Pilot-based Channel Estimation in OFDM Systems

Master Thesis

Muhammad Saad Akram

S001008

Abstract

The channel estimation techniques for pilot-based OFDM systems are investigated. The channel estimation is studied for different pilot densities (2, 4, 6, and 10) in frequency-domain and fixed pilot allocation in temporal-domain for low delay spread and high delay spread channels. The estimation of channel in frequency-domain is based on interpolation, polynomial-based generalized linear model (or least-squares method) and Wiener filtering (or Linear Minimum Mean Square Error) methods. The estimation of channel in temporal-domain is done using linear interpolation. We have compared the performance of all these schemes by measuring raw bit-error-rate with QPSK modulation on Rayleigh fading channel based on Clarks scattering model.

Index terms – Channel estimation, OFDM, pilot allocation, Clarks scattering channel model, LS method, LMMSE method, Wiener Filtering method.

Preface

This thesis is written to fulfill the requirements of Masters of Science degree. The project has been carried out over a period of 6 months from February 5, 2007 to August 6, 2007 at Modem System Design group at Nokia Denmark. The workload corresponds to 32.5 ECTS points. The supervisors for this project are Zoltan Safar from Nokia and Ole Winther from Informatics and Mathematical Modelling (IMM) at the Technical University of Denmark (DTU).

I would like to take this opportunity to express my gratitude to the supervisors for their guidance and help throughout the project. In addition, I would also like to extend special thanks to Pedro Højen-Sørensen and Thierry Bellier for taking their time to discuss various aspects of the report. Finally, I would like to thank Engineering Manager Niels Mørch for providing me the opportunity to pursue a thesis in his group.

Copenhagen, August 6, 2007

_____

Muhammad Saad Akram

Student ID: s001008

Email: saad.akram@gmail.com

Contents

Abbreviations

2G	2nd Generation
3G	3rd Generation
3GPP	3rd Generation Partnership Project
A/D	Analog-to-Digital
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CDMA	Code Division Multiple Access
D/A	Digital-to- Analog
EDGE	Enhanced Data rates for Global Evolution
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
GERAN	GSM EDGE Radio Access Network
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications (originally: Groupe Spécial Mobile)
HSPA	High Speed Packet Access
ICI	Inter-Carrier Interference
ISI	Inter-Symbolic Interference
LMMSE	Linear Minimum Mean Square
LOS	Line Of Sight
LPF	Low Pass Filter
LS	Least Squares
LTE	Long Term Evolution
LTI	Linear Time Invariant
MIMO	Multiple-Input Multiple-Output
MSE	Mean Square Error
NLOS	Non-Line Of Sight
OFDM	Orthogonal Frequency Division Multiplex
OFDMA	Orthogonal Frequency Division Multiple Access
PA	Pedestrian type-A channel
PB	Pedestrian type-B channel
P/S	Parallel-to-Serial

PSD	Power spectrum Density
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RMS	Root Mean Square
SC-FDMA	Single Carrier Frequency Division Multiple Access
S/P	Serial-to-Parallel
SVD	Singular Value Decomposition
TDD	Time Division Multiple Access
TDMA	Time Division Multiple Access
TR	Technical Report
TS	Technical Specification
UMTS	Universal Mobile Telecommunications System
UTRA	Universal Terrestrial Radio Access
UTRAN	Universal Terrestrial Radio Access Network
VA	Vehicular type-A channel
W-CDMA	Wireless Code Division Multiple Access
WiBro	Wireless Broadband
WiFi	Wireless Fidelity
WiMAX	Worldwide interpretability for Microwave Access
WLAN	Wireless Local Area Network

1 Introduction

Recently, a worldwide convergence has occurred for the use of Orthogonal Division Frequency Multiplexing as an emerging technology for high data rates. In particular, the wireless local network systems such as WiMax, WiBro, WiFi etc., and the emerging fourth-generation (or the so-called 3.9G) mobile systems are all OFDM based systems. OFDM is a digital multi-carrier modulation scheme, which uses a large number of closely-spaced orthogonal sub-carriers that is particularly suitable for frequency-selective channels and high data rates [1], and [2]. This technique transforms a frequency-selective wide-band channel into a group of non-selective narrow-band channels, which makes its robust against large delay spreads by preserving orthogonality in the frequency domain. Moreover, the introduction of a so-called cyclic prefix at the transmitter reduces the complexity at receiver to FFT processing and one tap scalar equalizer at the receiver [2].

The simplified equalization at receiver, however, requires knowledge of the channel over which the signal is transmitted. To facilitate the estimation of the channel in an OFDM system (such as WiMax, WiBro, WiFi, and 3.9/4G), known signals or pilots could be inserted in the transmitted OFDM symbol. Different methods can then be applied to estimate the channel using these known pilots. The focus of this report is to investigate performance of different channel estimators for an OFDM-based 3.9G system. The outcome of the report is to recommend a channel estimation method for implementation and future study.

The report is organized as follows. Section 2 is an investigation into the propagation aspects of the radio channel. Understanding the characteristics of radio channel is critical to investigation of channel estimators. Section 3 provides a mathematical model of an OFDM system. Section 4 provides the theoretical framework of different channel estimators, and Section 5 provides the performance simulations of the implemented channel estimators. Section 6 sums up the conclusion.

2 Propagation Aspects of Radio Channel

In wire-line communication, the data transmission is primarily corrupted by statistically independent Gaussian noise, as known as the classical additive white Gaussian noise (AWGN). In absence of interference, the primary source of performance degradation in such wire-line channels is thermal noise generated at the receiver. Reliable communication in wireless or radio channels, however, becomes a difficult task as the transmitted data is not only corrupted by AWGN, but also suffers from inter-symbol interference (ISI), in addition to (large-scale and small-scale) fading as well as interference from other users.

To master the art of wireless communications, one must understand the propagation characteristics of a radio channel. The fading in radio propagation can be classified into two groups; large-scale fading and small-scale fading as illustrated in Figure 2-1 [3], [4], [5], [6] and [7]. **Large-scale fading** manifests itself as the average signal power attenuation or path loss due to motion over large areas as shown in blocks 1, 2 and 3 in Figure 2-1. **Small-scale fading** refers to the dramatic changes in the signal amplitude and phase that occur due to small changes in the spatial separation between the transmitter and the receiver. As indicated by blocks 4, 5 and 6 in Figure 2-1, small-scale fading manifests itself in two mechanisms namely, time-spreading of the signal (or channel dispersion) and time-variant nature of the channel. The signal time-spreading (signal dispersion) and time-variant nature of the channel may be examined in two domains, time and frequency, as indicated in Figure 2-1, block 7, 10, 13 and 16. For signal dispersion nature, we categorize the fading degradation types as frequency selective and frequency non-selective (flat) as illustrated in blocks 8, 9, 11, and 12. For time-variant nature, we categorize the fading degradation types as fast fading and slow fading, as shown in blocks 14, 15, 17, and 18. The labels indicating Fourier transforms and Duals will be explained in later sections.

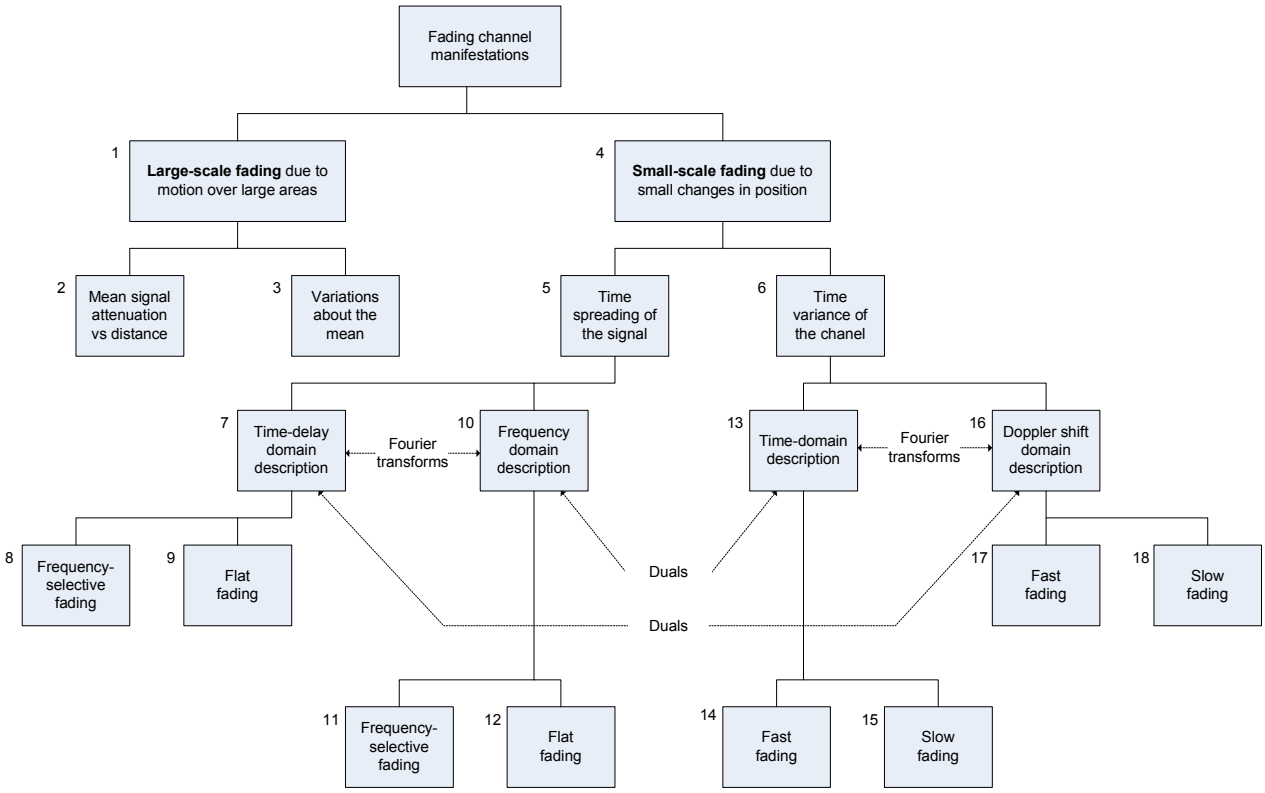


Figure 2-1. Fading Channel manifestations

In Figure 2-2, the received signal power $r(t)$ versus antenna displacement is plotted for a mobile radio. The received signal $r(t)$ can be partitioned into two component random variables as [4]

$$r(t) = m(t) \times r_0(t) \quad (2\text{-}1)$$

where

$m(t)$ is the large-scale fading component or path-loss, and

$r_0(t)$ is the small-scale fading component or multipath fading.

In Figure 2-2, it can be readily identified that small-scale fading is superimposed on the large-scale fading. The typical antenna displacement between the small-scaling nulls is approximately a half-wavelength. In Figure 2-2 (b), the large-scale fading $m(t)$ has been removed to view the small-scaling fading $r_0(t)$, about some constant power.

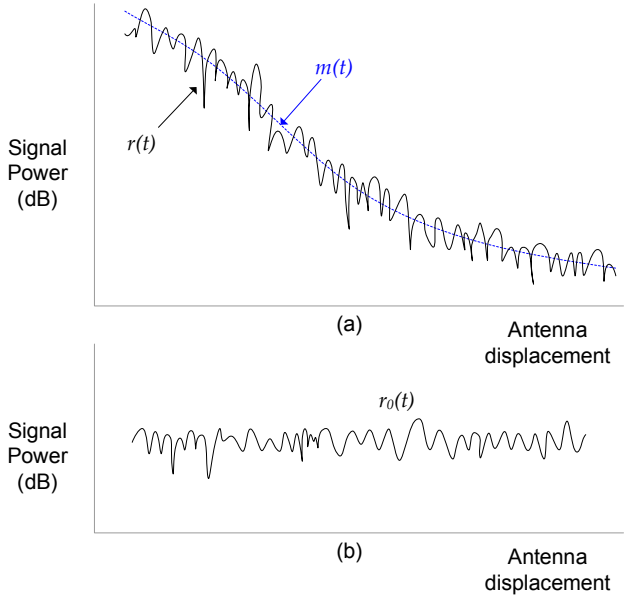


Figure 2-2. Large-scale fading and small-scale fading

Large-scale fading is responsible for path-loss in wireless communications, and large-scale fading models typically find applications in mobile network planning and understanding free space wireless communication over large areas [3]. In most practical wireless communication systems, the radio communications is far more complex then free-space situation, and is best explained by small-scale fading models. We now explore the various aspects of small-scale fading in coming sections.

2.1 Small-scale Fading

In a typical wireless communications system, the transmitted signal typically undergoes refractions, shadowing and various reflections due to the presence of various objects (buildings, trees, etc.) in the channel [5]. As a consequence the waves emitted by the receiver arrive at the receiver antenna over multiple paths, a phenomenon known as multipath propagation (illustrated in Figure 2-3). The complete set of propagation paths between transmitter and receiver forms the multipath channel. Each path can be characterized by three parameters: delay, attenuation and phase shift. The path

delay depends on the path length and on the speed at which a wave is propagating in the different media along the path, while the attenuation and phase shift is caused by fading.

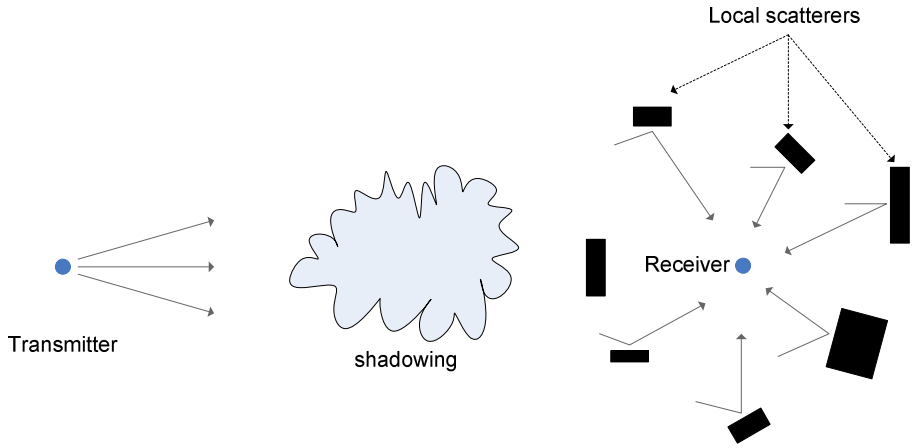


Figure 2-3. Multipath radio propagation

Two distinctions have been made between line-of-sight (LOS) and non-line of sight (NLOS) setups. In LOS setup, there is a line of sight between the transmitter and receiver. The received signal is made up of several multiple reflective waves as well as a significant LOS component. In such a case, the envelope of the received signal has Rician distribution, and the fading is referred as *Rician fading* [5], [6]. In NLOS case, there is no direct line of sight between the transmitter and receiver, so all incoming waves have been reflected at least once. The multipath propagation scenario in Figure 2-3 is a NLOS case. In such a set-up, the envelope of the received signal can be best described by a Rayleigh distribution, and the fading is known as *Rayleigh fading* [5], [6].

The discrete time-variant channel impulse response $h(\tau,t)$ of the complex multipath channel can be written as [8]

$$h(\tau,t)=\sum_{n} a_n(t)\cdot e^{-j2\pi f_c \tau_n(t)}\cdot \delta[\tau-\tau_n(t)] \qquad (2\text{-}2)$$

where

$a_n(t)$ is the attenuation factor for the signal received on the n-th path time instant t,

$\tau_n(t)$ is the propagation delay at the n-th path at time t

$e^{-j2\pi f_c \tau}$ is the phase rotation of the signal component at delay τ at carrier frequency f_c

$\delta[.]$ is the Dirac delta function

Applying a few constraints on multipath channel, it can be easily shown that the received signal envelope has Rician distribution for LOS multipath propagation, and Raleigh distribution for NLOS multipath propagation [7], [9].

As indicated in Figure 2-1, block 4, 5 and 6, small-scale fading manifests itself in two mechanisms

- Time-spreading of the underlying pulses within the signal
- Time-variance of the channel due to motion of the receiver (or occasionally a transmitter).

Figure 2-4 summarizes these mechanisms for small-scale fading in two domains (time or time-delay and frequency or Doppler shift). Note that any mechanism described in time domain can be equally well described in the frequency domain. For example, the time-spreading mechanism will be characterized as multipath delay spread in time domain, and in the frequency domain as channel coherence bandwidth. Similarly, time variant mechanism will be described as channel coherence time, and in Doppler shift domain as Doppler spread. These mechanisms will be examined in greater detail in the coming sections.

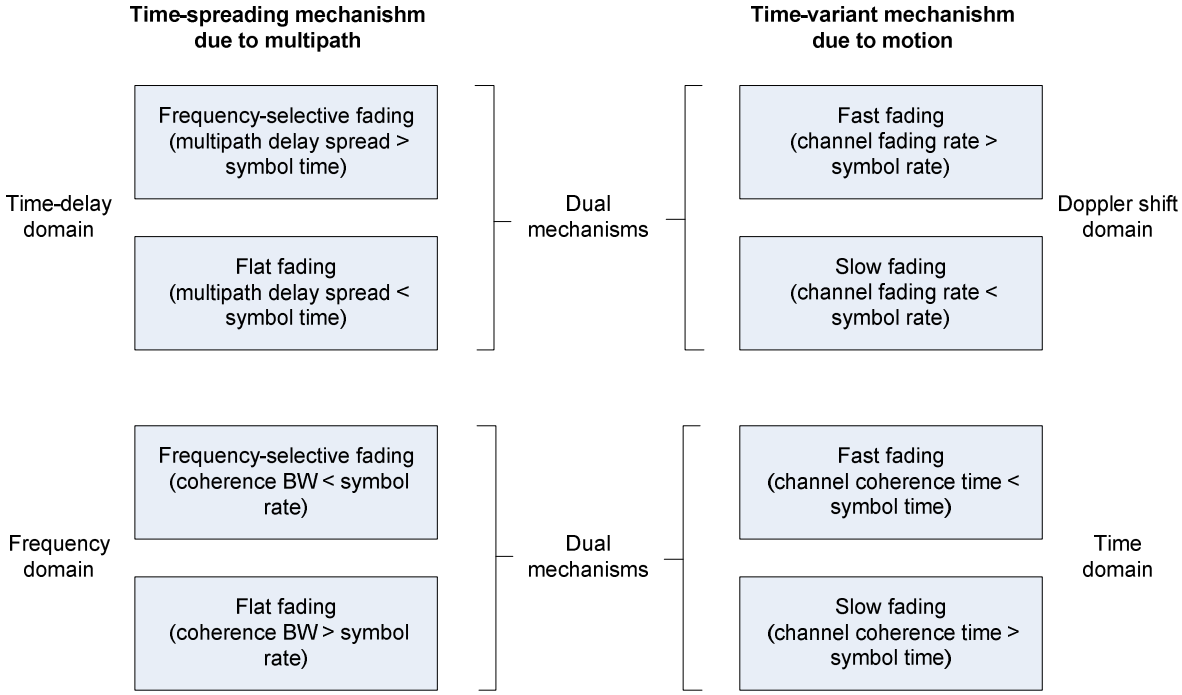


Figure 2-4. Small-scale fading manifestation in various domains [4]

2.1.1 Multipath Intensity Profile (signal time-spreading in time delay domain)

Given a channel impulse response $h(\tau,t)$, we can calculate the autocorrelation function, $S(\tau;t_1,t_2)$, of the channel impulse response as [8]

$$S(\tau;t_1,t_2)=\frac{1}{2}E\{h^*(\tau;t_1)\cdot h(\tau;t_2)\} \quad (2\text{-}3)$$

where

$E\{\cdot\}$ is the expectation operator,

$h(\tau,t_1)$ and $h(\tau,t_2)$ are the channel impulse response at time instants t_1 and t_2, respectively.

Assuming that the channel is wide-sense-stationary (WSS), we re-write Equation (2-3) as

$$S(\tau;\Delta t)=\frac{1}{2}E\{h^*(\tau;t)\cdot h(\tau;t+\Delta t)\} \quad (2\text{-}4)$$

where

$S(\tau;\Delta t)$ gives the average power output of the channel as a function of time delay τ and difference Δt in observation time.

If we now let $\Delta t = 0$, the the resulting function $S(\tau;0) = S(\tau)$ is called the *multipath intensity profile* or *delay power spectrum* of the channel. $S(\tau)$ is simply the average power output of the channel as a function of time delay τ as shown in Figure 2-6a on Page 16.

An important characterization of channel when studying multipath intensity profile is *excess delay*. Excess delay is defined as the signal propagation delay that exceeds the delay of the first signal arrival at the receiver. The values of excess delay can be quantified by the multipath intensity profile into three categories

- *maximum excess delay*, T_m, is the time between the first and the last received component of a transmitted pulse, during which the multipath signal power falls below that of the strongest component.
- *average delay*, μ_τ, is mean delay of the multipath channel expressed as [9]

$$\mu_\tau = \frac{\int_0^{+\infty} \tau \cdot S(\tau) d\tau}{\int_0^{+\infty} S(\tau) d\tau} \tag{2-5}$$

  where $S(\tau)$ is multipath intensity profile of the channel.

- *root-mean-square (RMS) delay* , σ_τ, is defined as the standard deviation value of the delay of reflections, weighed proportional to the energy of the waves, and expressed as [9]

$$\sigma_\tau = \sqrt{\frac{\int_0^{+\infty} (\tau - \mu_\tau)^2 \cdot S(\tau) d\tau}{\int_0^{+\infty} S(\tau) d\tau}} \tag{2-6}$$

  The value of σ_τ is commonly used to give rough indication of maximum data rate that can be reliably supported by the channel, when no special measures, such as equalization, are taken.

The relationship between the maximum excess delay T_m and symbol time T_s determines the type of fading degradation experienced by the receiver as indicated in Figure 2-1. The multipath channel is said to exhibit *frequency-selective fading* if $T_m > T_s$. This condition occurs because the received multipath components of the symbol extend beyond the symbol's time duration, resulting in channel-induced ISI. Various signal processing techniques (e.g., equalization, rake receiver, etc.) exist for mitigating the effect of channel induced ISI. A channel is said to experience *frequency non-selective* or *flat fading* if $T_m < T_s$. In such a case, all multipath components of the symbol arrive within the symbol duration, and hence are not resolvable. Although there is no channel induced ISI as a result, there is performance degradation since irresolvable phasor components can add up destructively to yield a substantial reduction in SNR.

2.1.2 Spaced-Frequency Correlation Function (signal time-spreading in frequency domain)

An analogous characterization of the time-variant multipath channel can be provided in frequency domain by taking the Fourier transform of $h(\tau,t)$ as [8]

$$H(f,t)=\int_{-\infty}^{+\infty} h(\tau;t)\cdot e^{-j2\pi f\tau}d\tau \qquad (2\text{-}7)$$

where

$H(f,t)$ is *channel transfer function*.

We can calculate the autocorrelation function, $R(\Delta f;t_1,t_2)$, of the channel impulse response as [8]

$$R(\Delta f;t_1,t_2)=\frac{1}{2}E\{H^*(f_1;t_1)\cdot H(f_2;t_2)\} \qquad (2\text{-}8)$$

where

$\Delta f = f_2 - f_1$ is the frequency difference,

$H(f_1,t_1)$ and $H(f_2,t_2)$ are the channel transfer function at frequency-time pair (f_1,t_1) and (f_2,t_2) respectively.

Assuming the channel is wide sense stationary (WSS), we re-write Equation (2-8) as

$$R(\Delta f;\Delta t)=\frac{1}{2}E\{H^*(f_1;t)\cdot H(f_2;t+\Delta t)\} \qquad (2\text{-}9)$$

where

Δt is time difference between channel transfer function observations, and

$R(\Delta f;\Delta t)$ is called *spaced-frequency, spaced-frequency correlation function* [8] of the channel.

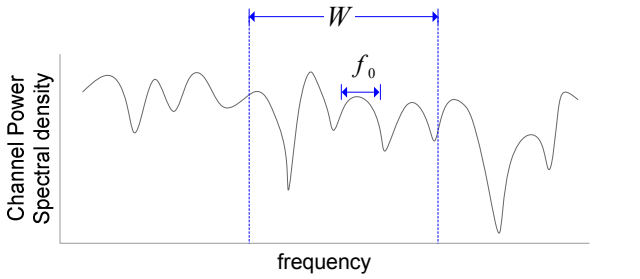
By letting $\Delta t = 0$ in Equation (2-10), we get $R(\Delta f;0)=R(\Delta f)$ which is called the *spaced-frequency correlation function*. $R(\Delta f)$ provides a measure correlation of the channel transfer function as a function of frequency difference. Illustrated in Figure 2-6b on Page16, the space-frequency function $R(\Delta f)$ is the direct Fourier transform of multipath intensity profile $S(\tau)$, expressed as [8]

$$R(\Delta f)=\int_{-\infty}^{+\infty} S(\tau)\cdot e^{-j2\pi\Delta f\tau}d\tau \qquad (2\text{-}11)$$

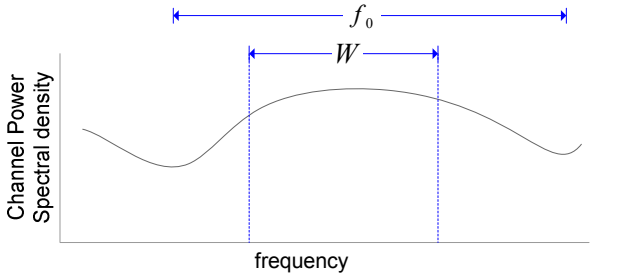
Since $R(\Delta f)$ is the autocorrelation function in the frequency variable, it provides a measure of the frequency coherence of the channel. The *coherence bandwidth* f_0 of the channel is defined as the statistical measure of the range of frequencies over which the channel passes all spectral components with approximately equal gain and linear phase. Due to Fourier transform relationship between $R(\Delta f)$ and $S(\tau)$, the coherence bandwidth and excess delay are reciprocally related

$$f_0 \propto \frac{1}{T_m} \qquad (2\text{-}12)$$

The relationship between coherence bandwidth f_0 and bandwidth of the transmitted signal W determines the type of fading degradation. The channel is said to experience frequency-selective fading when $f_0 < W$, whereas *frequency non-selective* or *flat fading* occurs when the $f_0 > W$. These two types of fading are illustrated in Figure 2-5.



(a) Typical frequency-selective fading channel $(f_0 < W)$



(b) Typical flat fading channel $(f_0 > W)$

Figure 2-5. Relationship between coherence bandwidth and signal bandwidth

2.1.3 Spaced-Time Correlation Function (time variance in time domain)

The spaced-time correlation function $R(\Delta t)$ provides channel correlation information at various instants in time domain. Illustrated in Figure 2-6d on Page 16., the spaced-time correlation function can be calculated as [8]

$$R(\Delta t) = \int_{-\infty}^{+\infty} S(\tau; \Delta t) d\tau \qquad (2\text{-}13)$$

where

$S(\tau; \Delta t)$ is the autocorrelation function of the channel impulse response as given in Equation (2-4).

Spaced-time correlation function provides a coherence of the channel in time domain. We can define coherence time T_0 as the measure of the expected time duration over which channel's response is essentially invariant. The spaced-time correlation function and coherence time T_0 provide a measure of the rapidity of channel variation. If $T_0 < T_s$, where T_s is the symbol time, then the channel is said to experience *fast fading*, whereas if $T_0 > T_s$ then the channel is said to be experience *slow fading*. During fast fading the channel fading characteristics vary multiple times during a symbol. On the contrary, the channel fading characteristics remain almost constant during the channel in slow fading.

