-L-d+1} &= [\mathbf{Y}_J \ \mathbf{A}_J]\boldsymbol{\theta} && (7.6) \\ &= \mathbf{H}\boldsymbol{\theta}, && (7.7) \end{aligned}$$

where

$$\hat{\mathbf{x}}_{N-L-d+1} = [\hat{x}(t), \hat{x}(t-1), \cdots, \hat{x}(t+L-N+d)]^T, \tag{7.8}$$

$$\mathbf{Y}_J = \begin{bmatrix} y(t-L) & y(t-L-1) & \cdots & y(t-L-d+1) \\ y(t-L-1) & y(t-L-2) & \cdots & y(t-L-d) \\ \vdots & \vdots & \vdots & \vdots \\ y(t-N+d) & y(t-N+d-1) & \cdots & y(t-N+1) \end{bmatrix} \tag{7.9}$$

$$\mathbf{A}_J = \begin{bmatrix} e^{j\omega_1(t+L)} & e^{j\omega_2(t+L)} & \cdots & e^{j\omega_p(t+L)} \\ e^{j\omega_1(t+L-1)} & e^{j\omega_2(t+L-1)} & \cdots & e^{j\omega_p(t+L-1)} \\ \vdots & \vdots & \vdots & \vdots \\ e^{j\omega_1(t+2L-N+d)} & e^{j\omega_2(t+2L-N+d)} & \cdots & e^{j\omega_p(t+2L-N+d)} \end{bmatrix}, \tag{7.10}$$

Then the LS estimate of $\boldsymbol{\theta}$ is obtained by solving

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \arg\min_{\boldsymbol{\theta}} \|\mathbf{y}_{N-L-d+1} - \hat{\mathbf{x}}_{N-L-d+1}\|^2 \\ &= \arg\min_{\boldsymbol{\theta}} \|\mathbf{y}_{N-L-d+1} - \mathbf{H}\boldsymbol{\theta}\|^2, && (7.11) \end{aligned}$$

where

$$\mathbf{y}_{N-L-d+1} = [y(t), y(t-1), \cdots, y(t-N+L+d)]^T. \tag{7.12}$$

Finally, the LS estimate of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{y}_{N-L-d+1}. \tag{7.13}$$

and the channel prediction is

$$\hat{x}(t+L) = [\mathbf{y}_d^T \ \hat{\mathbf{a}}(L)^H]\hat{\boldsymbol{\theta}}. \tag{7.14}$$

Note that although the parameter estimation is the LS solution, the channel predictor is still an LMMSE predictor. So (7.14) is named Joint LMMSE predictor. In the terminology of the previous chapter, it is a conditional LMMSE predictor, since $\boldsymbol{\theta}$ is assumed to be given and s is modeled as deterministic. As before, the frequency estimate can be obtained using Unitary ESPRIT or IVM.

7.3 Performance Evaluation of Joint LMMSE Predictors Using Measured Data

In this study, most parameter settings are the same as in Section 6.4.2 except the following:

- IVM is used for frequency estimation, where the temporal displacement is 6;
- The order of standard LP is 4 and fixed.
- The order of the LP in Joint LMMSE predictor is 3 and fixed.
- The number of sinusoids is 1, where only the strongest path is selected in this test;

In Fig. 7.1, the overall NMSE's of different prediction methods using measurement data in urban area is presented vs. the level of confidence as before. In this case we see that the LP still performs (slightly) better than the others. This might due to the non station of the model parameters, and the poor performance of the model selection method, since the strong sinusoids might not have consistent parameters, while those consistent sinusoids is not necessarily the strongest, especially in an urban environment. But the Joint LMMSE predictor is better than the Combined LMMSE and LP predictor methods. This improvement is due to the joint estimate of the model parameters.

In Fig. 7.2, the measurements from the suburban area are used, where the Joint LMMSE predictor was found to be the best. This might be because the strongest path is better described by the sinusoidal model than the non-parametric method.

To further investigate the relation between the prediction accuracy and the SNR of the data block, the NMSE is plotted as a function of the mean power of the data block. They are given in Fig. 7.3 and Fig. 7.4 for urban and suburban area respectively. It can be seen in both figures that the NMSE's of all methods decrease with the increase of the mean powers (or SNR), which agrees with our intuition. From these two figures, the absolute mean square errors can be easily read out, which is the sum of the coordinates of the dots on the curves. For example, in Fig. 7.3, the MSE for a block of data with mean power of -55 dB is about $-55 - 20 = -75$ dB. The discontinuity and the non smoothness of these curves are due to the small number of measurements and non uniformly distributed mean powers.

