PILOT SYMBOL ALLOCATION FOR LIMITED FEEDBACK OFDM AND MIMO-OFDM SYSTEMS

by

Ali Yazdan Panah
B.A.Sc, Sharif University of Technology, 2004

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTERS OF APPLIED SCIENCE
in the School
of
Engineering Science

SIMON FRASER UNIVERSITY
Summer 2007

APPROVAL

Name: Ali Yazdan Panah

Degree: Masters of Applied Science

Title of thesis: Pilot Symbol Allocation for Limited Feedback OFDM and MIMO-OFDM Systems

Examining Committee: Dr. Ivan Bajic
Chair

Dr. Rodney G. Vaughan,
Professor, Engineering Science
Senior Supervisor

Dr. Jim Cavers,
Professor, Engineering Science
Supervisor

Dr. Jie Liang, Internal Examiner

Date Approved: June 19, 2007

Abstract

Next generation wireless communication protocols must provide for higher bandwidths, more suitable quality of services and the ability to accommodate more users. Orthogonal frequency division multiplexing (OFDM), as a broadband modulation technique, is one of the prime candidates for simultaneously reaching these goals. Whether, however, we achieve these promises fully or at least partially in practice depends on a sensible design of transceiver signal processing algorithms. One such class of algorithms that is of considerable importance is the process of accurate channel estimation. Another class of techniques, that has only recently been the subject of intense research, is the use of channel state information and feedback to improve various performance measures of wireless links. The objective of this thesis is to find ways of combining these subjects in the hopes of optimizing channel estimation in OFDM systems with feedback. To reach this goal we begin by presenting a brief background in conventional open-loop OFDM channel estimation. Next we derive a general mathematical model for OFDM channel estimation using non uniformly spaced pilots in the presence of feedback. The effort here will be to maximize the average signal-to-noise ratio at the data frequencies while allowing for acceptable channel estimation errors. Our formulation leads to an interesting optimization problem in which we present two simple sub-optimum closed form solutions. Next, and for sake of completeness, we reformulate our problem and tackle it as a combinatorial optimization problem. Here, we seek the help of evolutionary strategies and the genetic algorithm to solve the original problem. Further, we look at more elaborate compression methods, and namely vector quantization, for reducing the stringent feedback required by our solution(s). Finally, and with practicality in mind, we extend what we have learned to modern multiple antenna systems such as candidate models for the IEEE 802.16 WiMAX protocol. At each stage we present comprehensive computer simulations and, when feasible, theoretical confirmation.

To all my teachers,... and to the best of them; to my family.

"If we knew what it was we were doing, it would not be called research, would it?"

— Albert Einstein

Contents

List of Tables

List of Figures

Preface

As engineers we enjoy creating problems. My thesis is essentially about the creation and solution of a problem in wireless communications. To do so I will follow what I believe to be the systematic steps in solving problems in engineering. I begin by presenting a problem in using pilot symbols for the channel estimation of OFDM systems with feedback. To do so, I first motivate the reader by giving an introduction to OFDM (Chapter 1) and to the various issues regarding channel estimation for OFDM (Chapter 2). Next, I mathematically formulate and classify a problem in pilot allocation as an optimization problem of combinatorial nature (Chapter 3). I proceed to give a solution to my problem and in doing so introduce the methodology used in solving the problem (Chapter 4). No engineering problem is ultimately beneficial unless it finds direct or indirect applications in our everyday lives. Hence in Chapter 6 I extend and combine my proposed solution to more modern digital communication technologies such as multiple-input multiple-output (MIMO) systems and space-time block codes (STBC). Such technologies have already been considered for integration into next generation (4G) wireless telecommunication protocols.

Finally I would like to note that the following journal paper has been approved for publication (under minor revisions) based on the results and analysis given in Chapter 3 of this thesis.

[J1] **A. Yazdan-Panah**, B. N. Makouei and R. G. Vaughan, "Non-Uniform Pilot Symbol Allocation for Closed-Loop OFDM," *(approved on June 13^{th}, 2007) IEEE Transactions on Wireless Communications.*

Other publications as of 2007 include:

[C1] **A. Yazdan-Panah**, B. N. Makouei and R. G. Vaughan, "Pilot Feedback Equalization for Time-Varying OFDM Systems" in *Proc. IEEE Canadian Conference on Electrical and Computer Engineering 2007*, Vancouver, Canada.

[C2] **A. Yazdan-Panah**, B. N. Makouei and R. G. Vaughan, "An Expectation - Maximization Solution to Interpolated OFDM Systems," in *Proc. IEEE Canadian Conference on Electrical and Computer Engineering 2007,* Vancouver, Canada.

[C3] **A. Yazdan-Panah**, B. N. Makouei and R. G. Vaughan, "OFDM with Cyclic-Pilot Time Diversity" in *Proc. 66th IEEE Vehicular Technology Conference (VTC)*, Oct. 2007, Baltimore, Maryland U.S.A.

[C4] **A. Yazdan-Panah** and R. G. Vaughan, "Fault Tolerance of Quantized Unitary Precoders," in *Proc. IEEE Workshop on Signal Processing Systems (SiPS)*, Banff Canada, Oct. 2006.

[C5] **A. Yazdan-Panah** and R. G. Vaughan, "Quantized Unitary Precoders for Limited Feedback MIMO Systems," in *Proc. Wireless Personal Multimedia Communications Symposia (WPMC)*, San Diego California, Sep. 2006.

[C6] **A. Yazdan-Panah** and R .G. Vaughan, "Diversity Performance of Quantized Unitary Precoders," in *Proc. International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 449-454, Vancouver Canada, Jul. 2006.

[C7] **A. Yazdan-Panah**, B. N. Makouei and F. Marvasti, "Comparison Between Several Methods of PPM Demodulation Based on Iterative Techniques," in *Proc. 11th International Conference on Telecommunications (ICT)*, pp. 554-559, vol. 3124, Fortaleza Brazil, Aug. 2004.

List of Acronyms

AWGN	Additive white Gaussian noise
BICM	Bit-interleaved coded modulation
CFR	Channel frequency response
CIR	Channel impulse response
CP	Cyclic prefix
CSI	Channel state information
CSIT	Channel state information at transmitter
DAB	Digital audio broadcasting
DFT	Discrete Fourier transform
DL	Downlink
DVB-T	Digital video broadcasting - terrestrial
FDD	Frequency division duplexing
FFT	Fast Fourier transform
GA	Genetic algorithm
GLA	Generalized Lloyd algorithm
HIPERLAN	High performance local area networks
IEEE	Institute of electrical and electronic engineers
MIMO	Multiple-input multiple-output
MLI	Maximum likelihood interpolator
MPSK	M-ary phase shift keying
MSE	Mean square error
OFDM	Orthogonal frequency-division multiplexing
PSACE	Pilot symbol-aided channel estimation
PSAM	Pilot symbol-aided modulation
SGA	Simple Genetic Algorithm
STBC	Space-time block code
TDD	Time division duplexing
UL	Up link
WLAN	Wireless local area networks
WSS	Wide-sense stationary

List of Notations

$\mathbf{x}$	matrix or vector (both upper and lower case)
$\hat{\mathbf{x}}$	an estimate of $\mathbf{x}$
$\circledast$	discrete convolution
$(\cdot)^*$	conjugate
$(\cdot)^T$	transpose
$(\cdot)^H$	conjugate-transpose
$(\cdot)^{-1}$	square matrix inverse
$(\cdot)^\dagger$	non-square matrix pseudo-inverse
$\lfloor \cdot \rfloor$	lowest rounded integer
$\|\cdot\|_2$	vector 2-norm
$\|\cdot\|_F$	matrix Frobenius norm
$\mathfrak{Re}\{\cdot\}$	real part
$\mathfrak{Im}\{\cdot\}$	imaginary part
$\mathbb{C}^N, \mathbb{Z}^N, \mathbb{R}^N$	N-dimensional complex, integer, and real vector spaces
$\mathbf{X}_{(m,n)}$	$(m,n)^{th}$ element of matrix $\mathbf{X}$
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbf{A} \otimes \mathbf{B}$	kronecker product
$\mathbb{E}_{\mathbf{x}}\{\cdot\}$	expectation with regard to random variable $\mathbf{x}$
$tr\{\cdot\}$	trace of a matrix (sum of diagonal elements)
$vec\{\cdot\}$	set vectorizing operator
$diag\{\cdot\}$	vector containing the diagonal elements of a matrix
$\mathcal{CN}(0,1)$	complex normal distribution

Chapter 1

OFDM Preliminaries

As a candidate for next generation (4G) high-speed digital wide-band communications, orthogonal frequency division multiplexing (OFDM), first introduced in [40], has already been used in European digital audio broadcasting (DAB) [35], digital video broadcasting (DVB) systems [34], high performance radio local area networks (HIPERLAN) and the IEEE 802.11 family of wireless local area networks (WLAN) [14]. OFDM is a particularly attractive scheme for broadband systems which encounter large channel delay spreads. In this chapter, we review OFDM modulation starting with the general broadband mobile channel model. After presenting the OFDM baseband transmitter and receiver models, we review a channel estimation method commonly referred to as pilot symbol-aided channel estimation (PSACE), also known as pilot symbol-aided modulation (PSAM) in the context of OFDM signaling.

1.1 Channel Model

Consider a single-input single-output (SISO) wireless link operating over a multi-path time-dispersive channel with bandwidth B. Such a channel is conventionally modeled in *continuous-time* with a wide-sense-stationary uncorrelated scattering (WSSUS) impulse response

$$h(t;\tau) = \sum_{l=1}^{L} \alpha_l(t)\delta(\tau - \tau_l) \tag{1.1}$$

where τ_l is the l^{th} path delay and $\alpha_l(t)$ is the corresponding complex gain. Sampling this channel at $1/B \approx T_s$ intervals and normalizing the delays by the symbol duration T_s

($\tau_1 = 0, L = \lfloor \tau_{max}/T_s \rfloor$), we have in *discrete-time* for $t = k$ and $j = 1, 2, \ldots, L$

$$h_k(j) = \sum_{l=1}^{L} \alpha_l(k)\delta(l - j). \tag{1.2}$$

With adequate scattering, and due to the central limit theorem, the complex path gains can be modeled as wide-sense stationary (WSS) independently complex Gaussian random processes with non time-varying moments give by ($l = 1, 2, \ldots, L$)

$$\mathbb{E}\{\alpha_l\} = 0 \tag{1.3}$$

$$\sigma_{\alpha_l}^2 = \mathbb{E}\left\{\alpha_l^2\right\} = p(l, \zeta) \tag{1.4}$$

$$\mathbb{E}\{\alpha_n \alpha_m\} = 0, \quad (n \neq m) \tag{1.5}$$

where $p(l, \zeta)$ is the *power delay profile* and is adopted to be an exponential function of the form $p(l, \zeta) = Ke^{\left(\frac{-l}{\zeta L}\right)}$. The parameter ζ may be used to control the rate of decay (time constant) while K is a normalizing constant to assure $\sum_{l=1}^{L} |p(l, \zeta)| = 1$ for any ζ. Using (1.2) and (1.5), such a normalization lead to the statistical normalization of

$$\mathbb{E}\left\{\sum_j |h_k(j)|^2\right\} = \mathbb{E}\left\{\|\mathbf{h}_k\|_2^2\right\} = \sum_{l=1}^{L} \sigma_{\alpha_l}^2 = 1 \tag{1.6}$$

where $\mathbf{h}_k$ is the length-L channel impulse response (CIR) vector at time k. The length-N channel frequency response (CFR) vector is related to the CIR through a subset of the unitary $N \times N$ DFT matrix $\mathbf{F}$, where $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_N]$ and for $m, n = 1, 2, \ldots, N$

$$\mathbf{F}_{m,n} = \mathbf{f}_n(m) = \frac{1}{\sqrt{N}} e^{-j2\pi(n-1)(m-1)/N}. \tag{1.7}$$

Defining $\mathbf{G} = \sqrt{N}\,[\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_L]$ we have

$$\mathbf{H}_k = \mathbf{G}\mathbf{h}_k. \tag{1.8}$$

Using (1.6), (1.8) and the fact that $\mathbf{G}^H\mathbf{G} = N\mathbf{I}_L$, we have the statistical normalization in the frequency domain as

$$\begin{aligned} \mathbb{E}\left\{\|\mathbf{H}_k\|_2^2\right\} &= \mathbb{E}\left\{\|\mathbf{G}\mathbf{h}_k\|_2^2\right\} \\ &= N\mathbb{E}\left\{\|\mathbf{h}_k\|_2^2\right\} \\ &= N. \end{aligned} \tag{1.9}$$

Correlation Functions

The OFDM channel described above is correlated in time and in frequency. Using (1.5), the frequency correlation function at the n^{th} sub-carrier is

$$\begin{aligned}
\mathbf{R}_f(n) &= \mathbb{E}\{H(n)H^*(n+m)\} \\
&= \mathbb{E}\left\{\left(\sum_l h(l)e^{-j2\pi lm/N}\right)\left(\sum_{l'} h^*(l')e^{j2\pi l'(m+n)/N}\right)\right\} \\
&= \mathbb{E}\left\{\sum_i |h(i)|^2 e^{j2\pi n/N}\right\} \\
&= \sum_i \sigma^2_{\alpha_i} e^{j2\pi in/N} \qquad (1.10)
\end{aligned}$$

Hence in vector form we have

$$\mathbf{R}_f = \mathbf{G}^H \mathbf{\Sigma_h} \qquad (1.11)$$

where $\mathbf{\Sigma_h} = [\sigma^2_{\alpha_1}, \sigma^2_{\alpha_2}, \ldots, \sigma^2_{\alpha_L}]^T$. The time correlation function, on the other hand, is a function of the Doppler frequency, following the Jakes model:

$$\mathbb{E}\{h_i(l)h^*_{i+k}(l)\} = \sigma^2_{\alpha_l} J_0(2\pi f_d N T_s k) \qquad (1.12)$$

Using (1.12) and (1.8) the time correlation function at the n^{th} sub-carrier is

$$\begin{aligned}
\mathbf{R}_t(k) &= \mathbb{E}\{H_i(n)H^*_{i+k}(n)\} \\
&= \mathbb{E}\left\{\left(\sum_l h_i(l)e^{-j2\pi ln/N}\right)\left(\sum_l h^*_{i+k}(l)e^{j2\pi ln/N}\right)\right\} \\
&= \mathbb{E}\left\{\sum_l h_i(l)h^*_{i+k}(l)\right\} \\
&= \sum_l \mathbb{E}\{h_i(l)h^*_{i+k}(l)\} \\
&= J_0(2\pi f_d N T_s k)\sum_l \sigma^2_{\alpha_l} \\
&= J_0(2\pi f_d N T_s k). \qquad (1.13)
\end{aligned}$$

Since $\sum_l \sigma^2_{\alpha_l} = 1$ from (1.6). $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind with $J_0(0) = 1$ and $J_0(x) = \frac{1}{\pi}\int_0^\pi \cos(x\sin\theta)\, d\theta$. T_s is the symbol rate and f_d is the channel Doppler frequency given by $f_d = f_c v/c$, where v is the receiver speed, f_c is the carrier frequency and c is the prorogation speed of light. Here we assume that the channel is

quasi-static, changing only from OFDM symbol to symbol. Therefore, we define the *OFDM normalized Doppler frequency* as $F_d = f_d T_s N$. This is, in effect, the Doppler felt by a N successive modulated symbols (or one OFDM symbol).

1.2 Transceiver Model

Consider at time k, the transmission of N baseband complex independent and identically distributed (i.i.d.) symbols $s(n)$, $n = 1, 2, \ldots, N$ from a single antenna at the transmitter ($M_t = 1$). The symbols are drawn from a M-PSK constellation $\mathcal{M}$ so the power is constrained by $\mathbb{E}\{|s(n)|^2\} = 1$. Denote the transmitted vector by $\mathbf{s}_k = [s_k(1), s_k(2), \ldots, s_k(N)]^T$ with correlation matrix given by $\mathbf{R_s} = \mathbb{E}\left\{\mathbf{s}_k^H \mathbf{s}_k\right\} = \mathbf{I}_N$. Then the *total power* is constrained by $tr\{\mathbf{R_s}\} = N$. We describe the OFDM transmitter and receiver separately in more detail below.

1.2.1 The Transmitter

The OFDM transmitter is shown in Fig. 1.1. The transmission stream is first multiplexed into N parallel streams, followed by a unitary N-point inverse discrete Fourier transform (DFT). This yields the vector $\mathbf{x}_k = [x_k(1), x_k(2), \ldots, x_k(N)]^T$ given by

$$\mathbf{x}_k = \mathbf{F}^H \mathbf{s}_k \tag{1.14}$$

where $\mathbf{F}$ is defined in (1.7). In practice N is chosen to be a power of 2 to allow for efficient implementation using the inverse fast Fourier transform (IFFT). Note that owing to the central limit theorem for large N, since the elements of $\mathbf{s}_k$ are i.i.d. distributed, the elements of $\mathbf{x}_k$ can be assumed to be Gaussian. Assuming the L-tap channel model of (1.2), a prefix is constructed consisting of the last $L-1$ symbols of $\mathbf{x}_k$. This prefix is conventionally called the *cyclic prefix*(CP) (or guard interval), and as we will see, will eliminate the need for an equalizer at the OFDM receiver. Hence the final transmitted signal is a length $(N + L - 1)$ vector

$$\tilde{\mathbf{x}}_k = [\overbrace{x_k(N-L+2), \ldots, x_k(N)}^{cyclic\ prefix}, \overbrace{x_k(1), x_k(2), \ldots, x_k(N)}^{data}]^T. \tag{1.15}$$

The elements of this vector are transmitted serially through a frequency selective channel $\mathbf{h}_k$ in a total OFDM symbol duration time of $T_s^{OFDM} = (N + L - 1)T_s = (N + L - 1)/B$, where B is again the channel bandwidth.

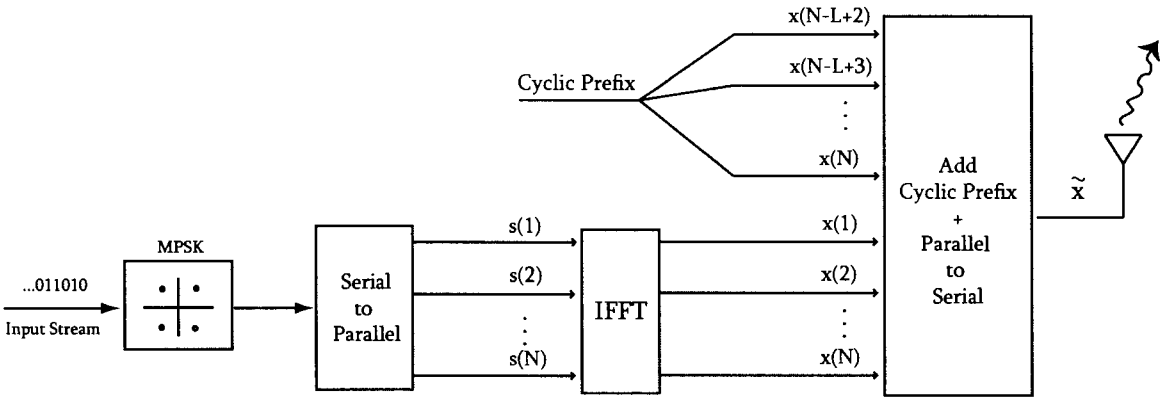

Figure 1.1: OFDM baseband transmitter model

1.2.2 The Receiver

The corresponding OFDM receiver is shown in Fig. 1.2. The transmitted vector $\tilde{\mathbf{x}}_k$ is convolved with the L-tap CIR as it propagates through the multi-path channel $\mathbf{h}_k$, resulting in the length $(N + 2L - 2)$ received vector $\tilde{\mathbf{y}}_k$.

$$\tilde{\mathbf{y}}_k = \mathbf{h}_k \circledast \tilde{\mathbf{x}}_k + \tilde{\mathbf{n}}_k \tag{1.16}$$

or

$$\tilde{y}_k(i) = \sum_q h_k(q)\tilde{x}_k(i-q) + \tilde{n}_k(i) \tag{1.17}$$

This convolution may be written in terms of a Toeplitz matrix representation of the channel

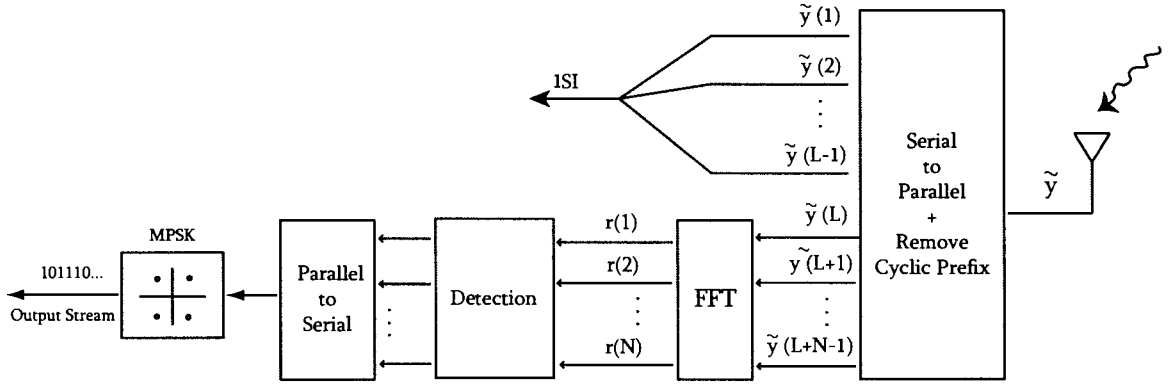

Figure 1.2: OFDM baseband receiver model

at time k

$$\tilde{\mathbf{y}}_k = \tilde{\mathbf{C}}_k^{(toeplitz)} \tilde{\mathbf{x}}_k + \tilde{\mathbf{n}}_k \tag{1.18}$$

where $\tilde{\mathbf{n}}_k$ is the zero-mean Gaussian noise vector with correlation matrix $\sigma_n^2 \mathbf{I}_{(N+2L-2)}$ and

$$\tilde{\mathbf{C}}_k^{(toeplitz)} = \begin{bmatrix} h_k(1) & 0 & 0 & 0 & \cdots & 0 \\ h_k(2) & h_k(1) & 0 & 0 & \cdots & 0 \\ h_k(3) & h_k(2) & h_k(1) & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & 0 \\ h_k(L) & h_k(L-1) & h_k(L-2) & h_k(L-3) & \cdots & 0 \\ 0 & 0 & \vdots & \vdots & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}_{(N+2L-2)\times(N+L-1)} \tag{1.19}$$