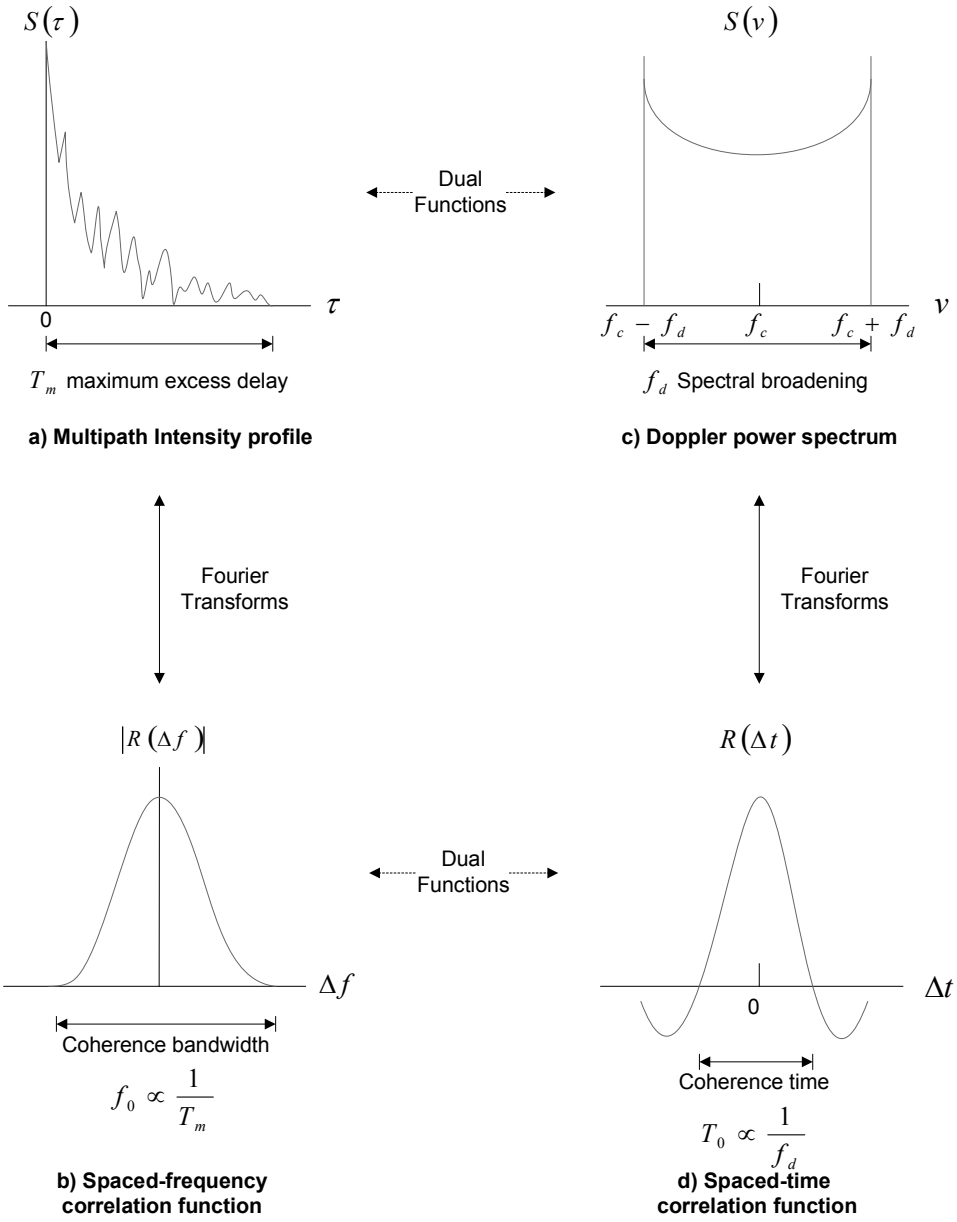


Figure 2-6. Relationship between channel correlation function and power density function

2.1.4 Doppler Power Spectrum (time-variance in Doppler shift domain)

The Doppler power spectrum $S(v)$ provides knowledge of spectral broadening of a frequency impulse passing through a channel, as shown in Figure 2-6c. It can be calculated by taking the Fourier transform of the time-spaced correlation function in Equation (2-13) as [8]

$$S(v) = \int_{-\infty}^{+\infty} R(\Delta t) \cdot e^{-j2\pi v \Delta t} d\Delta t \qquad (2\text{-}14)$$

where v is the frequency shift relative to the carrier frequency. The width of the Doppler power spectrum is referred to as *spectral broadening* or *Doppler spreading*, denoted as f_d. Note that the

Doppler spread f_d and coherence time T_0 are reciprocally related due to the Fourier relationship between Doppler power spectrum and spaced-time correlation function.

$$f_d \propto \frac{1}{T_0} \tag{2-15}$$

The channel is said to be fast fading if the bandwidth of the signal W is less than the Doppler spread i.e., $W < f_d$. Conversely, the channel is referred as slow fading if $W > f_d$.

2.1.5 Duality of fading channel manifestations

In Figure 2-6, duality between multipath intensity function $S(\tau)$ and Doppler power spectrum $S(v)$ is identified. It means that the two functions exhibit similar behavior across time domain and frequency domain. Just as $S(\tau)$ identifies expected power of the received signal as a function of time delay, $S(v)$ identifies expected power of the received signal as a function of frequency.

Similarly, spaced-frequency correlation function $R(\Delta f)$ and spaced-time correlation function $R(\Delta t)$ are duals. It implies that just as $R(\Delta f)$ represents channel correlation in frequency, $R(\Delta t)$ represents the channel correlation function in time in a similar fashion.

2.2 Simulating Multipath Fading Channels

To investigate the performance of a wireless communication system, a fading channel simulator must be implemented. Various methods exist for implementing propagation aspects of a fading channel. Before proceeding with the implementation methods for channel simulators, we take a look at the finite impulse response (FIR) filter description of multipath fading channel.

The impulse response of a multipath fading channel, given in Equation (2-2), can be described as a time-varying FIR filter as illustrated in Figure 2-7 as [5]

$$h(\tau;t) = h_0(t)\delta(\tau - \tau_0(t)) + h_1(t)\delta(\tau - \tau_1(t)) + \cdots + h_{L-1}(t)\delta(\tau - \tau_{L-1}(t)) \tag{2-16}$$

where

$h(\tau;t)$ denotes the time-varying impulse response of the L path fading channel,

$h_i(t)$ and $\tau_i(t)$ denote the time-varying complex gain and excess delay of the i'th path.

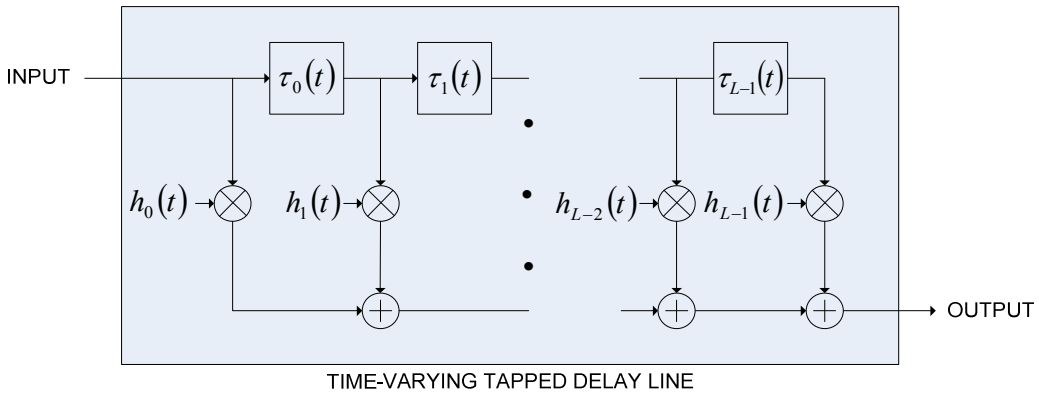


Figure 2-7. Clark's Scattering Model as a time-varying FIR filter

The multipath fading channel can thus easily implemented as a FIR filter if we can generate the time varying complex gains $h_i(t)$ and the excess delay $\tau_i(t)$. The task becomes slightly easier by fixing the excess delay of the channel paths i.e., making the excess delay time-invariant. This simplifies the task of implementing a fading channel to generating the time-varying complex gains $h_i(t)$.

2.2.1 Filtered Gaussian Noise

A simple method to generate a time-varying complex gain $h_i(t)$ is to filter two independent white Gaussian noise sources with low-pass filters [5], as shown in Figure 2-8. The PSD of $g_R(t)$ and $g_I(t)$ is determined by the squared amplitude response of the low pass filters, each with the same transfer function $G(f)$. Simple (typically first-order) low-pass filters are used for $G(f)$. Note that for first-order low pass filter models the fading process as a Markov process.

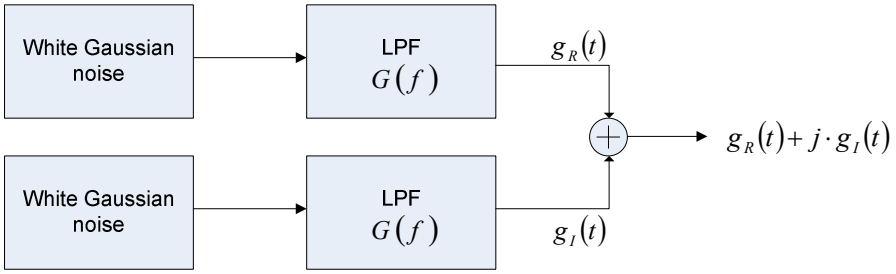


Figure 2-8. Filtered Gaussian noise based fading simulator for a single-tap

To produce a Raleigh faded envelope, the two different noise sources must be zero-mean and the same variance. The simplicity of the model has some limitation as only rational forms of Doppler spectrum can be produced by the model, whereas, a typical fading environment has non-rational Doppler spectrum. To approximate non-rational Doppler spectrum, a higher order filter is required. Unfortunately, a higher order filter has a long impulse response, and this increases the simulation time significantly.

2.2.2 Sum of Sinusoids method (Clark's Model)

A simple and efficient fading simulation method based on scatters was first proposed by Clark, commonly known as Clark's 2-D isotropic scattering model [5], [10] and [11]. The model assumes a number of independent scatterers placed randomly on co-centric circles as shown in Figure 2-9. Each scatterer i reflects the transmitted signal with a given amplitude a_i and phase φ_i, assumed to be independent throughout the duration of transmission such that [10]

$$E\{a_i \cdot a_k\} = 0 \text{ and } E\{\varphi_i \cdot \varphi_k\} = 0 \text{ for } i \neq k \qquad (2\text{-}17)$$

The mobile is assumed to be in origin of the circles and moving with velocity v. In Figure 2-9, $R_i(x)$ represents the distance between mobile and i-th scatter. θ_i represents the angle between the direction of mobile movement and the i-th scatterer. The time difference $\Delta\tau_n$ between the two co-centric circles represents the relative time difference between the recognizable signal paths. The model assumes that scatterer angles θ_is and phase φ_is are evenly distributed in interval $[-\pi,+\pi]$. Also note that the model only accounts for microscopic movements of the mobile, and does not account for macroscopic movements.

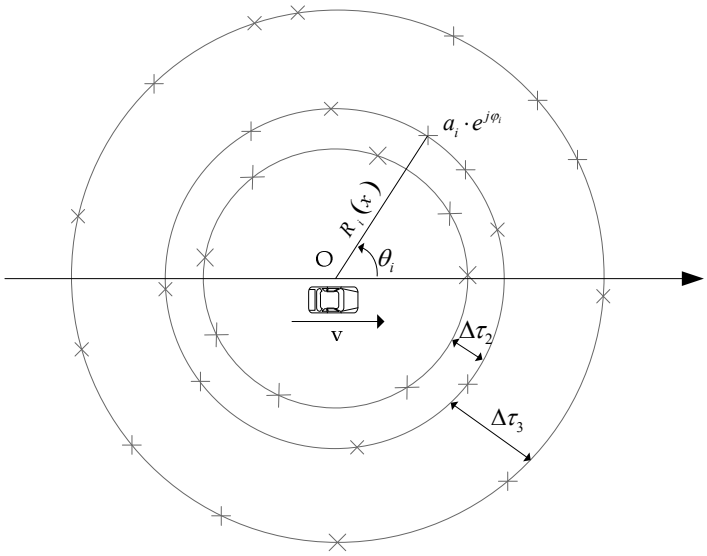


Figure 2-9. Jakes 2D isotropic scattering environment

The signal contribution from a single scatterer, as shown in Figure 2-10, can be written as [10]

$$a_i \cdot e^{j\varphi_i} \cdot e^{-j\frac{2\pi}{\lambda}R_i(x)} \tag{2-18}$$

where

$a_i \cdot e^{j\varphi_i}$ is the complex magnitude of the signal associated with i-th scatterer,

$R_i(x)$ is the distance from the point of observation to the position of scatterer, and

λ is the wavelength of the signal carrier frequency.

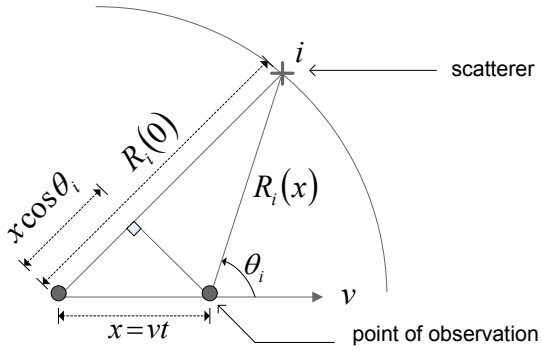


Figure 2-10. Single scatter in Jakes model

We can approximate the distance between the point-of-observation and the position of scatterer as

$$R_i(x) \approx R_i(0) - x\cos\theta_i \tag{2-19}$$

where

θ_i is the angle between the point of observation and the position of the scatterer,

x is the distance traveled by receiver by the time signal reaches $R_i(x)$ from the i th scatterer

The assumption holds since the velocity of the mobile receiver v is much smaller then the speed of wave propagation (velocity of light c). Inserting the approximation in Equation (2-18), we get

$$a_i \cdot e^{j\varphi_i} \cdot e^{-j\frac{2\pi}{\lambda}R_i(x)} \approx a_i \cdot e^{j\varphi_i} \cdot e^{-j\frac{2\pi}{\lambda}(R_i(0)-x\cos\theta_i)} \tag{2-20}$$

Using the relationship between the wavelength λ and frequency f of the carrier wave

$$f = \frac{c}{\lambda}$$

We rewrite Equation (2-20) as

$$a_i \cdot e^{j\varphi_i} \cdot e^{-j\frac{2\pi f}{c}(R_i(0)-x\cos\theta_i)} \tag{2-21}$$

$$= a_i \cdot e^{j\varphi_i} \cdot e^{-j\omega\left(\frac{R_i(0)-x\cos\theta_i}{c}\right)}$$

$$= a_i \cdot e^{j\varphi_i} \cdot e^{-j\left(\omega\tau_i-\omega\frac{vt}{c}\cos\theta_i\right)}$$

$$= a_i \cdot e^{j\varphi_i} \cdot e^{-j\omega\tau_i} \cdot e^{j\omega_d t} \tag{2-22}$$

where

ω is the angular carrier frequency related to carrier frequency f as $\omega = 2\pi f$,

ω_d represents the Doppler shift f_d in frequency, $\omega_d = 2\pi f_d$

τ_i is the delay from the i -th scatterer

Equation (2-22) implies that each scatter creates a delay τ_i, depending upon the distance between position of scatter and point of observation, and a Doppler shift, depending upon the velocity v and carrier frequency.

The contribution from N scatters in each of the co-centric circles is the sum of the contribution forming the channel transfer function as

$$H(\omega,x) = \underbrace{e^{j\omega\tau_0}\sum_{i=1}^{N}\underbrace{a_i \cdot e^{j\varphi_i} \cdot e^{j\left(2\pi\frac{x}{\lambda}\cos\theta_i\right)}}_{\text{contribution from scatter } i}}_{\text{contribution from first circle of scatters}} + \cdots + \underbrace{e^{j\omega\tau_n}\sum_{i=1}^{N}a_i \cdot e^{j\varphi_i} \cdot e^{j\left(2\pi\frac{x}{\lambda}\cos\theta_i\right)}}_{\text{contribution from nth circle of scatters}} \tag{2-23}$$

The time-domain channel impulse response can be found by taking the inverse Fourier transform of the channel transfer function and using the relationship $x = vt$, resulting in

$$h(\tau,t) = \left(\sum_{i=1}^{N} a_i \cdot e^{j\varphi_i} \cdot e^{j\left(2\pi\frac{vt}{\lambda}\cos\theta_i\right)}\right)\delta(t-\tau_0) + \cdots + \left(\sum_{i=1}^{N} a_i \cdot e^{j\varphi_i} \cdot e^{j\left(2\pi\frac{vt}{\lambda}\cos\theta_i\right)}\right)\delta(t-\tau_n) \tag{2-24}$$

$$= \sum_n h_n(t)\delta(t-\tau_n) \tag{2-25}$$

The channel impulse can be understood as a series of time delayed impulses, where the n-th impulse $\delta(t-\tau_n)$ is weighed with a factor $h_n(t)$. Note that impulse weighing factor $h_n(t)$ is the n-th complex tap defined as

$$h_n(t) = \sum_{i=1}^{N} a_i e^{j\varphi_i} e^{j\left(2\pi\frac{v}{\lambda}t\cos\theta_i\right)} = \sum_{i=1}^{N} a_i e^{j\varphi_i} e^{j(2\pi f_d t\cos\theta_i)} \tag{2-26}$$

where f_d is the maximum Doppler frequency given as $f_d = v/\lambda$.

Since phase φ_i and scatterer angle θ_i are independent and identically-distributed random variables in interval $[-\pi,+\pi]$, their sum will converge to complex Gaussian for a large N due to central limit theorem. Hence, for a large number of scatterers, the channel tap $h_n(t)$ will be Rayleigh in nature [10]. The amplitude a_i is assumed to be constant for all the scatters in a channel tap $h_n(t)$, and set to

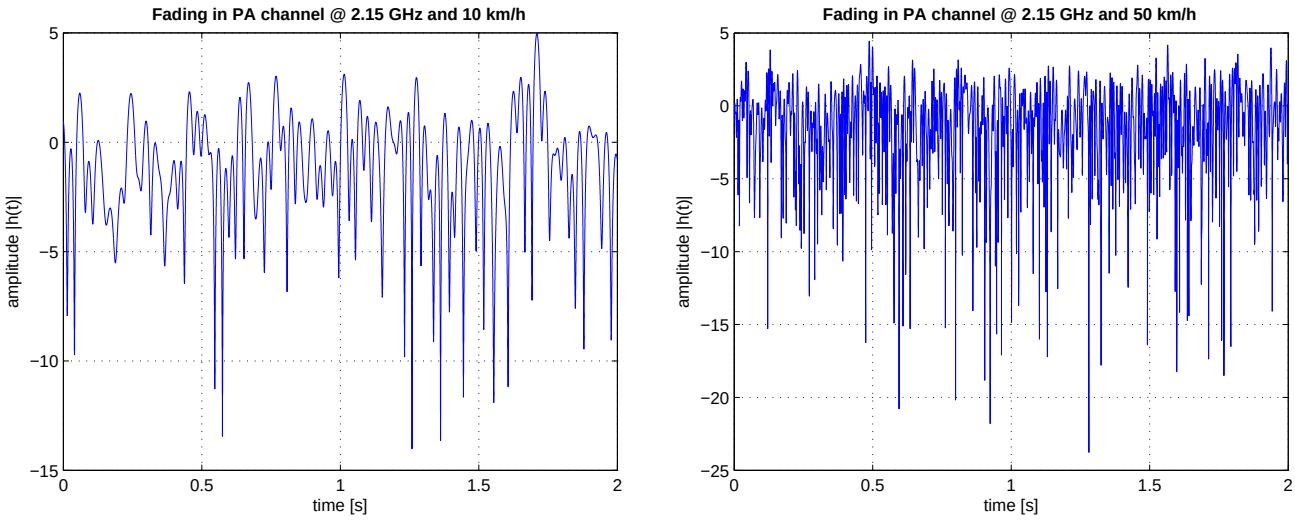
$$a_i = \sqrt{\frac{P_n}{N}} \tag{2-27}$$

where P_n is the average power of the received signal for n-th path.

The channel tap $h_n(t)$ can thus be given as

$$h_n(t) = \sqrt{\frac{P_n}{N}} \sum_{i=1}^{N} e^{j\varphi_i} e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i t\right)} \tag{2-28}$$

In Figure 2-11, fading in a Clark's isotropic scattering environment is plotted for different velocities for a single-tap. The carrier frequency of 2.15 GHz with 10 and 50 km/h. The figures clearly illustrates that the rapidity of fading increases proportional to the increase in the velocity of the mobile.



(a) 10 km/h (b) 50 km/h

Figure 2-11. Fading in Clark's scattering model at 2.15 GHz at different velocities

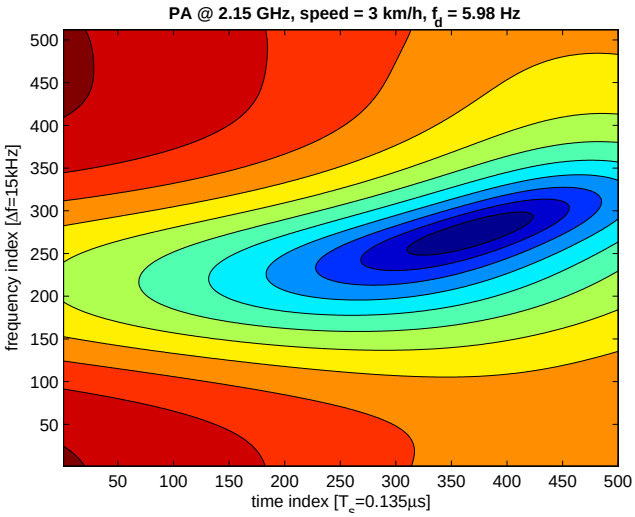
2.3 Fading Manifestations of the Clark's Scattering Model

We now look at the fading manifestations of the Clark's scattering model of the simulated channel. This is dealt in three sections:

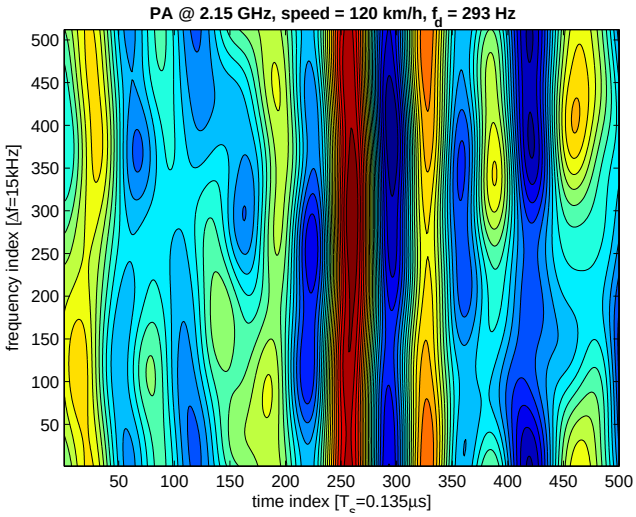
- Channel-Impulse response and Channel Transfer function
- Spaced-Time Correlation and Doppler Power Spectrum
- Spaced-Frequency Correlation and Power Delay Profile

2.3.1 Channel-Impulse response and Channel Transfer function

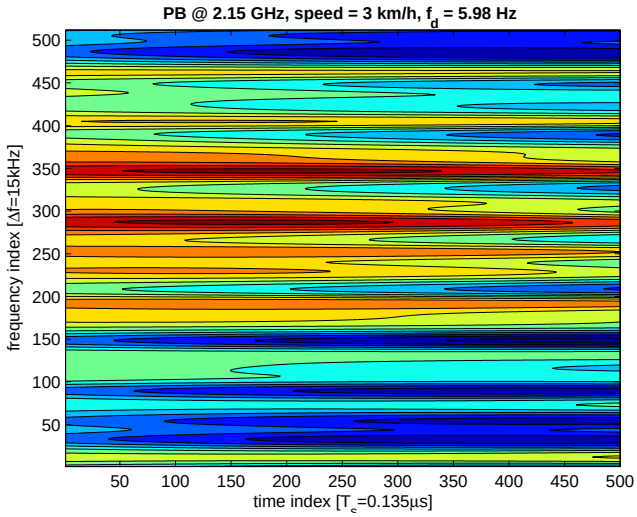
Figure 2-12 shows a realization of time-varying channel transfer function for a channel with low and high delay spread at low and high velocity. The channel transfer function varies more rapidly in frequency for channels with large delay spread (large time dispersion). Similarly, the rate at which time-varying channel transfer function varies with time is higher for high velocities (larger Doppler spread). The low delay spread and low Doppler spread channel (Figure 2-12a) is the most attractive for noise-reduction point-of-view since averaging can be done in both time and frequency domains. Conversely, the high delay spread and high Doppler spread channel (Figure 2-12d) is the least attractive for noise-reduction point-of-view.



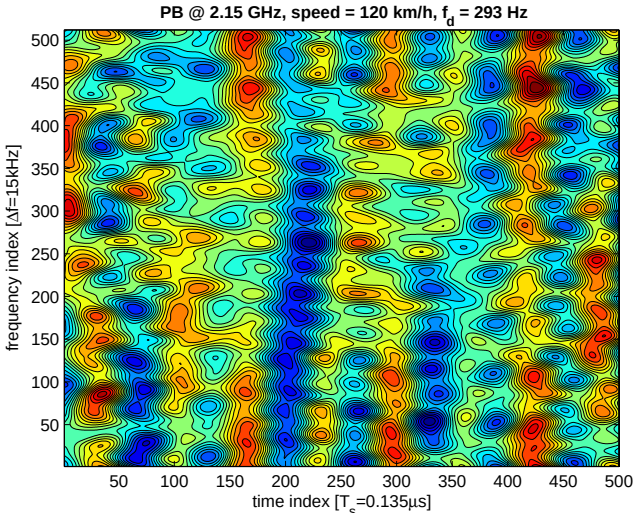
(a) low Delay Spread and low Doppler Spread



(b) low Delay Spread and high Doppler Spread



(c) high Delay Spread and low Doppler Spread



(d) high Delay Spread and high Doppler Spread

Figure 2-12. Time-varying channel transfer function

2.3.2 Spaced-Time Correlation and Doppler Power Spectrum

The spaced-time correlation function, $R(\Delta t)$, of the Clark's scattering model for single-tap channel is derived in Appendix A.1. The derivation shows that the spaced-time correlation of the single-tap channel is

$$R(\Delta t) = P \cdot J_0\left(2\pi \frac{v}{\lambda} \Delta t\right) \tag{2-29}$$

where

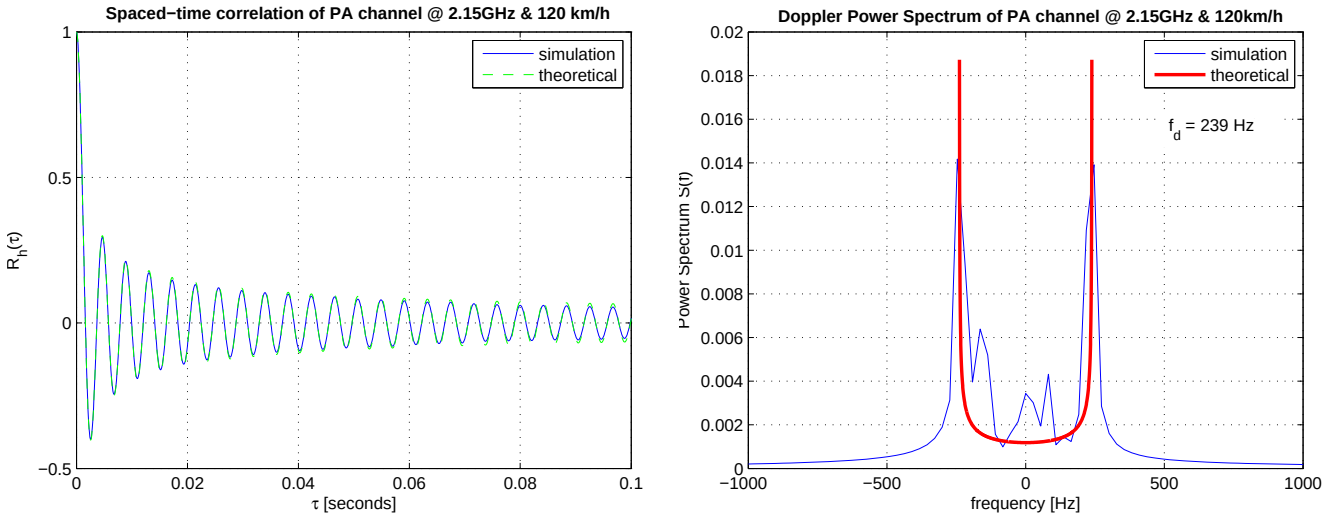
P is the expected power of the channel-tap, and

$J_0(\cdot)$ is the zero-th order Bessel Function of the first kind.

The Doppler-power spectrum $S(v)$ of the one-tap channel can be calculated by taking the Fourier transform of the spaced-time correlation function of $R(\Delta t)$ as [5]

$$S(v) = \begin{cases} \dfrac{P_n}{\pi f_d \sqrt{1-(v/f_d)^2}} & , \quad |v| < f_d \\ 0 & otherwise \end{cases} \tag{2-30}$$

In Figure 2-13a, the theoretical and simulated spaced time correlation function is plotted based on Clarks scattering model. A carrier frequency of 2.15 GHz and a single-tap channel was used for the fading simulation. The speed of the mobile was set to 120 km/h. It may be concluded that there is a strong correspondence between the theoretical and simulated time correlation function. In Figure 2-13b, the theoretical and simulated Doppler power spectrum is for the same conditions. A Doppler shift of 239 Hz is evident for both theoretical and simulated channel.



(a) Spaced-Time Correlation Function (b) Doppler Power Spectrum

Figure 2-13. Fading characteristics of Clark's scattering model

2.3.3 Frequency-Spaced Correlation and Power Delay Profile

We can also calculate the frequency-spaced correlation of the mobile channel, given its power delay profile or multipath intensity profile. The power delay profile, h_{PDP}, of a channel may be represented as a series of L impulses

$$h_{PDP} = P_0\delta(t-\tau_0)+P_1\delta(t-\tau_1)+\cdots+P_L\delta(t-\tau_{L-1}) \quad (2\text{-}31)$$

where

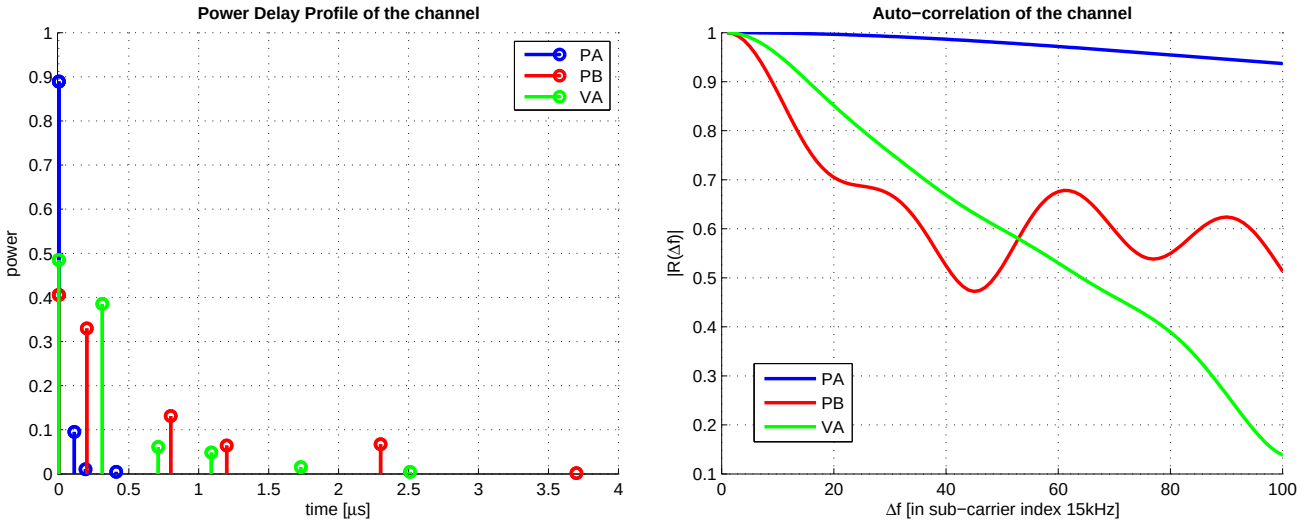
P_n is the expected power of the received signal at the n'th impulse

In Appendix A.2, we show that the frequency spaced spectrum, $R(\Delta f)$ of such a channel is

$$R(\Delta f) = \sum_{i=1}^{L} P_i \cdot e^{j2\pi\Delta f} \quad (2\text{-}32)$$

In Figure 2-14a, we plot power delay profile of different 3GPP standardized channels i.e., Pedestrian-A (PA), Pedestrian-B (PB) and Vehicular A (VA). The expected power and delay of these channels' taps are presented in Table 3 in Appendix A.3, We note the following:

- PA has a shorter excess delay than PB and VA, and therefore will have larger coherence bandwidth than PB and VA. This is also illustrated in Figure 2-14b, where we plot the magnitude of the spaced -frequency correlation functions for the power delay profiles of PA, PB, and VA.
- PB and VA are more frequency selective than PA. This follows automatically by observing the coherence bandwidth of the respective channels.