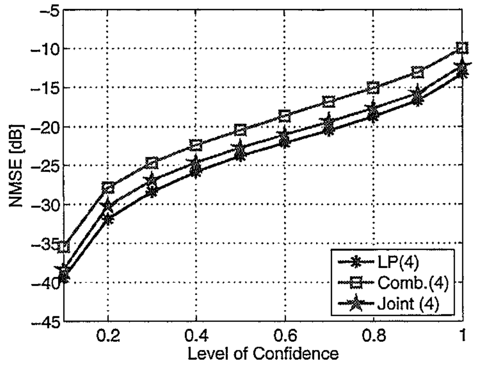

Figure 7.1: NMSE vs. Level of Confidence (urban area, $d = 4$)

In Fig. 7.5 and 7.6, the NMSE's are given as functions of the power of the channel to be predicted, $|x(t+L)|^2$, for urban and suburban measurements respectively. Similar results are found in these figures. This is mainly due to the high correlation between $|x(t+L)|^2$ and the mean power of the data block.

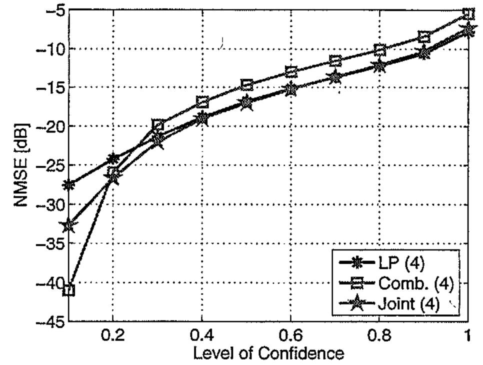

Figure 7.2: NMSE vs. Confidential Level (suburban area, $d = 4$)

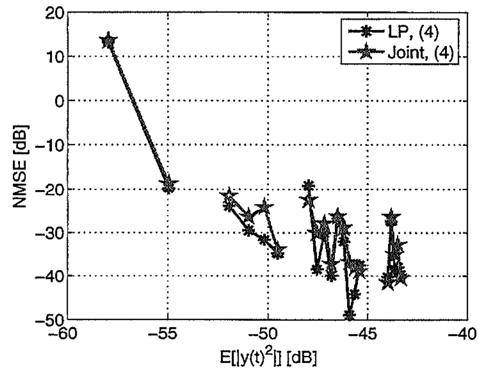

Figure 7.3: NMSE vs. mean power (Urban area, $d = 4$)

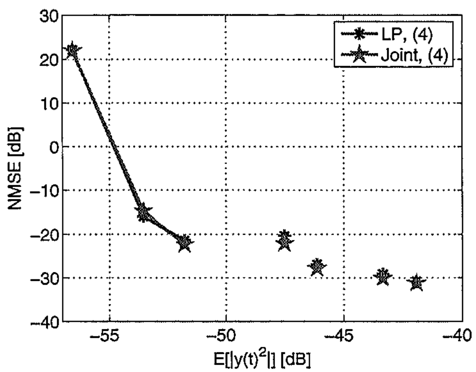

Figure 7.4: NMSE vs. mean power (suburban area, $d = 4$)

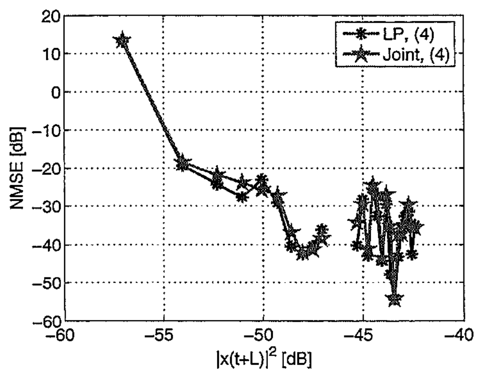

Figure 7.5: NMSE vs. prediction power (Urban area, $d = 4$)

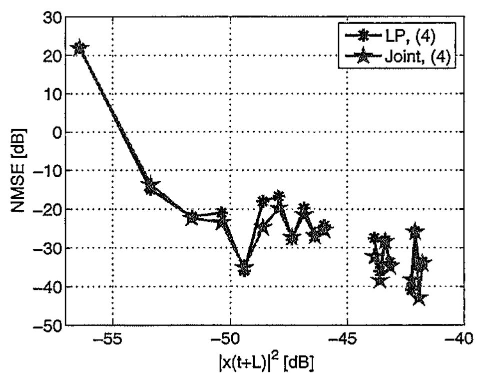

Figure 7.6: NMSE vs. prediction power (suburban area, $d = 4$)

Chapter 8

Conclusions and Future Works

8.1 Conclusions

The problem of the prediction of flat Rayleigh fading channels in wireless communications is studied in the frame of sinusoidal modeling in this thesis.