The receiver next ignores the first $L-1$ symbols of $\tilde{\mathbf{y}}_k$ (the CP containing ISI) and gathers the next N samples $\mathbf{y}_k = [\tilde{y}_k(L), \tilde{y}_k(L+1), \ldots, \tilde{y}_k(N+L-1)]^T$, which from (1.19) satisfies

$$\mathbf{y}_k = \tilde{\mathbf{C}}_k \tilde{\mathbf{x}}_k + \mathbf{n'}_k \tag{1.20}$$

with

$$\tilde{\mathbf{C}}_k = \begin{bmatrix} h_k(L) & \cdots & h_k(2) & h_k(1) & 0 & 0 & \cdots & 0 \\ 0 & h_k(L) & \cdots & h_k(2) & h_k(1) & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \vdots & 0 & h_k(L) & \cdots & h_k(2) & h_k(1) & 0 \\ 0 & 0 & 0 & 0 & h_k(L) & \cdots & h_k(2) & h_k(1) \end{bmatrix}_{N\times(N+L-1)} \tag{1.21}$$

From (1.15), the first $L-1$ samples of $\tilde{\mathbf{x}}_k$ are identical to the last $L-1$ samples on account of the CP. Therefore (1.20) may be simplified to

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{n'}_k \tag{1.22}$$

where

$$\mathbf{C}_k = \begin{bmatrix} h_k(1) & 0 & \cdots & 0 & 0 & h_k(L) & \cdots & h_k(2) \\ h_k(2) & h_k(1) & 0 & \cdots & 0 & 0 & \ddots & \vdots \\ \vdots & h_k(2) & h_k(1) & 0 & 0 & \ddots & 0 & h_k(L) \\ h_k(L) & \vdots & h_k(2) & \ddots & 0 & \ddots & 0 & 0 \\ 0 & h_k(L) & \vdots & \ddots & h_k(1) & \ddots & \ddots & 0 \\ \vdots & 0 & h_k(L) & \ddots & h_k(2) & h_k(1) & 0 & 0 \\ \vdots & \vdots & 0 & \ddots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & h_k(L) & \cdots & h_k(2) & h_k(1) \end{bmatrix} \tag{1.23}$$

Note that $\mathbf{C}_k = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N]$ is a *circulant* matrix hence has the eigen-decomposition of

$$\mathbf{C}_k = \mathbf{F}^H \mathbf{\Lambda}_k \mathbf{F} \tag{1.24}$$

where $\mathbf{F}$ is the unitary DFT matrix of (1.7) and $\mathbf{\Lambda}_k$ is a diagonal matrix with elements given by the Fourier transform (non-unitary) of the first column of $\mathbf{C}$

$$\mathbf{\Lambda}_k = \mathbf{G}\mathbf{c}_1 \tag{1.25}$$

Since $\mathbf{c}_1$ is in fact the (zero padded) CIR, we have from (1.8)

$$\mathbf{\Lambda}_k = \mathbf{G}\mathbf{c}_1 = \mathbf{H}_k \tag{1.26}$$

Taking the fast Fourier transform (FFT) of (1.22)

$$\begin{aligned} \mathbf{R}_k &= \mathbf{F}\mathbf{y}_k \\ &= \mathbf{F}\mathbf{F}^H \mathbf{\Lambda}_k \mathbf{F} \overbrace{\mathbf{F}^H \mathbf{s}_k}^{\mathbf{x}_k} + \mathbf{F}\mathbf{n}'_k \\ &= \mathbf{\Lambda}_k \mathbf{s}_k + \mathbf{n}_k \end{aligned} \tag{1.27}$$

where we have used (1.24) and $\mathbf{F}^H\mathbf{F} = \mathbf{I}_N$ which also results in $\mathbf{n}_k = \mathbf{F}\mathbf{n}'_k$ having the same distribution as $\mathbf{n}'_k$ with a correlation matrix of $\sigma_n^2 \mathbf{I}_n$. Using (1.26) this result may be rewritten as

$$\mathbf{R}_k = \mathbf{S}_k \mathbf{H}_k + \mathbf{n}_k \tag{1.28}$$

or using (1.8)

$$\mathbf{R}_k = \mathbf{S}_k \mathbf{G} \mathbf{h}_k + \mathbf{n}_k \tag{1.29}$$

where $\mathbf{G} = \sqrt{N}\mathbf{F}$ and $\mathbf{S}_k = diag\{s_k(1), s_k(2), ..., s_k(N)\}$ is an $N \times N$ diagonal matrix of transmitted symbols taken from the constellation space $\mathcal{M}$, $\mathbf{R}_k = diag\{r_k(1), r_k(2), ..., r_k(N)\}$ is the received vector and $\mathbf{H}_k = [H_k(1), H_k(2), ..., H_k(N)]^T$ is the vector representing the CFR at time k.

1.3 Chapter Summary

We reviewed the concept of baseband OFDM modulation from a mathematical point of view. We showed how ISI in broadband communications may be completely eliminated by using a cyclic prefix and the discrete Fourier transform. We also reviewed the wide-sense-stationary mobile channel model. This model will be used extensively in subsequent chapters where the focus will be primarily on channel estimation in time and in frequency.

Chapter 2

OFDM Channel Estimation

Pilot-symbol aided channel estimation (PSACE), also known as pilot-symbol aided modulation (PSAM), was first analyzed by Cavers in [8] as a simple and effective means of estimating multi-path mobile channels. In PSAM, a set of known symbols commonly referred to as *pilot tones* are multiplexed into the transmission stream prior to transmission and used at the receiver to estimate, and if need be, track the communications channel. Since its introduction, PSAM has been rigorously analyzed and optimized both in terms of average error rates [8], [37], [9], [27], [6] and average throughput performances [29], [30] for slow and fast-fading environments. The robustness of the technique and its simplicity have resulted in a multitude of applications for the use of PSAM in coherent digital and wireless communications. One such application is the channel detection of systems using orthogonal frequency division multiplexing (OFDM) where the wideband channel is divided into several parallel narrowband sub-channels [5], [40]. Multiplexing known data symbols into the time-frequency transmission grid gives the OFDM receiver the ability not only to estimate the mobile channel (see e.g. [28], [12], [38]), but with proper design, also to track its fluctuations (see e.g. [36], [20] and references therein).

OFDM channel estimation for time-frequency multiplexed pilot tones is generally a two step process. The first step involves the *statistical estimation* of the channel at the pilot-tones. Statistical methods generally include classical linear least mean squares (LLMS) and minimum mean square error (MMSE) estimators [38] with adaptive techniques reported in [31]. The channel within the sub-bands is subsequently determined by *interpolation* such as linear, second order, and time domain interpolations [13]. Although appealing in terms of computational complexity, such non-statistical interpolation techniques, suffer considerable

performance losses (mainly due to irreducible error floors). In particular, the method proposed in [33] provides channel estimates based on piecewise-constant and piecewise-linear interpolations between pilots. It is simple to implement, but it needs a large number of pilots to get satisfactory performance [26].

A different approach is to directly incorporate the interpolation into the estimation process leading to a one step channel estimation process. Maximum likelihood interpolation (MLI) and minimum mean square error interpolation (MMSEI), exploit certain channel statistics such as time and or frequency correlation functions in the entire time or frequency band to estimate the channel in a *single step*. The average error rate performance of both these approaches are compared by simulation in [13], [26] and theoretically analyzed in [27], [37], [26] and [10], concluding the superiority of the latter approach (see e.g. the results in [26]). We therefore choose the maximum likelihood interpolator (MLI) as the OFDM channel estimator throughout this work unless explicitly stated otherwise.

2.1 The Maximum Likelihood Interpolator

In this section we follow the analysis of [26] for PSAM-OFDM. The derivations and notations here will be used in subsequent sections when we consider non-uniformly located pilot symbols.

The receiver of Fig. 1.2 must obtain a reliable estimate of the channel frequency response to perform coherent detection. To this end we assume that known pilot symbols are multiplexed into the data stream, and channel estimation is performed by interpolation between the pilots locations at the receiver. A total of N_s data carrying symbols $\{s_d \in \mathcal{M};\ 1 \leq d \leq N_s\}$ are inserted in the OFDM symbol at data locations given by the *data index vector* $\mathbf{n_d} = vec\{n_i\}_{i=1}^{N_s} \in \mathcal{I}_d$ where $\mathcal{I}_d$ is the finite set of all possible $\mathbf{n_d}$ vectors. Also, N_p pilots, $\{s_p \in \mathcal{M};\ 1 \leq p \leq N_p\}$, are multiplexed into the OFDM symbol at known locations given by the *pilot index vector* $\mathbf{n_p} = \{n_p(1), \ldots, n_p(N_p)\} = vec\{n_i\}_{i=1}^{N_p} \in \mathcal{I}_p$. We assume that all sub-carriers are active (used for transmission) so $N = N_s + N_p$. The N_p-dimensional output vector at the pilot locations is $\mathbf{R}^{\langle p \rangle} = [R(\mathbf{n_p}(1)), R(\mathbf{n_p}(2)), \ldots R(\mathbf{n_p}(N_p))]^T$ and from (1.29)

$$\mathbf{R}^{\langle p \rangle} = \mathbf{S}^{\langle p \rangle}\mathbf{B}\mathbf{h} + \mathbf{n}^{\langle p \rangle} \tag{2.1}$$

where $\mathbf{B}$ is an $N_p \times L$ subset matrix of $\mathbf{G}$ with entries

$$\mathbf{B}_{p,k} = e^{-j2\pi(\mathbf{n_p}(p)-1)(k-1)/N}. \quad 1 \leq p \leq N_p \ , \quad 1 \leq k \leq L \tag{2.2}$$

Pilot symbols are taken from the MPSK constellation $\mathcal{M}$ so $|s_p|^2 = 1$. Pre-multiplying both sides of (2.1) by $\left(\mathbf{S}^{\langle p \rangle}\right)^H$ produces

$$\mathbf{Y}^{\langle p \rangle} = \mathbf{Bh} + \mathbf{w}^{\langle p \rangle}. \tag{2.3}$$

The noise vector $\mathbf{w}^{\langle p \rangle} = \left(\mathbf{S}^{\langle p \rangle}\right)^H \mathbf{n}^{\langle p \rangle}$ has the same distribution as $\mathbf{n}^{\langle p \rangle}$. The goal is to estimate $\mathbf{H}$ from the observation vector $\mathbf{Y}^{\langle p \rangle}$. The maximum likelihood solution is based on the assumption that $\mathbf{h}$ is deterministic but unknown. Using (2.3) and the fact that $\mathbf{w}^{\langle p \rangle}$ is Gaussian distributed, the solution is (zero-forcing solution)

$$\hat{\mathbf{h}} = \arg\min_{\mathbf{h}} ||\mathbf{Y}^{\langle p \rangle} - \mathbf{Bh}||_2^2 = \mathbf{D}^{-1}\mathbf{B}^H\mathbf{Y}^{\langle p \rangle} \tag{2.4}$$

where $\mathbf{D}$ is an $L \times L$ square matrix

$$\mathbf{D} = \mathbf{B}^H\mathbf{B}. \tag{2.5}$$

Using (1.8) and the invariance property of the maximum likelihood estimator [26] the CFR estimation is

$$\widehat{\mathbf{H}} = \mathbf{G}\widehat{\mathbf{h}}. \tag{2.6}$$

Finally substituting (2.4) into (2.6) and using (2.3) we get

$$\widehat{\mathbf{H}} = \mathbf{Gh} + \mathbf{GD}^{-1}\mathbf{B}^H\mathbf{w}^{\langle p \rangle}. \tag{2.7}$$

Defining

$$\mathbf{P} = \mathbf{GD}^{-1}\mathbf{B}^H \tag{2.8}$$

as the $N \times N_p$ *interpolating matrix* and using (1.8), (2.7) we have

$$\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{Pw}^{\langle p \rangle}. \tag{2.9}$$

Substituting (2.9) into (1.29) we get for the output

$$\begin{aligned}\widehat{\mathbf{R}} &= \mathbf{S}\widehat{\mathbf{H}} + \mathbf{n} \\ &= \mathbf{S}\left(\mathbf{H} + \mathbf{P}\mathbf{w}^{\langle p \rangle}\right) + \mathbf{n} \\ &= \mathbf{S}\mathbf{H} + \overbrace{\underbrace{\mathbf{S}\mathbf{P}\mathbf{w}^{\langle p \rangle}}_{\text{interpolation noise}} + \underbrace{\mathbf{n}}_{AWGN}}^{\text{total noise}}.\end{aligned} \tag{2.10}$$

This expression shows two distinct sources of noise. One is the AWGN inherent in the receiver and the second is AWGN at the pilot locations distributed to all sub-carriers via the interpolation. As before, $\mathbf{n}$ is white Gaussian noise with covariance matrix give by (same as correlation since the noise is zero mean) $\mathbf{C}_{\mathbf{n}} = \mathbb{E}\left\{\mathbf{n}\mathbf{n}^H\right\} = \sigma_n^2\mathbf{I}_N$, with total average power $P_{\mathbf{n}} = tr\left\{\mathbb{E}\left\{\mathbf{n}\mathbf{n}^H\right\}\right\} = N\sigma_n^2$.

2.1.1 Cramer-Rao Lower Bound

Treating the channel $\mathbf{H}$ as deterministic but unknown it is easy to see from 2.9 that the MLI is unbiased

$$\mathbb{E}\left\{\widehat{\mathbf{H}}\right\} = \mathbb{E}\left\{\mathbf{H} + \mathbf{P}\mathbf{w}^{\langle p \rangle}\right\} = \mathbf{H} \tag{2.11}$$

since the noise $\mathbf{w}$ is zero mean on all sub-carriers. The covariance of the estimate is

$$\begin{aligned}\mathbf{C}_{\widehat{\mathbf{H}}} &= \mathbb{E}\left\{\left(\widehat{\mathbf{H}} - \mathbb{E}\left\{\widehat{\mathbf{H}}\right\}\right)\left(\widehat{\mathbf{H}} - \mathbb{E}\left\{\widehat{\mathbf{H}}\right\}\right)^H\right\} && (2.12)\\ &= \mathbb{E}\left\{\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H\right\} && (2.13)\\ &= \mathbb{E}\left\{\left(\mathbf{P}\mathbf{w}^{\langle p \rangle}\right)\left(\mathbf{P}\mathbf{w}^{\langle p \rangle}\right)^H\right\} && (2.14)\\ & && (2.15)\end{aligned}$$

We would be hard pressed to further simplify this expression since the interpolating matrix $\mathbf{P}$ is a function of the pilot index $\mathbf{n}_p$ and hence a function of the channel. Assuming equi-spaced pilots, $\mathbf{P}$ is constant and using the definitions in (2.8) and (2.5) we have

$$\mathbf{C}_{\widehat{\mathbf{H}}} = \mathbf{P}\mathbb{E}\left\{\mathbf{w}^{\langle p\rangle}\left(\mathbf{w}^{\langle p\rangle}\right)^H\right\}\mathbf{P}^H \tag{2.16}$$

$$= \sigma_n^2\mathbf{P}\mathbf{P}^H \tag{2.17}$$

$$= \sigma_n^2\mathbf{G}\mathbf{D}^{-1}\mathbf{B}^H\mathbf{B}\mathbf{D}^{-H}\mathbf{G}^H \tag{2.18}$$

$$= \sigma_n^2\mathbf{G}\mathbf{D}^{-1}\mathbf{G}^H \tag{2.19}$$

Similarly for $\mathbf{h}$ we have

$$\mathbf{C}_{\widehat{\mathbf{h}}} = \sigma_n^2\mathbf{D}^{-1} \tag{2.20}$$

Denote by $\mathfrak{Re}\{\mathbf{h}\} = \mathbf{h}_r$ and $\mathfrak{Im}\{\mathbf{h}\} = \mathbf{h}_i$ the real and imaginary parts of $\mathbf{h}$ and define $\mathbf{\Phi} = (\mathbf{h}_r\mathbf{h}_i^T)^T$. The components of the Fisher matrix are give by the log-likelihood function of (2.3)

$$\mathbf{\Xi}_{(i,j)} = -\mathbb{E}\left\{\frac{\partial\ln(\mathcal{L}(\mathbf{Y}^{\langle p\rangle};\mathbf{\Phi}))}{\partial\mathbf{\Phi}(i)\partial\mathbf{\Phi}(j)}\right\} \tag{2.21}$$

where $\mathcal{L}(\mathbf{Y}^{\langle p\rangle};\mathbf{\Phi})$ is the likelihood of the received vector at the pilot locations given $\mathbf{\Phi}$

$$\mathcal{L}(\mathbf{Y}^{\langle p\rangle};\mathbf{\Phi}) = \frac{1}{(\pi\sigma_n^2)^{N_p}} \times e^{-\frac{1}{\sigma_n^2}\left[\mathbf{Y}^{\langle p\rangle}-\mathbf{B}\mathbf{h}\right]^H\left[\mathbf{Y}^{\langle p\rangle}-\mathbf{B}\mathbf{h}\right]} \tag{2.22}$$

Substituting (2.22) into (2.21) yields

$$\mathbf{\Xi} = \frac{2}{\sigma_n^2}\begin{bmatrix} \mathfrak{Re}\{\mathbf{D}\} & -\mathfrak{Im}\{\mathbf{D}\} \\ \mathfrak{Im}\{\mathbf{D}\} & \mathfrak{Re}\{\mathbf{D}\} \end{bmatrix} \tag{2.23}$$

and

$$\mathbf{\Xi}^{-1} = \frac{\sigma_n^2}{2}\begin{bmatrix} \mathfrak{Re}\{\mathbf{D}^{-1}\} & -\mathfrak{Im}\{\mathbf{D}^{-1}\} \\ \mathfrak{Im}\{\mathbf{D}^{-1}\} & \mathfrak{Re}\{\mathbf{D}^{-1}\} \end{bmatrix}. \tag{2.24}$$

The Cramer-Rao lower bound (CRLB) is by definition given by $\mathsf{CRLB}(\mathbf{h}) = tr\{\mathbf{\Xi}^{-1}\}$. Using the fact that $\mathbf{D}$ is Hermitian, we have from (2.24)

$$\mathsf{CRLB}(\mathbf{h}) = \sigma_n^2 tr\left\{\mathbf{D}^{-1}\right\} \tag{2.25}$$

Comparing this result with the covariance matrix of the channel given by (2.20), we conclude that the covariance of the MLI coincides with the CRLB [26].

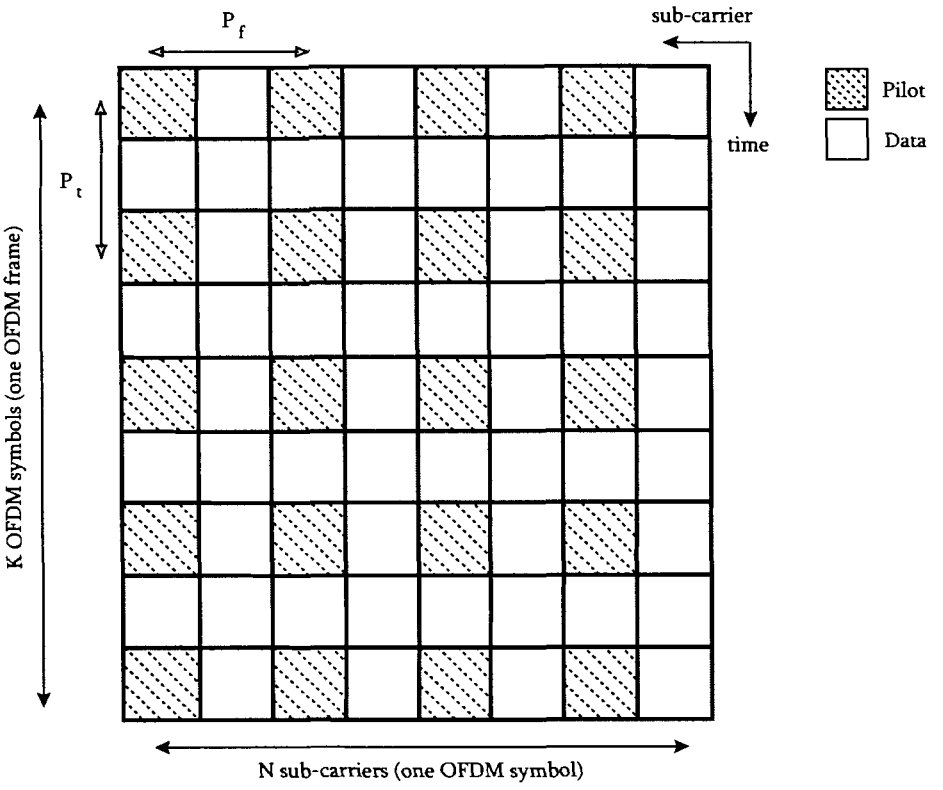

Figure 2.1: An example of comb-type pilot symbol allocation

2.1.2 Performance of Equi-Spaced Pilots

Consider the case where the pilot symbols are equi-spaced in frequency. In other words if N is the total number of sub-carriers per OFDM symbol and N_p denotes the number of pilots per OFDM symbol then, $\mathbf{n_p}(i+1) - \mathbf{n_p}(i) = \lfloor N/N_p \rfloor$, for any i where $1 \leq i \leq N-1$. An example of such an arrangement is shown in Fig. 2.1 where $N_p = 4$ pilots are uniformly inserted in between $N = 8$ sub-carriers of every other OFDM symbol. Such an arrangement of pilots is typically called "comb-type pilot allocation". The parameter P_f in Fig. 2.1 denotes the horizontal (frequency) pilot separation which in the uniform case equals N/N_p and P_t denotes the vertical (time) pilot separation which is necessary due to the Doppler occurred from a moving receiver. These parameters will be quantified in the forthcoming simulations on a need be basis.