(a) Power delay profile

(b) Frequency-spaced correlation

Figure 2-14. Frequency-spaced correlation function for various channel

3 The OFDM Principle

In this chapter we look at the principles of an orthogonal division multiplexing (OFDM) system. Since our objective is to investigate channel estimation methods for OFDM systems, it is essential to acquire a solid understanding of OFDM systems before proceeding with the channel estimation investigation.

3.1 Multi-carrier Modulation

Multi-carrier modulation was first proposed in 60's and forms the basis of the OFDM modulation technique. In multi-carrier modulation, the available bandwidth W is divided into number of N_c sub-bands or sub-carriers, each with a width of $\Delta f = \frac{W}{N_c}$. This subdivision is illustrated in Figure 3-1.

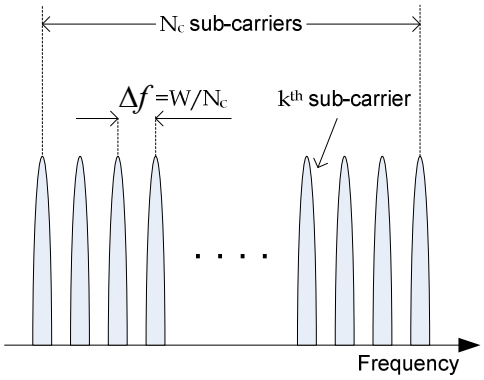


Figure 3-1. Subdivision of bandwidth into N_c sub-carriers

Instead of transmitting the data symbols serially, the multi-carrier transmitter partitions the data into blocks of N_c data symbols that are transmitted in parallel by modulating the N_c carriers. The symbol duration for a modulated carrier is $T_s = \frac{1}{W}$.

The multi-carrier signal can be written as a set of modulated carriers as [12]

$$s(t) = \sum_{k=0}^{N_c-1} x_k \psi_k(t) \qquad (3\text{-}1)$$

where

x_k is the data symbol modulating the k^{th} sub-carrier

$\psi_k(t)$ is the modulation waveform at the k^{th} sub-carrier

$s(t)$ is the multi-carrier modulated signal

The process of generating a multi-carrier modulated signal is illustrated in the Figure 3-2.

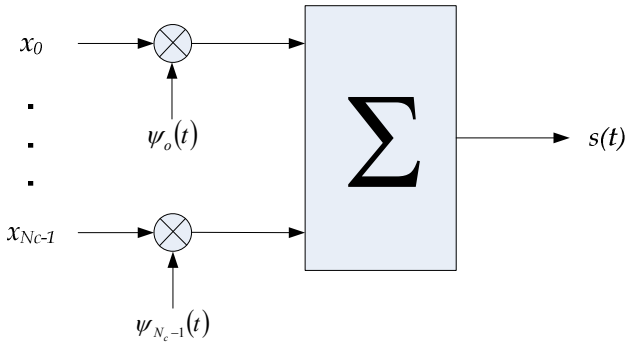


Figure 3-2. Multi-carrier modulation

A number of steps can be taken when designing a multi-carrier system to mitigate the effects of fading.

- In time domain, the data symbol duration can be made much longer than the maximum excess delay of the channel. This can be done either by choosing $T_s >> \tau_{\max}$.
- In frequency domain, the bandwidth of the sub-carriers can be made small compared to the coherence bandwidth of the channel $B_{coh} >> W/N_c$. The sub-bands then experience flat-fading, which reduces the equalization to a single complex multiplication per carrier.

We note that above mentioned steps are two sides of the same coin, depending upon the time or frequency domain.

3.2 Orthogonal Frequency Division Multiplexing

Multi-carrier modulations that use orthogonal waveform for modulating the sub-carriers are called orthogonal frequency division multiplex (OFDM) schemes. Since the sub-carriers are modulated by orthogonal waveforms, the sub-carriers are permitted to have overlapping spectrum, thus achieving higher spectrum efficiency.

A general set of orthogonal waveforms is given by [1], [12]

$$\psi_k(t) = \begin{cases} \frac{1}{\sqrt{T_s}} e^{j2\pi f_k t} & t \in [0, T_s] \\ 0 & otherwise \end{cases} \quad (3\text{-}2)$$

where

f_k is the frequency of the k^{th} sub-carrier, with $k = 0,1,\cdots,N_c - 1$

Since the waveform $\psi_k(t)$ is restricted in the time window $[0, T_s]$, the spectrum of the k^{th} sub-carrier is a sinc function in the frequency domain as illustrated in Figure 3-3 for $N_c = 3$ and $N_c = 64$. Furthermore, although the spectrum is overlapping, it does not cause interference at carrier locations due to orthogonal nature of the sub-carriers. The sub-carriers are indeed orthogonal since

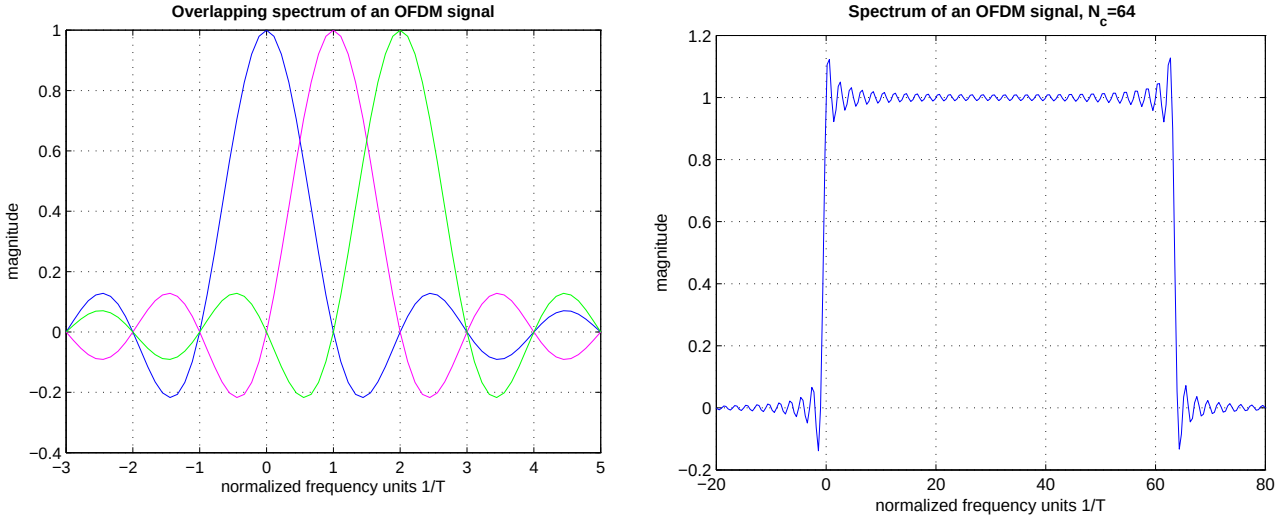
$$\int_0^{T_s} \psi_k(t) \cdot \psi_l^*(t) dt = \delta[k-l] \tag{3-3}$$

where

$\psi^*(t)$ is the complex conjugate function of $\psi(t)$ waveform, and

$\delta[k-l] = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases}$ is the kronecker delta function

Figure 3-3a also illustrates the significance of frequency synchronization in OFDM systems. The spectrum of the other carriers is only non-zero at the exact carrier locations, and orthogonality is lost in the absence of carrier synchronization. The Gibbs phenomenon is also evident by looking at the edges of the spectrum for $N_c = 64$ in Figure 3-3b.



(a) 3 sub-carriers (b) 64 sub-carriers

Figure 3-3. Overlapping spectrum of an OFDM signal

The demodulation in OFDM system is based on the orthogonality of the sub-carriers, and consists of a bank of N_c matched filters. Using correlator implementation of matched filters, we write [12]

$$y_k = \int_0^{T_s} s(t) \cdot \psi_k^*(t) dt \tag{3-4}$$

where

y_k is the demodulated data symbol from the k^{th} sub-carrier

$s(t)$ is the transmitted signal as given in Equation (3-1).

The demodulation in OFDM can be implemented as shown in the Figure 3-4. The figures used to depict an OFDM system so far illustrate the use of oscillators to implement the OFDM transmitter and correlators at the receiver. Implementation of such a system becomes quite complex as the number of sub-carriers increases. It can be shown that replacing the transmitter oscillators by IDFT and matched filter by DFT achieves the same effect [13].

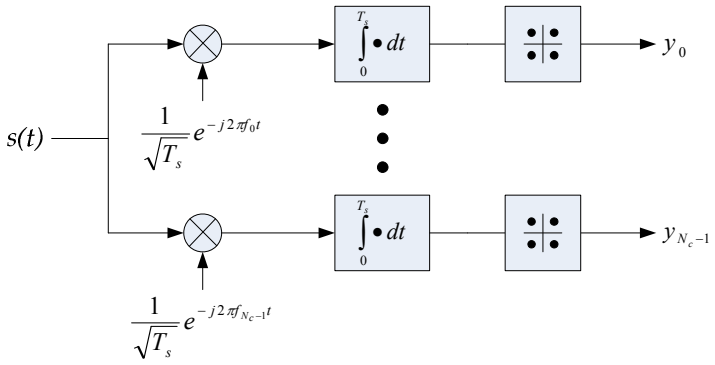


Figure 3-4. OFDM demodulation using correlators

3.3 Cyclic prefix

Intersymbolic-interference (ISI) is induced in a signal when it passes through a frequency-selective channel. In OFDM systems, it causes the loss of orthogonality of the sub-carriers, resulting in inter-carrier interference (ICI). The concept of cyclic prefix (CP) was introduced to combat this problem [14]. Cyclic prefix is a copy of the last part of the OFDM symbol that is pre-appended to the transmitted symbol, as shown in Figure 3-5, and removed before demodulation.

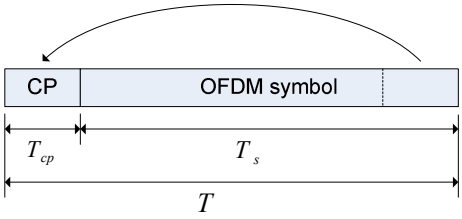


Figure 3-5. Cyclic prefix

The length of cyclic prefix should be at least as long as the significant part of the channel impulse response experienced by the transmitted signal. This way the benefit of the cyclic prefix is two-fold

- ISI is avoided since cyclic prefix acts as guard space between successive symbols.
- Linear convolution with the channel impulse response is converted into a cyclic convolution. Since cyclic convolution in time domain corresponds to a scalar multiplication in the frequency domain, the sub-carriers remain orthogonal and there is no ICI.

The advantages of CP are not without a cost. The transmitted energy required to transmit the signal increases with the length of cyclic prefix. This SNR loss (SNR_{loss}) due to the insertion of CP is given as [12]

$$SNR_{loss} = -10\log_{10}\left(1 - \frac{T_{cp}}{T}\right) \quad (3\text{-}5)$$

where

T_{cp} is the length of the cyclic prefix, T_s is the symbol time, and

$T = T_{cp} + T_s$ is the length of the transmitted symbol,

3.4 OFDM System Model

In Section 3.2, it was stated that an OFDM system can be constructed using DFT and IDFT operations and the cyclic prefix. A discrete time model of such an OFDM system is illustrated in Figure 3-6 [12].

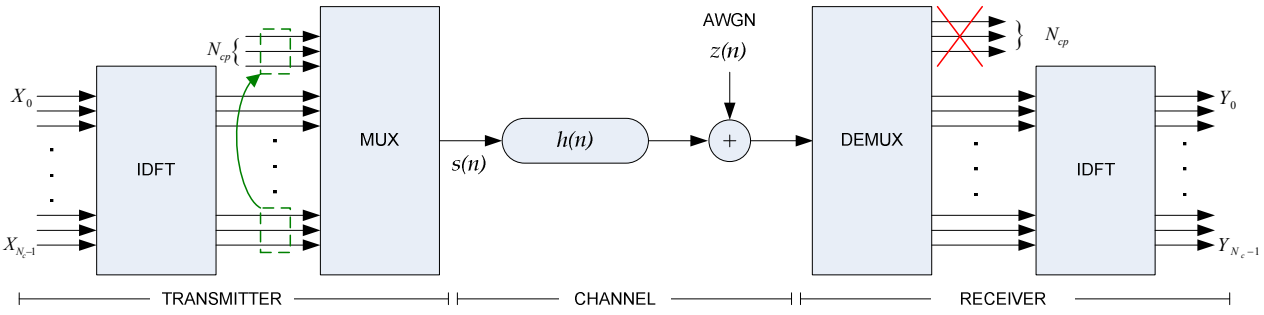


Figure 3-6. Discrete time model of an OFDM system

In the transmitter, the data bits are grouped in blocks of N_c data symbols. These groups, called OFDM symbols, are collected in a data vector $\mathbf{X} = \begin{pmatrix} X_0 & X_1 & \cdots & X_{N_c-1} \end{pmatrix}^T$. An IDFT is performed on each data vector $\mathbf{X}$ and a cyclic prefix of length N_{cp} is added. The resulting complex baseband signal for a single signaling period can written as [12]

$$s(n) = \begin{cases} \frac{1}{N_c} \sum_{k=0}^{N_c-1} X_k e^{j2\pi k(n-N_{cp})/N_c} & if \quad n \in [0, N_c + N_{cp} - 1] \\ 0 & otherwise \end{cases} \tag{3-6}$$

The transmitted signal $s(n)$ passes through a linear time-invariant discrete channel $h(n)$ and is further corrupted by zero-mean additive white Gaussian noise (AWGN) $z(n)$ with variance σ_z^2. The time-invariant constraint ensures that the channel is constant during one OFDM symbol period. Additionally, we assume that the transmitter and receiver are perfectly synchronized. Assuming that the cyclic prefix is longer than channel impulse response i.e., $h(n) = 0$ for $n < 0$ and $n > N_{cp} - 1$, we can write the received signal $r(n)$ as

$$r(n) = \sum_{\eta=0}^{N_{cp}-1} h(\eta) \cdot s(n-\eta) + z(n) \tag{3-7}$$

At the receiver, cyclic prefix is removed from the incoming sequence $r(n)$, resulting in a vector $\mathbf{r} = \begin{bmatrix} r(N_{cp}) & r(N_{cp}+1) & \cdots & r(N_{cp}+N_c-1) \end{bmatrix}^T$. The received data symbol Y_l is then obtained by performing a N_c-point DFT on this vector. Y_l is thus given as

$$Y_l = \sum_{n=0}^{N_c-1} r(N_{cp}+n) \cdot e^{-j2\pi l \cdot n/N_c} \tag{3-8}$$

Substituting for $r(n)$ from (3-7), we get

$$Y_l = \sum_{n=0}^{N_c-1}\left(\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot s\left(N_{cp}+n-\eta\right)+ z\left(N_{cp}+n\right)\right)\cdot e^{-j2\pi l\cdot n/N_c} \tag{3-9}$$

$$= \sum_{n=0}^{N_c-1}\left(\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot s\left(N_{cp}+n-\eta\right)\right)\cdot e^{-j2\pi l\cdot n/N_c} + \sum_{n=0}^{N_c-1} z(n)\cdot e^{-j2\pi l\cdot n/N_c} \tag{3-10}$$

Substituting for $s(n)$ from Equation (3-6), we obtain

$$Y_l = \underbrace{\sum_{n=0}^{N_c-1}\left(\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot\left(\frac{1}{N_c}\sum_{k=0}^{N_c-1} X_k e^{j2\pi k(n-\eta)/N_c}\right)\right)\cdot e^{-j2\pi l\cdot n/N_c}}_{\widetilde{Y}_l} + \underbrace{\sum_{n=0}^{N_c-1} z(n)\cdot e^{-j2\pi l\cdot n/N_c}}_{Z_l} \tag{3-11}$$

We now introduce $\widetilde{Y}_l$ as purely signal term and Z_l as purely noise term, and rewrite the Equation (3-11) as

$$Y_l = \widetilde{Y}_l + Z_l \tag{3-12}$$

Focusing on $\widetilde{Y}_l$, **the signal only term**, below

$$\widetilde{Y}_l = \sum_{n=0}^{N_c-1}\left(\frac{1}{N_c}\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot\sum_{k=0}^{N_c-1} X_k e^{j2\pi k(n-\eta)/N_c}\right)\cdot e^{-j2\pi l\cdot n/N_c} \tag{3-13}$$

$$= \sum_{n=0}^{N_c-1}\left(\frac{1}{N_c}\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot\sum_{k=0}^{N_c-1} X_k e^{j2\pi k\cdot n/N_c} e^{-j2\pi k\cdot\eta/N_c}\right)\cdot e^{-j2\pi l\cdot n/N_c} \tag{3-14}$$

$$= \frac{1}{N_c}\sum_{k=0}^{N_c-1} X_k\left(\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot e^{-j2\pi k\cdot\eta/N_c}\right)\sum_{n=0}^{N_c-1} e^{-j2\pi l\cdot n/N_c} e^{j2\pi k\cdot n/N_c} \tag{3-15}$$

$$= \frac{1}{N_c}\sum_{k=0}^{N_c-1} X_k\left(\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot e^{-j2\pi k\cdot\eta/N_c}\right)\sum_{n=0}^{N_c-1}\left(e^{j2\pi(k-l)/N_c}\right)^n \tag{3-16}$$

For $k = l$

$$\widetilde{Y}_l = \frac{1}{N_c}\sum_{l=0}^{N_c-1} X_l\left(\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot e^{-j2\pi l\cdot\eta/N_c}\right)\cdot N_c \tag{3-17}$$

For $k \neq l$ we have

$$\widetilde{Y}_k = \frac{1}{N_c}\sum_{k=0}^{N_c-1} X_k\left(\sum_{\eta=0}^{N_{cp}-1} h(\eta)\cdot e^{-j2\pi k\cdot\eta/N_c}\right)\sum_{n=0}^{N_c-1}\left(e^{j2\pi(k-l)/N_c}\right)^n \tag{3-18}$$

Since $\left|e^{j2\pi(k-l)/N_c}\right| < 1$, we can simplify the summation over n using geometric series as

$$\sum_{n=0}^{N_c-1}\left(e^{j2\pi(k-l)/N_c}\right)^n = \frac{1-\left(e^{j2\pi(k-l)/N_c}\right)^{N_c}}{1-e^{j2\pi(k-l)/N_c}} \tag{3-19}$$

$$= \frac{1 - e^{j2\pi(k-l)}}{1 - e^{j2\pi(k-l)/N_c}} \tag{3-20}$$

Since $k - l$ is always an integer and $e^{j2\pi m} = 1$ for integer values of m, it follows that

$$\sum_{n=0}^{N_c-1} \left(e^{j2\pi(k-l)/N_c} \right)^n = 0 \tag{3-21}$$

which leads to

$$\widetilde{Y}_l = 0 \text{ for } k \neq l \tag{3-22}$$

We now conclude that signal only term can be written as

$$\widetilde{Y}_l = X_l \cdot H_l \tag{3-23}$$

where H_l, the frequency response of the channel, is given as

$$H_l = \sum_{\eta=0}^{N_{cp}-1} h(\eta) \cdot e^{-j2\pi l \cdot \eta / N_c} \tag{3-24}$$

We now turn our attention to Z_l, **the purely noise term**, only

$$Z_l = \sum_{n=0}^{N_c-1} z(n) \cdot e^{-j2\pi l \cdot n / N_c} \tag{3-25}$$

Since Fourier transformation is a linear transformation and $z(n)$ is zero-mean AWGN with variance σ^2, the purely noise term Z_l is Gaussian in nature. We now find the mean of Z_l

$$E[Z_l] = E\left[\sum_{n=0}^{N_c-1} z(n) \cdot e^{-j2\pi l \cdot n / N_c} \right] \tag{3-26}$$

$$= \sum_{n=0}^{N_c-1} E[z(n)] \cdot e^{-j2\pi l \cdot n / N_c} = 0 \tag{3-27}$$

We now consider the correlation:

$$E[Z_l \cdot Z_k^*] = E\left[\sum_{n=0}^{N_c-1} z(n) \cdot e^{-j2\pi l \cdot n / N_c} \cdot \sum_{m=0}^{N_c-1} z^*(m) \cdot e^{j2\pi k \cdot n / N_c} \right] \tag{3-28}$$

$$= E\left[\sum_{n=0}^{N_c-1} \sum_{m=0}^{N_c-1} z(n) z^*(m) \cdot e^{-j2\pi l \cdot n / N_c + j2\pi k \cdot n / N_c} \right] \tag{3-29}$$

$$= \sum_{n=0}^{N_c-1} \sum_{m=0}^{N_c-1} E[z(n) z^*(m)] \cdot e^{-j2\pi l \cdot n / N_c + j2\pi k \cdot m / N_c} \tag{3-30}$$

Since $z(n)$ is white i.e.,

$$E\left[z(n)z^*(m)\right]=\sigma_z^2\delta[n-m], \tag{3-31}$$

The correlation of the noise in Equation (3-28) simplifies to

$$E\left[Z_l\cdot Z_l^*\right]=\sum_{n=0}^{N_c-1}\sigma_z^2\cdot e^{-j2\pi l\cdot n/N_c}\cdot e^{j2\pi k\cdot n/N_c} \tag{3-32}$$

$$=\sigma_z^2\cdot\sum_{n=0}^{N_c-1}\left(e^{j2\pi(k-l)/N_c}\right)^n \tag{3-33}$$

$$=\begin{cases}N_c\sigma_z^2 & for\ k=l\\ 0 & for\ k\neq l\end{cases} \tag{3-34}$$

Thus the transformation of zero-mean AWGN noise with $e^{-j2\pi l\cdot n/N_c}$ leads to a zero-mean AWGN noise with a scaled variance.

We can now re-write the received data symbol Y_k in Equation (3-8) as

$$Y_k=X_k\cdot H_k+Z_k \tag{3-35}$$

We thus conclude that the received data symbol Y_k on each sub-carrier k equals the data symbol X_k, that was transmitted on that sub-carrier, multiplied with the corresponding frequency-domain channel coefficient H_k in addition to the transformed noise contribution Z_k, as illustrated in Figure 3-7. It is however important to remember the conditions under which the above equation is valid i.e., the transmitter and receiver are perfectly synchronized, and that the cyclic prefix is longer than channel impulse response i.e., $h(n)=0$ for $n<0$ and $n>N_{cp}-1$.

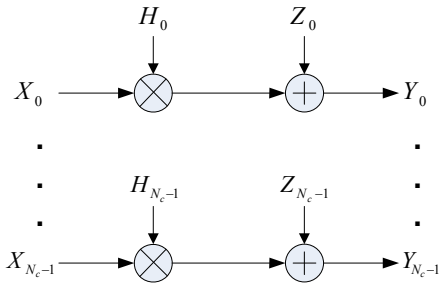


Figure 3-7. OFDM system interpreted as parallel Gaussian channels

The transmitted data symbols X_k can estimated by applying a single tap equalizer on the received data symbols Y_k, followed by a slicer. In the equalizer the receiver divides each received symbol Y_k by its corresponding channel coefficient H_k. The slicer output can then be rounded towards the nearest symbol in the modulation alphabet, thus providing a hard estimate $\hat{X}_k$ of the transmitted data symbols.

For more compact notation, a matrix equivalent may be used. For a single OFDM symbol, it equals

$$\mathbf{Y}=DIAG(\mathbf{H})\cdot\mathbf{X}+\mathbf{Z} \tag{3-36}$$

where

$DIAG(\mathbf{H})$ is a diagonal matrix with elements of $\mathbf{H}$

$\mathbf{H}$ is the frequency domain channel coefficient vector $\mathbf{H} = \begin{bmatrix} H_0 & H_1 & \cdots & H_{N_c-1} \end{bmatrix}^T$

$\mathbf{X}$ is the transmitted data symbol vector $\mathbf{X} = \begin{bmatrix} X_0 & X_1 & \cdots & X_{N_c-1} \end{bmatrix}^T$

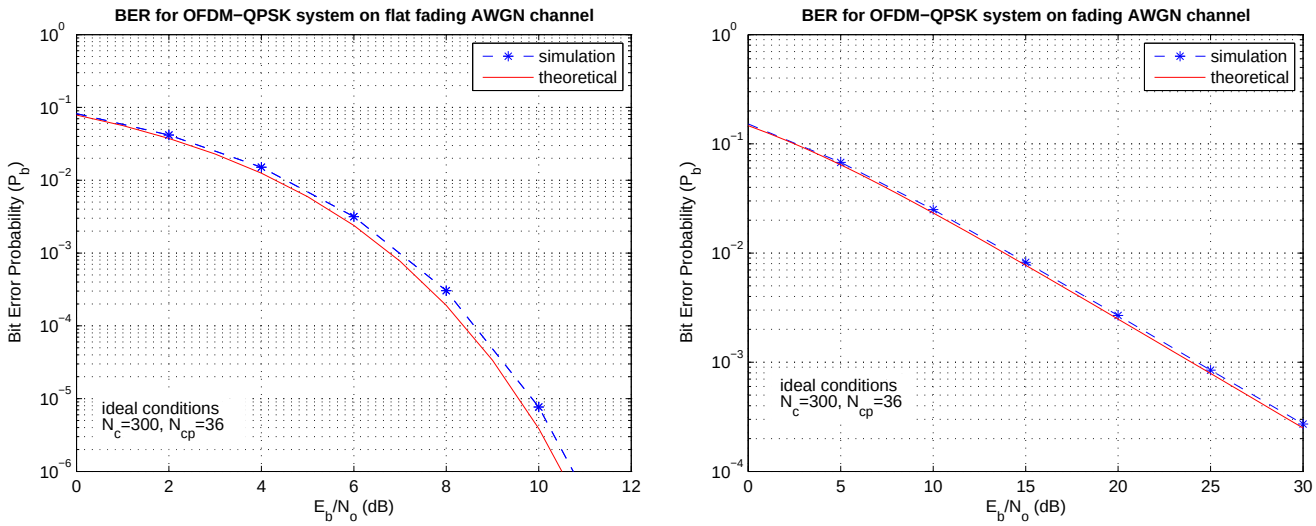
$\mathbf{Y}$ is the received data symbol vector $\mathbf{Y} = \begin{bmatrix} Y_0 & Y_1 & \cdots & Y_{N_c-1} \end{bmatrix}^T$, and

$\mathbf{Z}$ is the frequency domain scaled AWGN vector $\mathbf{Z} = \begin{bmatrix} Z_0 & Z_1 & \cdots & Z_{N_c-1} \end{bmatrix}^T$

3.5 OFDM System Performance

In the previous section, it was shown that the received data symbol transmitted on k-th sub-carrier is given by the corresponding transmitted symbol multiplied with the channel frequency response, sampled at the k-th sub-carrier frequency plus noise. For linear channel time-invariant (LTI) channel with AWGN, this corresponds to a parallel set of AWGN channels with equal SNR. As a consequence, the raw bit-error-rate (BER) performance will be identical to that of a single carrier modulation over AWGN, except for the SNR loss due to cyclic prefix.

Figure 3-8 presents the BER performance versus E_b/N_o for an OFDM-QPSK system with 300 carriers (N_c) and a cyclic prefix N_{cp} of 36. In Figure 3-8a, BER performance is plotted for an AWGN channel, while in Figure 3-8b BER performance is plotted for a frequency-selective (PA) channel. The SNR loss (0,3dB) due to cyclic prefix can also be seen in the figures.