A stochastic sinusoidal model to Rayleigh fading channel is proposed based on the nature of multi-path propagation environment in wireless communications. The average of the conditional power spectrum of this model is proved to be the Jake's model.

Estimation methods of the model parameters are reviewed. The CRLB for the frequency estimate is derived. An algorithm to calculate the compressed CRLB is proposed when frequencies are closely separated.

The Jake's model is confirmed by the NMDS in both urban and suburban measurements. The tracking of the time variation property of the model parameters is made. It was found that the model parameters belonging to strong sinusoidal signals are more stationary in suburban environments than in urban environments.

Based on the proposed sinusoidal modeling, a regularized LS estimate of the signal amplitudes is proposed, which is named LMMSE amplitude estimation, given frequency estimates. Further a Conditioned LMMSE channel prediction is proposed. By modeling the frequency estimation error as IID Gaussian random variables, an Unconditioned LMMSE prediction is proposed, where the modeling of frequency errors contributes an exponential weighting on the data. The performances of these model-based LMMSE predictors are evaluated by simulations and measurement data, where a simple SVD based model selection algorithm is proposed. In the simulations, the LMMSE prediction methods performs much better than the standard LP. The Unconditioned LMMSE predictor outperforms the Conditioned LMMSE predictor. But the LP predictor is found to be better using measurement data. The worse performances of LMMSE prediction methods are mainly due to the non station of the model parameters, especially in urban environments. A Combined LMMSE and LP predictor is proposed to fix the problem of modeling errors. It outperforms the Conditioned and Unconditioned LMMSE prediction methods, but is still worse than the standard LP.

A subjective study of the LMMSE prediction methods to nearby tones are performed by simulations. The Unconditioned LMMSE predictor outperforms the Conditioned LMMSE predictor in case of closely separated frequencies. An SVD based the i^{th} biggest gradient rank estimation method is proposed. It works well in suburban measurements, but poorly in urban environment.

A Joint Moving Average and Sinusoidal (JMAS) model is proposed for channel prediction methods. It splits the signal into two parts, which are described by an MA processing and sinusoidal model jointly. Based on the JMAS model, the Joint LMMSE predictor is proposed. Its performance is evaluated by using the measurement data. It is found that the Joint LMMSE predictor outperforms the standard LP in suburban environment, but under performs in urban area. This is probably due to the poor performance of the model selection algorithm in urban environment.

8.2 Discussions and Future Works

The Joint LMMSE predictor has been shown to be a promising channel prediction technique. So far it works well in suburban and probably in rural area with the proposed simple model selection method. Still it is possible to improve its performance in an urban area by improving the model selection algorithm. In the design of the Joint LMMSE predictor with fixed order of linear regressors, the order allocation to MA and sinusoidal parts plays an important role. Further study is needed to find good strategies for selecting the orders.

The channel prediction methods which have been studied so far, focus on narrow band channels or the channels in a single frequency bin in OFDM systems. In wide band systems or OFDM system, the channels most likely experience frequency selective fading. Typical channel estimation techniques in OFDM systems are allocating pilot symbols over a number of equally separated frequency bins. Channel prediction can be performed on these pilot bins. Then the frequency domain interpolation can be used to predict the non-pilot bins.

In the sinusoidal modeling based channel prediction methods, the Doppler frequencies are the core model parameters. In reality, the Doppler frequency is due to the multi-path propagation, and the moving mobile terminal or reflecting objects. But the fundamental physical quantity is the wave length of the sinusoidal signals. It is reasonable to assume the structures of the multi-path propagations of the neighboring OFDM bins to be highly correlated. So they might have highly correlated model parameters, *i.e.* Doppler frequencies. This guess makes it interesting to investigate the correlation of the model parameters between frequency bins. Such a knowledge might help to improve the prediction accuracy in OFDM systems.

Regarding the measurement data, the study has been performed on the strongest tap in the channel impulse response. It is also interesting to have similar studies on the second and following taps. It would of course be beneficial to take into account the correlation (of any) that exists between the different taps. This gives a 2D structure, even in the SISO case. It might lead to very computationally costly algorithms in the SIMO and MIMO case, but it might improve performance.