Mean Square Error and the Optimality of Uniform Pilots

To see the optimality of uniformly allocated pilots we must go back to the ZF solution of (2.4). From (2.4) and (2.5) we have

$$\widehat{\mathbf{h}} = \mathbf{h} + \mathbf{B}^{\dagger}\mathbf{w}^{\langle p \rangle} \tag{2.26}$$

where $\mathbf{B}^{\dagger} = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$ is by definition the psedo-inverse of $\mathbf{B}$. Such an inversion requires that $\mathbf{B}$ be full rank or $N_p \geq L$. This means the number of pilots must be greater than the number of channel taps. We will call the minimum number of pilots, $N_p = L$, the Nyquist pilot rate (NPR). The *total* channel mean square error (MSE) is

$$\begin{aligned}
\overline{\mathsf{MSE}}(\mathbf{h}) &\stackrel{\text{def}}{=} \mathbb{E}\left\{||\widehat{\mathbf{h}} - \mathbf{h}||_2^2\right\} \\
&= \mathbb{E}\left\{tr\left\{\mathbf{B}^{\dagger}\mathbf{w}^{\langle p \rangle}(\mathbf{w}^{\langle p \rangle})^H(\mathbf{B}^{\dagger})^H\right\}\right\} \\
&= tr\left\{\mathbb{E}\left\{\mathbf{w}^{\langle p \rangle}(\mathbf{w}^{\langle p \rangle})^H(\mathbf{B}^{\dagger})^H\mathbf{B}^{\dagger}\right\}\right\} \\
&= tr\left\{\sigma_n^2\mathbf{I}_{N_p}\mathbb{E}\left\{(\mathbf{B}^{\dagger})^H\mathbf{B}^{\dagger}\right\}\right\} \\
&= \sigma_n^2\mathbb{E}\left\{tr\left\{(\mathbf{B}^H\mathbf{B})^{-1}\right\}\right\} \\
&= \sigma_n^2\mathbb{E}\left\{tr\left\{\mathbf{D}^{-1}\right\}\right\} \\
&= \sigma_n^2 tr\left\{\mathbf{D}^{-1}\right\}
\end{aligned} \tag{2.27}$$

The expectation was dropped at the last step based on the assumption that information regarding the channel (also known as channel state information[1] (CSI)) is not available at the transmitter. In any case, this result is basically the derivation of (2.20) in the previous section[2]. Note that $\mathbf{D}$ depends on the pilot locations, i.e. $\mathbf{n_p}$ thus the objective is to find $\mathbf{n_p}$ that minimizes the channel MSE, or from (2.27) a pilot allocation that minimizes the trace of the inverse of $\mathbf{D}(\mathbf{n_p})$.

$$\mathbf{n_p}^{opt} = \arg\min_{\mathbf{n_p}\in\mathcal{I}_p} tr\left\{\mathbf{D}(\mathbf{n_p})^{-1}\right\} = \arg\min_{\mathbf{n_p}\in\mathcal{I}_p}\sum_{k=1}^{L}\lambda_k^{-1} \tag{2.28}$$

where λ_k is the k^{th} eigenvalue of $\mathbf{D}$. Also from the definition in (2.2) we have

[1] Also known as channel state information at transmitter (CSIT)

[2] Note from (2.27) and (2.25) how $\overline{\mathsf{MSE}}(\mathbf{h}) = \mathsf{CRLB}(\mathbf{h})$.

$$\mathbf{D}(\mathbf{n_p})_{(s,q)} = \sum_{m=1}^{N_p} e^{j2\pi(s-q)\mathbf{n_p}(m)/N} \tag{2.29}$$

therefore the diagonal entries of $\mathbf{D}$ are unity regardless of $\mathbf{n_p}$. Hence we have the constraint that $\sum_{k=1}^{L} \lambda_k = N_p L$ and the optimization in (2.28) can be rewritten as

$$\begin{aligned} \text{optimize} : \mathbf{n_p}^{opt} &= \arg \min_{\mathbf{n_p} \in \mathcal{I}_p} \sum_{k=1}^{L} \lambda_k^{-1} \\ \text{subject to} : \sum_{k=1}^{L} \lambda_k &= N_p L \end{aligned} \tag{2.30}$$

Since $\mathbf{D}$ is non negative definite, all the eigenvalues are non negative. Hence the minimization occurs when all the eigenvalues are equal which can only occur for $\mathbf{D}(\mathbf{n_p}^{opt}) = N_p \mathbf{I}_L$ or equi-spaced pilots (also called uniform allocation). Substituting this result into (2.27) we get

$$\overline{\mathsf{MSE}} = \sigma_n^2 \left(\frac{L}{N_p} \right) \tag{2.31}$$

In Appendix A. we derive the same result by considering the per-carrier MSE. Note that as expected with uniform pilot allocation, the channel MSE is not a function of the sub-carrier index and for the worst case of $N_p = L$ the MSE equals the noise variance σ_n^2. Note that without loss of generality we may set $\mathbf{n_p}(1) = 1$ since as long as $\mathbf{n_p}(i+1) - \mathbf{n_p}(i) = \lfloor N/N_p \rfloor$, the orthogonality of $\mathbf{D}(\mathbf{n_p})$ holds regardless of the location of initiating pilot[3]. An immediate result from (2.31) is:

"MLI with the lowest possible number of pilots, $N_p = L$, incurs a 3 dB penalty in SNR due to interpolation and AWGN."

Proof: for $N_p = L$ from (2.31) we have: $\overline{\mathsf{MSE}} = \sigma_n^2 \left(\frac{L}{N_p} \right) = \sigma_n^2$, therefore the total noise on each sub-carrier is $2\sigma_n^2$ which translates to a 3 dB loss in signal power. In Chapter 3 we show how uniform pilots also maximize the average SNR per OFDM symbol (see Fig 3.2).

[3]In effect since we do not consider suppressed (virtual) sub-carriers there is no "windowing" problem associated with the channel estimation process and the location of the first and last pilots is irrelavent to the either edges of the OFDM symbol.

2.1.3 The Interpolating Functions

From (2.4) and (2.6), the estimated channel on the n^{th} sub-carrier is in general a weighted sum of the received signal on the pilot location

$$\widehat{\mathbf{H}}(n) = \sum_{m=1}^{N_p} \mathbf{Y}^{(p)}(m) p(n,m) \tag{2.32}$$

with

$$p(n,m) = \mathbf{P}_{(n,m)} = \sum_{k=1}^{L} \left[\mathbf{D}^{-1}\mathbf{B}^H\right]_{(k,m)} e^{-j2\pi(n-1)(k-1)/N} \tag{2.33}$$

$$= \sum_{k=1}^{L} \mathbf{B}^{\dagger}_{(k,m)} e^{-j2\pi(n-1)(k-1)/N} \tag{2.34}$$

For the uniform case we know from the results of (2.30) that $\mathbf{D}^{-1} = \frac{1}{N_p}\mathbf{I}_L$. Hence $\mathbf{B}^{\dagger} = \frac{1}{N_p}\mathbf{B}^H$ and (2.33) can be further simplified using the definition of $\mathbf{B}$ in (2.2)

$$\begin{aligned}
p(n,m) &= \frac{1}{N_p}\sum_{k=1}^{L} \mathbf{B}^{*}_{(m,k)} e^{-j2\pi(n-1)(k-1)/N} \\
&= \frac{1}{N_p}\sum_{k=1}^{L} e^{j2\pi(\mathbf{n_p}(m)-1)(k-1)/N} \times e^{-j2\pi(n-1)(k-1)/N} \\
&= \frac{1}{N_p}\sum_{k=1}^{L} e^{j2\pi(\mathbf{n_p}(m)-n)(k-1)/N} \\
&= \frac{1}{N_p}\left(\frac{e^{\frac{j2\pi L}{N}(\mathbf{n_p}(m)-n)} - 1}{e^{\frac{j2\pi}{N}(\mathbf{n_p}(m)-n)} - 1}\right) \\
&= \frac{1}{N_p}\left(\frac{e^{\frac{j2\pi L}{N}(\frac{N}{N_p}(m-1)-n+1)} - 1}{e^{\frac{j2\pi}{N}(\frac{N}{N_p}(m-1)-n+1)} - 1}\right) \qquad 1 \le n \le N, \quad 1 \le m \le L,
\end{aligned} \tag{2.35}$$

where for uniform allocation we used $\mathbf{n_p}(m) = \frac{N}{N_p}(m-1)+1$. Note that for the $N_p = L$ case, from the definition of $\mathbf{P}$ in (2.8) we have $\mathbf{P} = \mathbf{G}\mathbf{D}^{-1}\mathbf{B}^H = \mathbf{G}\mathbf{B}^{-1}$ hence using the fact that $\mathbf{B}^{-1} = \frac{1}{N_P}\mathbf{B}^H$ we have

$$\mathbf{P}^H\mathbf{P} = \mathbf{B}^{-H}\mathbf{G}^H\mathbf{G}\mathbf{B}^{-1} \quad (2.36)$$

$$= N\mathbf{B}^{-H}\mathbf{B}^{-1} \quad (2.37)$$

$$= \left(\frac{N}{N_p}\right)\mathbf{I}_L \quad (2.38)$$

which shows that the columns of the interpolating matrix, which form a total of L *interpolating functions*, are orthogonal. Fig. 2.2 shows an illustration of these functions for this case.

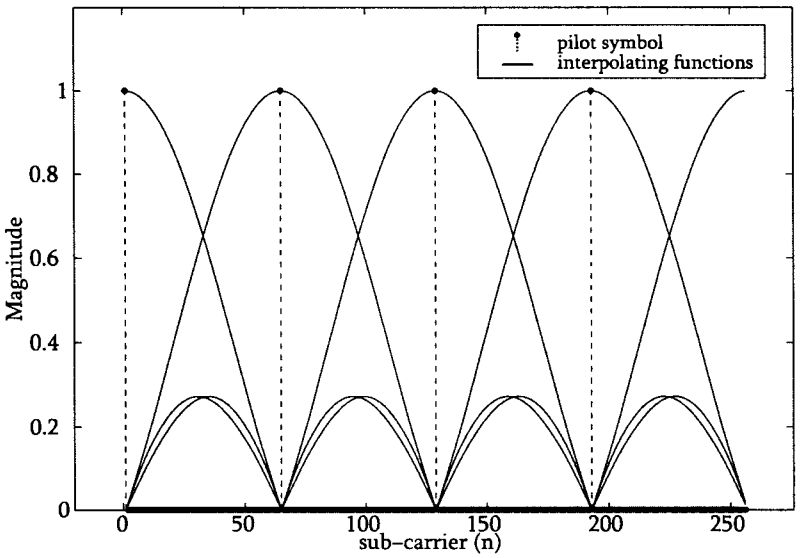

Figure 2.2: Interpolating functions for $N = 256$, $L = 4$ and $N_p = L = 4$.

Simulation

To assess the performance of an OFDM system using uniform PSACE we simulated the baseband system models of Fig. 1.1 and Fig. 1.2. At each time instance, N baseband symbols are randomly drawn from a MPSK constellation space denoted by $\mathcal{M}$. These symbols form $\mathbf{S}_k = diag\{s_k(1), s_k(2), ..., s_k(N)\}$ and the underlying OFDM system equation of (1.29). As before, the subscript k represents discrete-time and is an indication of the k^{th} OFDM symbol. A group of K subsecutive OFDM symbols will be called an OFDM frame.

Each OFDM symbol in a frame experiences a different channel $\mathbf{H}_k$ as it propagates through the wireless environment. However for simplicity and to avoid inter-channel interference, the channel is assumed to be constant at each time k (quasi-static). The channel model is essentially that of which is explained in Section 1.1. For consistency, the power decay factor of the power delay profile $p(\tau, \zeta)$ is set to $\zeta = 2$ here and in all subsequent simulations[4]. For accuracy, the channel itself is simulated using a computationally efficient modification of the "Smith method" for generating correlated Rayleigh random variables as explained in [41]. Fig. 2.3 shows an example of the time evolution of a 4-tap frequency selective channel used in the simulations. The signal-to-noise ratio (SNR) at each sub-carrier is from (1.29) and $\mathbf{R}_k(n) = \mathbf{S}_k(n)\mathbf{H}_k(n) + \mathbf{n}_k(n)$ given by

$$\mathsf{SNR} \stackrel{\text{def}}{=} \mathsf{SNR}_k(n) = \frac{\mathbb{E}\left\{|\mathbf{S}_k(n)\mathbf{H}_k(n)|^2\right\}}{\mathbb{E}\left\{|\mathbf{n}_k(n)|^2\right\}} = \frac{1}{\sigma_n^2}, \tag{2.39}$$

since $\mathbb{E}\left\{|\mathbf{S}_k(n)|^2\right\} = 1$, $\mathbb{E}\left\{|\mathbf{H}_k(n)|^2\right\} = 1$. Fig. 2.4 shows the performance of PSACE-OFDM using uniform pilot allocation for various number of pilot symbols N_p. The figure on the top left shows the channel MSE as a function of SNR from both simulation and from (2.31). The plot on the top right shows the average symbol error rate (SER) performance as a function of SNR[5]. The channel is assumed to be constant within the OFDM symbol and changing independently from one OFDM symbol to the next, i.e. $\mathbb{E}\left\{\mathbf{H}_k^H \mathbf{H}_{k+1}\right\} = 0$. In terms of the parameters of Fig. 2.1, this would account to letting $P_t = 0$ meaning that *each* OFDM symbol has an arrangement of comb-type pilots. QPSK modulation is used for both plots and $\geq 10^4$ channel realizations are simulated per SNR value. The total number of sub-carriers is $N = 512$ and the channel length is constant at $L = 4$ taps. From the SER curves we see roughly a 1.5 dB improvement in SNR by using $N_p = 2L$ pilots instead of $N_p = L$ and another 1 dB improvement in using $N_p = 4L$. Interestingly while the MSE curves show a steady decrease in MSE as N_p increases, the SER curves show diminishing returns. This is of course due to the fact that the SER is a complicated non-linear function of N_p and SNR while the MSE is, from (2.31), inversely related to both the SNR and to N_p. The bottom figures are examples of the resulting MLI output for $N_p = L = 4$.

[4]Such a tapered power profile is more consistent with indoor environments such as the WLAN channel model.

[5]Here, by symbol we mean individual symbols drawn from the constellation $\mathcal{M}$, i.e. $s_k(n)$, and not the entire OFDM symbol, i.e. $\mathbf{S}_k$. The latter is more commonly referred to as vector error rate.

2.2 Chapter Summary

We gave a complete analysis of maximum likelihood interpolation for channel estimation for OFDM systems. We presented a general mathematical model for the interpolation process and analyzed the case of equi-spaced pilots in detail. We derived the channel MSE and showed that it coincides with the CRLB for unbiased estimators. We showed how the entire interpolation process can be simply modeled by a projection matrix denoted by $\mathbf{P}$ operating on the pilot locations. The major results in these chapter are summarized below, including some addition comments:

- **The Method:** MLI is a simple yet effective means of estimating an L-tap frequency selective channel. The maximum likelihood solution is essentially the solution of a set of linear equations in the channel gains $\mathbf{h}_k(l)$ for $1 \leq l \leq L$. By allocating a total of N_p pilots for each OFDM symbol, we arrive at a total of N_p equations in L unknowns[6]. The solution is found in the *linear least mean square sense* by using a projection matrix (or interpolating matrix) and the pseudo-inverse.

- **Number of Pilots:** The MLI computes a matrix inverse (or pseudo-inverse) of a full rank matrix which requirs the number of equations to be greater than the number of unknowns, i.e. $N_p \geq L$.

- **Position of Pilots:** If CSI is not made available to the transmitter the optimum pilot positions in terms of minimizing the channel MSE is uniformly spaced pilot allocation (equi-spaced pilots).

- **MSE:** For optimum pilots the MSE is uniform across the sub-carriers and is equal to $\sigma_n^2\left(\frac{L}{N_p}\right)$, where σ_n^2 is the noise variance on each sub-carrier. Also, MLI with the lowest possible number of pilots, $N_p = L$, incurs a 3 dB penalty in SNR due to interpolation and AWGN.

- **Performance:** Clearly at any SNR level, both the SER and the channel MSE decrease as the number of uniform pilots increases. What's important to note however is that the SER curves show diminishing returns at higher N_p values while the MSE curves do not.

[6]More precisely $2N_p$ equations in $2L$ if we consider the fact that the channel gains are complex.

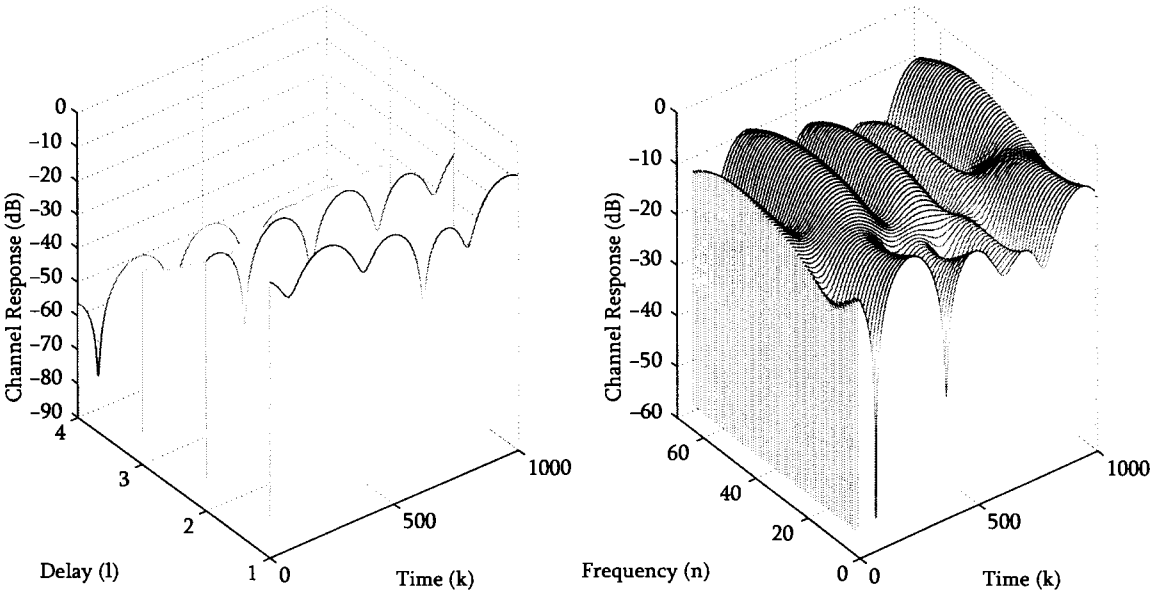

Figure 2.3: (left) An example of 1000 realizations of a wireless channel impulse response (time-delay grid) with $L = 4$, $N = 64$ and normalized Doppler frequency of $F_d = 0.002$. (right) Corresponding time-frequency response.

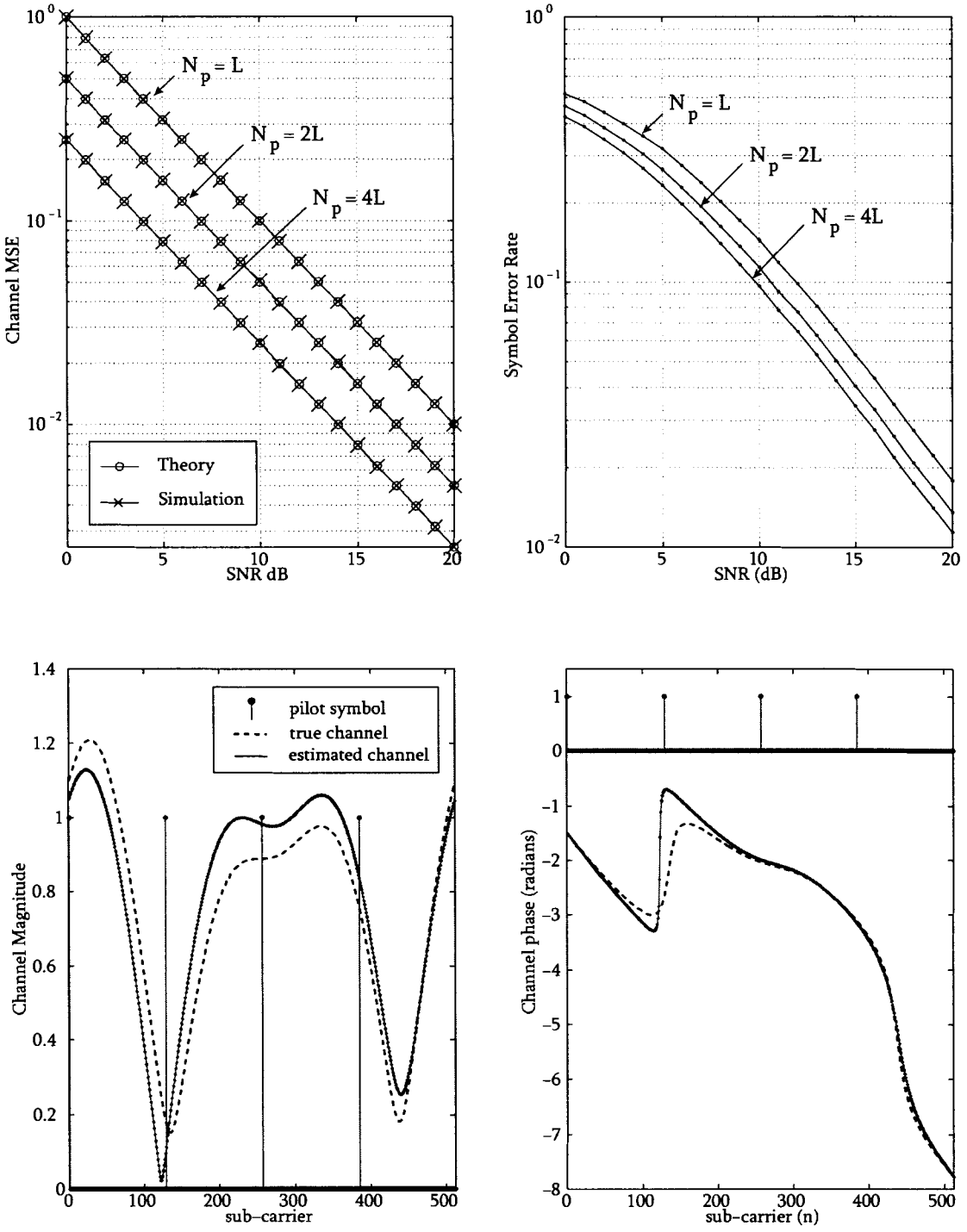

Figure 2.4: Performance of uniform pilot allocation for various N_p. (top left) Channel MSE vs. SNR. (top right) Average SER vs. SNR. (bottom left) Channel magnitude with $N_p = L = 4$. (botton right) Channel phase (unwrapped radians).