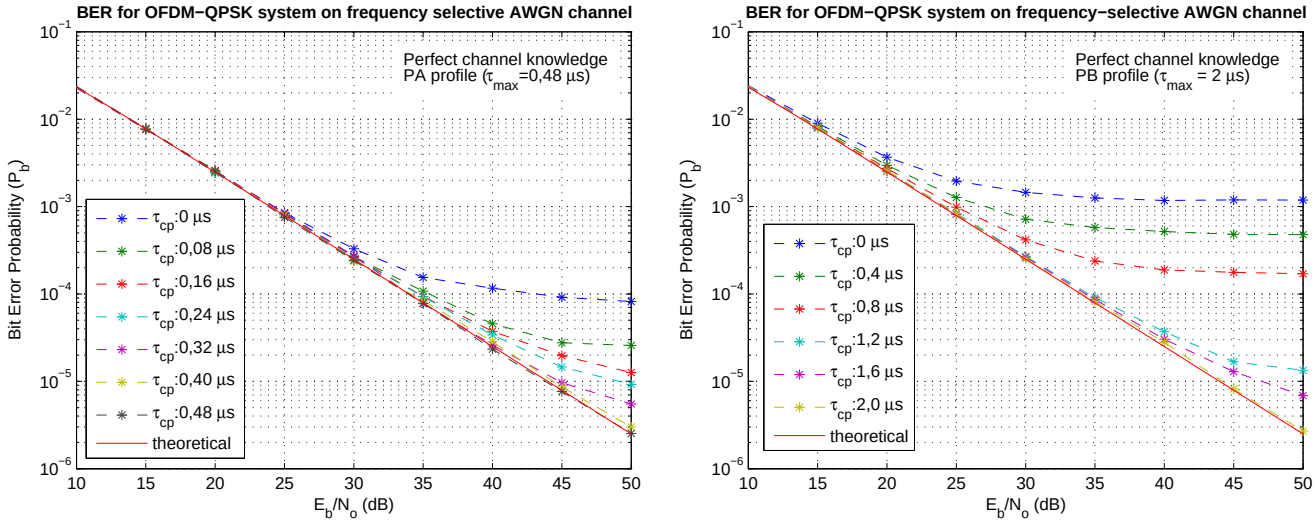
(a) AWGN channel (b) frequency selective fading channel (PA)

Figure 3-8. OFDM BER performance for AWGN

3.5.1 Impact of Cyclic prefix

Figure 3-9 illustrates the impact of cyclic prefix on BER performance versus E_b/N_o for OFDM_QPSK system on frequency selective (PA and PB) channels. We recall that excess delay of PB channel is larger than that of PA. It is evident from the figures that optimal BER performance is achieved when

cyclic prefix is longer than the excess delay of the channel. We also note that degradation due to cyclic prefix is most evident at higher SNR values because noise dominates the system performance at low SNR. Additionally, degradation is also higher when maximum excess delay ($\tau_{\max}$) of the channel is much larger than cyclic prefix length (τ_{cp}).



(a) PA channel (max. excess delay = 0,48 μ s)

(b) PB Channel (max. excess delay = 2,0 μ s)

Figure 3-9. Impact of cyclic prefix on OFDM performance

3.6 3GPP Long Term Evolution (LTE) and OFDM

The OFDM system used for our investigation of channel estimation methods is a 3GPP based system. 3GPP (3rd Generation Partnership Project) is the organization responsible for producing globally applicable Technical Specifications (TSs) and Technical Requirements (TRs) for mobile communication systems based on GSM/UMTS core networks and the radio access technologies they support (i.e., Universal Terrestrial Radio Access (UTRA)). In December 2004, 3GPP initiated a feasibility study on UTRA and UTRAN Long Term Evolution (LTE). The objective of the study was "to develop a framework for the evolution of the 3GPP radio-access technology towards a high-data-rate, low-latency and packet-optimized radio access technology" [15]. The study marks a step in the direction of Evolved UTRAN (EUTRAN) as shown in Figure 3-10.

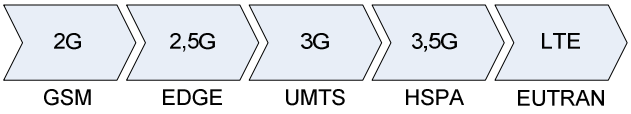


Figure 3-10. Evolution of mobile communications systems over time in 3GPP

3GPP LTE project will result in release 8 of the UMTS standard. The modifications and extensions proposed by the 3GPP LTE project are geared towards upgrading the UMTS system towards OFDM based wireless broadband internet system with voice and other services built on top [15], [16]. The LTE EUTRAN system is designed to be compatible and co-exist with the existing wireless standards such as GSM+EDGE/UMTS+HSPA.

3.6.1 EUTRAN Radio Interface

Six possible radio interface schemes were assessed in a report [17] for EUTRAN. In December 2005, the options were narrowed down to the use of Orthogonal Frequency Division Multiplexing (OFDM) in downlink and Single Carrier – Frequency Division Multiple Access (SC-FDMA) in the uplink. Other physical layer decisions include

- QPSK, 16QAM and 64QAM data-modulation scheme supported in uplink and downlink
- Use of Multiple Input Multiple Output (MIMO) schemes, with the possibility for up to four antennas at the mobile side, and four antennas at the Cell site.
- Re-use of channel coding (turbo codes) expertise from the UTRAN
- Support for FDD and TDD radio access schemes

The radio interface specifications [18] for the EUTRAN are summarized in Table 1.

	Downlink	**Uplink**
Duplex	TDD, FDD	TDD, FDD
Multiplexing	OFDMA	SC-FDMA
Modulation	QPSK, 16QAM, 64QAM	QPSK, 16QAM, 64QAM
Peak Rate	100 Mbps	20 Mbps
Spectrum	1.25 MHz, 2.5 MHz, 5 MHz, 10 MHz, 15 MHz and 20 MHz	
Mobility	Optimal performance for 0–15 km/h, moderate performance for 50–120 km/h and operational performance for 120–350 km/h	
MIMO	Up to 4x4	

Table 1. LTE Radio Interface Specifications

3.6.2 OFDM System Parameters in LTE Air Interface

A general description of physical layer of the LTE system is provided in [18] and [19]. The frame structure of the LTE system consists of a 10 ms long radio frame, containing 20 slots of 0.5 ms duration as illustrated in Figure 3-11. The frame structure is valid for both TDD and FDD transmissions.

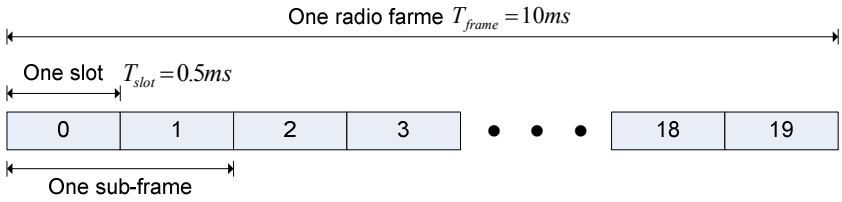


Figure 3-11. EUTRAN generic frame structure

The transmitted signal on each downlink time slot is defined by a resource grid of N_{BW}^{DL} sub-carriers and N_{symbl}^{DL} OFDM symbols as illustrated in Figure 3-12. Each element in the resource grid is called resource element and is uniquely identified by the index pair (k,l) where k and l are indices in

frequency and time domains, respectively. The frequency separation between the sub-carriers is set to 15 kHz.

The value of N_{BW}^{DL} depends upon the transmission bandwidth of the cell and fulfills $72 \le N_{BW}^{DL} \le 2048$. The value of N_{symbl}^{DL} is set to either 7 or 6, depending upon the cyclic prefix length used in the time slot. The longer cyclic prefix is used in channels with large excess delay such as hilly terrain. A resource block is defined as N_{symbl}^{DL} consecutive OFDM symbols in the time domain and $N_{\mathrm{BW}}^{\mathrm{RB}} = 12$ consecutive sub-carriers in the frequency domain. A resource block thus consists of $N_{\mathrm{symb}}^{\mathrm{DL}} \times N_{\mathrm{BW}}^{\mathrm{RB}}$ resource elements, and is the minimum amount of resource elements allocated to a user in downlink.

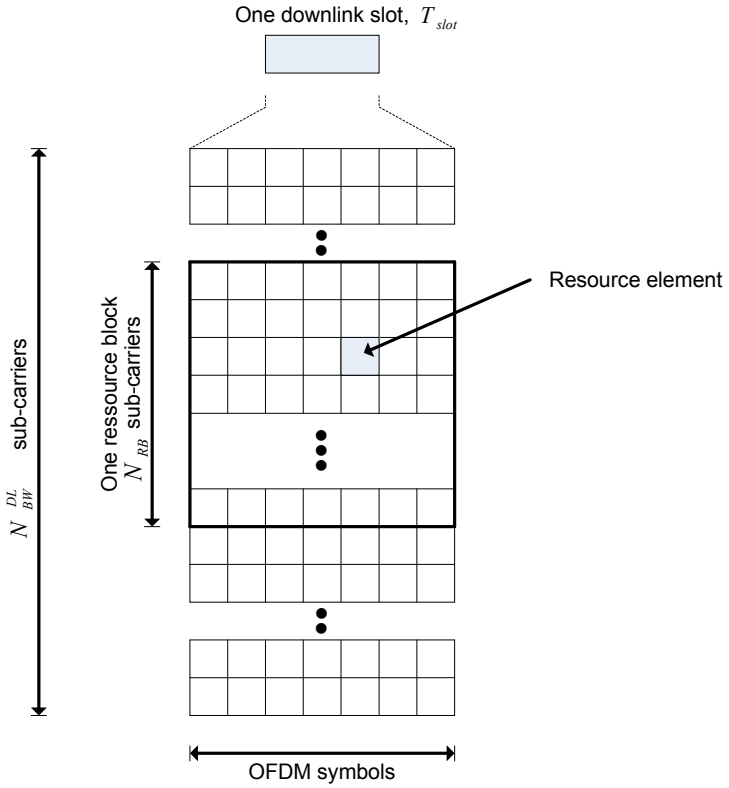


Figure 3-12. Downlink resource grid (generic frame structure, normal cyclic prefix)

OFDM downlink parameters in presented in Table 2. To take an example, the bandwidth of a 5MHz cell is divided into 512 sub-carriers, each separated with 15 kHz. Of the 512 sub-carriers, only 300 sub-carriers are occupied with data transmission, while the remaining sub-carriers serve as guard against adjacent bands. The sub-carriers are arranged symmetrically around the carrier frequency (0 Hz after down-conversion) or DC, and do not include the DC. Each time slot in the cell is 0.5 ms long, and contains 7 OFDM symbols when normal cyclic prefix is used. In such a configuration the cyclic prefix is 4.69/36 ms for the first 6 OFDM symbols, and 5.21/40 ms for the last OFDM symbol. When extended cyclic prefix is used in a cell, the 0.5ms time slot consists of 6 OFDM symbols and each OFDM symbol contains a cyclic prefix of length 16.67/128 ms.

Transmission BW		1.25 MHz	2.5 MHz	5 MHz	10 MHz	15 MHz	20 MHz
One slot duration		0.5 ms					
Sub-carrier spacing		15 kHz					
FFT size		128	256	512	1024	1536	2048
Number of occupied sub-carriers		72	150	300	600	900	1200
Number of OFDM symbols per slot (normal / extended CP)		7/6					
CP length (μ s)	**normal**	(4.69/9)×6, (5.21/10)×1	(4.69/18)×6, (5.21/20)×1	(4.69/36)×6, (5.21/40)×1	(4.69/72)×6, (5.21/80)×1	(4.69/108)×6, (5.21/120)×1	(4.69/144)×6, (5.21/160)×1
	extended	16.67/32	16.67/64	16.67/128	16.67/256	16.67/384	16.67/512

Table 2. OFDM system parameters in EUTRAN

4 Channel Estimation

While evaluating OFDM system performance in previous sections, we assumed perfect knowledge of the channel for equalization. While perfect channel knowledge can be used to find the upper limit of OFDM system performance, such perfect channel knowledge is not available in real-life and needs to be estimated. Channel estimation can be done in various ways: with or with the help of a parametric model, with the use of frequency and/or time correlation properties of the wireless channel, blind or pilot (training) based, adaptive or not.

Non-parametric methods attempt to estimate the quantities of interest (for example the frequency response) without relying on a specific channel model. Conversely, parametric estimation assumes a certain channel model, determines the parameters of this model and infers the quantities of interest. Spaced-time and spaced-frequency correlations are specific properties of channel that can be incorporated in the estimation method, improving the quality of estimate. Pilot based estimation methods are the most commonly used methods which are applicable in systems where the sender emits some known signal. Blind estimation, on the other hand, relies on some properties of the signal (e.g., cylco-stationarity of the signal) and is rarely used in practical OFDM systems. Adaptive channel estimation methods are typically used for rapidly time-varying channel.

In this chapter, we investigate pilot-based non-adaptive channel estimation methods. We do, however, exploit channel correlation properties in one of the channel estimation methods. Figure 4-1 presents an OFDM system that utilizes pilot-based channel estimation method for equalization at the receiver end. At the *transmitter*, binary data is mapped to a specific modulation (QPSK, 16QAM, 64QAM) and the modulated data undergoes serial-to-parallel (S/P) conversion, forming a vector of $\left(N_0 - N_{pilot}\right)$ symbols [12], [20]. N_0 is the number of occupied sub-carriers and N_{pilot} is the number of pilot sub-carriers. Known N_{pilot} pilot symbols are then inserted into the modulated data, forming frequency-domain transmitted data X_k of length N_o. N_c-point Inverse Fourier transform (IFFT) is performed on zero-padded X_k to generate time-domain vector $x(n)$. Cyclic prefix of N_{cp} is then pre-appended to $x(n)$ forming $x_g(n)$ vector of $N_c + N_{cp}$ symbols. After parallel to serial conversion (P/S),

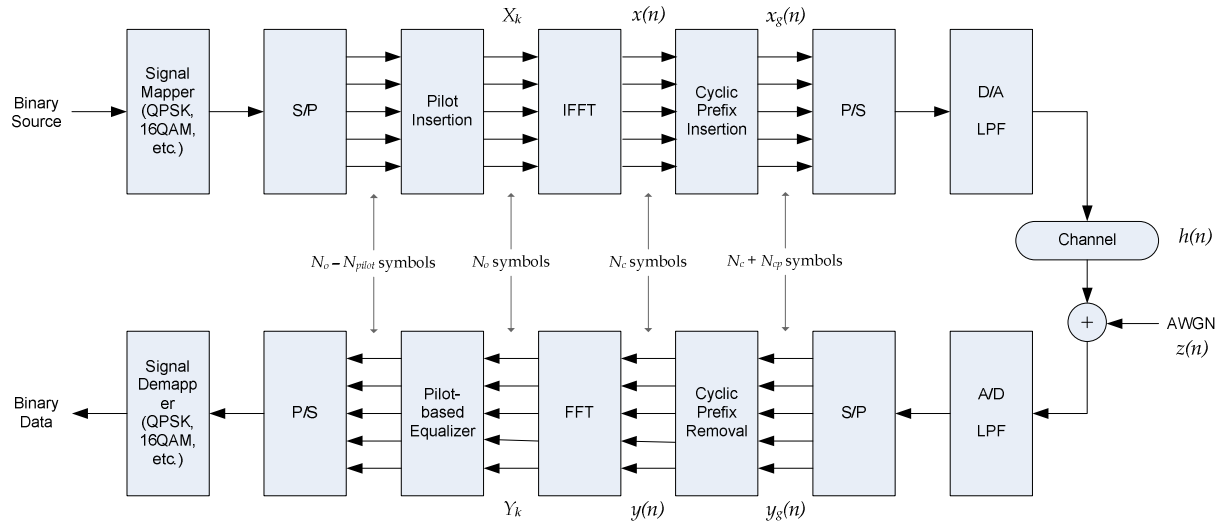


Figure 4-1. Pilot-based OFDM system model

D/A conversion and low-pass filtering, $x_g(n)$ is transmitted over a linear time-invariant AWGN channel $h(n)$ with noise $z(n)$.

At the *receiver*, the opposite set of operations is performed. After D/A conversion and low-pass filtering, the received signal goes through a S/P conversion and is stored in $y_g(n)$. Cyclic prefix is then removed from the $y_g(n)$, creating a N_c-dimensional received symbol vector $y(n)$. N_c-point Fast Fourier Transform (FFT) is then applied on $y(n)$, creating frequency based Y_k vector. Zero-padding is dropped from Y_k, shortening its length from N_c to N_o. Equalization is applied using a pilot-based channel estimation method, and the pilot symbols are removed from the equalized signal. The equalized data then undergoes a P/S conversion and demodulation, creating estimates of the transmitted binary data.

In Section 3.4, we showed the relationship between transmitted signal X_k and received Y_k as

$$Y_k = H_k \cdot X_k + Z_k \tag{4-1}$$

where

Z_k is the frequency-domain noise at the k-th sub-carrier frequency, and

H_k is the channel transfer function at the k-th sub-carrier frequency

To estimate the channel, pilots symbols are needed. We assume that every p-th sub-carrier contains known pilot symbols (X_{pk}). Using the known pilots symbols (X_{pk}) and the received symbols (Y_{pk}) at those pilot sub-carriers, we can calculate the raw channel estimate $(\hat{H}_{pk})$ at pilots as

$$\hat{H}_{pk} = \frac{Y_{pk}}{X_{pk}} + \frac{Z_{pk}}{X_{pk}} = H_{pk} + Z'_{pk} \tag{4-2}$$

where

Z_{pk} is the noise contribution at the pk-th sub-carrier, Z'_{pk} is a scaled noise contribution at that sub-carrier.

Different methods can then be applied to estimate the channel over all sub-carrier frequencies and not just at pilot sub-carrier frequencies. Figure 4-2 looks at details of the channel estimation method described above. After FFT and zero-padding removal, the received signal at pilot locations is extracted from the signal. Raw channel is estimated at pilot sub-carrier frequencies using known pilots. Different methods can then be applied to estimate channel at all frequencies $\hat{H}_{pk}$ from raw channel estimate. In the subsequent sections, we investigate the impact of pilot allocation on channel estimation performance, and explore the different methods used for channel estimation.

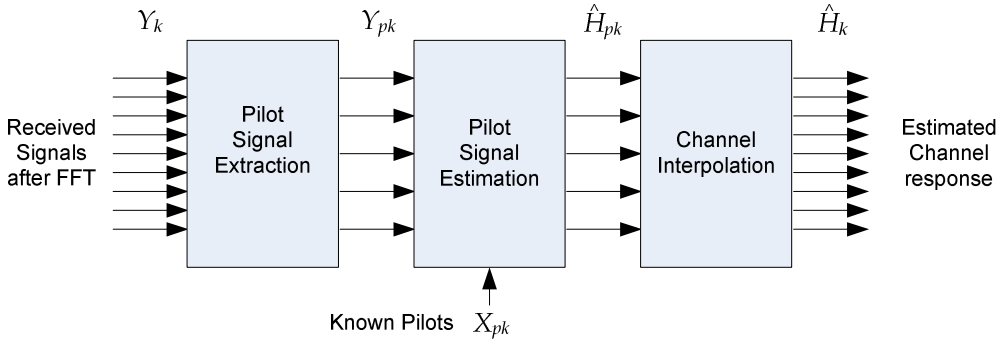


Figure 4-2. Frequency domain channel estimation

4.1 Time-Frequency Interpretation and Pilot Allocation

The OFDM transmission in Figure 4-1 can also be described using a two-dimensional lattice in the time-frequency domain. Figure 4-3 shows such a time-frequency domain representation where EUTRAN OFDM symbols are transmitted sequentially. As shown in the Figure 4-3, an OFDM symbol consists of N_c sub-carriers, symmetrically arranged around the carrier frequency (or DC). Note that not only N_o out of N_c sub-carriers are used for data transmission; no data is transmitted on DC and edge sub-carriers (to avoid interference in adjacent bands). The possible values for N_o and N_c are listed in Table 2 on Page 38.

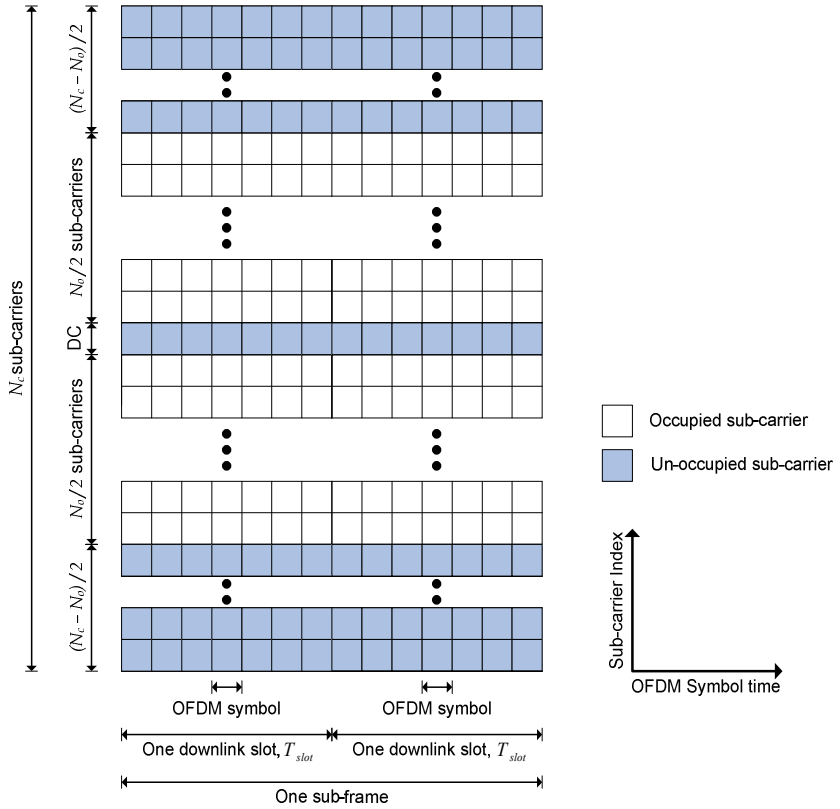
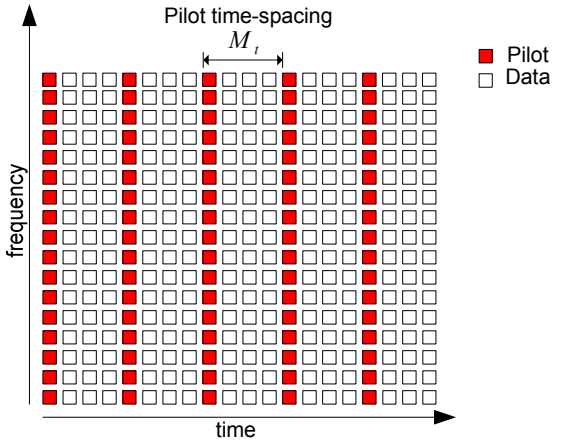


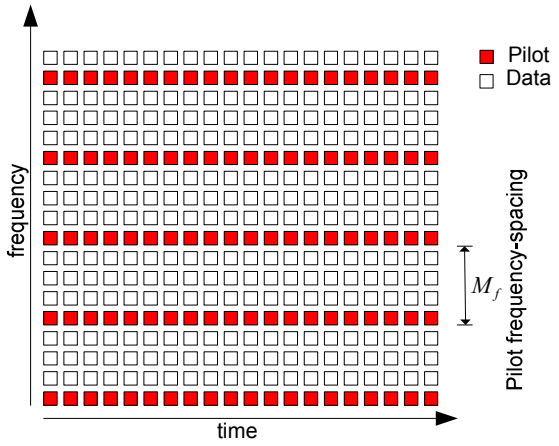
Figure 4-3. Time-Frequency lattice for OFDM transmission (generic frame structure, normal cyclic prefix)

Different possibilities exist for allocating pilots in the time-frequency domain of an OFDM system. We discuss three such possibilities as illustrated in Figure 4-4.

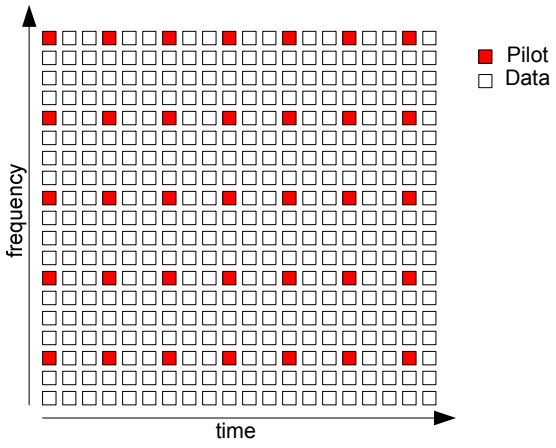
- An entire OFDM symbol may be allocated as pilot as shown in Figure 4-4a. Such an allocation will be highly beneficial for channel estimation in highly frequency-dispersive and low Doppler channels at the expense of sacrificing data rate. In Appendix A.5 we show that the raw channel estimate in an all frequency pilot is indeed the least-squares solution to channel in frequency domain [21].
- Pilots may be transmitted on individual sub-carriers during the entire transmission period as shown in Figure 4-4b. Such a strategy will be advantageous in moderately frequency-selective and high Doppler channels [21].
- Pilots may be allocated in spaced intervals in time and frequency as illustrated in Figure 4-4c and Figure 4-4d. Depending upon the time-frequency pilot spacing and channel properties, such an allocation strategy will work well in both high frequency-selective and high Doppler channels.



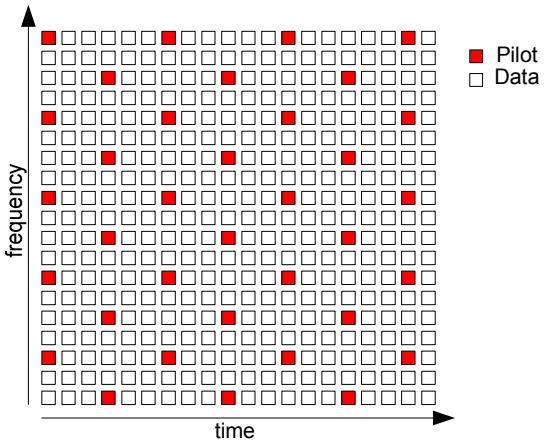
(a) All frequency-time spaced pilot allocation



(b) Frequency spaced - all time pilot allocation



(c) Frequency spaced - time spaced pilot allocation



(d) Frequency spaced - time spaced pilot allocation

Figure 4-4. Different possibilities for pilot allocation

Note that frequency-spacing, M_f, and time-spacing, M_t, between the pilots should be reasonable enough to fully estimate the channel. We note that for ideal reconstruction, the frequency-spacing M_f between the pilots should fulfill the requirements of the 2-D sampling theorem as [22]

$$M_f < \frac{N_c}{2W\tau_{\max}} \tag{4-3}$$

where

W is the bandwidth of the OFDM system,

$\tau_{\max}$ is the maximum excess delay of the channel, and

N_c is the number of sub-carriers in the OFDM system

Similarly, time-spacing M_t between the pilots should fulfill [22]

$$M_t < \frac{W}{2\left(N_c + W \cdot T_{cp}\right)\left(f_d + v\right)} \tag{4-4}$$

where

T_{cp} is the length of the cyclic prefix,

f_d and v are the the maximum Doppler spread and Doppler shift, respectively.

Figure 4-5 shows the pilot allocation in EUTRAN generic frame structure and normal cyclic prefix setup [19]. The frequency spacing between pilots is 6 sub-carriers, and time-spacing between pilots is three within a time slot and two with the incoming OFDM symbol as shown in the Figure 4-5.

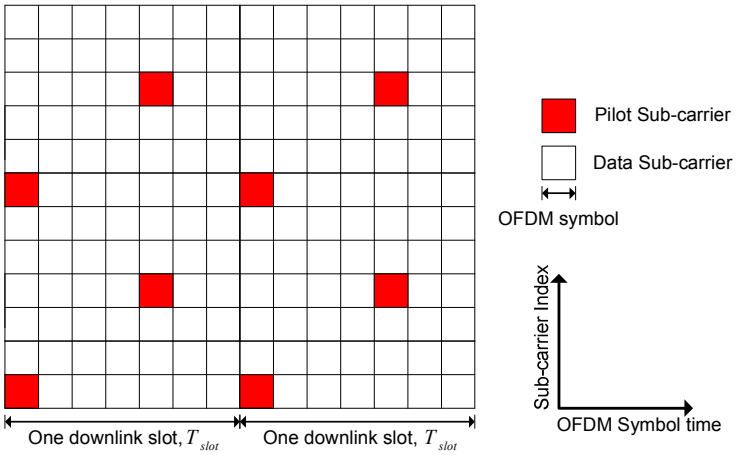


Figure 4-5. Pilot allocation in EUTRAN (generic frame structure, normal cyclic prefix)

4.2 Channel Estimation Methods

Channel estimation in OFDM is a two-dimensional (2-D) problem i.e., channel needs to be estimated in time-frequency domain as illustrated in Figure 4-6. Hence 2-D methods could be applied to estimate the channel from pilots. However, due to the computational complexity of 2-D estimators, the scope of channel estimators can be limited to one-dimensional (1-D). The idea behind 1-D estimators is to

estimate the channel in one dimension (say frequency) and later estimate the channel in the second dimension (say time), thus obtaining a 2-D channel estimate.