In practice, an adaptive channel prediction based on sinusoidal modelling is of interest to track the time varying model parameters and to reduce the calculation complexity. Keeping track of the parameter's time evolution can also help deciding which signal components to use in the predictor. This aspect will be studied in the future work.

Appendix A

Instrumental Variable Method for Frequency Estimation in Colored Noise

The classical so-called eigenvector (EV) techniques, such as ESPRIT, MUSIC etc., can produce highly biased frequency estimates in colored noise or interference. An Instrumental Variable Method (IVM) to frequency estimation in colored noise is proposed in [SVO94].

Assume the noise $\mathbf{e}$ in signal model (2.11) is colored with zero mean and covariance matrix $E[\mathbf{e}\mathbf{e}^H] = \mathbf{Q}$, which is a general positive definite matrix. Then the covariance matrix $\mathbf{R}_{yy}$ in (3.47) becomes

$$\mathbf{R}_{yy} = \mathbf{A}\mathbf{S}\mathbf{A}^H + \mathbf{Q}. \tag{A.1}$$

The signal subspace and the noise subspace cannot be separated by the eigenvalue decomposition as in (3.51). However if the colored noise $e(t)$ could be modelled as a q^{th} order Moving Average (MA) process, which is

$$e(t) = u(t) + b_1 u(t-1) + \cdots + b_q u(t-q) \tag{A.2}$$

where $u(t)$ is complex white Gaussian noise with zero mean and variance σ_u^2. The noise covariance matrix $\mathbf{Q} = E[\mathbf{e}\mathbf{e}^H]$, which is a banded Hermitian Toeplitz matrix with first row of $[q_0, q_1, \cdots, q_q, 0, \cdots, 0]^T$, where $q_k = E[e(t)e(t-k)^*]$, and $q_k = 0$, when $k > q$. Motivated by this property, the signal subspace can be separated from the noise subspace by using the IVM.

Define the data vector $\boldsymbol{\zeta}_y$ and instrumental vector $\boldsymbol{\psi}_y$ as

$$\begin{aligned} \boldsymbol{\zeta}_y &= [y(\tau), y(\tau-1), \cdots, y(\tau-K+1)]^T, & \text{(A.3)} \\ \boldsymbol{\psi}_y &= [y(\tau-M), y(\tau-M-1), \cdots, y(\tau-M-L+1)]^T, & \text{(A.4)} \end{aligned}$$

where $L \geq K \geq q$, and M is the temporal displacement. The cross-correlation $\mathbf{R}_{\psi\zeta} = E[\boldsymbol{\psi}_y \boldsymbol{\zeta}_y^H]$ is

$$\mathbf{R}_{\psi\zeta} = \begin{bmatrix} r_{yy}(M) & r_{yy}(M-1) & \cdots & r_{yy}(M-K+1) \\ r_{yy}(M+1) & r_{yy}(M) & \cdots & r_{yy}(M-K+2) \\ \vdots & \vdots & \vdots & \vdots \\ r_{yy}(M+L-1) & r_{yy}(M+L-2) & \cdots & r_{yy}(M+L-K) \end{bmatrix} \tag{A.5}$$

$$= \begin{bmatrix} r_{xx}(M) & r_{xx}(M-1) & \cdots & r_{xx}(M-K+1) \\ r_{xx}(M+1) & r_{xx}(M) & \cdots & r_{xx}(M-K+2) \\ \vdots & \vdots & \vdots & \vdots \\ r_{xx}(M+L-1) & r_{xx}(M+L-2) & \cdots & r_{xx}(M+L-K) \end{bmatrix} + \begin{bmatrix} r_{ee}(M) & r_{ee}(M-1) & \cdots & r_{ee}(M-K+1) \\ r_{ee}(M+1) & r_{ee}(M) & \cdots & r_{ee}(M-K+2) \\ \vdots & \vdots & \vdots & \vdots \\ r_{ee}(M+L-1) & r_{ee}(M+L-2) & \cdots & r_{ee}(M+L-K) \end{bmatrix} \tag{A.6}$$

Since $r_{ee}(\tau) = 0$ for $\tau > q$. It can be seen that when $(M - K + 1) > q$, the second term in the last equation in (A.6) becomes zeros. This means that the sinusoidal signal subspace can be obtained from SVD of $\mathbf{R}_{\psi\zeta}$ without interference of the colored noise.