Chapter 3

Non Uniform Pilot Allocation

The previous chapter discussed the details of OFDM channel estimation with emphasis on open-loop systems where CSI is not available at the transmitter. We now consider a scenario in which the transmitter has some form of knowledge regarding the multi-path channel gains. The actual method in which this information is acquired at the transmitter depends on the type of duplexing we assume on the wireless link; namely time division duplexing (TDD) or frequency division duplexing (FDD). Each method has its own way of acquiring CSIT. The objective here is to somehow obtain the downlink channel prior to transmission.

- TDD : A situation where the uplink (UL)[1] and downlink (DL)[2] are co-located in frequency. Channel feedback conceptually does not apply here since assuming a relatively slow channel, i.e. low Doppler frequency, the DL channel may simply be estimated using PSAM on the UL.

- FDD : A situation where the UL and DL simultaneously transmit in time but are separated in frequency. To avoid co-channel interference a frequency guard interval is often inserted between the respective passbands. With the uplink-downlink separation being greater than the coherence bandwidth of the channel, the respective channels in the UL and DL will fade independently and the concept of *channel feedback* can be applied. Here, PSAM is **only** applied in the downlink direction. Once the channel is estimated at the receiver, it is relayed back to the transmitter through a dedicated

[1] Mobile to base station.

[2] Base station to mobile.

feedback bandwidth[3]. Again assuming a slow channel, with instantaneous error-free feedback, the transmitter proceeds to use the relayed information as the CSI for the next signaling period.

It is well known that with CSIT the performance of OFDM can be optimized in terms of symbol error rate (SER) or average capacity by power loading on the sub-carriers [23], [17]. Power loading techniques rely on the general understanding that if the transmitter has a-priori knowledge regarding the channel (channel state information), then it may pre-equalize the transmitted symbols to combat the adverse effects of the channel. The transmitter may for instance allocate more power to the sub-carrier experiencing deep fades. The authors in [17] derive the optimum (and sub-optimum) power loading algorithm for minimizing the average OFDM symbol error rate while the authors in [23] provide further insight into the feedback requirements of such a technique. In all cases however, it is assumed that the channel is either completely known at the receiver and transmitter without error or that the channel is estimated using PSAM and uniform pilot locations throughout the sub-carriers. As we illustrated in the previous chapter, for OFDM systems *without channel information feedback*, the optimal pilot tone allocation is equal-power uniformly located pilots (see [30] and [28] for similar proofs). For OFDM systems *with channel information feedback*, on the other hand, the optimal pilot tone allocation is optimum power loading on the sub-carriers while the pilot locations for the this case has yet to be investigated in literature and constitutes the main contribution of this chapter.

In this chapter we present an alternative approach to power loading for closed-loop OFDM systems *with channel information feedback* where instead of changing the actual sub-carrier powers we propose the use of non-uniformly located pilot tones with equal-power. Similar to conventional power loading, the pilot locations are allocated in an attempt to increase the average signal-to-noise ratio at the receiver output. We begin by analyzing the received power in a baseband OFDM system.

3.1 Average SNR Maximization

Since the pilots are not information carrying symbols, the average symbol error rate (SER) may be calculated only at the data locations. We propose a pilot allocation strategy based

[3]In addition to it being used for coherent detection

on the maximization of the average signal-to-noise ratio. Denoting the *data* locations by $\langle d \rangle$, the average output signal-to-noise ratio as defined on the *data* symbols at time k is from (2.10)

$$\overline{\mathsf{SNR}}_k = \frac{tr\left\{\mathbb{E}\left\{\mathbf{H}_k^{\langle d\rangle}\left(\mathbf{H}_k^{\langle d\rangle}\right)^H\right\}\right\}}{tr\left\{\mathbb{E}\left\{\mathbf{v}_k^{\langle d\rangle}\left(\mathbf{v}_k^{\langle d\rangle}\right)^H\right\}\right\} + tr\left\{\mathbb{E}\left\{\mathbf{n}_k^{\langle d\rangle}\left(\mathbf{n}_k^{\langle d\rangle}\right)^H\right\}\right\}}. \tag{3.1}$$

Note that the pilot index vector $\mathbf{n_p}$ is chosen based upon the channel estimate at time $k-1$ and used for transmission at time k. Therefore the receiver has the additional task of forward predicting the channel. The reliability, hence optimality of $\mathbf{n_p}$, depends on the accuracy of this prediction or on how fast the channel fluctuates in time. For small Doppler, the channel time variations is slow, i.e. $f_d T_s \ll 1$, so the channel at $k-1$ may be used at the receiver to attain a reliable prediction of the channel at k. Since the focus of this work is on pilot allocation, full Doppler analysis is not attempted here and we simply use $\widehat{\mathbf{H}}_k \cong \widehat{\mathbf{H}}_{k-1}$ as the channel prediction. Faster Doppler will obviously degradate the performance of the system as can be seen in Fig. 3.7. More importantly, such channel prediction eliminates the expectation operation over the channel and allows for the simplification of the numerator of (3.1)

$$\overline{\mathsf{SNR}}_k = \frac{\left(\widehat{\mathbf{H}}_k^{\langle d\rangle}\right)^H \widehat{\mathbf{H}}_k^{\langle d\rangle}}{tr\left\{\mathbb{E}\left\{\mathbf{v}_k^{\langle d\rangle}\left(\mathbf{v}_k^{\langle d\rangle}\right)^H\right\}\right\} + tr\left\{\mathbb{E}\left\{\mathbf{n}_k^{\langle d\rangle}\left(\mathbf{n}_k^{\langle d\rangle}\right)^H\right\}\right\}}. \tag{3.2}$$

This expression is simplified further in the Appendix, rewriting the final result (B.6)

$$\overline{\mathsf{SNR}}_k = \frac{\left(\widehat{\mathbf{H}}_k^{\langle d\rangle}\right)^H \widehat{\mathbf{H}}_k^{\langle d\rangle}}{\sigma_n^2\left(\|\mathbf{P}_k\|_F^2 - L\right) + (N - N_p)\,\sigma_n^2} \tag{3.3}$$

Again since $\widehat{\mathbf{H}}_k = \widehat{\mathbf{H}}_{k-1}$, (3.3) may be written as

$$\overline{\mathsf{SNR}}_k = \overline{\mathsf{SNR}}_{CCI}\frac{\left(\widehat{\mathbf{H}}_{k-1}^{\langle d\rangle}\right)^H \widehat{\mathbf{H}}_{k-1}^{\langle d\rangle}}{\left(\|\mathbf{P}_k\|_F^2 - L\right) + (N - N_p)} \tag{3.4}$$

where $\overline{\mathsf{SNR}}_{CCI} = 1/\sigma_n^2$ is the SNR for a system with *complete channel information* (CCI) i.e. no interpolation. We define the *average SNR gain* as

$$\Omega_k(\alpha,\beta) = \frac{\overline{\mathsf{SNR}}_k}{\overline{\mathsf{SNR}}_{CCI}} = \frac{\alpha\left(\widehat{\mathbf{H}}_{k-1}, \mathbf{n_d}\right)}{\beta(\mathbf{n_d})} \tag{3.5}$$

where α and β are scalar functions. Note that α is a function of the CFR as well as the vector $\mathbf{n_d}$. β on the other hand is only a function of $\mathbf{n_d}$[4].

$$\alpha\left(\widehat{\mathbf{H}}_{k-1}, \mathbf{n_d}\right) = \left(\widehat{\mathbf{H}}_{k-1}^{\langle d\rangle}\right)^H \widehat{\mathbf{H}}_{k-1}^{\langle d\rangle}$$

$$\beta(\mathbf{n_d}) = \|\mathbf{P}_k\|_F^2 + N - N_p - L.$$

Note also that the interpolating matrix is a direct function of the pilot locations, i.e. $\mathbf{P}_k = \mathbf{P}\left(\mathbf{n_d}\right)_k$ and that α and β can readily be rewritten as functions of $\mathbf{n_p}$

$$\alpha\left(\widehat{\mathbf{H}}_{k-1}, \mathbf{n_p}\right) = \widehat{\mathbf{H}}_{k-1}^H \widehat{\mathbf{H}}_{k-1} - \sum_{i=1}^{N_p} \left|\widehat{\mathbf{H}}_{k-1}(\mathbf{n_p}(i))\right|^2 \tag{3.6}$$

$$\beta(\mathbf{n_p}) = \left\|\mathbf{P}\left(\mathbf{n_p}\right)_k\right\|_F^2 + N - N_p - L \tag{3.7}$$

and $\beta(\mathbf{n_p})$ can be further simplified for the MLI explained in the previous section. Using (2.8) we have

$$\begin{aligned} \|\mathbf{P}\|_F^2 &= tr\left\{\mathbf{P}^H\mathbf{P}\right\} \\ &= tr\left\{\mathbf{B}\mathbf{D}^{-H}\mathbf{G}^H\mathbf{G}\mathbf{D}^{-1}\mathbf{B}^H\right\} \\ &= tr\left\{\mathbf{D}^{-1}\mathbf{B}^H\mathbf{B}\mathbf{D}^{-H}\mathbf{G}^H\mathbf{G}\right\} \\ &= Ntr\left\{\mathbf{D}^{-1}\right\} \end{aligned} \tag{3.8}$$

where we have used the trace identity $tr\{AB\} = tr\{BA\}$, the fact that $\mathbf{G}^H\mathbf{G} = N\mathbf{I}_L$, and the definition of $\mathbf{D}$ in (2.5). Substituting this result into (3.7)

$$\beta(\mathbf{n_p}) = Ntr\left\{\mathbf{D}(\mathbf{n_p})_k^{-1}\right\} + (N - N_p - L). \tag{3.9}$$

Since the average signal-to-noise ratio is a direct function of the pilot locations or $\mathbf{n_p}$, we seek to maximize (3.3), which is equivalent to maximizing the average SNR gain $\Omega_k(\alpha,\beta)$ defined in (3.5) since $\overline{\mathsf{SNR}}_{CCI}$ is not a function of $\mathbf{n_p}$. This optimization is summarized below:

[4]Since $\mathbf{n_d}$ is a function of the CFR, β is also *implicitly* a function of the CFR.

$$\boxed{\begin{array}{c} \text{optimize : } \mathbf{n_p}^{opt} = \arg \max_{\mathbf{n_p} \in \mathcal{I}_p} \left[\Omega_k(\alpha, \beta) = \frac{\alpha}{\beta} \right] \\ \text{subject to : } tr\left\{ \mathbf{D}(\mathbf{n_p})_k \right\} = N_p L \\ \text{where :} \\ \alpha = \widehat{\mathbf{H}}_{k-1}^H \widehat{\mathbf{H}}_{k-1} - \sum_{i=1}^{N_p} \left| \widehat{\mathbf{H}}_{k-1}(\mathbf{n_p}(i)) \right|^2 \\ \beta = N tr\left\{ \mathbf{D}(\mathbf{n_p})_k^{-1} \right\} + (N - N_p - L). \end{array}} \tag{3.10}$$

The expressions in (3.10) present a nonlinear optimization problem in $\mathbf{n_p}$. The solution set $\mathcal{I}_p$, is however finite and the optimum solution $\mathbf{n_p}^{opt}$ may be found by exhaustive search through this set. The cardinality of this set is $|\mathcal{I}_p| = \binom{N}{N_p}$ which can grow very large even for moderate values of N. We therefore seek suboptimal solutions with acceptable performance losses in terms of SNR.

3.1.1 The Case of No Feedback

Cioffi *et.al* [28] prove that for a transmitter without CSI, the optimum pilot allocation is equal power loading and uniform pilot distribution across the sub-carriers. In the context of our problem the SNR maximizing pilot allocation is given by (3.10) for the case where $\widehat{\mathbf{H}}_{k-1}$ is known at the receiver and used to search for $\mathbf{n_p}^{opt}$ which is subsequently relayed to the transmitter. Without CSI, however, $\widehat{\mathbf{H}}_{k-1}$ is random to the transmitter and must be averaged out of α in (3.6). So $\mathbb{E}\left\{ \alpha\left(\widehat{\mathbf{H}}_{k-1}, \mathbf{n_p} \right) \right\} = N - N_p$, which is not a function of $\mathbf{n_p}$. Therefore (3.10) is maximized by minimizing $\beta(\mathbf{n_p})$ or equivalently $tr\left\{ \mathbf{D}(\mathbf{n_p})_k^{-1} \right\}$ and

$$\mathbf{n_p}^{opt} = \arg \min_{\mathbf{n_p} \in \mathcal{I}_p} tr\left\{ \mathbf{D}(\mathbf{n_p})_k^{-1} \right\} = \arg \min_{\mathbf{n_p} \in \mathcal{I}_p} \sum_{k=1}^{L} \lambda_k^{-1} \tag{3.11}$$

where λ_k is the k^{th} eigenvalue of $\mathbf{D}$ with the constraint $\sum_{k=1}^{L} \lambda_k = N_p L$. The optimum eigenvalues can easily be shown to be all equal to N_p which means $\mathbf{D}(\mathbf{n_p}^{opt}) = N_p \mathbf{I}_L$ and $\mathbf{n_p}$ is uniform, i.e. $\mathbf{n_p}(i+1) - \mathbf{n_p}(i) = \lfloor N/N_p \rfloor$, $1 \leq i \leq N-1$, where by convention we set $\mathbf{n_p}(1) = 1$. Substituting this result into (3.10) and simplifying, we reach

$$\Omega_k^{max} = \Omega_k^{uniform} = \frac{N - N_p}{NL/N_p + N - N_p - L} \tag{3.12}$$

Substituting $N_p = L$ in (3.12) we get $\Omega_k^{max} = \Omega_k^{uniform} = 1/2$. This shows that a uniformly pilot-allocated open-loop OFDM system using coherent ML channel interpolation with a minimum number of pilots tones, incurs, on average, a 3 dB SNR loss compared to a system with complete channel information. Such an observation is also confirmed by computer simulation in Section IV (see Fig. 3.2).

3.1.2 The Case of Feedback - Suboptimal Solutions

The optimum gain factors α^{opt} and β^{opt} are coupled in (3.10) through a complicated expression in $\mathbf{n_p}$. We propose a decoupled sub-optimum strategy and two solutions where α and β are disjointly optimized to maximize Ω. The problem is stated as

Problem:
Find the optimum pilot index vector in terms of maximizing the average SNR gain Ω^{max} given that $\Omega^{max} = \Omega(\alpha^{opt}, \beta^{opt}) = \Omega(\alpha^{max}, \beta^{min})$.

Proposed Solution #1 - Decoupled Equal Power Optimization (DEPO)

Given a total of N_p pilot symbols suppose N_p^α pilots are allocated to the task of optimizing α and N_p^β are used for β so that $N_p = N_p^\alpha + N_p^\beta$.

i) Optimum α: From (3.10) and without regard to β:

$$\begin{aligned} \alpha^{opt} &= \alpha^{max} \\ &= \arg\max_{\mathbf{n_p} \in \mathcal{I}_p} \left[\widehat{\mathbf{H}}_{k-1}^H \widehat{\mathbf{H}}_{k-1} - \sum_{i=1}^{N_p} \left| \widehat{\mathbf{H}}_{k-1}(\mathbf{n_p}(i)) \right|^2 \right] \\ &= \arg\min_{\mathbf{n_p} \in \mathcal{I}_p} \sum_{i=1}^{N_p} \left| \widehat{\mathbf{H}}_{k-1}(\mathbf{n_p}(i)) \right|^2 \end{aligned} \tag{3.13}$$

which has the obvious solution of sorting the gain values (from smallest to greatest) $\left|\widehat{\mathbf{H}}_{k-1}\right|$, and selecting the first N_p^α index values. Denote such a solution as $\mathbf{n_p}^{opt(\alpha)}$.

ii) Optimum β: From (3.10) and using the result of (3.11), $\beta^{opt} = \beta^{min}$ at:

$$\mathbf{n_p}^{opt(\beta)} = \mathbf{n_p}^{uniform} \tag{3.14}$$

which is the uniform pilot allocation scheme (the case of no feedback).

iii) (Sub)optimum $\mathbf{n_p}$: To reach a compromise between $\mathbf{n_p}^{opt(\alpha)}$ and $\mathbf{n_p}^{opt(\beta)}$ set

$$\mathbf{n_p}^{subopt} = \left[(\mathbf{n_p}^{opt(\alpha)})^T \ (\mathbf{n_p}^{opt(\beta)})^T\right]^T \tag{3.15}$$

which is a length N_p vector that is comprised of both solutions (assuming no overlap). The values assigned to N_p^α and N_p^β may also be viewed as *optimizing weighting parameters.* Setting $N_p^\alpha = 0$, for instance, would result in the optimization process weighing totally in favor of β and vis versa. Fig. 3.8 illustrates an example of this pilot allocation for a length 4 channel at $1/\sigma_n^2 = 0$ dB and the curves in Fig. 3.5 show how $N_p^\alpha = N_p - L$ and $N_p^\beta = L$ minimize the error rate performance. *Note that in this case there is no power loading on the sub-carriers. The power is instead effectively distributed*[5] *by distributing the pilot locations.*

Proposed Solution #2 - Decoupled on/off Power Optimization (DOPO)

Consider again the previous proposed solution (DEPO) along with Fig. 3.5 and Fig. 3.8. This solution allocates L uniform pilots to N_p^β and the rest ($N_p^\alpha = N_p - L$) to the (locally) minimum channel gain locations. Note, however, that the entire set of pilots is ultimately used for maximum likelihood channel interpolation. Intuitively, since the pilot set N_p^α is occupying the channel minimum, hence fade locations, the better parts of the channel (higher gain) will be available for data transmission. On the other hand, the pilot set N_p^α is also consuming transmitter power that may have otherwise been used for data transmission. Note that there is an underlying trade-off present here since the mere presence of more pilots as a result of N_p^α is expected to yield better channel interpolation i.e. lower β values. To investigate this trade-off and to complement the DEPO solution, suppose that the pilot set N_p^α is completely eliminated and its power (which on average is equal to $(N_p - L)$ is evenly distributed amongst the rest of the sub-carriers. In other words the transmitter sends

[5] *Equalized in a sense.*

Table 3.1: Summary of simulation parameters and notations

Active sub-carriers	N
Channel length	L
Number of pilot-tones	N_p
Nyquist pilot rate	$N_p = L$
Normalized Doppler frequency	$F_d = f_d N T_s$
Average SNR ($\overline{\text{SNR}}$)	$1/\sigma_n^2$
Average SNR gain	Ω
Decoupled Equal Power Optimization	DEPO
Decoupled on/off Power Optimization	DOPO

L uniform pilots along with $(N_p - L)$ null symbols at $\mathbf{n_p}^{opt(\alpha)}$. We call this solution the decoupled on/off power optimization (DOPO) technique.

3.2 Numerical Results

Computer simulations are used to evaluate the average symbol error rate performance of the proposed feedback scheme in Fig. 3.1. MPSK modulation is used throughout the simulations with an L-tap time-dispersive channel model given by (1.1). The normalized Doppler frequency is $F_d = f_d N T_s$, where f_d is the channel Doppler frequency and T_s is the symbol rate. Simulations are carried out for a range of per carrier signal-to-noise ratios $(1/\sigma_n^2)$ and for various Doppler frequencies. For simulation purposes, Table. 3.1 summarizes the various notations used in this paper.

3.2.1 Determination of N_p^α and N_p^β (optimizing weighting parameters)

The value assigned to N_p^β can be determined by inspection. Note that since the maximum likelihood interpolator of (2.4) is essentially the solution of an *overdetermined* set of linear equations it can only be designed at or above the Nyquist rate ($N_p \geq L$). Simple inspection of (3.12) reveals that the decoupled β solution for uniform allocation has a global minimum at the uniform allocation vector. A logical choice for N_p^β in (3.15) is therefore L, which gives $N_p^\alpha = N_p - N_p^\beta = 12$ for $N_p = 16$. This observation is confirmed in Fig. 3.5, where

Table 3.2: Gain values for $N_p = 2L$

Average	DEPO	DOPO
α	59.55	63.52
β	100.2	116.0
Ω	0.5948	0.5477

the average SER of the DEPO solution is plotted as functions of N_p^α for BPSK modulation over $N = 128$ sub-carriers.

3.2.2 Symbol Error Rate

The optimum pilot allocation vector may be attained from the optimization problem of (3.10) by exhaustive search. Fig. 3.3 compares the SER performance of this optimum solution with the proposed solution #1 (DEPO) at $N_p = 2L$. The performance loss is roughly 1 dB at sufficiently low SER. The SER performances of the proposed solutions (DEPO and DOPO) are compared with uniform allocation and with each other in Fig. 3.4. The simulation is run for both $N_p = 2L$ and $N_p = 4L$ with QPSK modulation. The superiority of the DEPO method is an indication that for the DOPO method, the increase in β, more than offsets the increase in SNR gain it achieves through the increase in α. The aggregate effect is a loss in average Ω and also SER performance. The numerical gain values are shown in Table 3.2.