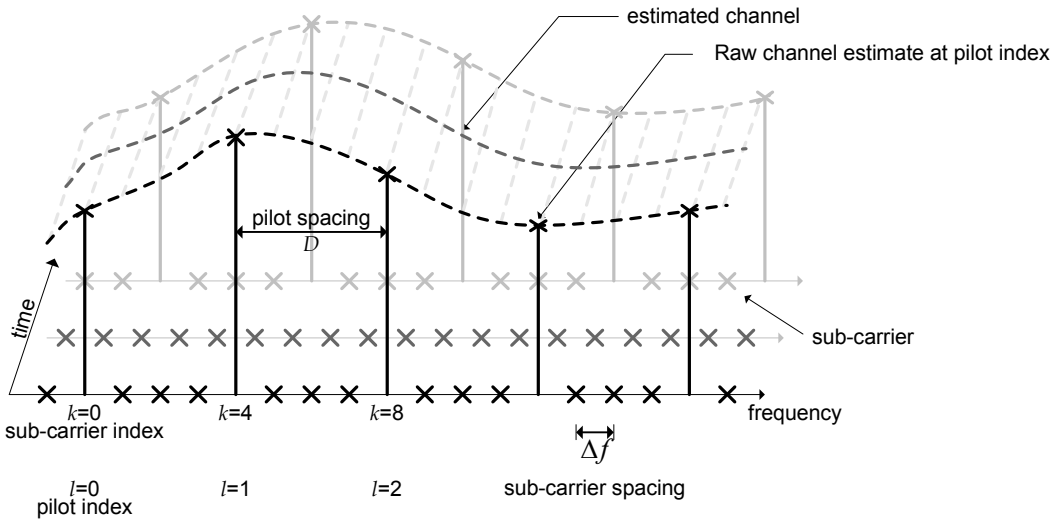


Figure 4-6. Channel estimation in time-frequency domain

Channel estimation methods, 2-D or 1-D dimensional, can be characterized into three types

- Simple linear interpolation
- Generalized linear network models based on orthogonal polynomials (least-squares method)
- Wiener filtering using second-order statistics of the channel (LMMSE method)

The focus of this report has been on 1-D methods, where we apply channel estimation in frequency-domain first, and later use temporal interpolation by either linear interpolation or simple averaging between two pilot positions. We note that channel estimate could also be done in reverse order i.e., time-domain estimation followed by frequency-domain estimation. This method was not investigated since it causes delay in the decoding of an OFDM symbol.

We now explore the channel estimation methods in the following sub-sections.

4.2.1 Linear Interpolation

Linear interpolation is the simplest method to estimate channel from raw channel estimates at pilot frequencies. This is done by linearly interpolating the raw channel estimates at the two nearest pilot sub-carriers. Although linear interpolation provides some limited noise reduction of the channel estimates at data locations, it is the simplicity of the solution that is attractive. Some additional gain may be obtained by averaging the interpolated channel estimates using a sliding window. We note that averaging window length (number of used pilots in window) is inversely proportional to the coherence bandwidth of the channel.. For a channel with high frequency coherence bandwidth (small excess delay), a large number of interpolated channel coefficients may be used for averaging and vice versa. We note that other options exist for interpolation-based methods such as Gaussian interpolation filters and cubic spline interpolation filters [23].

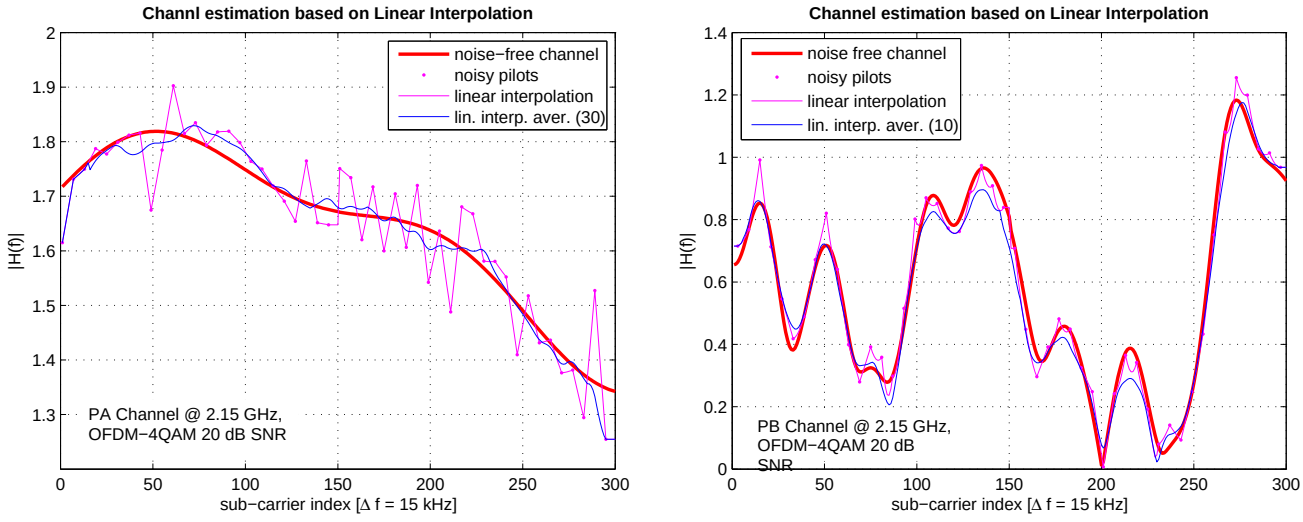
4.2.1.1 2-D Linear Interpolation

2-D linear interpolation method estimates the channel by interpolating between channel transfer function at 2-D frequency and time index pairs $\hat{H}_{k,l}$ and $\hat{H}_{m,n}$ [23]. Note that $\hat{H}_{k,l}$ is the raw channel estimate at sub-carrier frequency k and time l, and $\hat{H}_{m,n}$ is the raw channel estimate at sub-carrier frequency m and time n.

4.2.1.2 1-D Linear Interpolation

1-D linear interpolation method estimates the channel by interpolating between channel transfer function between $\hat{H}_{k,l}$ and $\hat{H}_{m,l}$ (fixed time), or between $\hat{H}_{k,l}$ and $\hat{H}_{k,n}$ (fixed frequency) raw channel estimates [20], [21]. The interpolation method implemented, we perform interpolation in frequency-domain first and later apply interpolation in time-domain. We note that linear interpolation is the channel estimation method implemented for channel estimation in time-domain.

Figure 4-7 illustrates channel estimation based on 1-D frequency-domain linear interpolation method for a high coherence bandwidth channel (Figure 4-7a) and low coherence bandwidth channel (Figure 4-7b). It is evident that simply using linear interpolation on raw channel estimates provides inferior channel estimate than using post averaging after linear interpolation. Thirty and ten sub-carriers were used for post averaging after linear interpolation in Figure 4-7a and Figure 4-7b respectively. Every 6th sub-carrier was used as a pilot and QPSK modulation was used in downlink. The carrier frequency was set to 2.15 GHz, the mobile speed was set to 120 km/h and the SNR over the channel was set to 20 dB.



(a) PA channel (high coherence bandwidth) (b) PB channel (low coherence bandwidth)

Figure 4-7. Channel estimation using linear interpolation

4.2.2 General Linear Model (Least Squares Method)

The generalized linear model [24] uses a set basis function to estimate the channel transfer function. The basis functions could either be 2-D or 1-D. We now briefly discuss the general linear model based on 2-D basis functions before proceeding to 1-D basis functions. Note that 1-D basis function

estimation was implemented in frequency-domain, and then interpolation was applied in time-domain to get the estimated channel.

4.2.2.1 2-D General Linear Models

The channel transfer function $H_{k,l}$ at frequency-time index (k,l) can be modeled as a linear weighed sum of 2–D basis function evaluated at the k-th sub-carrier frequency and l-th time index as

$$H_{k,l} = \sum_{i=0}^{N-1} \theta_i \cdot \Phi_i(f_k, t_l) \tag{4-5}$$

where

$\Phi_i(f_k, t_l)$ is the i-th basis function evaluated at that k-th sub-carrier frequency f_k and l-th time index

θ_i is the weighing factor of the basis function $\Phi_i(f_k, t_l)$,

N is the number of basis functions used in the linear model.

Different options for selecting the basis functions exist such as frequency-time polynomials [25], [26], [28], and [29] or other 2-D functions. Least-squares method could be applied to estimate the optimal weighing factor, requiring inversion of two-dimensional matrices. Due to the due the computational complexity required to estimate the optimal weighing factors, the general linear model using 2-D basis function was not investigated.

4.2.2.2 1-D General Linear Models

The channel estimation problem can also be solved using 1-D generalized linear model framework [25], [26], [28], and [29]. The channel transfer function H_k can be modeled as a linear weighed sum of some basis function evaluated at the k-th sub-carrier frequency as

$$H_k = \sum_{i=0}^{N-1} \theta_i \cdot \Phi_i(f_k) \tag{4-6}$$

where

$\Phi_i(f_k)$ is the i-th basis function evaluated at that k-th sub-carrier frequency f_k,

θ_i is the weighing factor of the basis function $\Phi_i(f_k)$,

N is the number of basis functions used in the linear model.

The generalized linear model can be used to re-write the raw channel estimates, $\hat{H}_{pk}$, in Equation (4-4) as

$$\hat{H}_{pk} = H_{pk} + Z_{pk} = \sum_{i=0}^{N-1} \theta_i \cdot \Phi_i(f_{pk}) + Z_{pk} \tag{4-7}$$

Where

Every p -th sub-carrier is a pilot,

H_{pk} is the actual channel transfer function at pk -th carrier,

Z_{pk} is the noise at the pk -th carrier,

and $\Phi_i(f_{pk})$ is the i -th basis function evaluated at that pk -th sub-carrier frequency f_k.

The raw channel observations can be collected in a matrix set of equations as

$$\underbrace{\begin{bmatrix} \hat{H}_0 \\ \hat{H}_p \\ \vdots \\ \hat{H}_{p(N_p-1)} \end{bmatrix}}_{\hat{\mathbf{H}}} = \underbrace{\begin{bmatrix} \Phi_0(f_0) & \Phi_1(f_0) & \cdots & \Phi_{N-1}(f_0) \\ \Phi_0(f_p) & \Phi_1(f_p) & \cdots & \Phi_{N-1}(f_p) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_0(f_{p(N_p-1)}) & \Phi_1(f_{p(N_p-1)}) & \cdots & \Phi_{N-1}(f_{p(N_p-1)}) \end{bmatrix}}_{\mathbf{G}} \cdot \underbrace{\begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_{N-1} \end{bmatrix}}_{\theta} + \underbrace{\begin{bmatrix} Z_0 \\ Z_p \\ \vdots \\ Z_{p(N_p-1)} \end{bmatrix}}_{\mathbf{Z}} \tag{4-8}$$

where

N_p is the number of pilot sub-carriers.

Using matrix notation, we re-write the above as

$$\hat{\mathbf{H}} = \mathbf{G} \cdot \boldsymbol{\theta} + \mathbf{Z} \tag{4-9}$$

A least-squares estimate $\hat{\boldsymbol{\theta}}_{LS} = \begin{bmatrix} \hat{\theta}_{LS\,0} & \cdots & \hat{\theta}_{LS\,N-1} \end{bmatrix}^T$ of the weighing matrix $\boldsymbol{\theta}$ is derived in Appendix A.6 as

$$\hat{\boldsymbol{\theta}}_{LS} = \left(\mathbf{G}^H \mathbf{G}\right)^{-1} \mathbf{G}^H \cdot \hat{\mathbf{H}} \tag{4-10}$$

The least-square estimate $\hat{\boldsymbol{\theta}}_{LS}$ can then be used to estimate the channel $\hat{H}_{LS,k}$ at regular sub-carrier frequencies f_k as

$$\hat{H}_{LS,k} = \sum_{i=0}^{N-1} \hat{\theta}_{LS_i} \cdot \Phi_i(f_k) \tag{4-11}$$

A number of different options exist while selecting the basis functions such as polynomials, Fourier series, discrete cosine transform function, discrete sine transform series etc. In our investigation, we limit the scope of basis functions to polynomials. We considered two types of polynomials for use in our linear model: orthogonal polynomial at spaced pilot intervals (see Appendix A.7), and Legendre polynomials in interval $[-1,+1]$ (see Appendix A.8). The orthogonal polynomial discussed in Appendix A.7 is referred to as Poly A in this report. We note that Legendre polynomials are not orthogonal with respect to inner product at spaced pilot intervals.

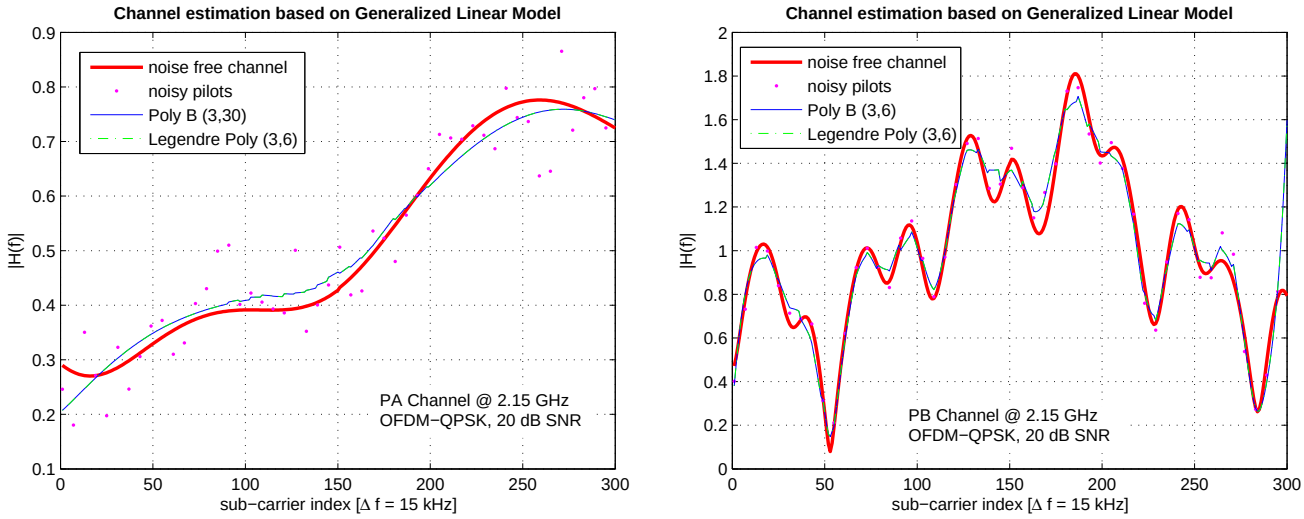
The reasons for using orthogonal polynomials are three-fold:

- The calculation of $\hat{\boldsymbol{\theta}}_{LS}$ remains numerically stable since the product $\mathbf{G}^H \mathbf{G}$ is a diagonal matrix for orthogonal polynomials
- Since $\mathbf{G}^H \mathbf{G}$ is a diagonal matrix, the inverse $\left(\mathbf{G}^H \mathbf{G}\right)^{-1}$ can be calculated at low computational complexity

- The degree of orthogonal polynomials can be increased without changes in any previous calculation results of $\hat{\theta}_{LS}$

When applying the polynomial-based channel estimation, we used a sliding window approach to estimate the channel. This was done to avoid using very high order polynomials to estimate a highly frequency-selective channel such as the channel in Figure 4-8b. The sliding window approach allows for local approximations to the channel transfer function using a window length larger than the polynomial order, and thus providing a better estimate. Two parameters need to be investigated when using the polynomial-based linear model: polynomial order and window length. A low order (2, 3) polynomial is recommended to avoid over-fitting to data points, and window length of atleast twice the polynomial order is suggested to minimize the effect of out-liers in data set.

Figure 4-8 illustrates channel estimation based on linear interpolation method using Poly A and Legendre for a high coherence bandwidth channel (Figure 4-8a) and low coherence bandwidth channel (Figure 4-8b). Notice that the estimated channel based on Poly A and Legendre are similar when the same parameters (window length and polynomial order) are used for interpolation. However, recall that Poly A being an orthogonal polynomial at pilot locations is much more efficient to calculate. In Figure 4-8a, the polynomial order was 3 and 30 raw channel estimate values were used to generate the estimate. In Figure 4-8b, the polynomial order was kept the same, while raw channel estimate values was decreased to 6. The pilot spacing was set to 6 and QPSK modulation was used in downlink. The carrier frequency was set to 2.15 GHz, mobile speed was 120 km/h and the SNR was set to 20 dB.



(a) PA channel (high coherence bandwidth) (b) PB channel (low coherence bandwidth)

Figure 4-8. Channel estimation using polynomial based generalized linear model

4.2.3 Wiener Filtering (LMMSE method)

Wiener filtering method uses knowledge of channel properties (spaced-frequency and spaced-time correlation) to estimate the unknown channel transfer function at non-pilot sub-carriers. These properties are assumed to be known at the receiver for the estimator to perform optimally.

4.2.3.1 2-D Wiener Filtering

In Section 2.1 Equation (4-10), we introduced the *spaced-frequency spaced-time correlation function* of the channel. The 2-D Wiener filtering method uses the knowledge of spaced-frequency spaced-time correlation function to minimize the expected mean-squared error between the actual and estimated channel [30], [31], [32], [23], [34], [35]. It can be shown that spaced-frequency spaced-time correlation function, $R(\Delta f, \Delta t)$, can be split up as [23]

$$R(\Delta f, \Delta t) = R(\Delta f) \cdot R(\Delta t) \tag{4-12}$$

where

$R(\Delta f)$ is the spaced-frequency correlation expressed in Equation (4-11), and

$R(\Delta t)$ is the spaced-time correlation expressed in Equation (4-13).

The optimal Wiener filter coefficients can then be estimated by minimizing the expected mean-squared error between the actual and estimated channel using the spaced-frequency spaced-time correlation functions, requiring inversion of two-dimensional matrices [30], [31], [32], [23], [34], [35]. Due to the computational complexity required to estimate the optimal Wiener coefficients, the 2-D Wiener filtering method was not investigated.

4.2.3.2 1-D Wiener Filtering

The Wiener filtering or Linear Minimum Mean Squares (LMMSE) estimator tries to minimize the expected mean-squared error between the actual and estimated channel. The theoretical framework for 1-D Wiener filtering is presented in and presented in [30], [31], [33], [34] and repeated in Appendix A.9, while we present the conclusions in this section.

The LMMSE $H_{LMMSE,k}$ estimate at k-th sub-carrier is calculated by filtering the of raw channel estimate vector $\hat{\mathbf{H}} = \begin{bmatrix} \hat{H}_0 & \hat{H}_p & \cdots & \hat{H}_{p(N_p-1)} \end{bmatrix}^T$ by a Wiener filter $\mathbf{c}_{LMMSE} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{N_p-1} \end{bmatrix}^T$ as follows

$$\hat{H}_{LMMSE,k} = \mathbf{c}_{LMMSE}^H \hat{\mathbf{H}} \tag{4-13}$$

where

c_i is the i-th filter coefficients, and $\hat{H}_{pi}$ is the i-th raw channel estimate.

The Wiener filter coefficients are calculated as (see Appendix A.9) [32]

$$\mathbf{c}_{LMMSE} = \left(\mathbf{R}_h + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{r} = \left(\mathbf{R}_{h,\sigma}\right)^{-1} \mathbf{r} \tag{4-14}$$

where

$\mathbf{R}_h$ is the autocorrelation matrix of the channel at pilot locations, σ^2 is the noise-variance per sub-carrier, and $\mathbf{r}$ is the cross-correlation vector of the channel at k-th sub-carrier and the channel at pilot locations.

Note that $\mathbf{R}_h$ is positive semi-definite, and given as

$$\mathbf{R}_h = \begin{bmatrix} E\{\hat{H}_0\hat{H}_0^*\} & E\{\hat{H}_0\hat{H}_p^*\} & \cdots & E\{\hat{H}_0\hat{H}_{p(N_p-1)}^*\} \\ E\{\hat{H}_p\hat{H}_0^*\} & E\{\hat{H}_p\hat{H}_p^*\} & \cdots & E\{\hat{H}_p\hat{H}_{p(N_p-1)}^*\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{\hat{H}_{p(N_p-1)}\hat{H}_0^*\} & E\{\hat{H}_{p(N_p-1)}\hat{H}_p^*\} & \cdots & E\{\hat{H}_{p(N_p-1)}\hat{H}_{p(N_p-1)}^*\} \end{bmatrix} \quad (4\text{-}15)$$

and $\mathbf{r}$

$$\mathbf{r} = \begin{bmatrix} E\{\hat{H}_0\hat{H}_k^*\} \\ E\{\hat{H}_p\hat{H}_k^*\} \\ \vdots \\ E\{\hat{H}_{p(N_p-1)}\hat{H}_k^*\} \end{bmatrix} \quad (4\text{-}16)$$

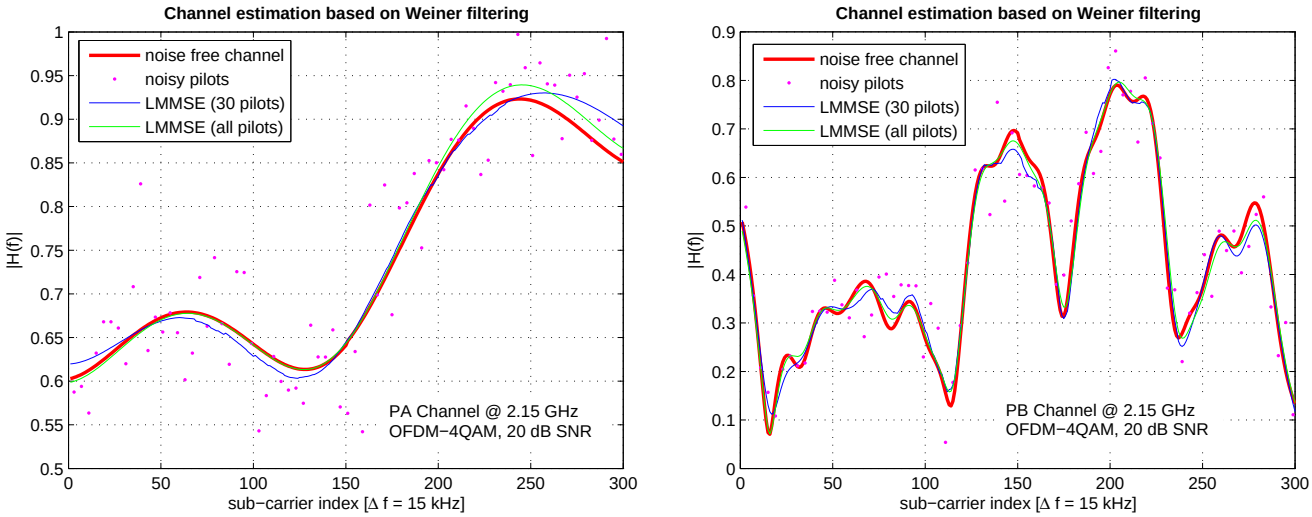
where

H_{pi} is the channel at the pi-th pilot location, and H_k is the channel at the k-th sub-carrier.

In order to calculate the LMMSE channel estimate, we should have prior knowledge to the spaced-frequency correlation function and the noise covariance of the transmission channel. Recall, that spaced-frequency correlation of the correlation can be calculated by taking Fourier transform of power delay profiles (PDPs) of the channel (see Appendix A.2). In our simulations we have assumed that perfect knowledge of power delay profile and noise variance of the channel is available prior to the LMMSE estimate. However, we note that incorrect power delay profile and noise variance will lead to performance degradation.

Calculating channel estimate using Wiener filtering includes inverting, $\mathbf{R}_{h,\sigma} = \mathbf{R}_h + \sigma^2\mathbf{I}$, the sum of autocorrelation matrix and diagonal noise matrix. This becomes a computationally demanding task when the size of matrix $\mathbf{R}_{h,\sigma}$ increases i.e., when we use a large number of pilots. Moreover, the memory requirement to save matrix $\mathbf{R}_h$ increase as the number of pilots increases. This can be dealt by using a sliding window approach for estimating the LMMSE channel estimate. By limiting the number of pilots to estimate, the size of the matrix $\mathbf{R}_{h,\sigma}$ can be reduced, hence reducing the computational complexity and memory requirements. Note that using a sliding window approach leads to degradation in LMMSE channel estimate using the entire data set. Alternatively, we can simplify the inversion of matrix $\mathbf{R}_{h,\sigma} = \mathbf{R}_h + \sigma^2\mathbf{I}$ by using singular-value-decomposition of $\mathbf{R}_h$ and the matrix inversion lemma. This is discussed in Appendix A.9.1.

Figure 4-9 illustrates channel estimation based on Wiener filtering method using all pilots and a sliding windowed approach using 30 pilots at a time for a low coherence bandwidth channel (Figure 4-9a) and a low coherence bandwidth channel (Figure 4-9b). We note that Wiener filter based on all pilots performs better than windowed Wiener filter based on 30 pilots. Decreasing the window size (i.e. number of pilots for channel estimation) will result in further degradation and vice versa. The pilot spacing was set to 5 and QPSK modulation was used in downlink. The carrier frequency was set to 2.15 GHz and the SNR over the channel was set to 20 dB.



(a) PA channel (high coherence bandwidth) (b) PB channel (low coherence bandwidth)

Figure 4-9. Channel estimation using Wiener filtering (LMMSE method)

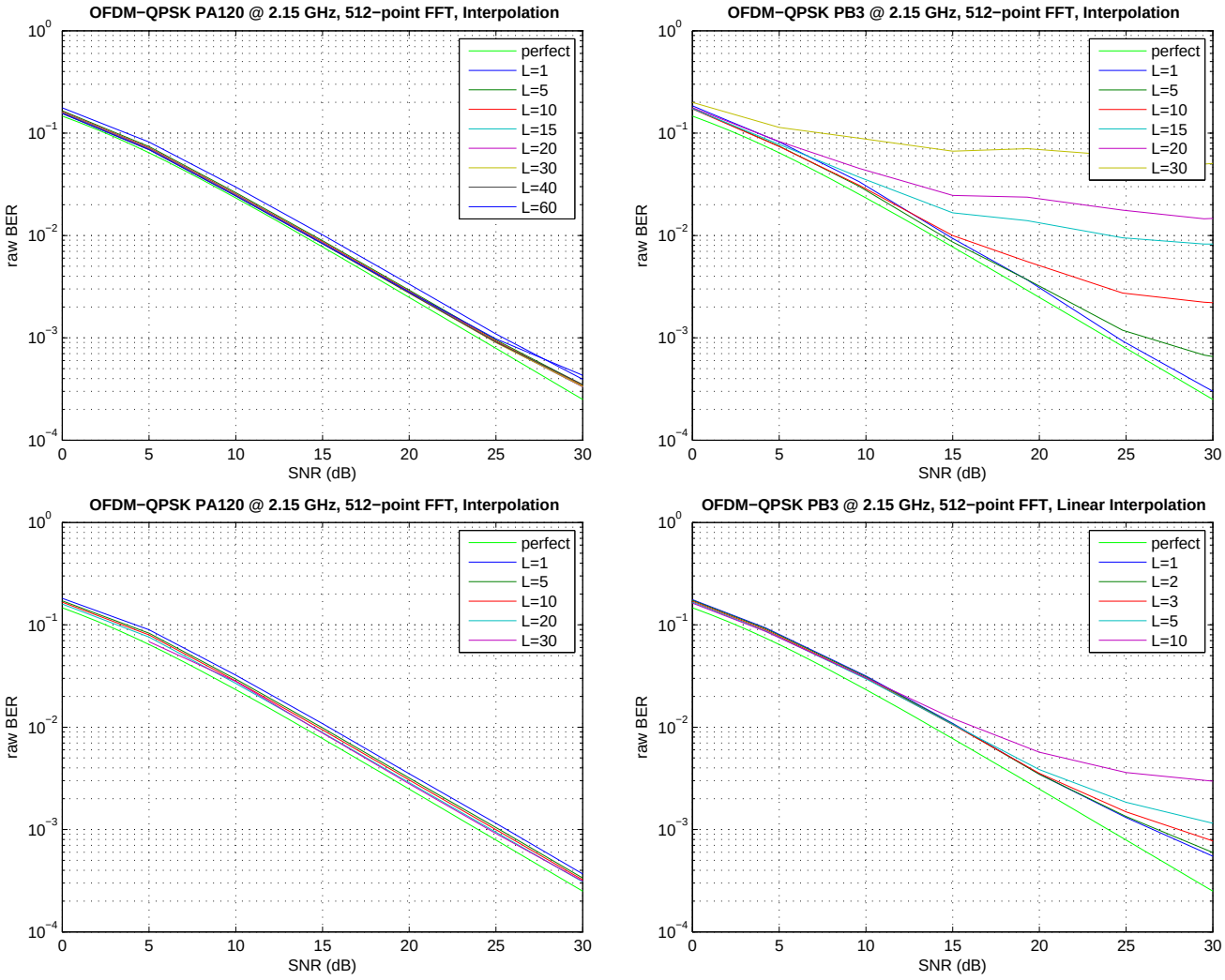
5 Simulation Results

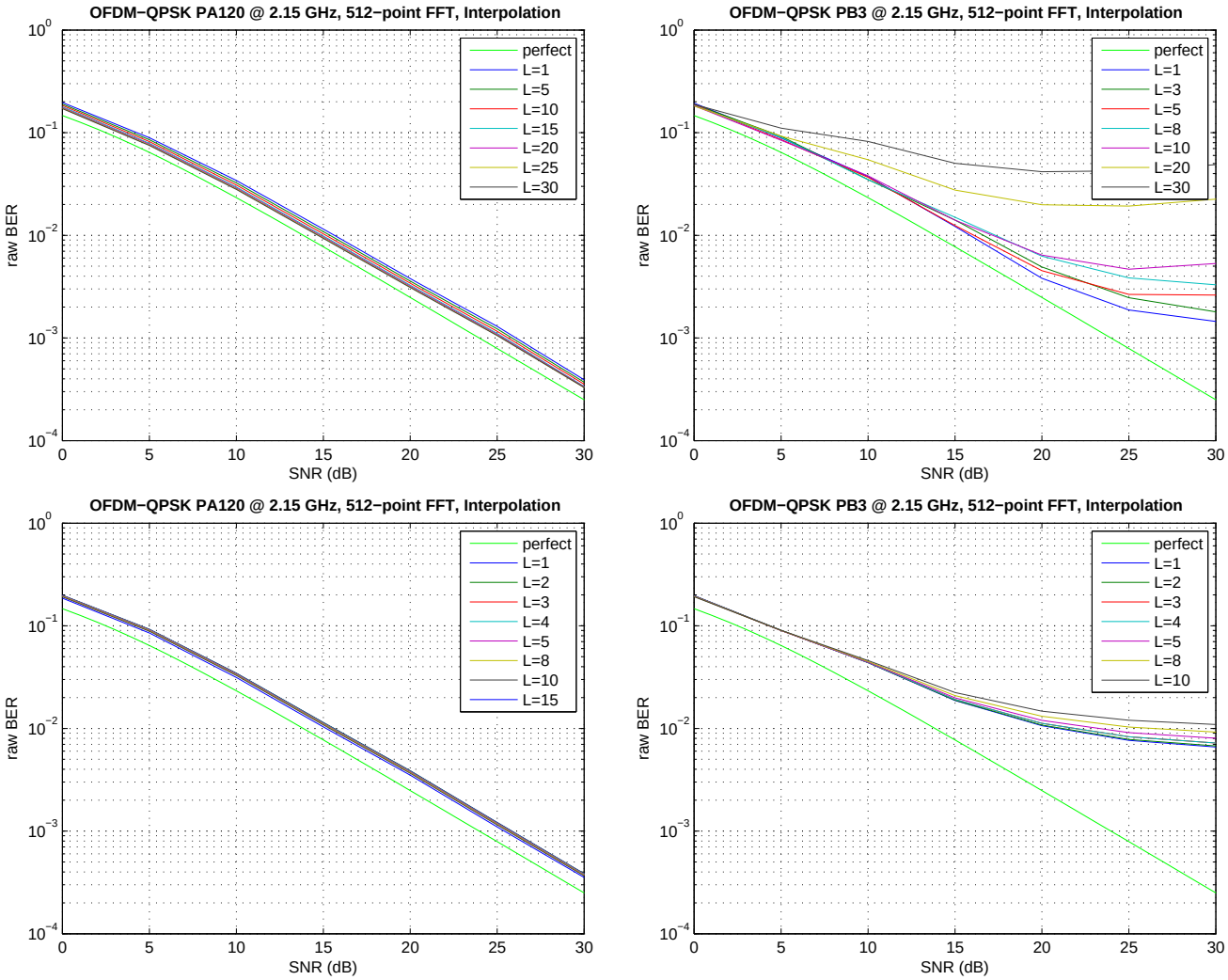
In previous section, we provide the theoretical framework for three methods i) linear interpolation, ii) polynomial-based general linear model, and iii) Wiener filtering. We now provide the simulation results, using raw bit-error-rate (BER), as a measure of performance for these methods. Raw BER is measured by counting the number of errors in estimated bits after equalization in the OFDM system.