In practice, assume $\mathbf{y} = [y(t), y(t-1), \cdots, y(t-N+1))]^T$ is observed, the estimate of $\mathbf{R}_{\zeta\psi}$ can be obtained as

$$\hat{\mathbf{R}}_{\zeta\psi} = \frac{1}{N}\mathbf{Z}^T\bar{\mathbf{Y}}, \tag{A.7}$$

$$\tag{A.8}$$

where

$$\mathbf{Z} = \begin{bmatrix} y(t-N+L) & \cdots & y(t-N+1) \\ \vdots & \vdots & \vdots \\ y(t-M-1) & \cdots & y(t-M-L) \end{bmatrix}, \tag{A.9}$$

$$\mathbf{Y} = \begin{bmatrix} y(t-N+M+L) & \cdots & y(t-N+M+L-K+1) \\ \vdots & \vdots & \vdots \\ y(t-1) & \cdots & y(t-K) \end{bmatrix}. \tag{A.10}$$

Both $\mathbf{Z}$ and $\mathbf{Y}$ are Toeplitz matrices. It is worth to note that, with limited number of observations, the IVM frequency estimate produces unbiased frequency estimates, but increased variance due to reduced number of effective number of observations. The minimum variance is obtained when $M = K + q - 1$.

Appendix **B**

Calculation of Compressed CRLB

The CRLB of the frequency estimate is derived in Section 3.4. But it turns out that this is not a easy task due to the nature of the stochastic sinusoidal modeling. In fact, if the random frequency components are close, the Fisher information matrix will be poorly conditioned. However, it is no problem with most practical frequency estimators. If two frequencies are too close, they will be represented by only one estimate, somewhere in between the two. This will not harm the prediction performance notably. Based on this consideration, a method to calculate the *compressed* CRLB of the frequency estimate is developed as follows:

1. Generate and sort p independent random Doppler frequencies with PDF (2.8);
2. Find the minimum frequency interval between the p frequencies, say $\Delta\omega_{min} = |\omega_i - \omega_{i+1}|$;
3. Calculate the Fisher information matrix and the covariance matrix of the frequency estimate $\mathbf{C}_\omega$.
4. Find the maximum standard deviation of the frequency estimate, $\sigma_{\omega,max}$;
5. If $\Delta\omega_{min} < \gamma\sigma_{\omega,max}$, let the mean frequency $(\omega_i + \omega_{i+1})/2$ substitute ω_i and ω_{i+1} and let $p = p - 1$, where γ is a design parameter;
6. Go back to step 2 and repeat till no more frequency pairs are merged.

The resulting frequencies in the above algorithm are considered as detectable frequencies denoted as $\boldsymbol{\omega}_D$. The CRLB calculated based on $\boldsymbol{\omega}_D$ is named the compressed CRLB. The best possible estimated frequencies are then generated as

$$\omega_{best} = \boldsymbol{\omega}_D + \mathbf{C}_\omega^{1/2}\mathbf{e}_\omega. \tag{B.1}$$

where $\mathbf{e}_\omega$ is the random frequency estimate error, distributed as $\mathcal{N}(0, \mathbf{I})$.

Appendix C

Equivalency of LMMSE and MMSE Predictors

For the given signal model in (2.11), the Conditioned LMMSE predictor of $x(t+L)$ is given in (5.23). Both the signal model and the predictor are represented here for convenience.

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{e}, \tag{C.1}$$

and

$$\hat{x}(t+L) = \hat{\mathbf{a}}(L)^H \mathbf{R}_\alpha^{-1} \hat{\mathbf{A}}^H \mathbf{y}. \tag{C.2}$$

In this appendix, the Conditioned LMMSE predictor in (C.2) is shown to be equivalent to the MMSE predictor given observations $\mathbf{y}$ and frequency estimates $\hat{\boldsymbol{\omega}}$.

Assume that the amplitude vector $\mathbf{s}$ is a Gaussian random vector with PDF, $\mathcal{CN}(\mathbf{0}, \sigma_s^2 \mathbf{I}_p)$. The MMSE prediction of $x(t+L)$ given $\mathbf{y}$ and $\hat{\boldsymbol{\omega}}$ is

$$\begin{aligned} \hat{x}(t+L) &= E[x(t+L)|\mathbf{y}, \hat{\boldsymbol{\omega}}], \\ &= E\left[\hat{\mathbf{a}}(L)^H \mathbf{s}|\mathbf{y}, \hat{\boldsymbol{\omega}}\right] \\ &= \hat{\mathbf{a}}(L)^H E\left[\mathbf{s}|\mathbf{y}, \hat{\boldsymbol{\omega}}\right], \end{aligned} \tag{C.3}$$