3.2.3 Doppler Frequency

Through computer simulations, the authors in [22] observed that certain non-uniform pilot allocations yield smaller SER at high Doppler frequencies. Fig. 3.6 shows the SER comparison of an OFDM system with uniform pilot allocation (open-loop) and the proposed decoupled solution (DEPO) for BPSK modulation. As the Doppler increases the channel prediction becomes more and more obsolete causing performance degradation for the proposed scheme. The proposed scheme is rather resilient to Doppler, showing about 1 dB loss for a twenty fold increase in Doppler frequency at high SNR with superior performance compared to the uniform allocation. Our simulation of the SER performance of the proposed scheme (DEPO) in Fig. 3.7 confirms this observation where the SER is plotted versus the normalized Doppler frequency.

3.3 Chapter Summary

In this chapter we presented a new perspective on the application of feedback for wireless OFDM systems. While conventional feedback schemes concentrate on equalizing the channel by redistributing the transmitter power through the sub-carrier we proposed the redistribution of pilot tone locations without the necessity of knowing the channel gains at the transmitter. To this end, we derived the average SNR gain at the OFDM receiver and proposed a combinatorial optimization problem in finding the SNR maximizing pilot tone locations. Simulations showed considerable improvements in terms of average error rates compared to open-loop OFDM systems. The feedback scheme also exhibited robust behavior under moderately high Doppler spreads. As noted in section 3.1 the number of feedback bits for the optimum pilot allocation is $\left\lceil \log_2 \binom{N}{N_p} \right\rceil$ bits per feedback slot, which can be large even for moderate values of N and N_p. In the next chapter we use the concept of *limited feedback communications* and the generalized Lloyd algorithm as a means of reducing this overhead.

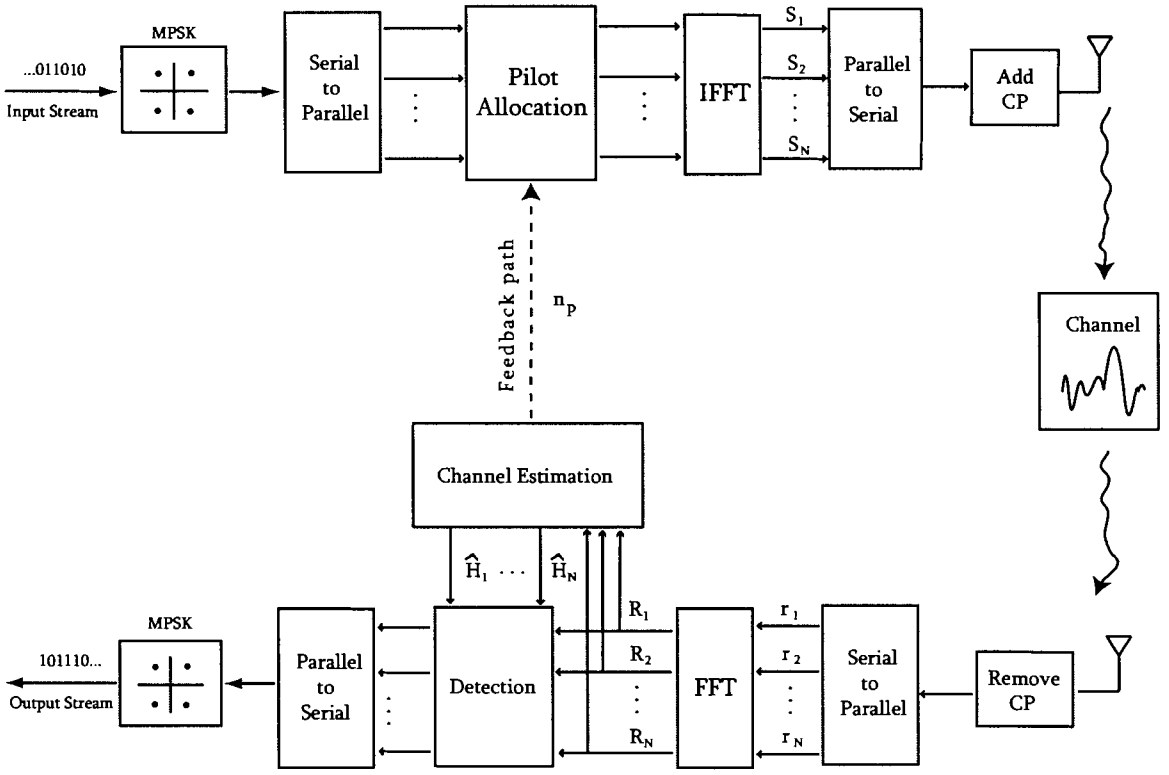

Figure 3.1: Closed-loop OFDM system model

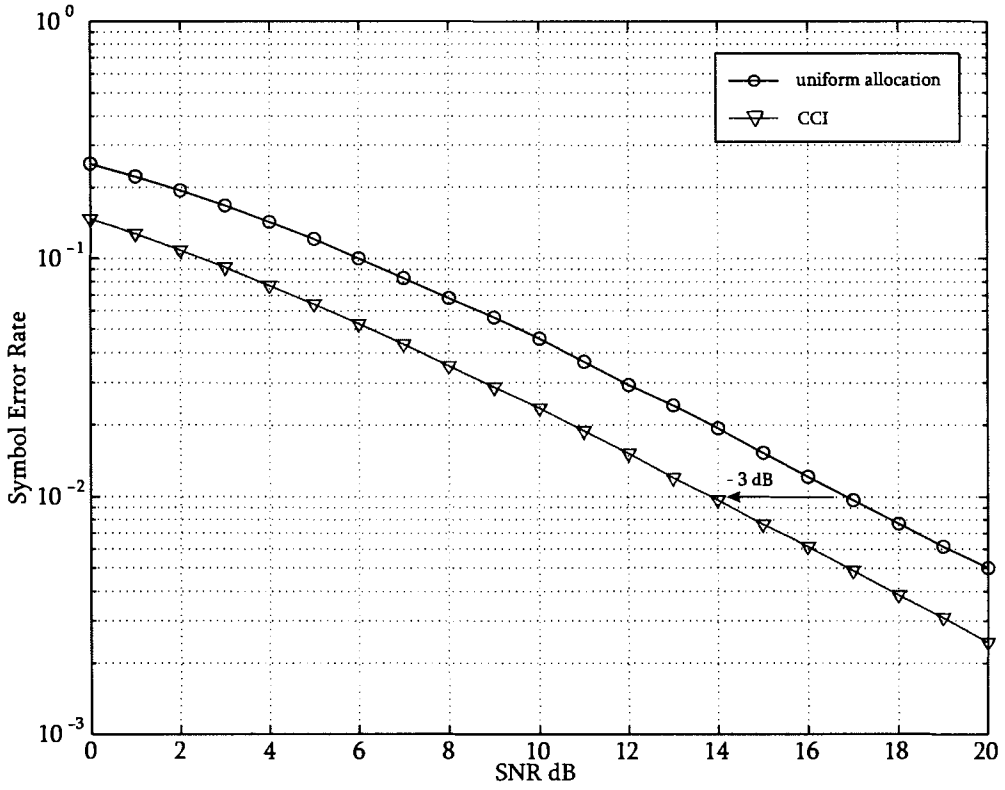

Figure 3.2: Average SER performance of uniform allocation compared to a system with complete channel information (CCI) with $N_p = 4$, $L = 4$, $N = 64$ and BPSK modulation. There is a 3 dB penalty in using uniform pilots at the Nyquist rate ($N_p = L$) and ML interpolation.

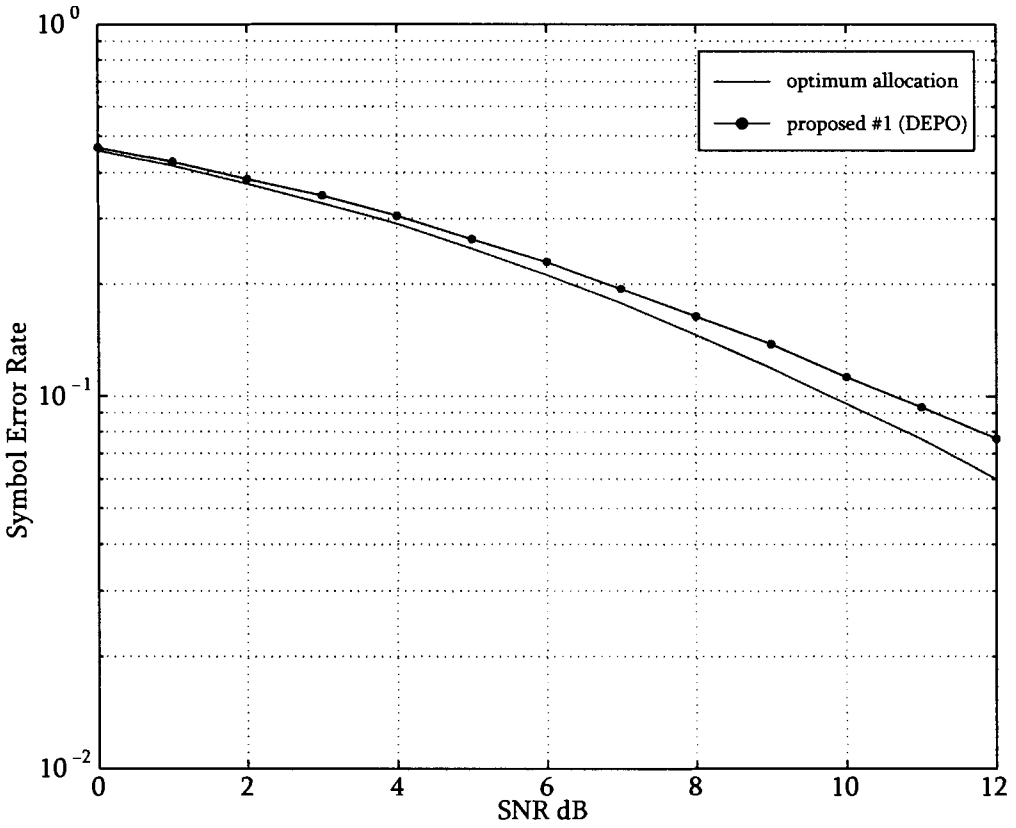

Figure 3.3: Average SER performance of optimum solution of (3.10) and proposed solutions #1 (DEPO) with $N_p = 4$, $L = 2$, $N = 16$ and QPSK modulation. The optimum solution was found by an exhaustive search at each SNR.

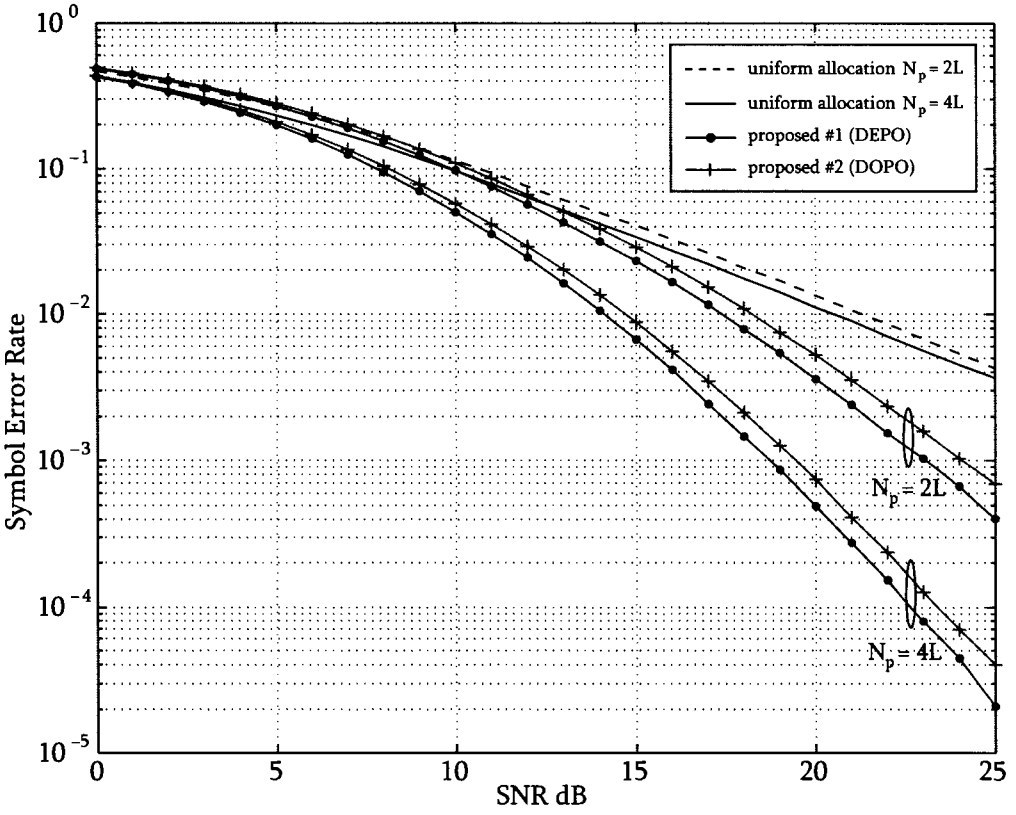

Figure 3.4: Average SER performance of proposed solutions #1 (DEPO) and #2 (DOPO) compared to uniform allocation (case of no feedback) with $L = 4$, $N = 64$ and QPSK modulation.

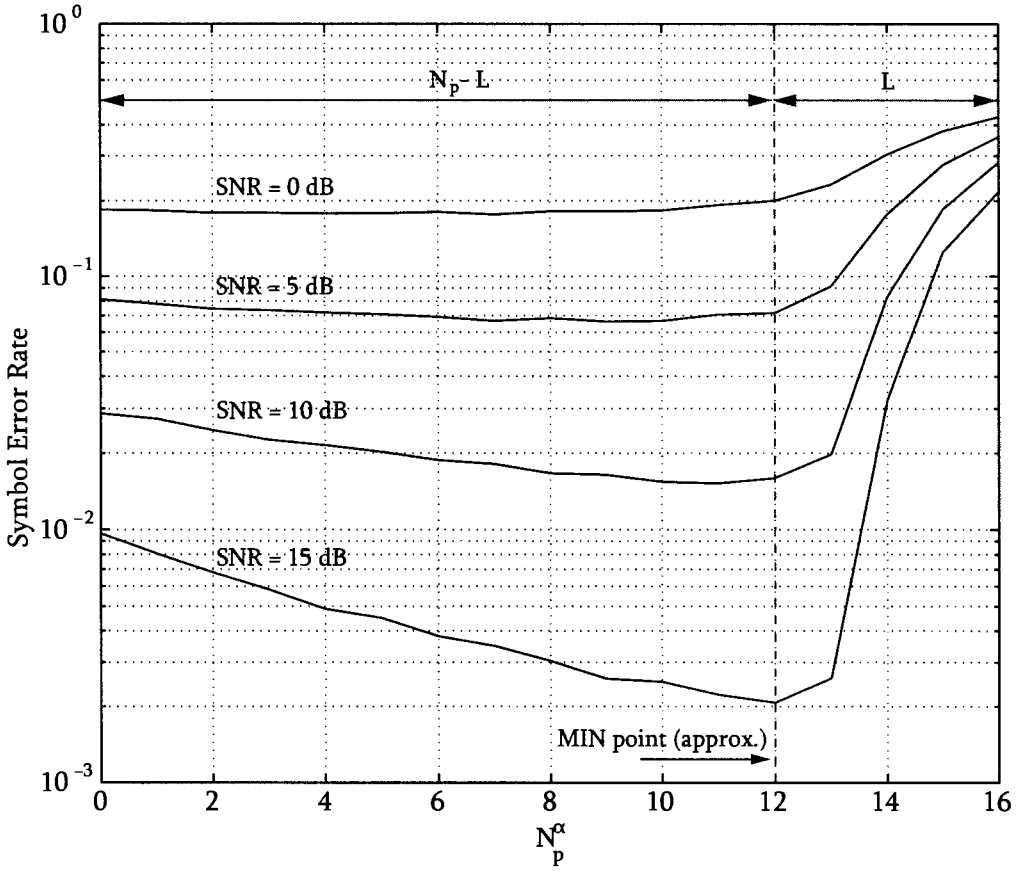

Figure 3.5: Average SER of proposed solution #1 (DEPO) as a function of $N_p^\alpha = N_p - N_p^\beta$ at various SNR with $N_p = 16$, $L = 4$, $N = 128$ and BPSK modulation. The point of minimum SER (point of interest) occurs at $(N_p^\alpha, N_p^\beta) = (12, 4)$

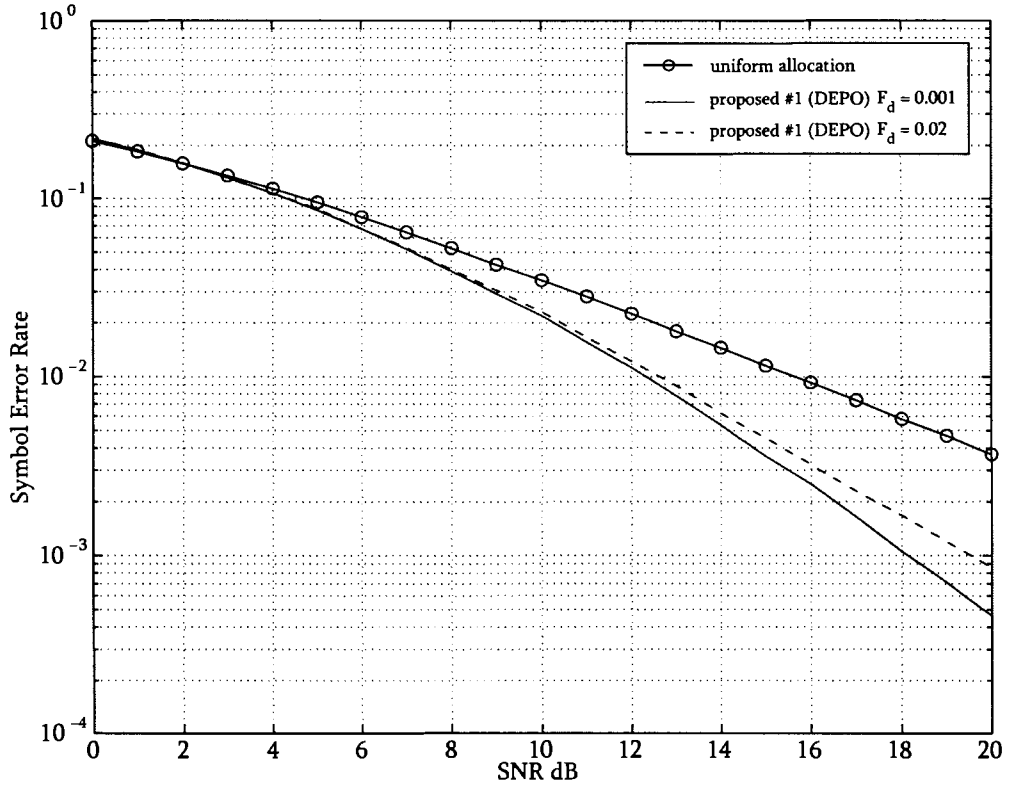

Figure 3.6: Average SER performance of the proposed solution #1 (DEPO) compared to uniform allocation at various (normalized) Doppler frequencies with $N_p = 8$, $L = 4$, $N = 64$ and using BPSK modulation.

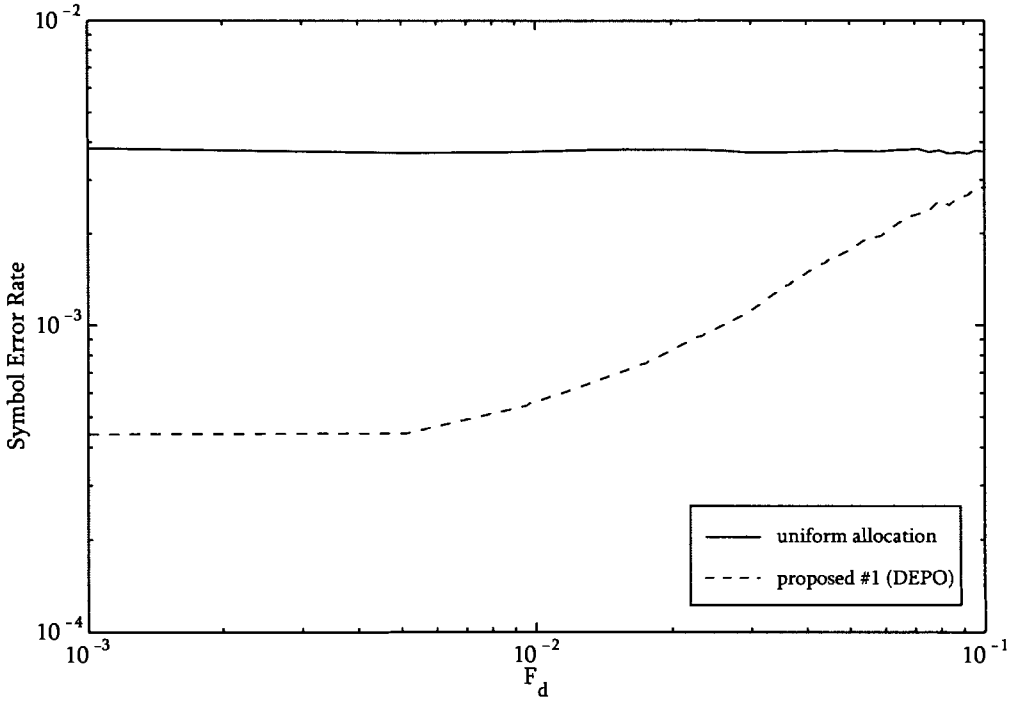

Figure 3.7: Average SER versus Doppler frequency at $1/\sigma_n^2 = 20$ dB with $N_p = 8$, $L = 4$, $N = 64$ and BPSK modulation. Note how the proposed system converges to uniform allocation at high Doppler due to the increasing uncertainty present in the forward channel prediction.