5.1 Linear Interpolation

Figure 5-1 shows the raw BER for linear interpolation method. The left hand-side plots show the performance over PA120, while the right hand-side plots show the performance over PB3 channel. The top-most figure shows performance for an OFDM system where every 2nd OFDM symbol was pilot. The lower most figure shows performance for an OFDM system where every 10th OFDM symbol. The frequency-pilot separation was varied from 2nd, 4th, 6th, and 10th. The interpolated samples were then averaged using a window of L -interpolated samples. Temporal-pilot separation was four.

We conclude that interpolation-based method works well for PA profile, although the performance degrades as the pilot separation is increased. The performance for PB profile degrades as the pilot separation and the averaging window length are increased.





PA120 channel PB3 channel

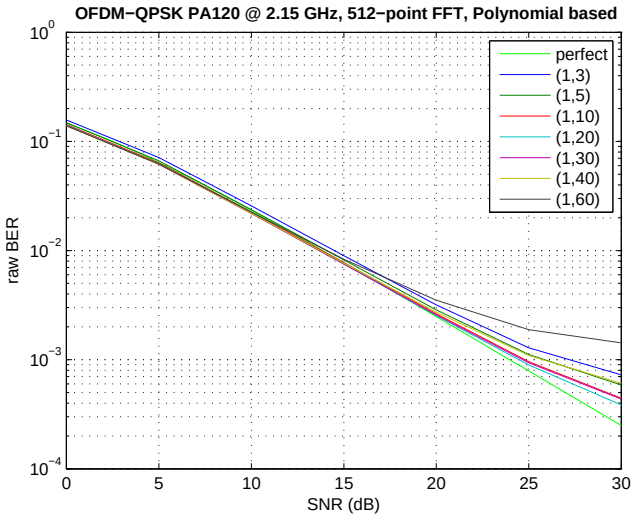
Figure 5-1. Simulation results for interpolation-based channel estimation

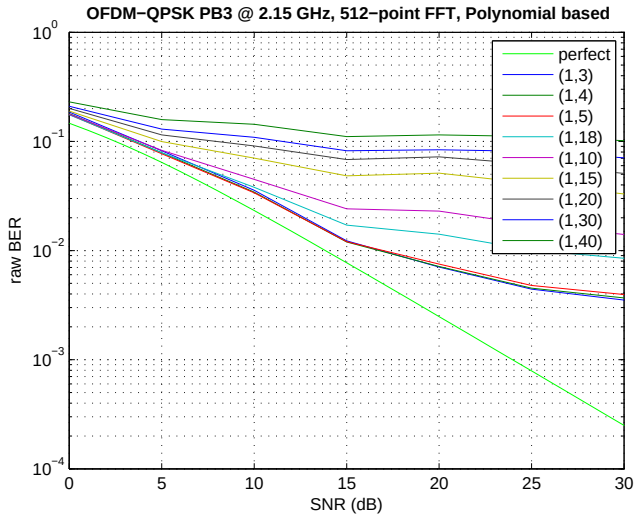
5.2 Polynomial-based Generalized Linear Model

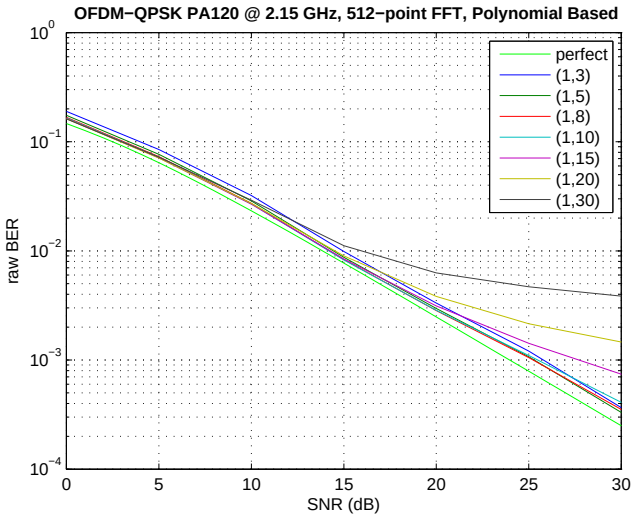
5.2.1 Polynomial Order 1

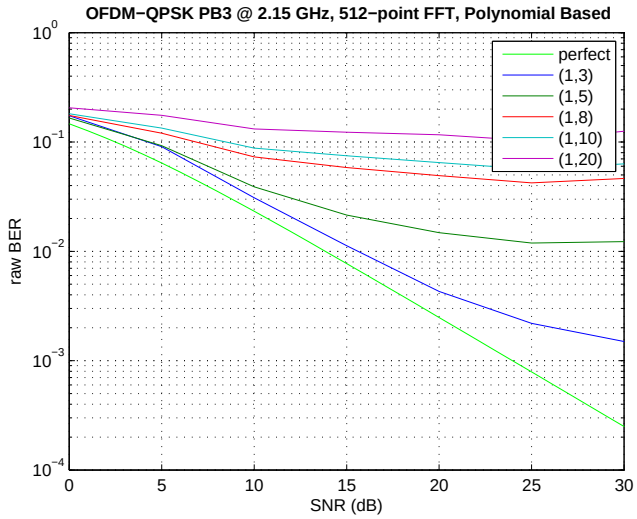
Figure 5-2 shows the raw BER for polynomial-based generalized linear model. The order of the polynomial was 1. The left hand-side plots show the performance over PA120, while the right hand-side plots show the performance over PB3 channel. The top-most figure shows performance for an OFDM system where every 2nd OFDM symbol was pilot. The lower most figure shows performance for an OFDM system where every 10th OFDM symbol. The pilot separation was varied from 2nd, 4th, 6th, and 10th. Note that L pilots were used to estimate the least-squares weighing factors for general linear model. Temporal-pilot separation was four.

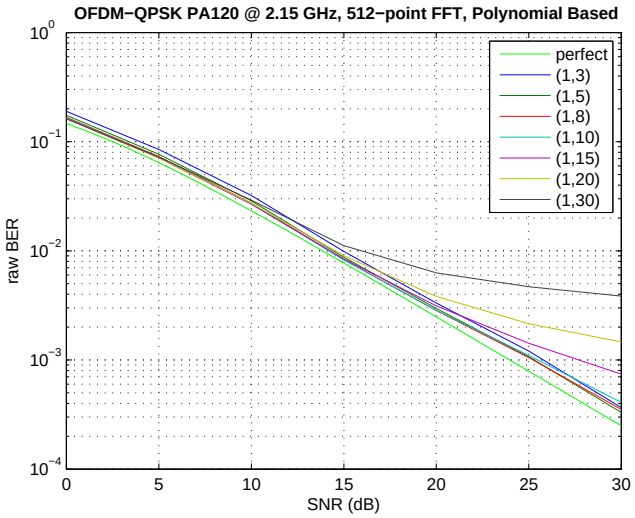
We conclude that polynomial-based method works well for PA profile, as long as the number of pilots used for least-squares is kept low. The performance for PB profile degrades as the pilot separation and the averaging window length are increased.

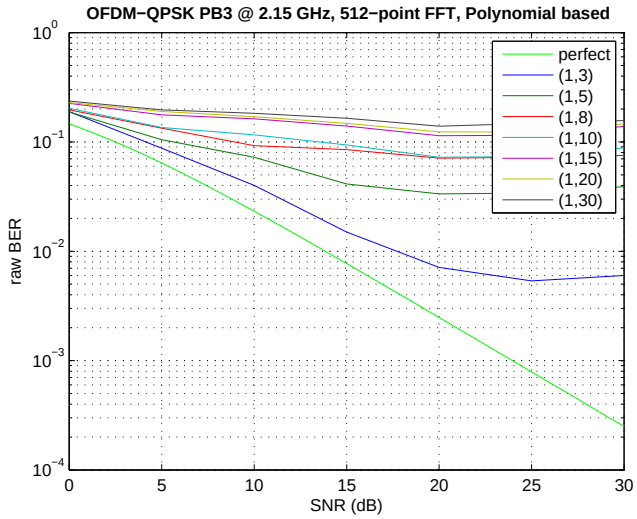


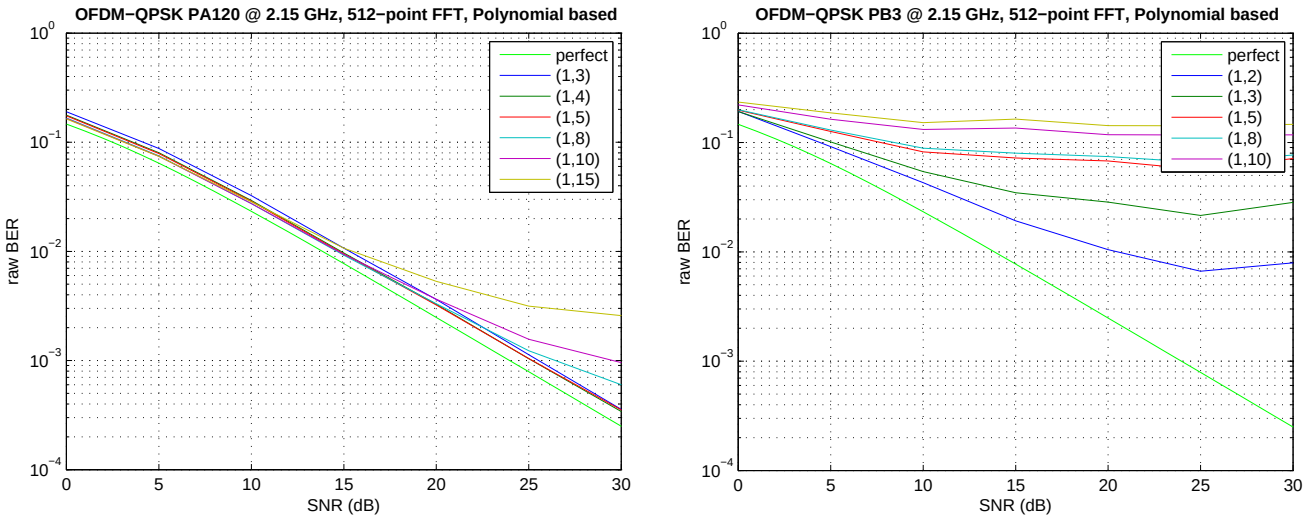












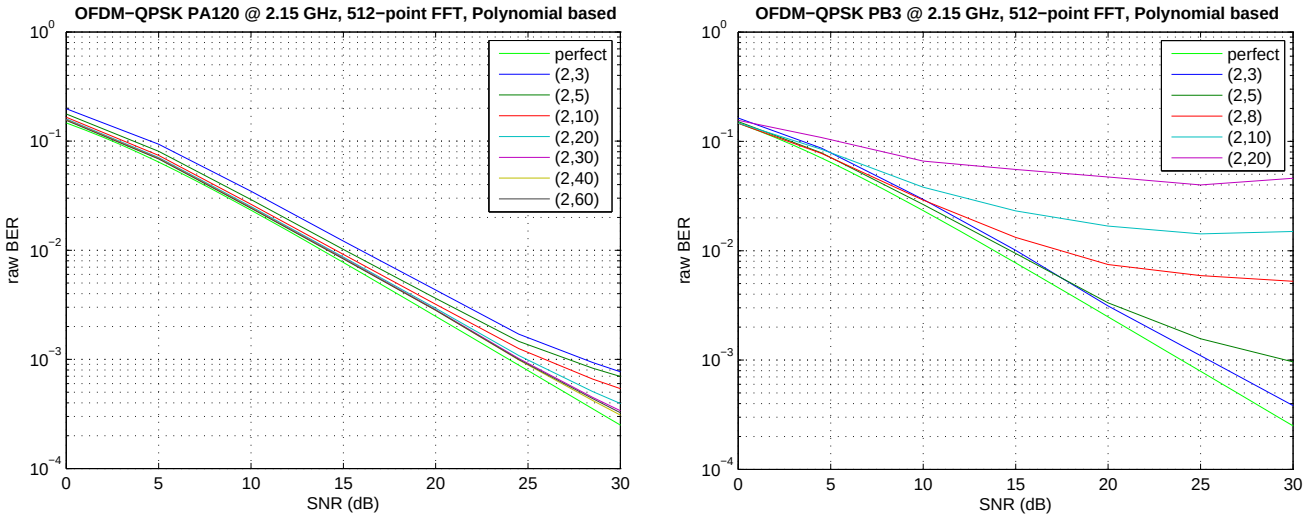
PA120 channel PB3 channel

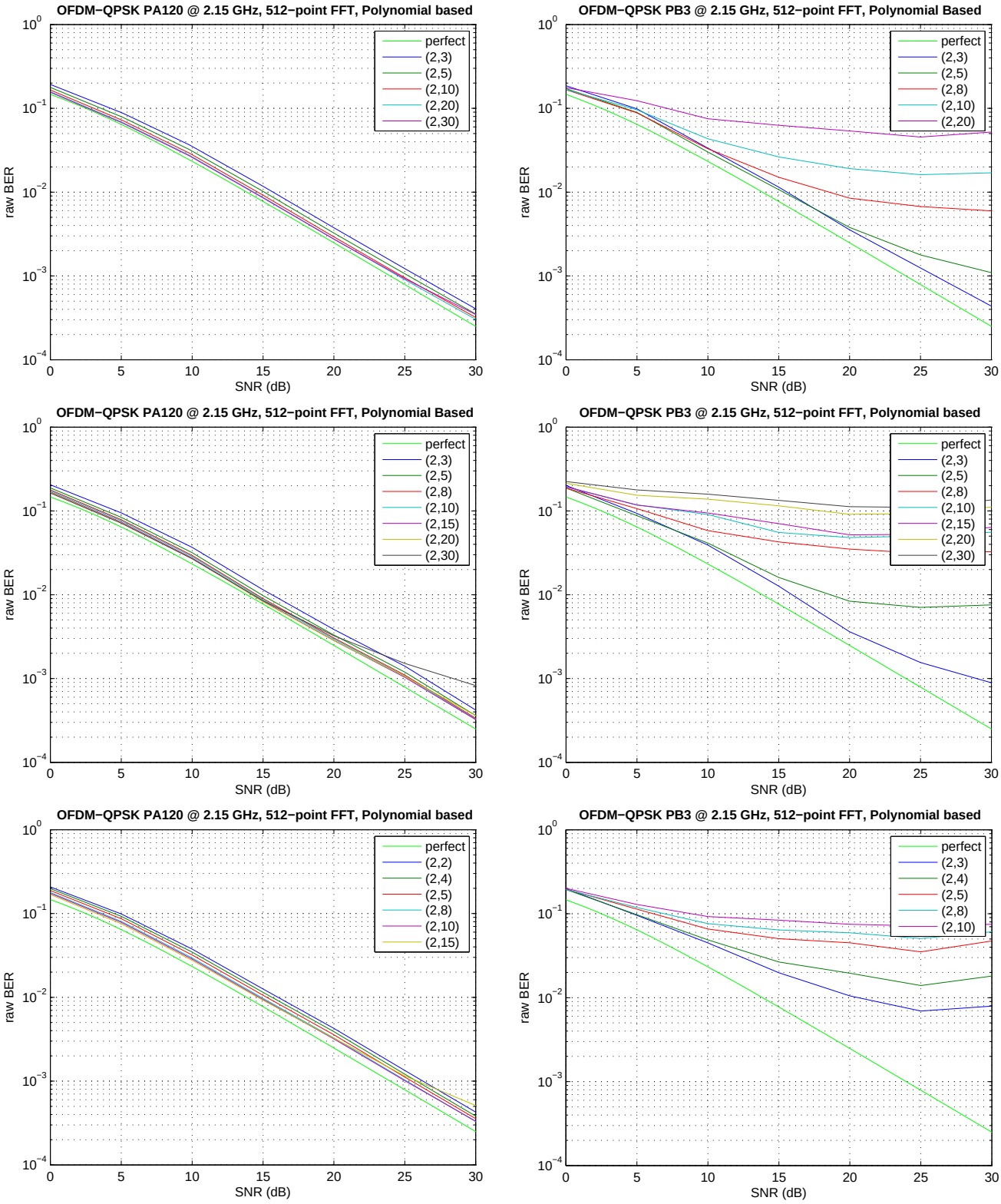
Figure 5-2. Simulation results for order 1 polynomial-based channel estimation

5.2.2 Polynomial Order 2

Figure 5-3 shows the raw BER for polynomial-based generalized linear model. The order of the polynomial was 2. The left hand-side plots show the performance over PA120, while the right hand-side plots show the performance over PB3 channel. The top-most figure shows performance for an OFDM system where every 2nd OFDM symbol was pilot. The lower most figure shows performance for an OFDM system where every 10th OFDM symbol. The pilot separation was varied from 2nd, 4th, 6th, and 10th. Note that L pilots were used to estimate the least-squares weighing factors for general linear model. Temporal-pilot separation was four.

We conclude that polynomial-based method works well for PA profile. The performance for PB profile degrades as the pilot separation and the averaging window length are increased.





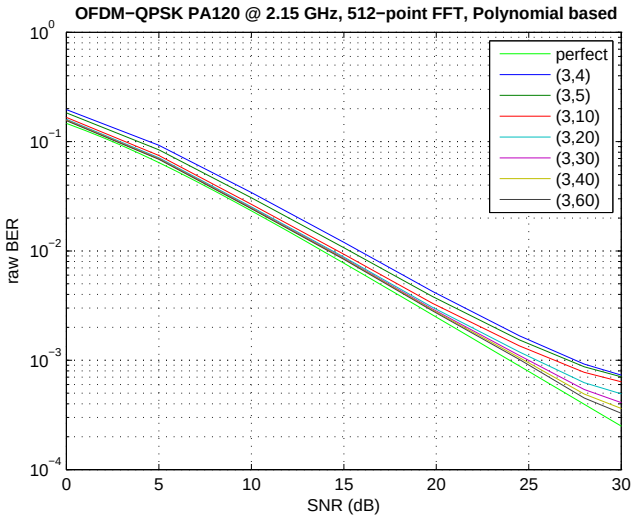
PA120 channel PB3 channel

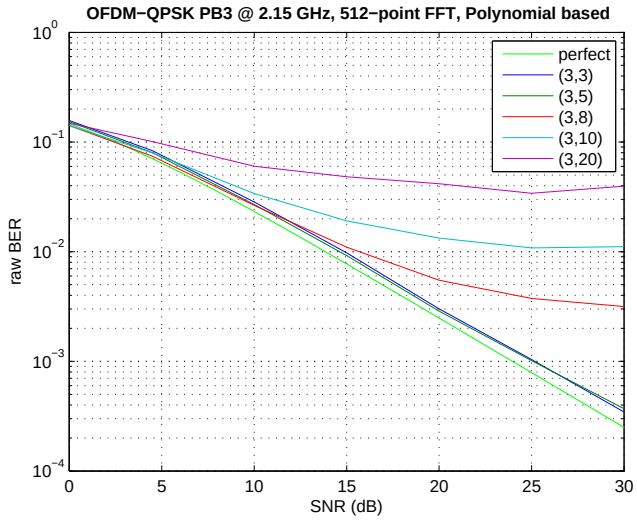
Figure 5-3. Simulation results for order 2 polynomial-based channel estimation

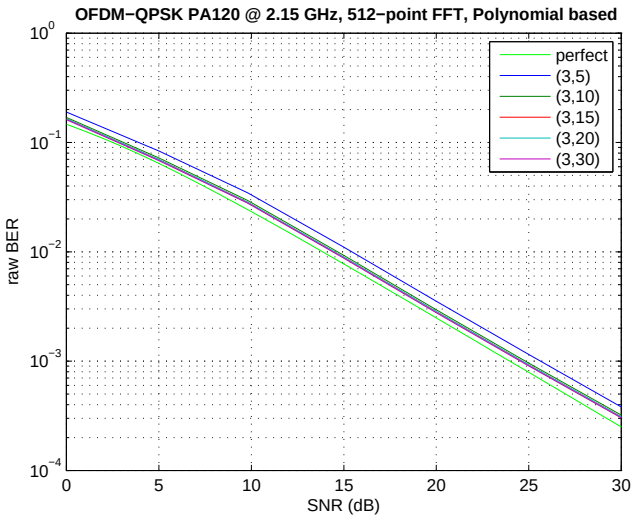
5.2.3 Polynomial Order 3

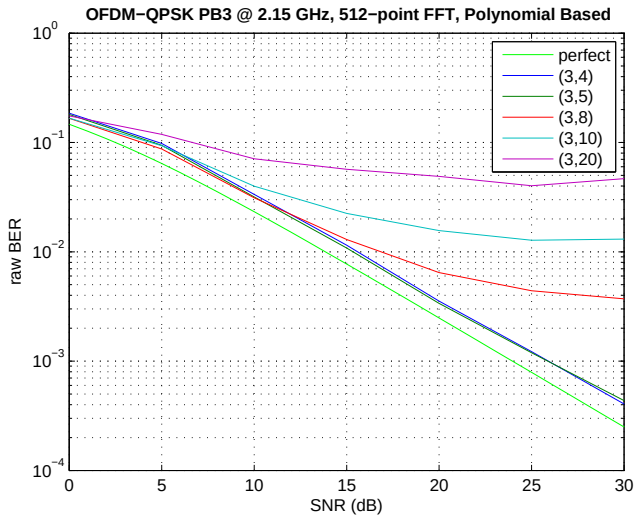
Figure 5-4 shows the raw BER for polynomial-based generalized linear model. The order of the polynomial was 3. The left hand-side plots show the performance over PA120, while the right hand-side plots show the performance over PB3 channel. The top-most figure shows performance for an OFDM system where every 2nd OFDM symbol was pilot. The lower most figure shows performance for an OFDM system where every 10th OFDM symbol. The pilot separation was varied from 2nd, 4th, 6th, and 10th. Note that L pilots were used to estimate the least-squares weighing factors for general linear model. Temporal-pilot separation was four.

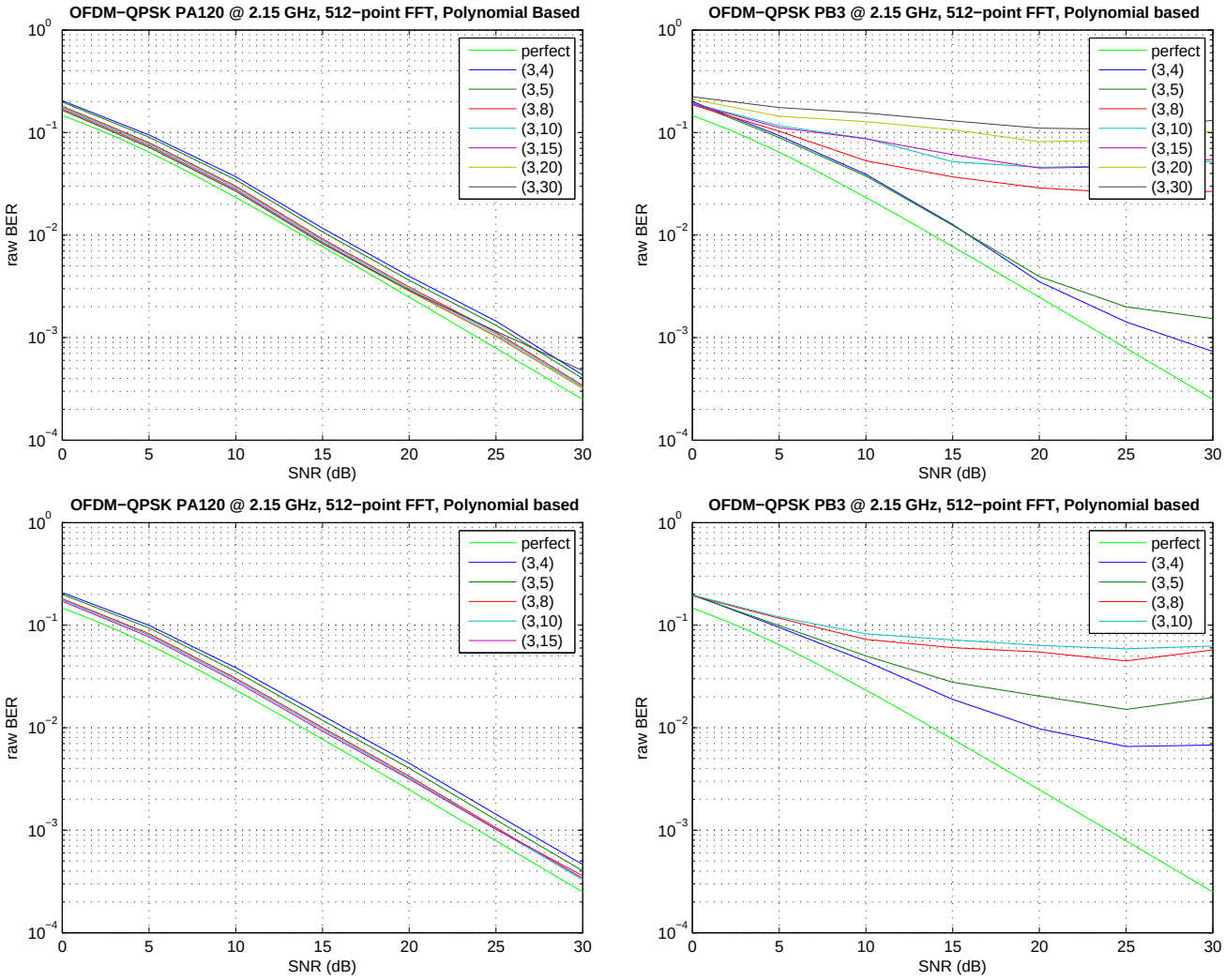
We conclude that polynomial-based method works well for PA profile. The performance for PB profile degrades as the pilot separation and the averaging window length are increased.