where $E[\mathbf{s}|\mathbf{y}, \hat{\boldsymbol{\omega}}]$ is the MMSE estimate of $\mathbf{s}$ given $\mathbf{y}$ and $\hat{\boldsymbol{\omega}}$. The Bayes's rule is used to find the PDF of $\mathbf{s}$ conditioned on $\mathbf{y}$ and $\hat{\boldsymbol{\omega}}$, $f(\mathbf{s}|\mathbf{y}, \hat{\boldsymbol{\omega}})$. We write

$$f(\mathbf{s}|\mathbf{y}, \hat{\boldsymbol{\omega}}) = \frac{f(\mathbf{s}, \mathbf{y}|\hat{\boldsymbol{\omega}})}{f(\mathbf{y}|\hat{\boldsymbol{\omega}})} = \frac{f(\mathbf{y}|\mathbf{s}, \hat{\boldsymbol{\omega}}) f(\mathbf{s})}{f(\mathbf{y}|\hat{\boldsymbol{\omega}})}. \tag{C.4}$$

The numerator in (C.4) can be expressed as

$$f(\mathbf{y}|\mathbf{s},\hat{\omega})f(\mathbf{s}) = \frac{1}{\pi^N\sigma_e^{2N}}e^{\frac{-(\mathbf{y}-\hat{\mathbf{A}}\mathbf{s})^H(\mathbf{y}-\hat{\mathbf{A}}\mathbf{s})}{\sigma_e^2}} \cdot \frac{1}{\pi^p\sigma_s^{2p}}e^{-\frac{\mathbf{s}^H\mathbf{s}}{\sigma_s^2}} \tag{C.5}$$

Define $\mathbf{R}_\alpha = \left(\hat{\mathbf{A}}^H\hat{\mathbf{A}} + \frac{\sigma_e^2}{\sigma_s^2}\mathbf{I}_p\right)$ as in (5.15) and complete the square, (C.5) can be written as

$$\begin{aligned} f(\mathbf{y}|\mathbf{s},\hat{\omega})f(\mathbf{s}) &= \frac{1}{\pi^p|\sigma_e^2\mathbf{R}_\alpha^{-1}|}e^{-(\mathbf{s}-\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y})^H(\sigma_e^2\mathbf{R}_\alpha^{-1})^{-1}(\mathbf{s}-\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y})} \\ &\quad \cdot|\alpha\mathbf{R}_\alpha^{-1}|\frac{1}{\pi^N\sigma_e^{2N}}e^{\frac{-(\mathbf{y}^H\mathbf{y}-\mathbf{y}^H\hat{\mathbf{A}}^H\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}\mathbf{y})}{\sigma_e^2}}, \end{aligned} \tag{C.6}$$

where $|\mathbf{R}_\alpha|$ is the determinant of $\mathbf{R}_\alpha$. In (C.6), the first term is the PDF of a complex Gaussian multivariate s with PDF of $\mathcal{CN}(\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y},\sigma_e^2\mathbf{R}_\alpha^{-1})$, while the second term is independent on s.

The denominator in (C.4) is

$$f(\mathbf{y}|\hat{\omega}) = \int f(\mathbf{y}|\mathbf{s},\hat{\omega})f(\mathbf{s})ds, \tag{C.7}$$

where the integrand is the same as the numerator in (C.4). So we have

$$\begin{aligned} f(\mathbf{y}|\hat{\omega}) &= \int \frac{1}{\pi^p|\sigma_e^2\mathbf{R}_\alpha^{-1}|}e^{-(\mathbf{s}-\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y})^H(\sigma_e^2\mathbf{R}_\alpha^{-1})^{-1}(\mathbf{s}-\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y})}ds \\ &\quad \cdot|\alpha\mathbf{R}_\alpha^{-1}|\frac{1}{\pi^N\sigma_e^{2N}}e^{\frac{-(\mathbf{y}^H\mathbf{y}-\mathbf{y}^H\hat{\mathbf{A}}^H\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}\mathbf{y})}{\sigma_e^2}}, \end{aligned} \tag{C.8}$$

where the first term integrates to one. Finally,

$$f(\mathbf{y}|\hat{\omega}) = |\alpha\mathbf{R}_\alpha^{-1}|\frac{1}{\pi^N\sigma_e^{2N}}e^{\frac{-(\mathbf{y}^H\mathbf{y}-\mathbf{y}^H\hat{\mathbf{A}}^H\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}\mathbf{y})}{\sigma_e^2}}. \tag{C.9}$$