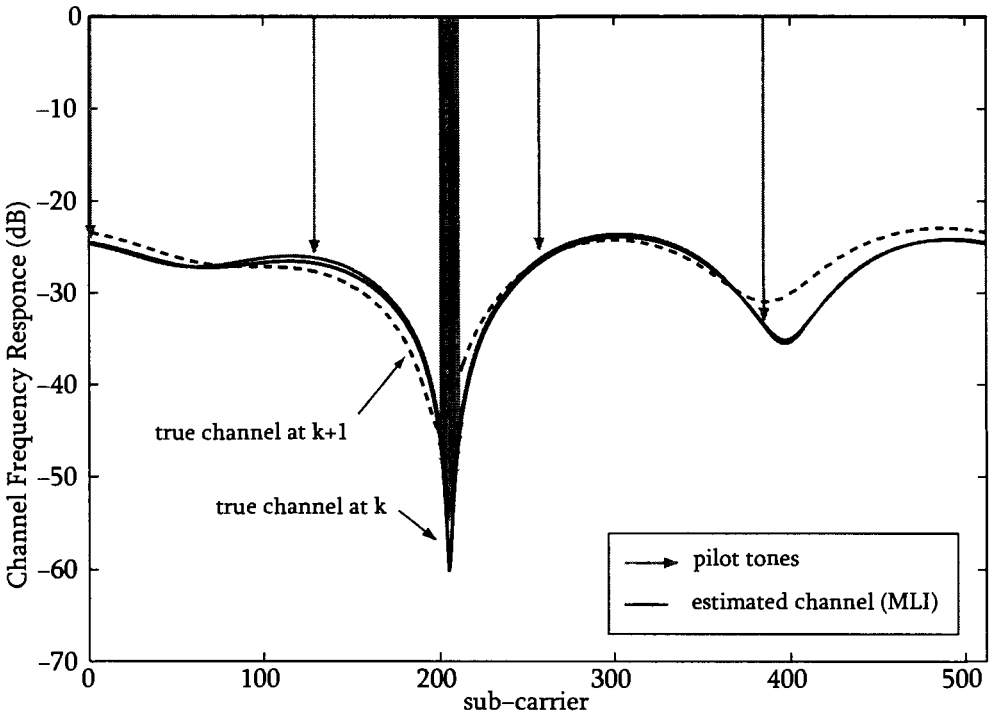

Figure 3.8: Actual (true) CFR and its estimate using the proposed pilot allocation (DEPO) at $1/\sigma_n^2 = 0$ dB with $N_p = 16(N_p^\alpha = 12, N_p^\beta = 4)$, $L = 4$, $N = 512$ and a Doppler of $F_d = 0.1$. The vertical lines indicate pilot tone locations. Notice how 4 pilots ($L = 4$) are allocated uniformly and the rest of the pilots at clustered at the channel minimum point. As the channel fluctuates in time (solid black to dashed black line) the channel minimum moves slightly to the left but is still partially covered by the clustered pilots allocated at time k.

Chapter 4

An Evolutionary Approach to Optimization

"As for a future life, every man must judge for himself between conflicting vague probabilities."

-Charles Darwin

In Chapter 3 we derived an optimization problem for determining the optimum pilot locations for SNR maximization. As noted there, the proposed problem is highly non-linear on the search domain and thus a closed form solution seems out of reach. The search domain of (3.10) is the N_p-dimensional space of positive integer numbers which we will denote here as $\mathbb{Z}_+^{N_p}$. The optimal pilot index vector we seek is some $\mathbf{n_p}^{opt} \in \mathbb{Z}_+^{N_p}$. For instance, $\mathbf{n_p} = \{1, 4, 19, 20\}$ corresponds to the selection of the 1,4,19 and 20^{th} subcarriers as pilot symbols in each OFDM symbol. Clearly we are dealing with a class of optimization problems commonly referred to as *"Integer Programming"* or *"Combinatorial Optimization"* in mathematical literature.

Combinatorial optimization is a diverse branch of applied mathematics that finds a multitude of applications in solving real life problems such as network and graph problems, rule-based scheduling problems, capital budgeting problems, etc. Given their integer nature, the description of combinatorial optimization problems is often an easy task. Solving these problems on the other hand can in general be extremely difficult. The difficulty arises from the fact that unlike linear programming (continuous), for example, whose feasible region is a convex set, in combinatorial problems, one must search a lattice of feasible points. Thus, unlike linear programming where, due to the convexity of the problem, we can exploit

the fact that any local solution is a global optimum, integer programming problems may have many local optima. There are several approaches in solving combinatorial optimization problems and a detailed examination of these techniques is beyond the scope of this work. Instead, here we will utilize a branch of techniques that aims to mimic the natural evolution of biological life referred to as *"Evolutionary Strategies"* [24], [2].

Evolutionary computation algorithms are stochastic optimization problems; they are conveniently presented using the metaphor of natural evolution: a randomly initialized population of individuals evolves following a crude parody of the Darwinian principle of *survival of the fittest* [2]. The genetic algorithm [25] is sub-class of these evolutionary algorithms that due to its simplicity has found a broad range of applications in function optimization. The principal idea behind the genetic algorithm (GA) is simple: the probabilistic simulation of biological evolution. After briefly introducing the method we will see how the GA is directly applicable in solving our optimization problem for non-uniform pilot locations.

4.1 Introduction to the Genetic Algorithm

The idea of the GA appeared first in 1967 in J. D. Bagleys P.hD. thesis "The Behavior of Adaptive Systems Which Employ Genetic and Correlative Algorithms" [3]. The theory and applicability was then strongly influenced by J. H. Holland (and later his students), who can be considered as the pioneers of genetic algorithms. An exhaustive treatment of the subject is beyond the scope of this work but since we are dealing with an integer maximization problem, we will examine the GA as it is applied to a general combinatorial maximization problem of the form:

Find an $x^{opt} \in X$ such that the function f is maximal on the search domain X, where $f : X \to \mathbb{R}$ is an arbitrary real valued (and assumed positive) function such that $f(x^{opt}) = \arg\max_{x \in X} f(x)$.

In the GA, the search space X is seen as a direct analogy to the set of competing individuals in the real world. Each element in this search space is viewed as a biological entity with a certain genome structure. All living beings consist of cells, and each cell contains identical sets of one or more *"chromosomes"*. Confining to this evolutionary terminology,

here, we also call each $x_i \in X$ a *"chromosome"*. Each chromosome has two attributes:

1. An *"evaluation"* which is $f(x_i)$ and
2. A *"fitness"* which is a measure of how the evaluation of x_i compares with the evaluation of all other chromosomes. The fitness is therefore dependent on the evaluations of all other chromosomes in the population and can simply be defined as the fractional quantity $f(x_i)/\overline{f}$, where $\overline{f}$ is the mean evaluation of the entire population[1].

Here we are dealing with a combinatorial problem, hence we can assume without loss of generality that $X \subseteq \mathbb{Z}_+^N$. In the real world, reproduction, adaptation and ultimately evolution is carried out on the level of genetic information (DNA and RNA). This information resides in long genetic sequences or *"genes"* and not in whole integer numbers. Consequently, GAs do not operate directly on the chromosomes in the search space X, but on some coded versions of them. Therefore before the GA can initiate, the search domain X must adequately be coded (discretized) in an *a priori* fashion. Conventionally, for discrete search spaces such as $\mathbb{Z}_+^N$, a simple binary representation suffices. For example the chromosomes $1, 3, 5, ...$ are represented by the strings (or genes) $0001, 0011, 0101, ...$,so on and so forth. The actual evolution of a genetic algorithm involves three stages: selection, crossover[2], and mutation.

- **Selection**: In this stage candidate chromosomes are selected for reproduction. The fitter the chromosome, the more likely it is to be selected.
- **Crossover**: This is the actual reproduction stage between pairs of chromosomes. In real life organisms sexual reproduction is the result of the amalgamation of parent genomes. Similarly here, parent candidate chromosomes are split at a random location[3], swapped and re-merged to create two new offspring[4]. To account for the fact that two parents might not crossover at all, we denote the crossover probability as p_C which is typically close to unity, $p_C \cong 1$. For example the two parents 0110100 and 1110110 may cross over at the second bit location from the right to produce the

[1] Fitness can also be assigned according to a strings rank in the population [4].

[2] Also known as mating or combination

[3] The point of crossover is irrelevant and is chosen from a uniform distribution.

[4] Such a powerful reproduction mechanism is one of the primary reasons why sexual organisms have adapted much faster than asexual ones.

offsprings 0110110 and 1110100. Once crossover has occurred, the two parents are eliminated from the population and are replaced by their offsprings.

- **Mutation**: Mutation is the mechanism that accounts for the environmental factors[5] that influence crossover or mating. In mutation, and after crossover, each bit within the child gene is randomly complemented with a probability of p_M. The probability of mutation is usually very small $p_M \ll 1$. For example the child string 0111110 mutated on the third bit from the right gives 0111010.

The GA we adopt here is the Simple Genetic Algorithm (SGA) where the genetic process is broken down into two steps [39]. In a SGA a group of chromosomes form a *current population.* Selection is subsequently applied to this population based on fitness values to create an *intermediate population.* Finally crossover and mutation (as described above) are applied to the intermediate population to arrive at a *next population* of chromosomes. Such a process is shown in its simplest form in Fig. 4.1. The selection method of choosing chromosomes from the current population for inclusion in the intermediate population is typically based on the concept of "remainder stochastic sampling". For each string x_k, the integer portion of the fitness $f(x_k)/\overline{f}$ determines how many copies of x_k are placed in the intermediate population (deterministic selection). Subsequently, all strings are copied into the intermediate population with a probability equal to the fractional part of $f(x_k)/\overline{f}$ (random selection). For example a string with fitness of 2.09 will be duplicated twice and will have a 9% chance of placing an additional copy. The SGA therfore work as follows:

1. Randomly generate an initial population of M chromosomes each corresponding to an N-bit string. These are candidate solutions to the optimization problem.

2. Calculate the fitness of each chromosome in the population using the optimization objective function.

3. Perform selection, as described above using remainder stochastic sampling, to reach an intermediate population.

4. Randomly pair the chromosomes in the intermediate population and perform crossover with probability p_C. If no crossover occurs duplicate the parent chromosomes exactly.

[5]Such as radiation.

5. Mutate the offsprings with probability $p_M \ll 1$. The resulting population is called the next population.

6. Replace the current population with the next population.

7. Evaluate the termination criteria and proceed to step 2 to continue the algorithm, otherwise terminate.

Each iteration of this process is called a *generation.* The entire set of generations is called a *run.* The termination criteria can simply be a total number of generations.

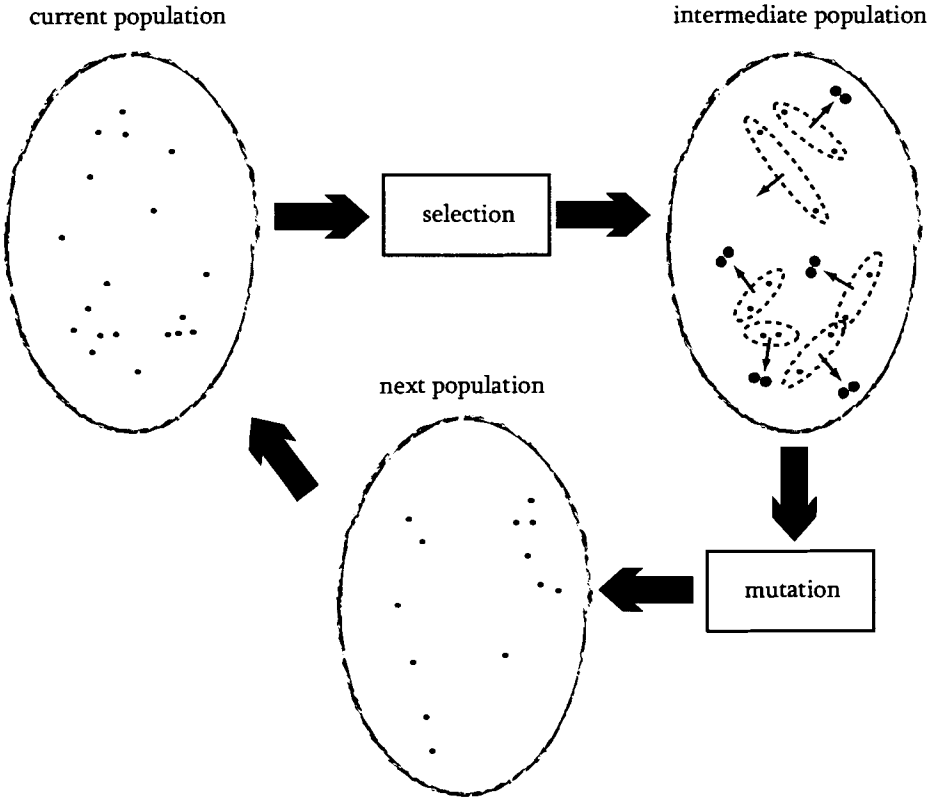

Figure 4.1: SGA diagram.

The science of evolutionary computation is a very diverse and versatile field. We have only brushed on the surface of a branch of evolutionary computational techniques, namely the genetic algorithm. Although we have not gone into the depth of the subject the iterative tool described above prepares us to employ the SGA to solve for the optimization problem

of Chapter 3 without resorting to brute-force approaches. As is evident from the discussion above, the search space of the problem must first be represented in "binary" form. We must therefore reformulate the derivations of Chapter 3.

4.2 Reformulation

Consider again the general baseband OFDM received signal model

$$\mathbf{R} = \mathbf{SGh} + \mathbf{n} \tag{4.1}$$

where $\mathbf{G} = \sqrt{N}[\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_L]$ and $\mathbf{f}_i$ is the i^{th} column of the $N \times N$ unitary DFT matrix. $\mathbf{S} = diag\{s(1), s(2), ..., s(N)\}$ is an $N \times N$ diagonal matrix of transmitted symbols taken from the MPSK constellation space $\mathcal{M}$. $\mathbf{R} = diag\{r(1), r(2), ..., r(N)\}$ is the baseband received vector and $\mathbf{h} = [h(1), h(2), ..., h(L)]^T$ is the vector representing the CIR at time k.

The transmission matrix $\mathbf{S}$ contains both pilot (known) and data symbols. The location of the pilot symbols, and hence the data symbols, can be uniquely determined through the pilot index vector $\mathbf{n_p} = [n_p(1), \ldots, n_p(N_p)]^T$. Alternatively, the pilots can be represented by a (diagonal) matrix $\mathbf{X} = diag\{x(1), x(2), \ldots, x(N)\}$, where $x(j) \in \{0, 1\}$ for $1 \leq j \leq N$. If the j^{th} sub-carrier is a pilot symbol then $x(j) = 1$ and $x(j) = 0$ otherwise. (4.1) can now be decoupled into pilot and data signals as

$$\begin{cases} \mathbf{R} = \mathbf{R}^{(pilot)} + \mathbf{R}^{(data)} \\ \mathbf{R}^{(pilot)} = \mathbf{XR} = \mathbf{SXGh} + \mathbf{Xn} \\ \mathbf{R}^{(data)} = (\mathbf{I} - \mathbf{X})\,\mathbf{R} = \mathbf{S}\,(\mathbf{I} - \mathbf{X})\,\mathbf{Gh} + (\mathbf{I} - \mathbf{X})\,\mathbf{n} \end{cases} \tag{4.2}$$

The maximum likelihood solution to the pilot locations is the zero forcing solution

$$\widehat{\mathbf{H}} = \mathbf{G}(\mathbf{SXG})^{\dagger}\mathbf{R}^{(pilot)} \tag{4.3}$$

substituting (4.2) into (4.3)

$$\widehat{\mathbf{H}} = \mathbf{H} + \underbrace{\mathbf{G}(\mathbf{SXG})^{\dagger}\mathbf{Xn}}_{interpolation\ noise} \tag{4.4}$$

substituting (4.4) into (4.2)

$$\widehat{\mathbf{R}}^{(data)} = \mathbf{S}(\mathbf{I}-\mathbf{X})\mathbf{H} + \left[\mathbf{S}(\mathbf{I}-\mathbf{X})\mathbf{G}(\mathbf{SXG})^{\dagger}\mathbf{X} + (\mathbf{I}-\mathbf{X})\right]\mathbf{n} \tag{4.5}$$

$$= \mathbf{R}^{(data)} + \mathbf{Jn} \tag{4.6}$$

where we define $\mathbf{J} = \mathbf{S}(\mathbf{I}-\mathbf{X})\mathbf{G}(\mathbf{SXG})^{\dagger}\mathbf{X}$.
Assuming a deterministic channel, the average SNR is from (4.5):

$$\overline{\mathsf{SNR}} = \frac{\overbrace{\|\mathbf{S}(\mathbf{I}-\mathbf{X})\mathbf{H}\|_F^2}^{signal\ power}}{\underbrace{\mathbb{E}\left\{\|\mathbf{Jn}\|_F^2\right\}}_{interpolation\ noise} + \underbrace{\mathbb{E}\left\{\|(\mathbf{I}-\mathbf{X})\mathbf{n}\|_F^2\right\}}_{AWGN}} \tag{4.7}$$

$$= \frac{\mathbf{H}^H\mathbf{H} - \mathbf{H}^H\mathbf{XH}}{N\sigma_n^2 tr\{\mathbf{J}^H\mathbf{J}\} + (N-N_p)\sigma_n^2} \tag{4.8}$$

$$= \overline{\mathsf{SNR}}_{CCI}\left(\frac{\mathbf{H}^H\mathbf{H} - \mathbf{H}^H\mathbf{XH}}{Ntr\{\mathbf{J}^H\mathbf{J}\} + N - N_p}\right) \tag{4.9}$$

where $\overline{\mathsf{SNR}}_{CCI} = 1/\sigma_n^2$ is the SNR for a system with *complete channel information* (CCI) i.e. no interpolation. We define the *average SNR gain* as

$$\Omega(\mathbf{X}) = \frac{\overline{\mathsf{SNR}}}{\overline{\mathsf{SNR}}_{CCI}} = \frac{\mathbf{H}^H\mathbf{H} - \mathbf{H}^H\mathbf{XH}}{Ntr\{\mathbf{J}^H\mathbf{J}\} + N - N_p} \tag{4.10}$$

The **J**-matrix may be further simplified using the facts that $\mathbf{S}^H\mathbf{S} = \mathbf{I}_N$, $\mathbf{G}^H\mathbf{G} = N\mathbf{I}_L$ and $\mathbf{XX} = \mathbf{X}$. With some mathematical manipulations we can show that the interpolation noise term can be written as:

$$Ntr\left\{\mathbf{J}^H\mathbf{J}\right\} = Ntr\left\{(\mathbf{G}^H\mathbf{XG})^{-1}\right\} - L \tag{4.11}$$

Hence going back to (4.10) and with $\Omega(\mathbf{X})$ as a global objective function, we have the following (integer) optimization problem:

$$\boxed{\begin{array}{c} \text{optimize}: \mathbf{X}^{opt} = \arg\max_{x_{ii}=\{0,1\}} \Omega(\mathbf{X}) \\ \text{subject to}: \ tr\{\mathbf{X}\} = N_p \\ \text{where}: \Omega(\mathbf{X}) = \dfrac{\mathbf{H}^H\mathbf{H} - \mathbf{H}^H\mathbf{XH}}{Ntr\{(\mathbf{G}^H\mathbf{XG})^{-1}\} + N - N_p - L} \end{array}} \tag{4.12}$$

Comparing this problem with (3.10), we clearly see that they are identical. Except, here, the optimization is over the N-dimensional binary search space instead of the N_p-dimensional integer search space $\mathbb{Z}_+^{N_p}$. In other words the coding or discretization step of the genetic algorithm has already been implemented and the SGA described in section 4.1 can directly be applied to solve for the optimal $\mathbf{X}$ variable.

4.3 More on Feasibility and on Constraints

Constraints:

The optimization of (4.12) is an NP-hard constrained combinatorial optimization problem. The constraint is obviously in the number of pilots and is denoted by $tr\{\mathbf{X}\} = N_p \geq L$. Interestingly this constraint does not directly enter the derivation of the objective function. Hence with little effort we can reformulate (4.12) and reach the more *"relaxed"* problem of:

$$\boxed{\begin{aligned} \text{optimize}: \mathbf{X}^{opt} &= \arg \max_{x_{ii}=\{0,1\}} \Omega(\mathbf{X}) \\ \text{subject to}: \; & tr\{\mathbf{X}\} \geq L \\ \text{where}: \Omega(\mathbf{X}) &= \frac{\mathbf{H}^H\mathbf{H} - \mathbf{H}^H\mathbf{X}\mathbf{H}}{Ntr\{(\mathbf{G}^H\mathbf{X}\mathbf{G})^{-1}\} + N - tr\{\mathbf{X}\} - L} \end{aligned}} \tag{4.13}$$

Note here that the number of pilots is not fixed to a specific value and is optimized along with the objective (cost) function. In other words not only do we seek the optimal pilot locations, we also seek the optimal number of pilots. We will refer to this problem as the *"relaxed"* optimization problem[6] .

Feasibility:

Given a constrained optimization problem such as (4.12), the optimum solution must be within the feasible region, i.e. $tr\{\mathbf{X}\} = N_p$. The simple genetic algorithm described above does not account for this constraint as it iterates to find the optimum solution. For example

[6]Note that we still have the general constraint that $tr\{\mathbf{X}\} \geq L$.

two parent chromosomes that both lie in the feasible region will not necessarily create offspring that also lie within the feasible region. The general approach to remedy this situation is to invoke a penalty for each child chromosome not in the feasible region. One way to penalize the offspring is to reduce it's fitness function by a value in accordance to its distance to the feasible region (see for example [32]). A more common approach would be to invoke the "death penalty" and completely discard such offspring. Given the simplicity, we will use the latter approach for our constrained optimization.