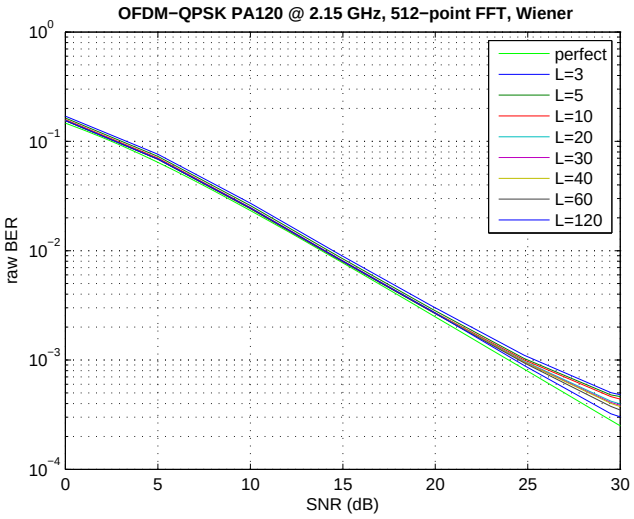
PA120 channel PB3 channel

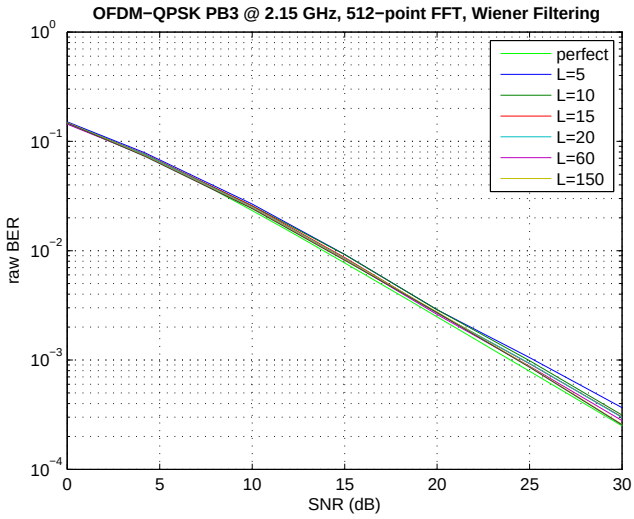
Figure 5-4. Simulation results for order 3polynomial-based channel estimation

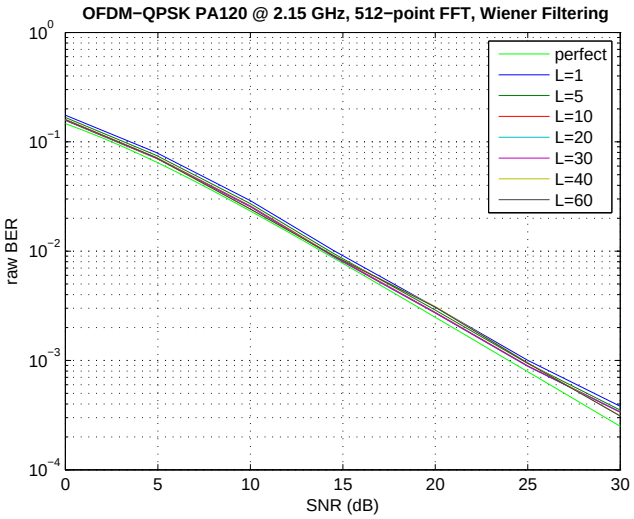
5.3 Wiener Filtering

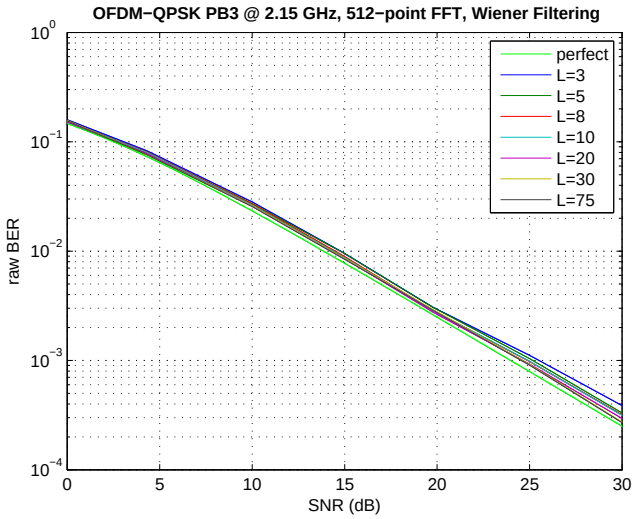
Figure 5-5 shows the raw BER for Wiener filtering model for channel estimation. The left hand-side plots show the performance over PA120, while the right hand-side plots show the performance over PB3 channel. The top-most figure shows performance for an OFDM system where every 2nd OFDM symbol was pilot. The lower most figure shows performance for an OFDM system where every 10th OFDM symbol. The pilot separation was varied from 2nd, 4th, 6th, and 10th. Note that L pilots were used to estimate the Wiener filter coefficients based on LMMSE method. Temporal-pilot separation was four.

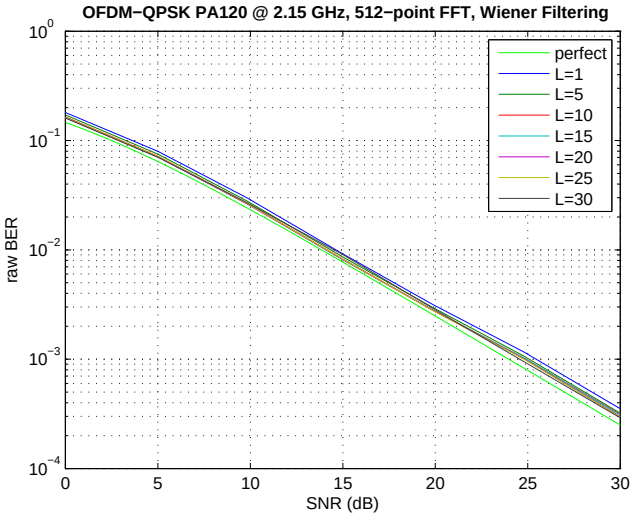
We conclude that Wiener filtering works well for both PA and PB channel profiles. The performance degrades slightly as the pilot separation is increased from 2 to 10, but is better than performance of the any other method (interpolation and polynomial-based methods).

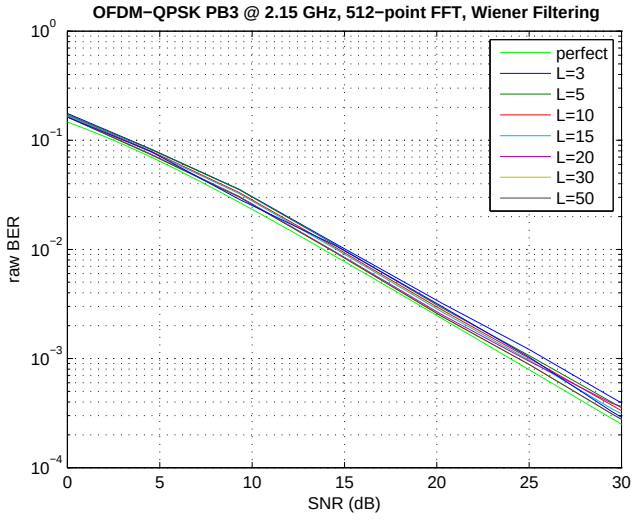


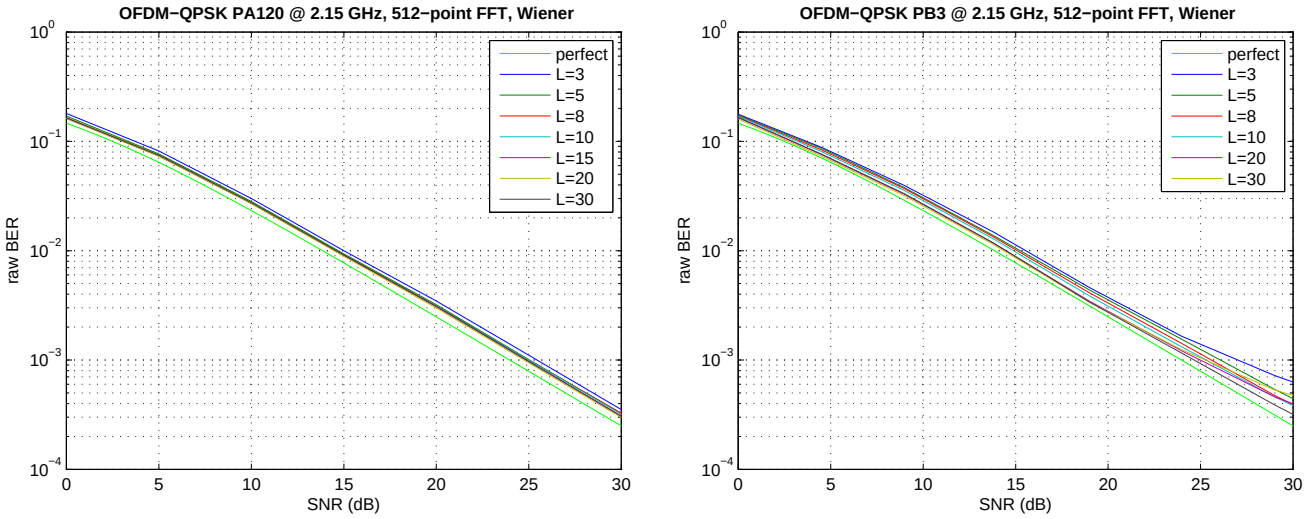












PA120 channel

PB3 channel

Figure 5-5. Simulation results for order 3polynomial-based channel estimation

6 Conclusion

In this report we have investigated three different frequency-based methods (interpolation, polynomial based model, and Weiner filtering) for channel estimation for 3.9/4G LTE, an OFDM based system. The performance of these methods was simulated on PA and PB channels, assuming perfect frequency synchronization and without receiver impairments. Raw BER was used a measure of performance evaluation. Interpolation was used for channel estimation in temporal-domain

From the simulations, we conclude that the *linear interpolation-based channel estimator* works well for a channel with high coherence bandwidth (PA), but fails for a channel with low coherence bandwidth (PB). We also observe that averaging after linear interpolation helps to improve performance, as long as the averaging window length does not exceed the coherence bandwidth. The *polynomial-based generalized linear model* works well as long as the right selection of polynomial order and number of pilots (for calculating the channel estimator) is done. The polynomial order of 2 seems to be the optimal choice, and the number of pilots seems to depend upon the coherence bandwidth of the channel. For high coherence bandwidth channel (PA), approximately 40 sub-carrier window (for calculating the channel estimator) seem to the optimal choice, while a low coherence bandwidth channel (PB), approximately 6 sub-carrier window seems to the optimal choice (for calculating the channel estimator). The *ideal Wiener filter method* outperformed both the linear interpolation and polynomial-based generalized linear model, especially at high SNR and low coherence bandwidth channel. The reason for good performance of Wiener filtering is because it automatically makes an ideal compromise between noise reduction and interpolation. We however note that the successful performance of the Wiener filter depends upon our knowledge of the spaced-frequency correlation of the channel; which has not been investigated yet.

We also conclude that the frequency-domain pilot density in an OFDM system plays an important role in the performance of a channel estimator. For an OFDM system with high pilot density (say every 2nd symbol is pilot), a low complexity channel estimator such as linear-interpolation with post averaging performs almost as well as any other advanced channel estimator. We also conclude that as pilot separation increases, the performance of channel estimator degrades. The Wiener filtering provides the least performance degradation as pilot separation is increased from 2 to 10. The pilot separation in temporal-domain was not investigated. Pilot separation in temporal-domain was set to 4.

We also note that linear interpolation-based method is least complex of channel estimators. The polynomial-based generalized linear model is more computational complex as it requires a matrix inversion and multiplication for each set of pilots. Since the same matrix is inverted for each set of pilot, the complexity can be reduced by saving the polynomial matrix inverse, and using the inverse matrix for channel estimation for each set of pilots. The Wiener filtering method is the most complex method for channel estimation, as it requires matrix inversion and matrix multiplication for each set of pilots.

Due to the gains obtained in highly frequency selective channels, we are inclined towards using Wiener filtering based estimator for channel estimation in OFDM-based systems provided that the receiver can handle the computational complexity of the receiver. The future steps will be to investigate the computation of channel statistics (noise-variance and spaced-frequency correlation) to justify the implementation of Wiener filter solution, and to explore sophisticated methods for temporal domain.

Appendix A

A.1 Spaced-time Correlation of Clark's Scattering Model

Consider impulse response of a single-tap channel in Clark's isotropic scattering model [5], [10] and [7]. The impulse response is given as

$$h(t)=\sqrt{\frac{P_n}{N}}\sum_{i=1}^{N}e^{j\varphi_i}e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i t\right)} \tag{6-1}$$

The spaced-time correlation function, $R_h(\tau)$, of such a channel is given as

$$R_h(\tau)=E\left\{h(t)\cdot h^*(t-\tau)\right\} \tag{6-2}$$

$$=E\left\{\left(\sqrt{\frac{P_n}{N}}\sum_{i=1}^{N}e^{j\varphi_i}e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i t\right)}\right)\cdot\left(\sqrt{\frac{P_n}{N}}\sum_{k=1}^{N}e^{j\varphi_k}e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_k (t-\tau)\right)}\right)^*\right\} \tag{6-3}$$

$$=E\left\{\sqrt{\frac{P_n}{N}}\sum_{i=1}^{N}e^{j\varphi_i}e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i t\right)}\cdot\sqrt{\frac{P_n}{N}}\sum_{k=1}^{N}e^{-j\varphi_k}e^{-j\left(2\pi\frac{v}{\lambda}\cos\theta_k (t-\tau)\right)}\right\} \tag{6-4}$$

$$=E\left\{\frac{P_n}{N}\sum_{i=1}^{N}\sum_{k=1}^{N}e^{j\varphi_i}e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i t\right)}\cdot e^{-j\varphi_k}e^{-j\left(2\pi\frac{v}{\lambda}\cos\theta_k (t-\tau)\right)}\right\} \tag{6-5}$$

$$=\frac{P_n}{N}\sum_{i=1}^{N}\sum_{k=1}^{N}E\left\{e^{j\varphi_i}e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i t\right)}\cdot e^{-j\varphi_k}e^{-j\left(2\pi\frac{v}{\lambda}\cos\theta_k (t-\tau)\right)}\right\} \tag{6-6}$$

Since the signal phase φ and the angle of scatterers θ are independent, we write

$$=\frac{P_n}{N}\sum_{i=1}^{N}\sum_{k=1}^{N}E\left\{e^{j\varphi_i}\cdot e^{-j\varphi_k}\right\}E\left\{e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i t\right)}\cdot e^{-j\left(2\pi\frac{v}{\lambda}\cos\theta_k (t-\tau)\right)}\right\}$$

Case $i\neq k$**:**

φ_i and φ_k are independent of each other .i.e.,

$$E\left\{e^{j\varphi_i}\cdot e^{-j\varphi_k}\right\}=E\left\{e^{j\varphi_i}\right\}\cdot E\left\{e^{-j\varphi_k}\right\} \tag{6-7}$$

Expectation $E\left\{e^{j\varphi_i}\right\}$ is

$$E\left\{e^{j\varphi_i}\right\}=\int_0^{2\pi}\frac{1}{2\pi}e^{j\varphi_i}d\varphi_i=\frac{1}{2\pi\cdot j}\left[e^{j\varphi_i}\right]_0^{2\pi} \tag{6-8}$$

$$=\frac{1}{2\pi\cdot j}\left[e^{j2\pi}-e^{o}\right]=0 \tag{6-9}$$

Hence for $i\neq k$,

$$R_h(\tau) = 0 \tag{6-10}$$

Case $i = k$ **:**

$$E\left\{e^{j\varphi_i} \cdot e^{-j\varphi_i}\right\} = E\left\{e^{0}\right\} = 1 \tag{6-11}$$

It implies that

$$R_h(\tau) = \frac{P_n}{N}\sum_{i=1}^{N} E\left\{e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i t\right)} \cdot e^{-j\left(2\pi\frac{v}{\lambda}\cos\theta_I (t-\tau)\right)}\right\} \tag{6-12}$$

$$= \frac{P_n}{N}\sum_{i=1}^{N} E\left\{e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i \tau\right)}\right\} \tag{6-13}$$

We note that

$$E\left\{e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i \tau\right)}\right\} = \int_0^{2\pi}\frac{1}{2\pi}e^{j\left(2\pi\frac{v}{\lambda}\cos\theta_i \tau\right)}d\theta_i = J_0\left(2\pi\frac{v}{\lambda}\tau\right) \tag{6-14}$$

is the zero-th order Bessel function.

Hence, the spaced-time correlation function is

$$R_h(\tau) = \frac{P}{N}\sum_{i=1}^{N} J_0\left(2\pi\frac{v}{\lambda}\tau\right) = P \cdot J_0\left(2\pi\frac{v}{\lambda}\tau\right) \tag{6-15}$$

A.2 Spaced-Frequency Correlation and Power Delay Profile

Consider a tapped-delay line channel with L taps

$$h(t,\tau)=\sum_{i=0}^{L-1}h_i(t)\cdot\delta(\tau-\tau_i) \tag{6-16}$$

where

$h_i(t)$ and τ_i denotes the complex gain and excess delay of the i-th path.

The power delay profile $h_{PDP}(\tau)$ of such a channel is

$$h_{PDP}(\tau)=\sum_{i=0}^{L-1}P_i\cdot\delta(\tau-\tau_i) \tag{6-17}$$

where

P_i is the expected power of the i-th path.

The time-varying channel transfer function for such a channel is the Fourier transform of impulse response [8]

$$H(t;f)=\sum_{i=0}^{L-1}h_i(t)\cdot e^{-j2\pi f\tau_i} \tag{6-18}$$

The frequency correlation function of the channel is defined as [8]

$$R(\Delta f)=E\left\{H(t;f)H^*(t;f+\Delta f)\right\} \tag{6-19}$$

$$=E\left\{\sum_{i=1}^{L}h_i(t)e^{-j2\pi f\tau_i}\cdot\sum_{k=1}^{L}h_k^*(t)e^{j2\pi(f+\Delta f)\tau_k}\right\} \tag{6-20}$$

$$=E\left\{\sum_{i=1}^{L}\sum_{k=1}^{L}h_i(t)h_k^*(t)\cdot e^{-j2\pi f\tau_i}e^{j2\pi f\tau_k}e^{j2\pi\Delta f\tau_k}\right\} \tag{6-21}$$

Since expectation is linear operator and channel amplitude is the only random variable, we get

$$=\sum_{i=1}^{L}\sum_{k=1}^{L}E\left\{h_i(t)h_k^*(t)\right\}e^{-j2\pi f\tau_i}e^{j2\pi f\tau_k}e^{j2\pi\Delta f\tau_k} \tag{6-22}$$

Since channel taps are independent $E\left\{h_i(t)h_k^*(t)\right\}=0$ for $i\neq k$, we get

$$R(\Delta f)=\sum_{i=1}^{L}P_i\cdot e^{j2\pi\Delta f} \tag{6-23}$$

where

P_i is the expected power of the i'th path.

A.3 3GPP Standardized Power Delay Profiles

3GPP defines three discrete power delay profiles for investigating wireless communication performance at 2.15 GHz ± 5MHz. These are pedestrian-A (PA), pedestrian-B (PB) and vehicular-A (VA) channel. The expected power of the taps and their corresponding delays are given in Table 3 [19].

Channel type	**Expected Power of taps [dB]**	**Path delay [μ s]**
PA	[0 -9.7 -19.2 -22.8]	[0 0.110 0.190 0.410]
PB	[0 -0.9 -4.9 -8.0 -7.8 -23.9]	[0 0.2 0.8 1.2 2.3 3.7]
VA	[0 0.31 0.71 1.09 1.73 2.51]	[0 -1.0 -9.0 -10.0 -15.0 -20.0]

Table 3. Power Delay profile for 3GPP channels

A.4 Orthogonality of OFDM waveforms

Two signals are orthogonal to each other if their inner product is zero. Consider two OFDM waveforms $\psi_k(t)$ and $\psi_l(t)$ at carrier frequencies f_k and f_l Hz respectively. The inner product between these two OFDM waveforms is given as

$$\int_0^T \psi_k(t)\cdot\psi_l^*(t)dt = \int_0^T \left(\frac{1}{\sqrt{T_s}}e^{j2\pi\cdot f_k t}\right)\cdot\left(\frac{1}{\sqrt{T_s}}e^{j2\pi\cdot f_l t}\right)^* dt \tag{6-24}$$

$$= \int_0^T \frac{1}{\sqrt{T_s}}e^{j2\pi\cdot f_k t}\cdot\frac{1}{\sqrt{T_s}}e^{-j2\pi\cdot f_l t}dt$$

$$= \frac{1}{T_s}\int_0^T e^{j2\pi\cdot(f_k-f_l)t}dt$$

for $k \neq l$,

$$= \frac{1}{T_s}\left[\frac{1}{j2\pi\cdot(f_k-f_l)}e^{j2\pi\cdot(f_k-f_l)t}\right]_0^T$$

$$= \frac{1}{T_s}\left\{\frac{1}{j2\pi\cdot(f_k-f_l)}e^{j2\pi\cdot(f_k-f_l)T} - \frac{1}{j2\pi\cdot(f_k-f_l)}\right\}$$

$$= \frac{1}{T_s}\frac{1}{j2\pi\cdot(f_k-f_l)}\left\{e^{j2\pi\cdot(f_k-f_l)T}-1\right\}$$

If the carrier spacing is $|f_k - f_l| = m\frac{1}{T}$, with $m = 1,2,\cdots$ the above simplifies the above to

$$= \frac{1}{T_s}\frac{1}{j2\pi\cdot(f_k-f_l)}\left\{e^{j2\pi m}-1\right\}$$

$$= 0$$

for $k = l$,

$$= \frac{1}{T_s}\int_0^T 1\ dt$$

$$= 1$$

Hence, the $\psi_k(t)$ and $\psi_l(t)$ at carrier frequencies f_k and f_l Hz respectively, are orthogonal.

A.5 Raw channel estimate for an all pilot OFDM symbol

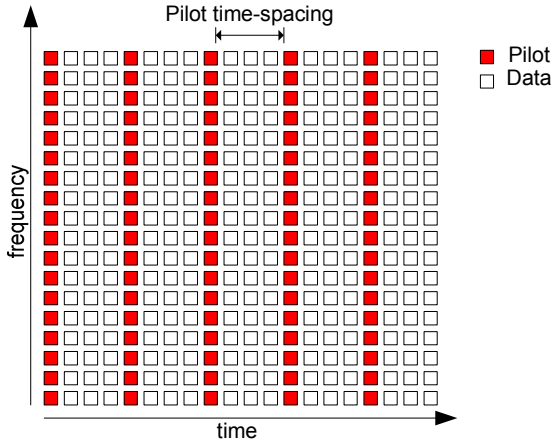


Figure 6-1. All pilot OFDM symbol transmission

We showed in Section 3.4 that the relationship between transmitted signal X_k and received Y_k is

$$Y_k = H_k \cdot X_k + Z_k \tag{6-25}$$

where

H_k is the channel transfer function at the k-th sub-carrier, and Z_k is the noise .

For an OFDM system with N_c carriers, the observations can be collected in as follows

$$\underbrace{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{N_c} \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & \vdots \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & X_{N_c} \end{bmatrix}}_{\mathbf{X}} \cdot \underbrace{\begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_{N_c} \end{bmatrix}}_{\mathbf{H}} + \underbrace{\begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{N_c} \end{bmatrix}}_{\mathbf{Z}} \tag{6-26}$$

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{H} + \mathbf{Z} \tag{6-27}$$

We now consider an OFDM system based on frequency comb-type pilot sequence [21] as shown in Figure 6-1. The least-squares estimate of such system is obtained by minimizing square distance between the received signal $\mathbf{Y}$ and the transmitted signal $\mathbf{X}$ as [36]

$$\min_{\mathbf{H}^H} J(\mathbf{H}) = \min_{\mathbf{h}^H} \left\{ \left| \mathbf{Y} - \mathbf{X}\mathbf{H} \right|^2 \right\} = \min_{\mathbf{h}^H} \left\{ (\mathbf{Y} - \mathbf{X}\mathbf{H})^H (\mathbf{Y} - \mathbf{X}\mathbf{H}) \right\} \tag{6-28}$$

$$= \min_{\mathbf{H}^H} \left\{ \mathbf{Y}^H\mathbf{Y} - \mathbf{Y}^H\mathbf{X}\mathbf{H} - \mathbf{H}^H\mathbf{X}^H\mathbf{Y} + \mathbf{H}^H\mathbf{X}^H\mathbf{X}\mathbf{H} \right\} \tag{6-29}$$

Differentiating the cost function with respect to $\mathbf{H}^H$ and finding the minima, we get

$$\frac{\partial}{\partial \mathbf{H}^H} J(\mathbf{H}) = -\mathbf{X}^H\mathbf{Y} + \mathbf{X}^H\mathbf{X}\mathbf{H} = 0 \tag{6-30}$$

$$\begin{aligned}\hat{\mathbf{H}}_{LS} &= \left(\mathbf{X}^H\mathbf{X}\right)^{-1}\cdot\mathbf{X}^H\mathbf{Y}\\ &= \mathbf{X}^{-1}\left(\mathbf{X}^H\right)^{-1}\cdot\mathbf{X}^H\mathbf{Y}\\ &= \mathbf{X}^{-1}\cdot\mathbf{Y}\\ &= \begin{bmatrix}\frac{Y_1}{X_1} & \frac{Y_2}{X_2} & \cdots & \frac{Y_{N_c}}{X_{N_c}}\end{bmatrix}^T\end{aligned} \qquad (6\text{-}31)$$

Therefore for an OFDM system with a comb-type frequency-based pilot, the least squares estimate of the channel at pilots can be estimated for a simple division as shown Equation (6-31). The channel estimate at the non-pilots can be estimated by using interpolation in time.

A.6 Derivation of Generalized Linear Model based Channel estimate (LS method)

This section is inspired from [24]. Given a generalized linear model for the channel transfer function $(H(f))$ we can write the raw channel estimate $\hat{H}_{pk}$ at pilot sub-carrier frequency f_{pk} as

$$\hat{H}_{pk} = H_{pk} + Z_{pk} = \sum_{i=0}^{N-1} \theta_i \cdot \Phi_i\left(f_{pk}\right) + Z_{pk} \tag{6-32}$$

where

$\Phi_i\left(f_{pk}\right)$ is the i-th basis function evaluated at that pk-th sub-carrier frequency f_k, θ_i is the weighing factor of the basis function $\Phi_i\left(f_{pk}\right)$, and Z_{pk} is the noise at the raw channel estimate. N is the order of the basis function. Every p-th sub-carrier is a pilot

We can collect a set of N_p raw channel estimate observations in the following observations as

$$\underbrace{\begin{bmatrix} \hat{H}_0 \\ \hat{H}_p \\ \vdots \\ \hat{H}_{p(N_p-1)} \end{bmatrix}}_{\hat{\mathbf{H}}} = \underbrace{\begin{bmatrix} \Phi_0(f_0) & \Phi_1(f_0) & \cdots & \Phi_{N-1}(f_0) \\ \Phi_0(f_p) & \Phi_1(f_p) & \cdots & \Phi_{N-1}(f_p) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_0\left(f_{p(N_p-1)}\right) & \Phi_1\left(f_{p(N_p-1)}\right) & \cdots & \Phi_{N-1}\left(f_{p(N_p-1)}\right) \end{bmatrix}}_{\mathbf{G}} \cdot \underbrace{\begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_{N-1} \end{bmatrix}}_{\theta} + \underbrace{\begin{bmatrix} Z_0 \\ Z_p \\ \vdots \\ Z_{p(N_p-1)} \end{bmatrix}}_{\mathbf{Z}} \tag{6-33}$$

Introducing matrices, we re-write the above as

$$\hat{\mathbf{H}} = \mathbf{G} \cdot \boldsymbol{\theta} + \mathbf{Z} \tag{6-34}$$

The least-squares estimate $\hat{\boldsymbol{\theta}}_{LS}$ of the weighing matrix $\boldsymbol{\theta}$ is calculated by minimizing the squared distance between the actual channel $\mathbf{H}$ and the modeled channel $\hat{\mathbf{H}}$ as

$$J(\boldsymbol{\theta}) = \left|\hat{\mathbf{H}} - \mathbf{H}\right|^2 = \left(\hat{\mathbf{H}} - \mathbf{H}\right)^H \cdot \left(\hat{\mathbf{H}} - \mathbf{H}\right) \tag{6-35}$$

$$= \left(\hat{\mathbf{H}} - \mathbf{G}\boldsymbol{\theta}\right)^H \cdot \left(\hat{\mathbf{H}} - \mathbf{G}\boldsymbol{\theta}\right) \tag{6-36}$$

$$= \hat{\mathbf{H}}^H\hat{\mathbf{H}} - \hat{\mathbf{H}}'^H\mathbf{G}\boldsymbol{\theta} - \boldsymbol{\theta}^H\mathbf{G}^H\hat{\mathbf{H}} + \boldsymbol{\theta}^H\mathbf{G}^H\mathbf{G}\boldsymbol{\theta} \tag{6-37}$$

Minimizing we get

$$\frac{\partial}{\partial \boldsymbol{\theta}^H} J\{\boldsymbol{\theta}\} = -\mathbf{G}^H\hat{\mathbf{H}} + \mathbf{G}^H\mathbf{G}\boldsymbol{\theta} = 0 \tag{6-38}$$

$$\hat{\boldsymbol{\theta}}_{LS} = \left(\mathbf{G}^H\mathbf{G}\right)^{-1}\mathbf{G}^H\hat{\mathbf{H}} \tag{6-39}$$

A.7 Regression using Orthogonal Polynomials

Using a set of $\{0 \quad 1 \quad \cdots \quad n\}$ orthogonal polynomials $\{\xi_0(f) \quad \xi_1(f) \quad \cdots \quad \xi_n(f)\}$, we can write an observed set of data $\{\hat{H}_p \quad \hat{H}_{2p} \quad \cdots \quad \hat{H}_{pN_p}\}$ as [24]

$$\underbrace{\begin{bmatrix} \hat{H}_p \\ \hat{H}_{2p} \\ \vdots \\ \hat{H}_{pN_p} \end{bmatrix}}_{\hat{\mathbf{H}}} = \underbrace{\begin{bmatrix} \xi_0(f_p) & \xi_1(f_p) & \cdots & \xi_n(f_p) \\ \xi_0(f_{2p}) & \xi_1(f_{2p}) & \cdots & \xi_n(f_{2p}) \\ \vdots & \vdots & \ddots & \vdots \\ \xi_0(f_{pN_p}) & \xi_1(f_{pN_p}) & \cdots & \xi_n(f_{pN_p}) \end{bmatrix}}_{\mathbf{G}} \cdot \underbrace{\begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}}_{\theta} + \underbrace{\begin{bmatrix} Z_p \\ Z_{2p} \\ \vdots \\ Z_{pN_p} \end{bmatrix}}_{\mathbf{Z}} \tag{6-40}$$

where

θ_i is the i'th weighing factor of the polynomial function $\xi_i(f_{pk})$ at pk-th sub-carrier frequency