Not surprisingly, this is exactly the second term in (C.6). Substituting (C.6) and (C.9) into (C.4), we get the complex Gaussian PDF as $\mathcal{CN}(\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y},\sigma_e^2\mathbf{R}_\alpha^{-1})$, which is given below.

$$f(\mathbf{s}|\mathbf{y},\hat{\omega}) = \frac{1}{\pi^p|\sigma_e^2\mathbf{R}_\alpha^{-1}|}e^{-(\mathbf{s}-\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y})^H(\sigma_e^2\mathbf{R}_\alpha^{-1})^{-1}(\mathbf{s}-\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y})}. \tag{C.10}$$

Then, the conditional expectation of s given $\mathbf{y}$ and $\hat{\omega}$ is

$$E[\mathbf{s}|\mathbf{y},\hat{\omega}] = \mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y}. \tag{C.11}$$

This is the same as the LMMSE estimate of s as in (5.15). Furthermore, the MMSE prediction of $x(t+L)$ in (C.3) is

$$E[x(t+L)|\mathbf{y},\hat{\omega}] = \hat{\mathbf{a}}(L)^H\mathbf{R}_\alpha^{-1}\hat{\mathbf{A}}^H\mathbf{y}, \tag{C.12}$$

which is identical to the LMMSE prediction (C.2). The MSE of (C.12) is obtained as

$$Var(x(t+L)|\mathbf{y},\hat{\omega}) = \sigma_e^2\hat{\mathbf{a}}(L)^H\mathbf{R}_\alpha^{-1}\hat{\mathbf{a}}(L). \tag{C.13}$$

Bibliography

[ABI74] T. N. E. Greville A. Ben-Israel. *Generalized Inverses Theory and Applications.* Wiley, New York, 1974.

[AJHF99] J. B. Andersen, J. Jensen, S. Holdt, and F. Frederiksen. Prediction of future fading based on past measurements. In *Proc. of IEEE VTC Fall*, volume 1, Amsterdam, Netherlands, 1999.

[BRS66] W. R. Burrus B. Rust and C. Schneeberger. A simple algorithm for computing the generalized inverse of a matrix. *Commun. ACM*, pages 381–387, 1966.

[CA01] M. Chen and H. Asplund. Measurements and models for direction of arrival of radio waves in los in urban microcells. In *Proc. of the 12th IEEE International Symposium of Personal, Indoor and Mobile Radio Communications*, San Diego, USA, Sept 2001.

[CEV05] M. Chen, T. Ekman, and M. Viberg. Two new approaches for channel prediction based on sinusoidal modeling. In *Proc. of IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, Jul. 2005.

[CV04] M. Chen and M. Viberg. Lmmse channel prediction based on sinusoidal modeling. In *Proc. of 3rd IEEE Sensor Array and Multichannel signal Processing Workshop*, Barcelona, Spain, Jul. 2004.

[DHHH00] A. Duel-Hallen, S. Hu, and H. Hallen. Long range prediction of fading signals: Enabling adaptive transmission for mobile radio channels. *IEEE Signal Processing Magazine*, 17(3):62–75, May 2000.

[Dur60] J. Durbin. The fitting of time series models. *Rev. Internat. Stati. Inst.*, 23:233–244, 1960.

[ECS+98] R.B. Ertel, P. Cardieri, K.W. Sowerby, T.S. Rappaport, and J.H. Reed. Overview of spatial channel models for antenna array communication systems. *IEEE Personal Communications*, 5(1):10–22, Feb. 1998.

[Ekm02] T. Ekman. *Prediction of Mobile Radio Channels, Modeling and Design.* Uppsala, Sweden, 2002.

[ESA02] T. Ekman, M. Sternad, and A. Ahlén. Unbiased power prediction of rayleigh fading channels. In *Proc. of VTC Fall*, Vancouver, Canada, Sept. 2002.

[Fal02] S. Falahati. *Adaptive Modulation and Coding in Wireless Communications with Feedback.* Göteborg, Sweden, Sept. 2002.

[Fle69] R. Fletcher. A technique for orthogonalization. *J. Inst. Math. Appl.*, 5:162–166, 1969.

[Hay96] M. H. Hayes. *Statistical Digital Signal Processing and Modeling.* John Wiley and Sons, Inc., 1996.

[HJ85] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, UK, 1985.

[HN95] M. Haardt and J. A. Nossek. Unitary esprit: How to obtain increased estimation accuracy with a reduced computational burden. *Proc. IEEE Trans. on Signal Processing*, 43(5):1232–1242, May 1995.