4.4 Numerical Results

We simulated a wireless OFDM link drawing baseband signals from an MPSK constellation space. For each channel realization the GA, as explained above, is run with the objective (cost) function of $\Omega(\mathbf{X})$ in mind. As noted before in order to properly estimate an L-tap channel, at least $N_p = L$ pilots are required in frequency. The structure of the optimization in (4.12) implicitly assumes this condition by assuming a fixed number of pilots with $N_p \geq L$. The optimization in (4.13), on the other hand, relaxes this condition and searches for the optimal number of pilots as well. To see the efficiency of the GA in solving both these combinatorial problems we conducted independent experiments, each explained below.

Constrained Single Channel Realization

Here we simulated a single realization of a frequency selective channel according to the model given in Chapter 1. The channel length is $L = 4$ and the total number of sub-carriers is fixed at $N = 32$. The CFR is shown in the top-left corner of Fig. 4.2. Also shown in Fig. 4.2 is the performance of the constrained optimization of (4.12). The GA was run using an initial population of $M = 10^2$ chromosomes and the probability of cross-over and mutation were set at $p_C = 0.7$ and $p_M = 10^{-3}$ respectively. The final result is shown in the top right corner of Fig. 4.2.

Interestingly the optimum solution tends to put more clustered pilots near the channel minimums (fades), much in the same way that the sub-optimum DEPO solution of Chapter 3 does. Also note that as we would expect, the total average objective (cost) function of the population increases with each iteration (or generation) indicating that the population of chromosomes are on average creating better (more fit) offspring. From the analysis in Chapter 3 and from (3.12) we know that had the pilots been allocated uniformly in

frequency the average objective (cost) function would be: $\Omega^{uniform} = \frac{N-N_p}{NL/N_p+N-N_p-L}$. Substituting with $N = 32, N_p = 8, L = 4$ leads to $\Omega^{uniform} = 0.4615$ which is also shown in Fig. 4.2 as a reference and for comparison. Note how the GA surpasses this value after roughly 50 generations. Also shown in this figure is the (instantaneous) variance of the objective function calculated at each generation. As expected, the variance of the objective function decreases with each generation which is a clear indication that the population fitness is saturating and the algorithm is converging. Here, the entire run consists of a total of $G = 100$ generations, however thresholding the objective function variance can be an alternative method of terminating the GA generations.

Relaxed Single Channel Realization

We again simulated a single frequency selective channel with length $L = 4$ and with $N = 32$ sub-carriers. The GA was first run under constrained conditions and then re-run using the same channel but with relaxed conditions. For the relaxed optimization the initial population was set to $M = 10^4$ chromosomes while the constrained optimization was initialized with $M = 10^6$ chromosomes. Both GAs were run with $p_C = 0.9$ and $p_M = 10^{-3}$ and a total of $G = 50$ generations. The results after the 50^{th} generation (final) are shown in the top plots of Fig. 4.3. As is evident from this figure, the relaxed optimization has filled nearly all the sub-carriers with pilots leaving only the maximum channel gain frequency open for data transmission. Although such an arrangement has led to considerably higher objective (cost) function (higher SNR gain) as is apparent from the bottom plot of Fig. 4.3, such a stringent condition of scarce data transmission sub-carriers renders this particular result useless in practical high data rate applications.

SER Performance

Finally, we simulated an OFDM system with $N = 16$ sub-carriers and a slowly time-varying channel of length $L = 2$. For each channel realization the GA was run to find the optimum pilot locations with $M = 10^4$, $p_C = 0.7$, $p_M = 10^{-3}$ and $G = 25$. The constrained optimization with $N_p = 4$ was considered and Fig 4.4 shows the resulting average SER calculated at each SNR. Clearly the GA improves upon uniform allocation and works very close to the optimum solution from brute force search.

4.5 Chapter Summary

By converting the search space of the optimum pilot locations into binary form, we were able to apply the genetic algorithm to find the best variable candidate. The genetic algorithm is a branch of evolutionary computations that aims to model the optimization using biological and genetic evolutions. Reformulating our problem into binary form revealed two optimization problems. One was the constrained optimization where we only seek the position of a fixed number of pilots. Much like the DEPO solution, we saw that the GA in this case tends to allocate pilots to the minimum channel gain frequencies (sub-carriers) while allowing for a number of dispersed pilots. For the relaxed problem, on the other hand, the final GA solution tends to fill nearly all the sub-carriers with pilots; keeping only the very high gain channel frequencies open for data. Finally, we simulated the SER performance of the constrained GA and compared the performance with a conventional brute force searching method.

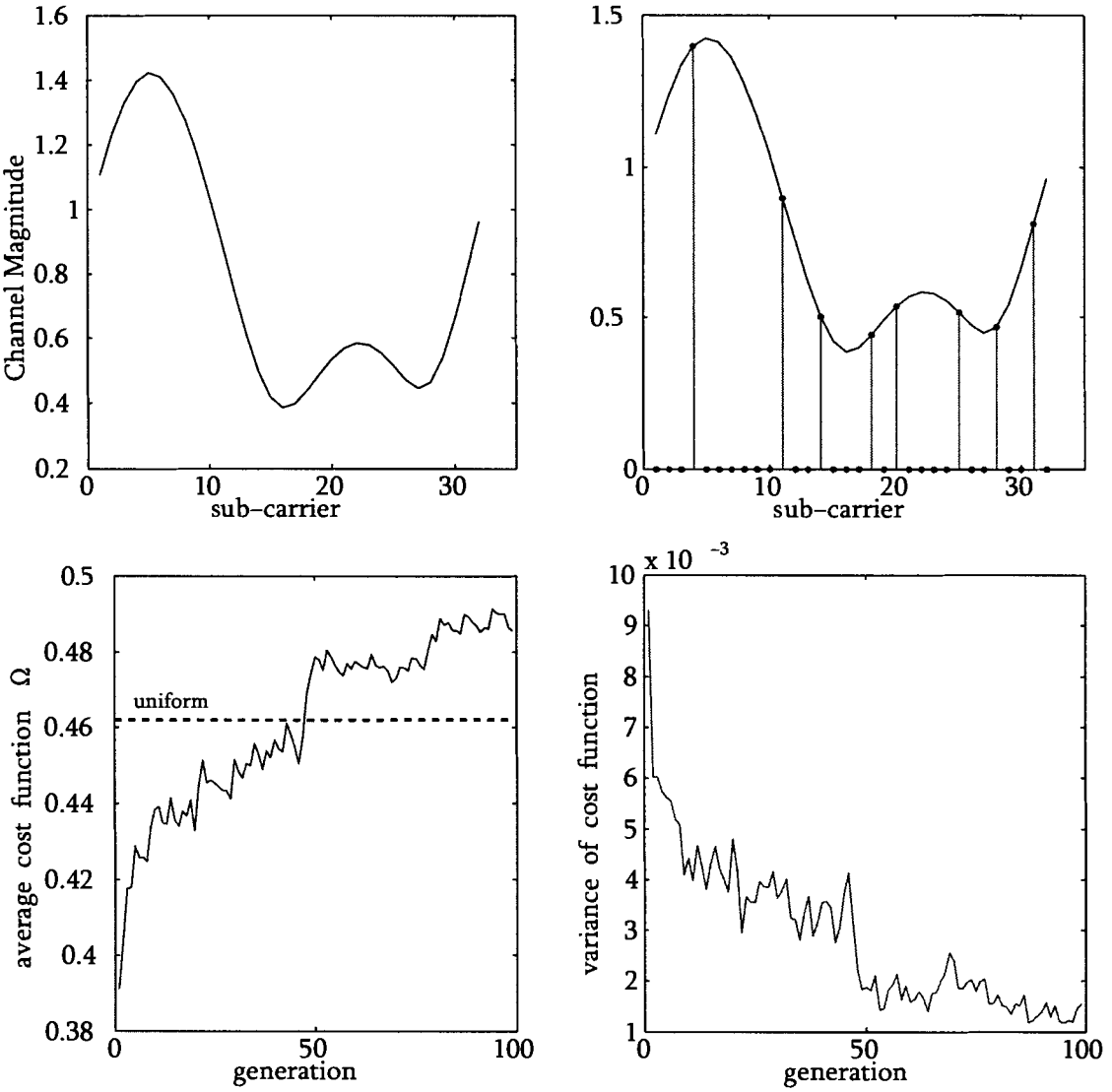

Figure 4.2: Constrained GA results on single channel realization. $N_p = 8$, $N = 32$ and $L = 4$. (top-left) Original CFR. (top-right) Pilot locations from constrained GA. (bottom-left) average population objective function per generation. (bottom-right) variance of objective function within population at each generation

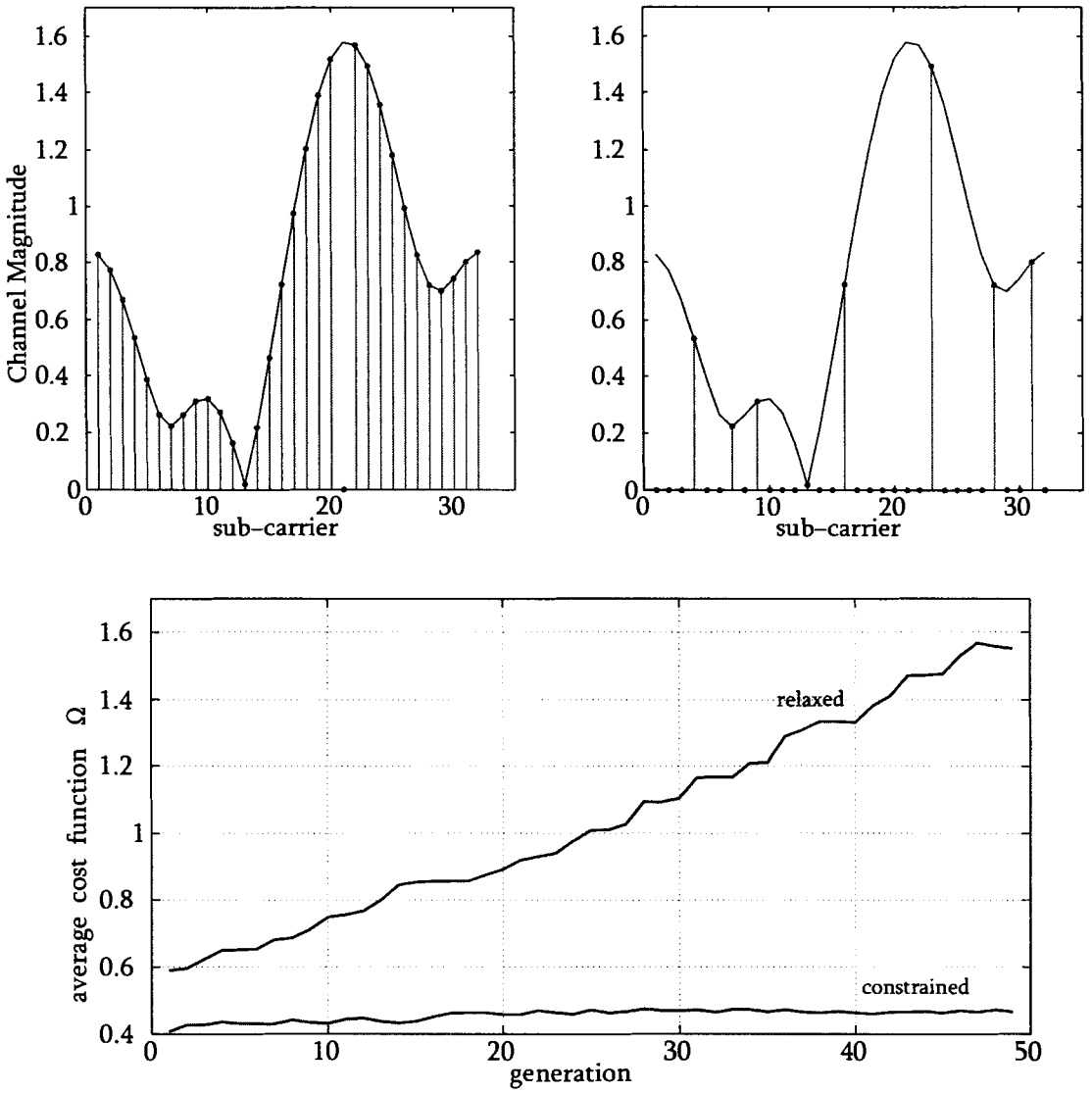

Figure 4.3: Relaxed GA results on single channel realization, $N = 32$ and $L = 4$. (top-left) Relaxed optimization (top-right) Constrained optimization. (bottom) Average cost function per generation for relaxed and constrained optimizations.

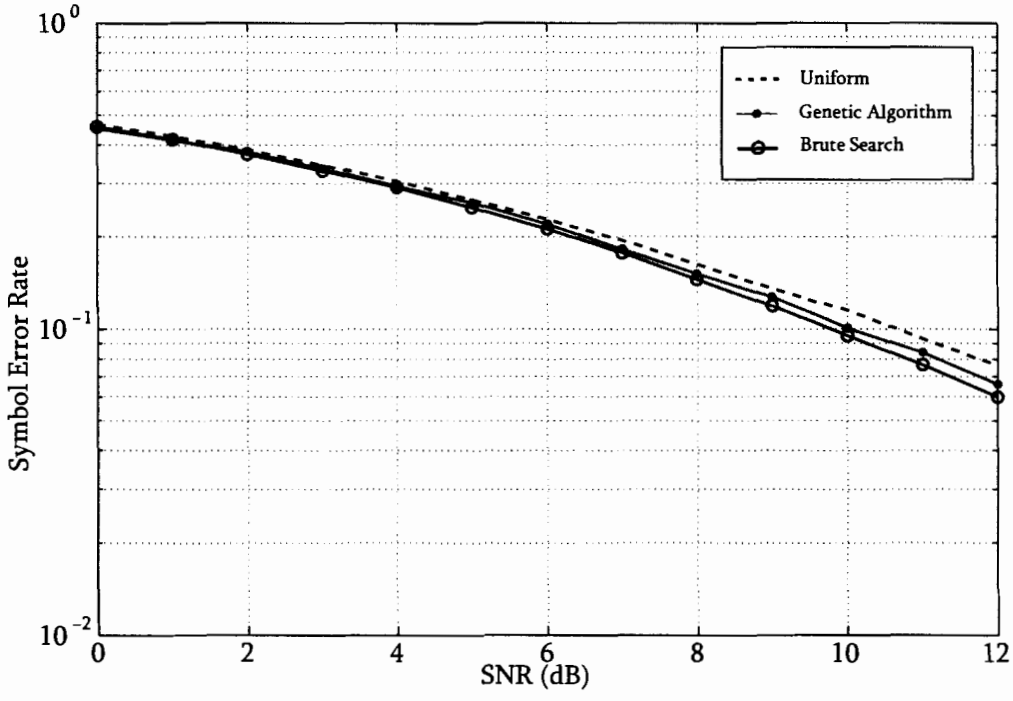

Figure 4.4: Average SER performance of (constrained) GA.

Chapter 5

Limited Feedback Communications

In Chapter 3 we derived the optimum pilot locations to maximize the average SNR at the OFDM receiver (see (3.10)). The idea is to simply allocate the pilots to the fading locations of the channel (maximize α) while maintaining an acceptable channel estimation error (minimize β). Although the closed form solution of the optimization in Eq. 3.10 was not derived, we presented two sub-optimum solutions; namely the DEPO and DOPO methods of Section 3.1.2 and compared their performances. We also looked at the actual optimum solution using evolutionary computations in the previous chapter. In this chapter, we change direction to take a closer look at the overhead associated with the actual feedback link. Our effort here will be to examine methods of reducing this overhead while maintaining acceptable levels of performance in terms of average SERs. To motivate our discussion we first answer the following questions:

1. *What exactly is the type of information that is fed back?*
2. *How much feedback is needed and how often does it need to be updated?*
3. *Can we possibly reduce the amount feedback with acceptable performance losses?*

5.1 Types of Information Feedback

The most obvious and perhaps most prevalent use of feedback in OFDM systems is the concept of *power loading*. In power loading the total available transmitter power is optimally (or at times sub-optimally) distributed within the sub-carriers (tones) to optimize a

performance criterion. The optimizing criterion is usually the minimization of SER or the maximization of capacity[1]. We will briefly examine both cases.

Consider again the general baseband received signal model of (1.28) at time k, $\mathbf{R}_k = \mathbf{S}_k\mathbf{H}_k + \mathbf{n}_k$. In power loading the transmitted signal on each sub-carrier, $s_k(n)$, is pre-multiplied by a weight $w_k(n)$ for $1 \leq n \leq N$. Therefore we can remodel the received signal as

$$\mathbf{R}_k = \mathbf{S}_k\mathbf{W}_k\mathbf{H}_k + \mathbf{n}_k \tag{5.1}$$

where $\mathbf{W}_k = diag\{w_k(1), w_k(2), ..., w_k(N)\} = diag\{\mathbf{w}_k\}$ is an $N \times N$ diagonal matrix of weights and $\mathbf{w}_k = [w_k(1), w_k(2), ...w_k(N)]^T$ is the weight vector.
The total *instantaneous* transmitted power is

$$P_T = tr\left\{\mathbf{S}_k\mathbf{W}_k\left(\mathbf{S}_k\mathbf{W}_k\right)^H\right\} = \sum_{n=1}^{N}|w_k(n)s_k(n)|^2 \tag{5.2}$$

and the *average* total power is

$$\mathbb{E}\left\{\sum_{n=1}^{N}|w_k(n)s_k(n)|^2\right\} = \sum_{n=1}^{N}|w_k(n)|^2\mathbb{E}\left\{|s_k(n)|^2\right\} = \sum_{n=1}^{N}|w_k(n)|^2, \tag{5.3}$$

since for MPSK modulation $\mathbb{E}\{|s_k(n)|^2\} = 1$. For fairness in comparisons with the non-power loaded case, we assume $\sum_{n=1}^{N}|w_k(n)|^2 = \|\mathbf{w}_k\|_2^2 = 1$. The challenge is to design the weight vector as a function of the channel $\mathbf{w}_k = \mathcal{F}(\mathbf{H}_k)$. We do not go into the details of power loading here since it is not the subject of this work. The results are however presented below for completeness.

5.1.1 SER Minimization

The authors in [17] show that the symbol error rate[2] minimizing solution is SNR equalization given by[3]

$$\mathbf{w}_k = \frac{\mathbf{g}_k}{\|\mathbf{g}_k\|_2} \tag{5.4}$$

where $\mathbf{g}_k = [|H_k(1)|^{-1}, |H_k(2)|^{-1}, ..., |H_k(N)|^{-1}]^T$ is the inverse channel gain response.

[1] Uncoded capacity.

[2] Vector error rate to be more precise.

[3] At high SNR.

5.1.2 Capacity Maximization

Capacity based power loading is usually combined with bit loading (i.e., varying the number of bits transmitted on each tone) in order to match the per-subcarrier capacity [23]. Since an OFDM system is essentially N flat-fading parallel sub-channels, the instantaneous capacity with power loading is

$$C(\mathbf{w}_k) = \sum_{n=1}^{N} \log_2 \left(1 + \frac{|w_k(n) H_k(n)|^2}{\sigma_n^2} \right). \tag{5.5}$$

The weight vector we seek is: $\mathbf{w}_k^{(opt)} = \arg\max_{\mathbf{w}_k \in \mathbb{R}_+^N} C(\mathbf{w}_k)$. When full channel knowledge is available at the transmitter, the optimal unquantized solution for power loading design is waterfilling [23]. An interesting observation here is how the above solutions contradict each other. The SER criteria allocates more power to fading sub-carriers while the capacity maximizing criteria (waterfilling) allocates more power to stronger channels.

5.1.3 Non Uniform Pilot Allocation

Note that regardless of the criteria and associated solutions above, the underlying approach in power loading is to distribute the transmitter power by directly increasing (or decreasing) the transmitted power on each sub-carrier. The non uniform pilot allocation introduced in Chapter 3 works differently.

For example, assume that a certain frequency is experiencing a deep fade. In an effort to protect the symbol from this fade, the SER criteria aims at boosting the power on this frequency. Our approach is different; we simply relocate the symbol altogether and consider sending a pilot symbol at that location. This task is done by properly choosing the pilot locations, leaving the more suitable (higher gain) locations for data transmission. In other words we ask *"Since we know where the fades are why bother to send data on those frequencies to begin with? Why not send pilots on those locations?"*. In doing so, we are in a sense indirectly equalizing our transmission stream. One must however be cautious since pilots in fades render very poor channel estimates. Hence we observed in Chapter 3, that the non-uniform pilot allocation is a delicate process of reaching an equilibrium of channel estimation error and SNR gain; an effort which ultimately concluded in performance enhancements.

Finally, note that the *information* given by feedback in our scheme lies not in the channel

gains (or inverse gains) but rather in the pilot locations. This information is given by the pilot location vector defined in Chapter 3 and denoted by $\mathbf{n}_p$. Such a structure has the immediate advantage of *"robustness"* . Any feedback scheme that relies directly on the channel gain values (such as the SER and capacity criteria above), will suffer from degradation due to information quantization. This is due to the continuous nature of the channel gains, $\mathbf{H}_k \in \mathbb{C}^N$. The pilot allocation feedback scheme on the other hand enjoys an inherent quantized information space since $\mathbf{n}_p \in \mathbb{Z}^{N_p}$ where $\mathbb{C}^N$ and $\mathbb{Z}^{N_p}$ are the complex and integer vector spaces, respectively.

5.2 Amount of Feedback

In Chapter 3 we presented three solutions to the SNR optimization problem of Eq. 3.10; optimum solution, DEPO, DOPO. Each has an associated feedback overhead per channel realization:

- ⋆ *Optimum Solution :* Requires an exhaustive search of the solution set $\mathcal{I}_p$. The cardinality of this set is $|\mathcal{I}_p| = \binom{N}{N_p}$ hence the number of feedback bits amounts to $\left\lceil \log_2 \binom{N}{N_p} \right\rceil$. For example a system with $N = 32$ sub-carriers and $N_p = 4$ pilots would require $\left\lceil \log_2 \binom{32}{4} \right\rceil = 16$ bits of feedback. Such an overhead is overwhelming even for the most flexible feedback standards!