We now introduce a new notation for the polynomial matrix

$$\mathbf{G} = \begin{bmatrix} \xi_0(f_p) & \xi_1(f_p) & \cdots & \xi_n(f_p) \\ \xi_0(f_{2p}) & \xi_1(f_{2p}) & \cdots & \xi_n(f_{2p}) \\ \vdots & \vdots & \ddots & \vdots \\ \xi_0(f_{pN_p}) & \xi_1(f_{npN_p}) & \cdots & \xi_n(f_{pN_p}) \end{bmatrix} = \begin{bmatrix} \xi_{0,0} & \xi_{1,0} & \cdots & \xi_{k,0} \\ \xi_{0,1} & \xi_{1,1} & \cdots & \xi_{k,1} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{0,n} & \xi_{1,n} & \cdots & \xi_{k,n} \end{bmatrix} \tag{6-41}$$

We note that orthogonal polynomials satisfy

$$\sum_{j=1}^{n} \xi_{i,j} = 0 \text{ and } j = \{1 \quad \cdots \quad n\} \tag{6-42}$$

and

$$\sum_{j=1}^{n} \xi_{\mu,j}\xi_{\gamma,j} = 0 \text{ for } \mu \neq \gamma \tag{6-43}$$

Which implies that $\mathbf{G}^H\mathbf{G}$ becomes a diagonal matrix given as

$$\mathbf{G}^H\mathbf{G} = \begin{bmatrix} n\xi_0^2 & 0 & \cdots & 0 \\ 0 & \sum \xi_{1j}^2 & & \\ \vdots & & \ddots & \\ 0 & \cdots & & \sum \xi_{kj}^2 \end{bmatrix} \tag{6-44}$$

Hence, the least-squares estimate becomes

$$\hat{\boldsymbol{\theta}}_{LS} = \left(\mathbf{G}^H\mathbf{G}\right)^{-1}\mathbf{G}^H\hat{\mathbf{H}} = \begin{bmatrix} \sum \hat{H}_j / N_p \cdot 1/\xi_0 \\ \sum \xi_{1j}\hat{H}_j / \sum \xi_{1j}^2 \\ \vdots \\ \sum \xi_{kj}\hat{H}_j / \sum \xi_{kj}^2 \end{bmatrix} \tag{6-45}$$

We notice the following

- Since $\mathbf{G}^H\mathbf{G}$ is a diagonal matrix, calculating its inverse is not computationally expensive. Similarly, the computation of $\hat{\boldsymbol{\theta}}_{LS}$ does not suffer from numerical instability

A.7.1 Determination of orthogonal polynomials

The polynomial of **0th degree** is [24]

$$\xi_0(f) = \xi_0 = 1 \tag{6-46}$$

The polynomial of **1st degree** is

$$\xi_1(f) = f + a \tag{6-47}$$

From Equation (6-42) we have

$$0 = \sum_{j=1}^{n} \xi_1(f_j) = \sum_{j=1}^{n} (f_j + a) = \sum_{j=1}^{n} f_j + na \tag{6-48}$$

It gives

$$a = -\frac{1}{n}\sum_{j=1}^{n} f_j = -\overline{f} \tag{6-49}$$

which gives

$$\xi_1(f) = f - \overline{f} \tag{6-50}$$

Similarly, the polynomial of **2nd degree** as a linear combination of 1, $\xi_1(f)$ and $\xi_1^2(f)$

$$\xi_2(f) = a_{02} + a_{12}(f - \overline{f}) + a_{22}(f - \overline{f})^2 \tag{6-51}$$

From Equation (6-42) we have

$$0 = \sum_{j=1}^{n} \xi_2(f_j) = na_{02} + a_{12}\sum_j (f_j - \overline{f}) + a_{22}\sum_j (f_j - \overline{f})^2 \tag{6-52}$$

$$\Rightarrow \frac{a_{02}}{a_{22}} = -\frac{1}{n}\sum_j (f_j - \overline{f})^2 \tag{6-53}$$

Similarly from (6-43) we have

$$0 = \sum_{j=1}^{n} \xi_1(f_j)\xi_2(f_j) = a_{02}\sum_j (f_j - \overline{f}) + a_{12}\sum_j (f_j - \overline{f})^2 + a_{22}\sum_j (f_j - \overline{f})^3 \tag{6-54}$$

$$= a_{12}\sum_j (f_j - \overline{f})^2 + a_{22}\sum_j (f_j - \overline{f})^3 \tag{6-55}$$

$$\Rightarrow \frac{a_{12}}{a_{22}} = -\frac{\sum_j (f_j - \overline{f})^3}{\sum_j (f_j - \overline{f})^2} \tag{6-56}$$

The **3rd**, **4th**, etc. **degree** polynomials are found analogously.

The polynomials generation can be summarized for equally spaced points as

$$\xi_0(f_m) = 1 \tag{6-57}$$

$$\xi_1(f_m) = f_m - \frac{n+1}{2} \tag{6-58}$$

$$\xi_{i+1}(f_m) = \xi_1(f_m)\xi_i(f_m) - \frac{i^2\left(n^2 - i^2\right)}{4\left(4i^2 - 1\right)}\xi_{i-1}(f_m), \text{ and } i = \{2,3,\cdots,n-1\} \tag{6-59}$$

where

n is the number of equally spaced observations, and

$f_m = \{f_1, f_2, \cdots, f_n\}$ is one of the equally spaced points, and f_m has been normalized.

Table 4 presents values for orthogonal polynomials $\xi_{k,m}$ for certain combinations of f_m, n and k (degree of polynomial). Note that values of orthogonal polynomials in Table 4 are whole numbers, since they have been normalized with λ. Table 4 has been taken from [24].

n	3		4			5				6					7				
m	$\xi_{1,m}$	$\xi_{2,m}$	$\xi_{1,m}$	$\xi_{2,m}$	$\xi_{3,m}$	$\xi_{1,m}$	$\xi_{2,m}$	$\xi_{3,m}$	$\xi_{4,m}$	$\xi_{1,m}$	$\xi_{2,m}$	$\xi_{3,m}$	$\xi_{4,m}$	$\xi_{5,m}$	$\xi_{1,m}$	$\xi_{2,m}$	$\xi_{3,m}$	$\xi_{4,m}$	$\xi_{5,m}$
1	−1	+1	−3	+1	−1	−2	+2	−1	+1	−5	+5	−5	+1	−1	−3	+5	−1	+3	−1
2	0	−2	−1	−1	+3	−1	−1	+2	−4	−3	−1	+7	−3	+5	−2	0	+1	−7	+4
3	+1	+1	+1	−1	−3	0	−2	0	+6	−1	−4	+4	+3	−10	−1	−3	+1	+1	−5
4			+3	+1	+1	+1	−1	−2	−4	+1	−4	−4	+2	+10	0	−4	0	+6	0
5						+2	+2	+1	+1	+3	−1	−7	−3	−5	+1	−3	−1	+1	+5
6										+5	+5	+5	+1	+1	+2	0	−1	−7	−4
7															+3	+5	+1	+3	+1
λ	1	3	2	1	$\frac{10}{3}$	1	1	$\frac{5}{6}$	$\frac{35}{12}$	2	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{7}{12}$	$\frac{21}{10}$	1	1	$\frac{1}{6}$	$\frac{7}{12}$	$\frac{7}{20}$

Table 4. Values of orthogonal polynomials

Table 4 is used to construct a linear model using orthogonal polynomials up-to degree $k = 3$ for $n = 4$ equally spaced observations

$$\hat{H}_{LS} = \hat{\theta}_{LS_0} + \lambda_1 \cdot \hat{\theta}_{LS_1} \cdot \xi_1(f) + \lambda_2 \cdot \hat{\theta}_{LS_2} \cdot \xi_2(f) + \lambda_3 \cdot \hat{\theta}_{LS_3} \cdot \xi_3(f) \tag{6-60}$$

where

$$\lambda_1 = 2, \quad \lambda_2 = 1, \quad \lambda_3 = \frac{10}{3} \tag{6-61}$$

The 1st degree polynomial is

$$\xi_1(f_m) = f_m - \frac{5}{2} \tag{6-62}$$

The 2nd degree polynomial is

$$\xi_2(f_m) = \left(f_m - \frac{5}{2}\right)^2 - \frac{1^2\left(4^2 - 1^2\right)}{4\left(4\cdot 1^2 - 1\right)}\left(f_m - \frac{5}{2}\right) \quad (6\text{-}63)$$

$$= f_m^2 - 5f_m + 5 \quad (6\text{-}64)$$

And 3rd degree polynomial is

$$\xi_3(f_m) = \left(f_m - \frac{5}{2}\right)\cdot\left(f_m^2 - 5f_m + 5\right) - \frac{2^2\left(4^2 - 2^2\right)}{4\left(4\cdot 2^2 - 1\right)}\left(f_m - \frac{5}{2}\right) \quad (6\text{-}65)$$

$$= f_m^3 - \frac{15}{2}f_m^2 + \frac{101}{6}f_m - \frac{65}{6} \quad (6\text{-}66)$$

Finally, the least-squares estimate $\hat{\boldsymbol{\theta}}_{LS}$ is calculated using Equation (6-45).

A.8 Legendre Polynomials

The Legendre polynomials, sometimes called Legendre functions of first kind, are solution to the Legendre differential equation [37].

$$\frac{d}{dx}\left[\left(1-x^2\right)\frac{d}{dx}P(x)\right]+n(n+1)P(x)=0 \tag{6-67}$$

The solution to the Legendre equation leads to Legendre polynomials. Each Legendre polynomial $P_n(x)$ is an n th-degree polynomial, which may be expressed using the so-called Rodrigues' formula:

$$P_n(x)=\frac{1}{2^n n!}\frac{d^n}{dx^n}\left[\left(x^2-1\right)^n\right] \tag{6-68}$$

An important property of Legendre Polynomials is that they are orthogonal over the interval $(-1,+1)$, satisfying

$$\int_{-1}^{+1} P_m(x)P_n(x)dx=\frac{2}{2n+1}\delta_{nm} \tag{6-69}$$

where

$\delta_{nm}=\begin{cases}1 & m=n\\ 0 & m\neq n\end{cases}$ denotes the Kronecker delta function.

We have Legendre polynomials to form the basis polynomials of LS-based channel estimation. The polynomials were selected for this purposed due their orthogonal nature. Figure 6-2 shows the first six Legendre polynomials over the interval $(-1,+1)$.

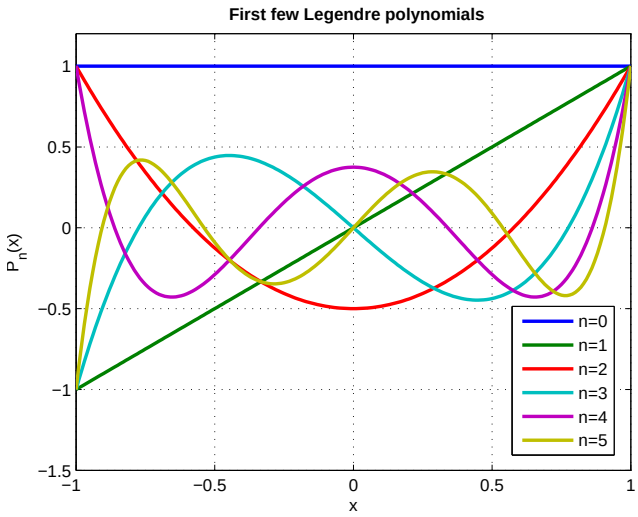


Figure 6-2. First few Legendre polynomials

Table 5 contains the first eleven Legendre polynomials which were used in the parametric channel estimation method [37].

n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}\left(3x^2-1\right)$
3	$\frac{1}{2}\left(5x^3-3x\right)$
4	$\frac{1}{8}\left(35x^4-30x^2+3\right)$
5	$\frac{1}{8}\left(63x^5-70x^3+15x\right)$
6	$\frac{1}{16}\left(231x^6-315x^4+105x^2-5\right)$
7	$\frac{1}{16}\left(429x^7-693x^5+315x^3-35x\right)$
8	$\frac{1}{128}\left(6435x^8-12012x^6+6930x^4-1260x^2+35\right)$
9	$\frac{1}{128}\left(12155x^9-25740x^7+18018x^5-4620x^3+315x\right)$
10	$\frac{1}{256}\left(46189x^{10}-109395x^8+90090x^6-30030x^4+3465x^2-63\right)$

Table 5. First few Legendre polynomials

A.9 Derivation of Wiener filtering based Channel estimate (LMMSE method)

Consider a set of N_p raw channel estimate as given in equation (4-4) [32]

$$\hat{H}_{pk} = H_{pk} + Z_{pk} \tag{6-70}$$

where

$\hat{H}_{pk}$ is a raw channel estimate at the pk -th sub-carrier frequency,

H_{pk} is the actual channel value at the pk -th sub-carrier frequency, and

Z_{pk} is the noise contribution at the pk -th sub-carrier frequency.

Given a set of observations on the raw channel estimate $\left(\hat{H}_{pk}\right)$, we can write the LMMSE estimate of the channel $\left(\hat{H}_{LMMSE,k}\right)$ as at the k th sub-carrier as

$$\hat{H}_{LMMSE,k} = \sum_{i=0}^{N_p-1} c_i \hat{H}_{pi} = \sum_{i=0}^{N_p-1} c_i \left(H_{pi} + Z_{pi}\right) \tag{6-71}$$

where

c_i is the i'th Wiener filter coefficient, and

$\hat{H}_{pi}$ is the i'th raw channel estimate

Using vectors, we re-write the above as

$$\hat{H}_{LMMSE,k} = \mathbf{c}^H \hat{\mathbf{H}} \tag{6-72}$$

where

$\mathbf{c}$ and $\hat{\mathbf{H}}$ are the column vectors, containing N_p Wiener filter coefficients and N_p raw channel estimates, respectively.

The Wiener filter coefficients are obtained by minimizing the expected mean-square error [36] given as

$$E\left\{\left|H_k - \hat{H}_{LMMSE,k}\right|^2\right\} = E\left\{\left(H_k - \hat{H}_{LMMSE,k}\right)\cdot\left(H_k - \hat{H}_{LMMSE,k}\right)^H\right\} \tag{6-73}$$

$$= E\left\{\left(H_k - \mathbf{c}^H\hat{\mathbf{H}}\right)\cdot\left(H_k - {}^H\hat{\mathbf{H}}\right)^H\right\} \tag{6-74}$$

$$= E\left\{H_k H_k^* - \mathbf{c}^H\hat{\mathbf{H}}H_k^* - H_k\hat{\mathbf{H}}^H\mathbf{c} + \mathbf{c}^H\hat{\mathbf{H}}\hat{\mathbf{H}}^H\mathbf{c}\right\} \tag{6-75}$$

$$= E\left\{H_k H_k^*\right\} - \mathbf{c}^H E\left\{\hat{\mathbf{H}}H_k^*\right\} - E\left\{H_k\hat{\mathbf{H}}^H\right\}\mathbf{c} + \mathbf{c}^H E\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\}\mathbf{c} \tag{6-76}$$

Minimizing we get

$$\frac{\partial}{\partial \mathbf{c}^H} E\left\{H_k - \hat{H}_{LMMSE,k}\right\} = -E\left\{\hat{\mathbf{H}} H_k^*\right\} + E\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\}\mathbf{c} = 0 \tag{6-77}$$

$$\mathbf{c}_{LMMSE} = \left(E\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\}\right)^{-1} E\left\{\hat{\mathbf{H}} H_k^*\right\} \tag{6-78}$$

We first consider $E\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\}$

$$E\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\} = E\left\{\left(\hat{\mathbf{H}} + \mathbf{Z}\right)\cdot\left(\hat{\mathbf{H}} + \mathbf{Z}\right)^H\right\} \tag{6-79}$$

$$= \underbrace{E\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\}}_{\mathbf{R}_h} + \underbrace{E\left\{\hat{\mathbf{H}}\mathbf{Z}^H\right\}}_{0} + \underbrace{E\left\{\mathbf{Z}\hat{\mathbf{H}}^H\right\}}_{0} + \underbrace{E\left\{\mathbf{Z}\mathbf{Z}^H\right\}}_{\mathbf{R}_z} \tag{6-80}$$

Since the channel and AWGN are not correlated, we write Equation (6-80) as

$$= \mathbf{R}_h + \mathbf{R}_z = \mathbf{R}_h + \boldsymbol{\sigma}^2\mathbf{I} \tag{6-81}$$

where $\mathbf{R}_h$ is the autocorrelation matrix of the channel and $\boldsymbol{\sigma}^2$ is the noise-variance per sub-carrier. Note that

$$\mathbf{R}_h = \begin{bmatrix} E\left\{\hat{H}_0\hat{H}_0^*\right\} & E\left\{\hat{H}_0\hat{H}_p^*\right\} & \cdots & E\left\{\hat{H}_0\hat{H}_{p(N_p-1)}^*\right\} \\ E\left\{\hat{H}_p\hat{H}_0^*\right\} & E\left\{\hat{H}_p\hat{H}_p^*\right\} & \cdots & E\left\{\hat{H}_p\hat{H}_{p(N_p-1)}^*\right\} \\ \vdots & \vdots & \ddots & \vdots \\ E\left\{\hat{H}_{p(N_p-1)}\hat{H}_0^*\right\} & E\left\{\hat{H}_{p(N_p-1)}\hat{H}_p^*\right\} & \cdots & E\left\{\hat{H}_{p(N_p-1)}\hat{H}_{p(N_p-1)}^*\right\} \end{bmatrix} \tag{6-82}$$

Similarly $E\left\{\hat{\mathbf{H}}H_k^*\right\}$ is given as

$$E\left\{\hat{\mathbf{H}}H_k^*\right\} = E\left\{\left(\mathbf{H} + \mathbf{Z}\right)\cdot H_k^*\right\} = \underbrace{E\left\{\mathbf{H}\cdot H_k^*\right\}}_{\mathbf{r}} + \underbrace{E\left\{\mathbf{Z}\cdot H_k^*\right\}}_{0} \tag{6-83}$$

Where $\mathbf{r}$ is the cross-correlation vector of the channel at k th sub-carrier and the channel at pilot locations.

$$\mathbf{r} = \begin{bmatrix} E\left\{\hat{H}_0\hat{H}_k^*\right\} \\ E\left\{\hat{H}_p\hat{H}_k^*\right\} \\ \vdots \\ E\left\{\hat{H}_{p(N_p-1)}\hat{H}_k^*\right\} \end{bmatrix} \tag{6-84}$$

Hence the Wiener filter coefficients that provide the LMMSE channel estimate is [32]

$$\mathbf{c}_{LMMSE} = \left(\mathbf{R}_h + \sigma^2\mathbf{I}\right)^{-1}\mathbf{r} \tag{6-85}$$

One we have calculated the Wiener filter coefficients; we can estimate the channel at k-th sub-carrier as

$$\hat{H}_{LMMSE,k} = \mathbf{c}_{LMMSE}^H\hat{\mathbf{H}} = \left(\left(\mathbf{R}_h + \sigma^2\mathbf{I}\right)^{-1}\mathbf{r}\right)\cdot\hat{\mathbf{H}} \tag{6-86}$$

A.9.1 Inversion of matrix $\mathbf{R}_h + \sigma^2\mathbf{I}$ using SVD and matrix inversion lemma

The matrix inversion lemma gives a method of estimating the

$$\left(\mathbf{A}+\mathbf{XYZ}\right)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{X}\left(\mathbf{Y}^{-1} + \mathbf{ZA}^{-1}\mathbf{X}\right)^{-1}\mathbf{ZA}^{-1} \quad (6\text{-}87)$$

where

$\mathbf{A}$, $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Y}$ are all matrices of the size $N\times N$, $N\times K$, $K\times K$ and $K\times N$.

We that this lemma could be applied to matrix inversion problem of $\left(\mathbf{R}_h + \sigma^2\mathbf{I}\right)^{-1}$, if we could re-write the $\mathbf{R}_h$ matrix as a product of three matrix. By taking the SVD of matrix $\mathbf{R}_h$, we write

$$\mathbf{R}_h = \mathbf{USV}^H \quad (6\text{-}88)$$

where

$\mathbf{U}$ is a $N\times N$ unitary matrix, $\mathbf{S}$ is a $N\times N$ matrix containing singular values on the diagonal and zeros on off-diagonal, and $\mathbf{V}^H$ is a $N\times N$ unitary matrix as well.

Using the SVD, we write the

$$\left(\mathbf{R}_h + \sigma^2\mathbf{I}\right)^{-1} = \left(\sigma^2\mathbf{I} + \mathbf{USV}^H\right)^{-1} \quad (6\text{-}89)$$

Which, using the matrix lemma, becomes

$$= \left(\sigma^2\mathbf{I}\right)^{-1} - \left(\sigma^2\mathbf{I}\right)^{-1}\mathbf{U}\left(\mathbf{S}^{-1} + \mathbf{V}^H\left(\sigma^2\mathbf{I}\right)^{-1}\mathbf{U}\right)^{-1}\mathbf{V}^H\left(\sigma^2\mathbf{I}\right)^{-1} \quad (6\text{-}90)$$

$$= 1/\sigma^2\cdot\mathbf{I} - 1/\sigma^2\cdot\mathbf{I}\cdot\mathbf{U}\left(\mathbf{S}^{-1} + \mathbf{V}^H 1/\sigma^2\cdot\mathbf{I}\cdot\mathbf{U}\right)^{-1}\mathbf{V}^H 1/\sigma^2\cdot\mathbf{I} \quad (6\text{-}91)$$

$$= 1/\sigma^2\cdot\mathbf{I} - 1/\sigma^2\cdot\mathbf{U}\left(\mathbf{S}^{-1} + 1/\sigma^2\cdot\mathbf{V}^H\mathbf{U}\right)^{-1}\mathbf{V}^H 1/\sigma^2 \quad (6\text{-}92)$$

We note that since $\mathbf{U}$ and $\mathbf{V}^H$ satisfy $\mathbf{V}^H\mathbf{U} = \mathbf{I}$, we get

$$= 1/\sigma^2\cdot\mathbf{I} - 1/\sigma^2\cdot\mathbf{U}\left(\mathbf{S}^{-1} + 1/\sigma^2\cdot\mathbf{I}\right)^{-1}\mathbf{V}^H 1/\sigma^2 \quad (6\text{-}93)$$

We note that $\left(\mathbf{S}^{-1} + 1/\sigma^2\cdot\mathbf{I}\right)$ is a diagonal matrix, and use the notation $\mathbf{S}_\sigma = \left(\mathbf{S}^{-1} + 1/\sigma^2\cdot\mathbf{I}\right)$ as

$$= 1/\sigma^2\cdot\mathbf{I} - 1/\sigma^2\cdot\mathbf{US}_\sigma^{-1}\mathbf{V}^H 1/\sigma^2 \quad (6\text{-}94)$$

Since $\widetilde{\mathbf{S}}$ is diagonal matrix, we can calculate its inverse by simply inverting the diagonal terms. Hence the problem of inverting the matrix has been reduced to simply matrix multiplications and addition, at the expense of calculating the SVD of the $\mathbf{R}_h$ matrix. However, this is not an issue if we have prior knowledge of $\mathbf{R}_h$ (and hence prior knowledge of SVD).

References

[1] Ove Edfors et al., “An introduction to orthogonal frequency-division multiplexing”, Research Report, Luleå University of Technology, Sweden, September 1996

[2] Merouance Debbah, “Short Introduction to OFDM”, White Paper, Mobile Communications Group, Institut Eurecom, February 2004.

[3] John Doble, “Introduction to Radio Propagation for Fixed and Mobile Communications”, Artech House Publishers, 1996

[4] Bernard Sklar, “Rayleigh Fading Channels in Mobile Digital Communication Systems Part I: Characterization”, *IEEE Communications Magazine*, Page(s):90 –100, July 1997.

[5] Gordan L. Stuber, “Principles of Mobile Communications”, 2nd Edition, Kluwer Academic Publishers, 2001.

[6] William C. Jakes Jr. , “Microwave Mobile Communications”, John Wiley & Sons, 1974

[7] William C. Y. Lee, “Mobile Communications Engineering”, McGraw-Hill, 1982

[8] John G. Proaksis, “” Digital Communications”, 4th Edition, McGraw-Hill, 2001.

[9] Suhas Mathur, “Small Scale Fading in Radio Propagation”, Department of Electrical Engineering, Rugters University, Lecture Notes for Wireless Communication Technologies, Spring 2005

[10] Johan Brøndum “General mobilradio-kanal til systemsimulering på Personal Computer”, Research Report, Institut of Eletroniske Systemer, Aalborg Universitet, July 1990.

[11] J. Bach Andersen, et al., “Propagation Aspects of Datacommunication over the Radio Channel – A tutorial”, *EUROCON 88: 8th European Conference on Electrotechnics*, Conference Proceedings on Area Communication (Cat. No.88CH2607-0) Page(s):301 – 305

[12] Marc Engels et al., “Wireless OFDM Systems, How to make them work?”, Kluwer Academic Publishers, 2002

[13] S. B. Weinstein and Paul Ebert, “Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform”, *IEEE Transactions on Communication Technology*, Vol. Com-19, No.5, Page(s):628 – 634, October 1971.

[14] A. Peled and A. Ruiz, “Frequency domain data transmission using reduced computational complexity algorithms”, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Denver CO USA, 1980, Page(s):964 – 967.

[15] 3GPP Web-site, “Long Term Evolution of the 3GPP radio technology”, http://www.3gpp.org/Highlights/LTE/LTE.htm

[16] 3GPP TR 25.913, “Requirements for evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved (E-UTRAN)”, Technical Specification Group Radio Access Network, (Release 7)

[17] 3GPP TR 25.814, “Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA)”, Technical Specification Group Radio Access Network, (Release 7)

[18] 3GPP TS 36.201, “LTE Physical Layer – General Description”, Technical Specification Group Radio Access Network (Release 8)

[19] 3GPP TS 36.211, “Physical Channels and Modulation”, Technical Specification Group Radio Access Network (Release 8)

[20] Sinem Coleri, et al., "Channel Estimation Techniques Based on Pilot Arrangement in OFDM Systems", *IEEE Transactions on Broadcasting*, Vol. 48, No.3, Page(s):223 – 229, September 2002.

[21] Meng-Han Hsieh and Che-Ho Wei, "Channel Estimation for OFDM Systems Based on Comb-Type Pilot Arrangement in Frequency Selective Fading Channels", *IEEE Transactions on Consumer Electronics*, Vol.44, No.1, Page(s):217 – 225, February 1998.

[22] M. Speth, et al., "Optimum Receiver Design for OFDM-based Broad-band Transmission – Part II: A Case Study", *IEEE Transactions on Communications*, Vol 49, No. 4, Page(s): 571 – 578, April 2001.

[23] Jae Kyoung and Song In Chol, "Performance of Channel Estimation Methods For OFDM Systems in a Multipath Fading Channels", IEEE Transactions on Consumer Electronics, Volume 46, Issue 1, Feb 2000 Page(s):161 - 170

[24] Bjarne Kjær Ersbøll, Knut Conradsen "An Introduction to Statistics Vol. 2", Lecture Notes, IMM, DTU, 2005.

[25] Ming-Xian Chang and Yu. T. Su "Model-Based Channel Estimation for OFDM Signals in Rayleigh Fading", IEEE Transactions on Communications, Vol. 50, No.4, April 2002, Page(s): 540–544.

[26] Ming-Xian Chang and Yu. T. Su "2D Regression Channel Estimation for Equalizing OFDM Signals", IEEE Vehicular Technology Conference Proceedings, 2000. VTC 2000-Spring Tokyo, Vol.1, Page(s):240–244

[28] Xiaowen Wang and K.J. Ray Liu "An Adaptive Channel Estimation Algorithm Using Time-Frequency Polynomial for OFDM with Fading Multipath Channels", EURASIP Journal on Applied Signal Processing, August 2002, Page(s): 818–830

[29] Haiyun Tang, Kam Y. Lau and Robert W. Brodersen, "Interpolation-Based Maximum Likelihood Channel Estimation Using OFDM Pilot Symbols", IEEE Global Telecommunications Conference, November 2002. GLOBECOM '02, Vol. 2, Page(s):1860–1864.

[30] Jan-Jaap van de Beek et al., "On Channel Estimation in OFDM Systems", *IEEE Vehicular Technology Conference*, Vol.2, Page(s):815-819, Chicago IL USA, September 1995

[31] Magnus Sandell and Over Edfors "A comparative study of pilot-based channel estimators for wireless OFDM", Research Report, Luleå University of Technology, Sweden, September 1996

[32] Over Edfors et al., "Analysis of DFT-Based Channel Estimators for OFDM", Wireless Personal Communications, Page(s):55-70, Kluwer Academic Publishers, 2000

[33] Sarah Kate Wilson, et al., "OFDM Channel Estimation by Singular Value Decomposition", IEEE Transactions on Communications, Vol. 46, No. 7, July 1998, Page(s):931 –939

[34] Michele Morelli and Umberto Mengali, "A Comparison of Pilot-Aided Channel Estimation Methods for OFDM Systems", IEEE Transactions on Signal Processing, Vol. 49, Issue 2, December 2001, Page(s): 3065–3073

[35] F. Shu, J. Lee, N. Wu and G.L. Zhao, "Time-Frequency Channel Estimating for Digital Amplitude Modulation Broadcasting Systems based on OFDM", IEEE Proceedings on Communications, Vol. 150, Issue 4, August 2003, Page(s): 259–264

[36] Dimitris G. Manolakis et al., "Statistical and Adaptive Signal Processing", McGraw-Hill, 2000

[37] Wikipedia, "Legendre Polynomials", http://en.wikipedia.org/wiki/Legendre_polynomials