[HW98] J. K. Hwang and J. H. Winters. Sinusoidal modeling and prediction of fast fading processes. In *Proc. of IEEE GLOBECOM*, Sydney, Australia, Nov. 1998.

[Jak74] W. C. Jakes. *Microwave Mobile Communications.* IEEE Press, Piscataway, NY, USA, 1974.

[Kay88] S. M. Kay. *Modern Spectral Estimation, Theory and Application.* Prentice-Hall, Englewood Cliffs, NJ, 1988.

[Kay93] S. M. Kay. *Fundamentals of Statistical Signal Processing - Estimation Theory.* Prentice-Hall, Englewood Cliffs, NJ, 1993.

[Kes86] S. B. Kesler. *Modern Spectrum Analysis* II. IEEE Press, 1986.

[KH95] R. Knopp and P.A. Humblet. Multiple-accessing over frequency-selective fading channels. In *Proc. of the 6th IEEE International Symposium of Personal, Indoor and Mobile Radio Communications*, Toronto, Canada, Sept. 1995.

[KN94] S. M. Kay and V. Nagesha. Maximum likelihood estimation of signals in autoregressive noise. *IEEE Transactions on Signal Processing*, 42(12):88–101, Jan. 1994.

[Kun78] S. Kung. A new identification and model reduction algorithm via singular value decomposition. In *Proc. of 12th Asilomar Conference on Circuits, Systems and Computers*, Pacific Grove, CA, 1978.

[KV96] H. Krim and M. Viberg. Two decades of array signal processing research-the parametric approach. *IEEE Signal Processing Magazine*, pages 67–94, Jul. 1996.

[Lev47] N. Levinson. The wiener RMS error criterion in filter design and prediction. *J. Math. Phys.*, 25:261–278, 1947.

[Lüt96] H. Lütkepol. *Handbook of Matrices.* John Wiley and Sons, Chichester, UK, 1996.

[Mar87] S. L. Marple. *Digital Spectral Analysis with Applications.* Prentice-Hall, 1987.

[MC91] S. Mohideen and V. Cherkassky. On recursive calculation of the generalized inverse of a matrix. *ACM Transactions on Mathematical Software*, 17(1):130–147, Mar. 1991.

[PRK85] A. Paulraj, R. Roy, and T. Kailath. Estimation of signal parameters via rotational invariance techniques - ESPRIT. In *Proc. of 19th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1985.

[Rap96] T. S. Rappaport. *Wireless Communications Principles and Practice.* Prentice-Hall Inc., 1996.

[RPK86] R. Roy, A. Paulraj, and T. Kailath. ESPRIT - a subspace rotation approach to estimation of parameters of cisoids in noise. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34(4):1340–1342, Oct. 1986.

[SEA01] M. Sternad, T. Ekman, and A. Ahlén. Power prediction on broadband channels. In *Proc. of VTC Fall*, Rhodes, Greece, May 2001.

[SK03] S. Semmelrodt and R. Kattenbach. Performance analysis and comparison of different fading forecast schemes for flat fading radio channels. In *Proc. of COST 273 TD(03)045*, 2003.

[SM97] P. Stoica and R. Moses. *Introduction to Spectral Analysis.* Prentice-Hall, Upper Saddle River, NJ, 1997.

[SS91] P. Stoica and T. Söderström. Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 39(8):1836–1847, Aug. 1991.

[SVO94] P. Stoica, M. Viberg, and B. Ottersten. Instrumental variable approach to array processing in spatially correlated noise fields. *IEEE Transactions on Signal Processing*, 42(3):121–133, Jan. 1994.

[SWS03] T. Svantesson, J. W. Wallace, and S. Semmelrodt. Performance evaluation of MIMO channel prediction algorithms using measurements. In *Proc. of the 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, Aug. 2003.

[TV01] P. D. Teal and R. G. Vaughan. Simulation and performance bounds for real-time prediction of the mobile multipath channel. In *Proc. of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, Singapore, Aug. 2001.

[Wal31] G. Walker. On periodicity in series of related terms. *Proc. of the Royal Society of London 131*, pages 518–532, 1931.

[YO02] K. Yu and B. Ottersten. Models for MIMO propagation channels, a review. *Special Issue on Adaptive Antennas and MIMO Systems, Wiley Journal on Wireless Communications and Mobile Computing*, 2002.

[Yul86] G. U. Yule. On the method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers. *Philos. Trans. of the Royal Society of London 226, (reprinted in [Kes86]*, pages 267–298, 1986.