- ⋆ *DEPO and DOPO Solutions :* These solutions depend directly on the OFDM channel $\mathbf{H}$ gains, which is dynamic by nature. Hence an exact numerical value of feedback is not tractable here. Note however that in these methods half the pilots are allocated uniformly (to reduce interpolation noise). Hence the search state is of the order of $\left\lceil \log_2 \binom{N-N_p/2}{N_p/2} \right\rceil$ which is strictly less than the optimum solution (9 bits for the example above).

5.3 Feedback Reduction

The Concept of Limited Feedback

With each channel realization a new pilot index vector must be determined at the receiver by solving the optimization problem of Eq. 3.10. This problem can either be optimally solved be exhaustive search (or the GA algorithm of Chapter 4) at the receiver or sub-optimally

determined by the DEPO (Sec. 3.1.2) or DOPO (Sec. 3.1.2) methods. Regardless of the method, once $\mathbf{n}_p$ has been chosen it must be relayed to the transmitter for pilot allocation. Since the feedback lies in a finite integer field of variables ($\mathbb{Z}_+^{N_p}$), the actual data values of $\mathbf{n}_p$, $\{n_p(1), ..., n_p(N_p)\}$, need not be transmitted from the receiver. Obviously both the transmitter and receiver have offline knowledge of this finite integer field and the receiver may simply transmit the index, i.e. j, of the chosen vector within this finite field using a total of $B = \log_2(j)$ bits. The optimum solution, for instance, would transmit $B = \left\lceil \log_2 \binom{N}{N_p} \right\rceil$ bits per feedback transmission.

Example: Assume that for $N_p = 4$ the channel is such that the optimum pilot index is determined to be $\mathbf{n}_p = [1, 8, 9, 15]$ for $N = 16$. The cardinality of $\mathcal{I}_p$ is $\binom{N}{N_p} = 1820$ which can be represented with 11 bits. The vector $\mathbf{n}_p = [1, 8, 9, 15]$ may for instance be represented by 00101000101, known before hand[4] by the transmitter and receiver[5].

The concept described above is generally referred to as *limited feedback communications* in literature. In limited feedback, the receiver and transmitter are equipped with identical look-up tables that are constructed offline. These look-up tables have the form of finite set codebooks. For our case, for example, the codebook can be denoted by a set of $Q = |\mathcal{I}_p|$ vectors of length N_p each

$$\mathcal{C} = \left\{ \mathbf{n}_p^{(1)}, \mathbf{n}_p^{(2)}, \ldots, \mathbf{n}_p^{(Q)} \right\}, \quad |\mathcal{C}| = Q. \tag{5.6}$$

The receiver first determines the optimum vector[6], $\mathbf{n}_p^{(opt)} = \mathbf{n}_p^{(q)}$, from this codebook and subsequently relays the index q to the transmitter using $\lceil \log_2 Q \rceil$ bits of feedback.

Reducing Feedback by Clustering

We propose the concept of codebook clustering as a means of reducing the feedback. The idea in codebook clustering is to partition the original codebook of (5.6) into a total of say Q' sub-sets where $Q' \leq Q$ and construct a new, lower order codebook such as:

$$\mathcal{C}' = \left\{ \tilde{\mathbf{n}}_p^{(1)}, \tilde{\mathbf{n}}_p^{(2)}, \ldots, \tilde{\mathbf{n}}_p^{(Q')} \right\}, \quad |\mathcal{C}'| = Q' \tag{5.7}$$

[4]With perhaps a look-up table.

[5]Note that this representation is different from the N-bit pilot representation of Chapter 4.

[6]Note that give the finite nature of $\mathbb{Z}_+^{N_p}$, $\mathbf{n}_p^{(opt)}$ is guaranteed to lie in the codebook.

where each $\tilde{\mathbf{n}}_p^{(i)}$ in the new codebook of (5.7), is a representation of several original $\mathbf{n}_p^{(j)}$ in the old codebook of (5.6). More precisely each $\tilde{\mathbf{n}}_p$ is a function of several $\mathbf{n}_p$

$$\tilde{\mathbf{n}}_p^{(i)} = \mathfrak{F}\left\{\mathbf{n}_p^{(j_1)}, \mathbf{n}_p^{(j_2)}, \ldots\right\}, \quad 1 \leq i \leq Q', \;\; 1 \leq j_1, j_2, \ldots \leq Q \tag{5.8}$$

Therefore to construct a clustered reduced-order codebook we must determine the operators and parameters of (5.8), which can be done in two steps:

1. Determine a partitioning or selection criteria to determine $\mathbf{n}_p^{(j_1)}, \mathbf{n}_p^{(j_2)}, \ldots$ for each representative $\tilde{\mathbf{n}}_p^{(i)}$.

2. Determine the mapping function $\mathfrak{F}\{\cdot\}$ thats acts on the partitioned vectors $\mathbf{n}_p^{(j_1)}, \mathbf{n}_p^{(j_2)}, \ldots$.

5.3.1 Vector Quantization - Generalized Lloyd Algorithm

Vector quantization is a technique often used in lossy data compression in which the basic idea is to code values from a multidimensional vector space into values from a discrete subspace of lower dimension. Vector quantization can be efficiently implemented using the well known generalized Lloyd algorithm (GLA) [16], [21] and [18] (with enhancements reported in [11]). The GLA is basically an extension of the theory of non-uniform quantization to vectors and matrices. Instead of giving a full treatment of the subject we look at how the GLA can be used in the context of our problem; codebook clustering. The GLA method is an iterative quantization method that converges to an optimum quantization while minimizing some user defined distortion function. The distortion function is design specific, and here we choose the minimum mean square error function given by [7]

$$\mathcal{D}(\mathcal{C}') = \mathbb{E}\left\{\min_{\mathbf{n}_p \in \mathcal{C}'} ||\mathbf{n}_p - \mathbf{n}_p^{(opt)}||_2^2\right\} \tag{5.9}$$

The design of an optimal quantizer is to seek the codebook that minimizes the average distortion over all possible codebooks. It can be easily shown that the optimal quantizer must satisfy the following two conditions [18]: First, it must be a nearest neighbor quantizer, i.e., it assigns to an arbitrary vector the codeword that is closest to it (the selection criteria above). Second, for a given partition of the feature space, it must satisfy the centroid condition, i.e., each codeword must be the centroid of the vectors that are mapped to it (the

[7] The elements of each $\mathbf{n}_p$ vector are sorted from smallest to greatest to render the Euclidean distance meaningful.

mapping function above or $\mathfrak{F}\{\cdot\}$).

We will design our vector quantizer to satisfy both these condition. The clustering algorithm is summarized below:

1. **Training:** Randomly generate a codebook of channel gains $\mathcal{H} = \left\{\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, ..., \mathbf{H}^{(M)}\right\}$ according to the complex Gaussian channel model.

2. For each channel within this codebook determine the corresponding optimum (or sub-optimum) pilot index vector and construct $\mathcal{C} = \left\{\mathbf{n}_p^{(1)}, \mathbf{n}_p^{(2)}, \ldots, \mathbf{n}_p^{(M)}\right\}$.

3. Of the M pilot index vectors of $\mathcal{C}$ randomly choose Q' to construct the initial[8] pilot index codebook $\mathcal{C}_0' = \left\{\tilde{\mathbf{n}}_{p,0}^{(1)}, \tilde{\mathbf{n}}_{p,0}^{(2)}, \ldots, \tilde{\mathbf{n}}_{p,0}^{(Q')}\right\}$.

4. Set $i = 1$ as the iteration counter.

5. **Clustering:** Partition the vectors of $\mathcal{C}$ into a total of Q' Voronoi regions (quantization regions) using a minimum distance criteria. The q^{th} region is defined as

$$\mathcal{Q}_q = \left\{\mathbf{n}_p \in \mathcal{C} | \ \ ||\mathbf{n}_p - \tilde{\mathbf{n}}_{p,i-1}^{(q)}||_2^2 \leq ||\mathbf{n}_p - \tilde{\mathbf{n}}_{p,i-1}^{(l)}||_2^2, \quad \forall l \neq q\right\}$$

6. Construct a new codebook $\mathcal{C}_i'$, with the q^{th} vector given from the q^{th} Voronoi region as

$$\begin{aligned} \tilde{\mathbf{n}}_{p,i}^{(q)} &= \arg \min_{\mathbf{n}_p \in \mathbb{Z}^{N_p}} \mathbb{E}\{||\mathbf{n}_p - \mathbf{x}||_2\} \quad \forall \mathbf{x} \in \mathcal{Q}_q \\ &\cong \frac{1}{|\mathcal{Q}_q|} \sum_{\mathcal{Q}} \mathbf{x} \end{aligned} \tag{5.10}$$

   where the last step is an approximation by choosing the center of mass (centroid) of the Voronoi region.

7. If $\mathcal{D}(\mathcal{C}_i') - \mathcal{D}(\mathcal{C}_{i-1}') \leq \epsilon$ or $i \geq I$, terminate the algorithm. Otherwise set $i = i + 1$ and go to step 5. The distortion function $\mathcal{D}(\cdot)$ is defined in (5.9) and can be approximated with the centroid similar to step 6, (5.10).

[8]This is one of the many ways of initialization and not necessarily the optimal.

8. Since the approximation of (5.10) leads to $\tilde{\mathbf{n}}_{p,i}^{(q)} \notin \mathbb{Z}_{+}^{N_p}$, the pilot index vectors of the final codebook will not necessarily lie in $\mathbb{Z}^{N_p}$ as required. Hence as a final step we round the codebook to the lowest (or highest) integer values[9]: $\mathcal{C}_I' = round\{\mathcal{C}_I'\}$.

Note that we leave the rounding operation for *after* the termination of the iterations to avoid excessive and unnecessary errors at the intermediate steps. In other words we *relax* the integer condition of $\mathbf{n}_p$ in the iterations but reinstate it after termination[10].

Another parameter of interest in the vector quantization process is the entropy of the quantizer at each iteration which gives a feeling of how many bits are necessary to reach a given precision for the quantizer. Since the total number of training symbols is M (step 1), at each iteration i, the probability of a training vector (step 2) belonging to the q^{th} Voronoi region is

$$p_{(m\mapsto q)}^{(i)} \stackrel{\text{def}}{=} Pr\left\{(\mathbf{n}_p^{(m)} \in \mathcal{C}) \in \mathcal{Q}_q\right\} = \frac{|\mathcal{Q}_q|^{(i)}}{M}$$

where $|\mathcal{Q}_q|^{(i)}$ is the cardinality of region q at the i^{th} iteration. The entropy is defined as

$$H^{(i)} = \sum_{q=1}^{Q'} -p_{(m\mapsto q)}^{(i)} \log_2 p_{(m\mapsto q)}^{(i)} \tag{5.11}$$

$$= \left(\frac{1}{M}\right) \sum_{q=1}^{Q'} -|\mathcal{Q}_q|^{(i)} \log_2 \left(\frac{|\mathcal{Q}_q|^{(i)}}{M}\right) \quad bits/vector \tag{5.12}$$

5.4 Numerical Results

5.4.1 Codebook Design

The GLA described above was run with an initial training set of $M = 5 \times 10^4$ randomly generated OFDM channels. A total of 4 separate codebook where designed with each case corresponding to $1, 2, 3$ and 4 bits of feedback, i.e. $Q' = \{1, 4, 8, 16\}$. Fig 5.2 shows the respective entropy and distortion functions calculated at each iteration for each codebook. The GLA terminated after 5 iterations due to the saturation in the entropies as is evident from Fig 5.2.

[9] Assuming here that the GLA has iterated a total of I times.

[10] We simulated both situations where the codebook is rounded per iteration and the case where the rounding is left to after the iterations. The results showed nearly identical results.

5.4.2 SER Performance

To test the performance of the GLA codebooks of Fig 5.2, we simulated an OFDM system using a QPSK modulation scheme with a channel length of $L = 4$. The total number of pilots (uniform and non-uniform) in all cases was set to twice the NR or $N_p = 8$. $N = 64$ sub-carriers are used and the channel is practically static in time by setting $F_d = 10^{-3} \ll 1$. The SER curve shows considerable degradation in using the quantized codebooks designed above. However in all cases the performance is still better than the uniform allocation. Note also, the vast amount of feedback reduction in using the codebooks. For example a pure DEPO solution (the unquantized curve of Fig. 5.3) would require on the order of 19 bits of feedback per channel change. Finally we note from Fig. 5.3 that the SNR conservation per increase in feedback is roughly 1 dB per bit, which might be of importance in power limited application. Given a total power constraint the system designer may allocated a codebook ranging from a few to several bits of feedback to meet these requirements.

5.5 Chapter Summary

In the light of the often huge feedback requirements in using (sub)optimal pilot allocation in OFDM systems, we suggested the use of vector quantization to reduce the feedback overhead. We tailored the GLA using a MMSE distortion function to design finite bit codebooks for the determination of the pilot locations. The designed codebooks converged after 5 iterations reaching acceptable output entropy rates in all cases. We also simulated the SER of such codebooks and compared the performance with uniform allocation and also unquantized solutions. We noted that with each one bit of feedback roughly 1 dB power may be conserved at the transmitter.

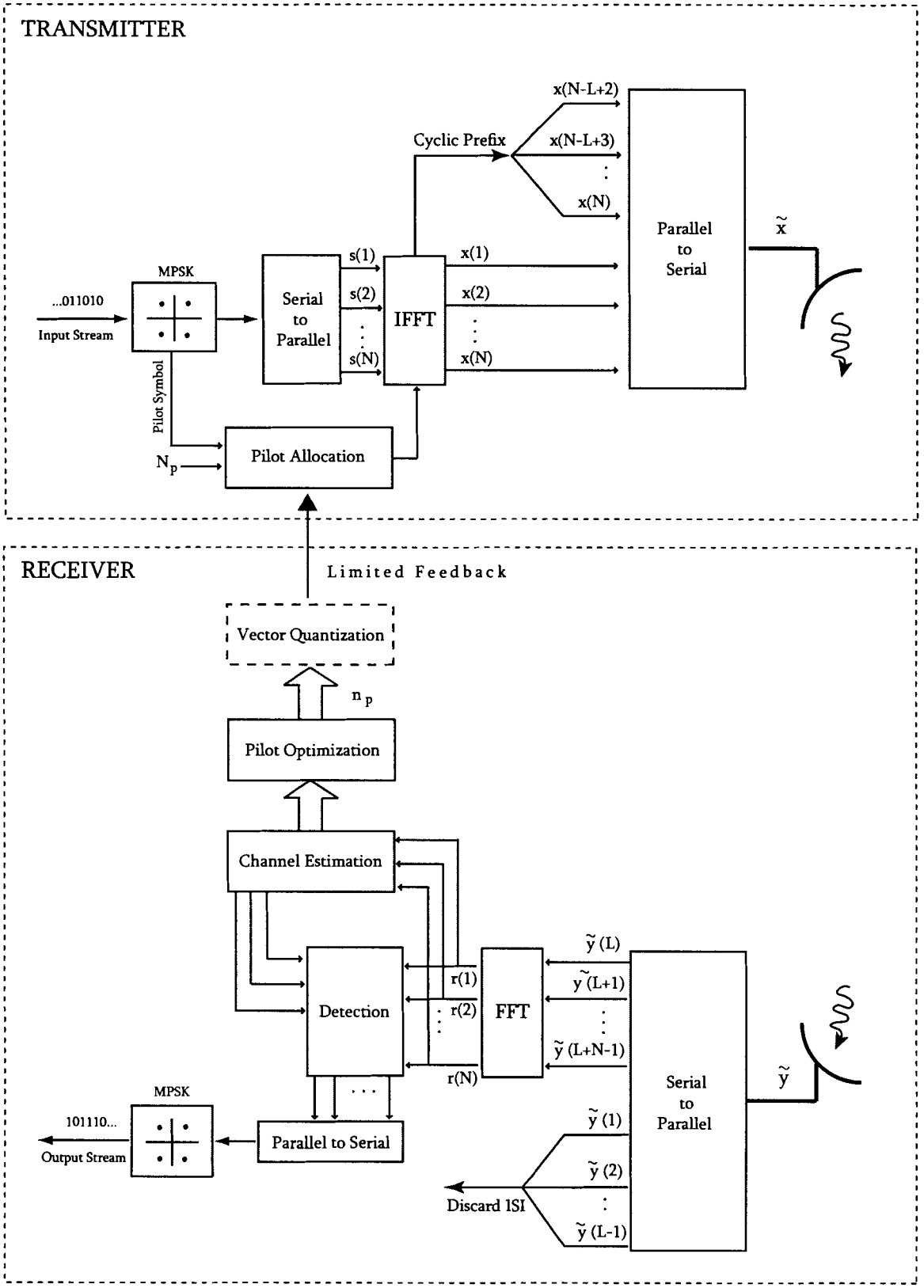

Figure 5.1: OFDM system model with limited feedback pilot allocation

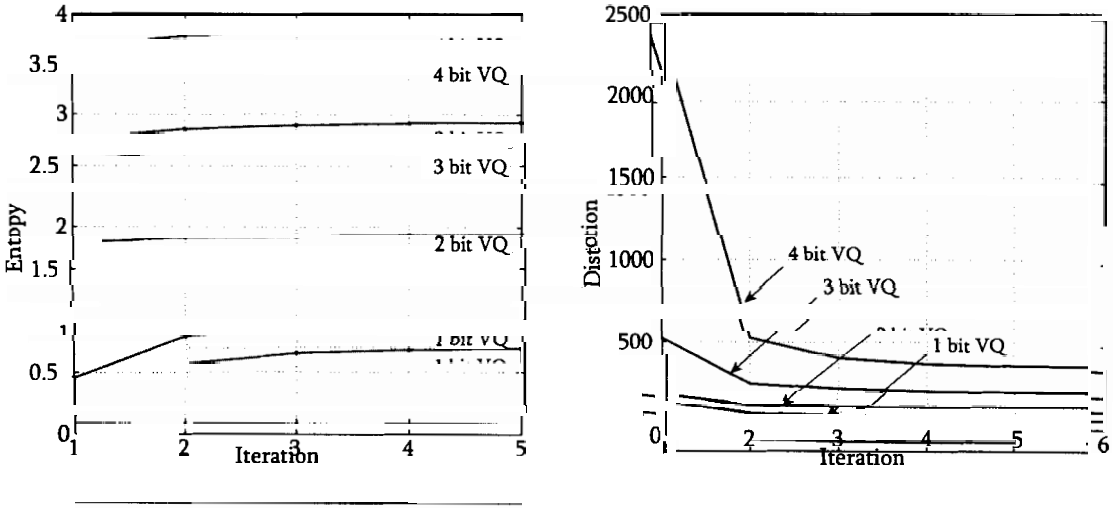

Figure 5.2: (left) Entropy. (right) Distortion.

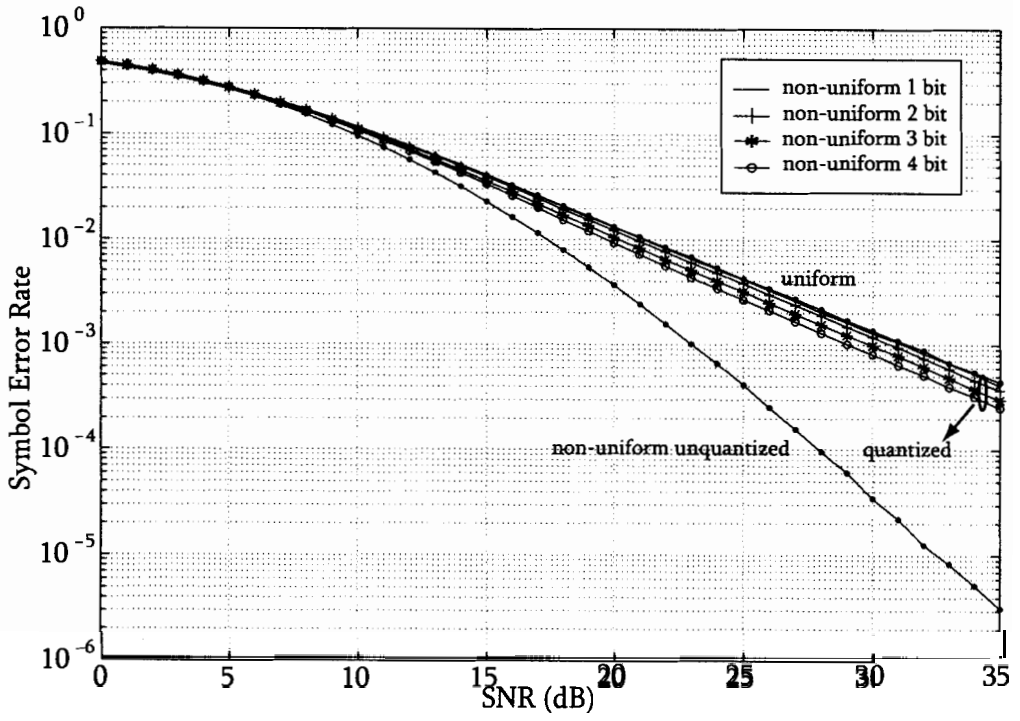

Figure 5.3: SER performance of non-uniform pilot allocation (DEPO) with limited feedback.

Chapter 6

Extension to Multiple Antennas

The multi-path nature of wireless channels often leads to the destructive addition of signals, leading to what is commonly referred to as deep fades. Signals received on, or near, the fade locations suffer from very high phase derivatives and low amplifications which often renders their correct detection hopeless. *"Diversity"* is one of the most sought out features in modern wireless communication systems for alleviating such a phenomenon[1]. OFDM-based systems, for instance, carry with them an inherent diversity in the frequency domain. Two sub-carriers separated by the coherence bandwidth will on average experience independent fading. From a simplistic point of view, hence, it may be desirable to transmit the same signal on these frequencies in the hopes that at least one will survive deep fading. In practical single-antenna OFDM systems, frequency diversity is obtained by coding and interleaving across the sub-carriers; a technique also known as bit-interleaved coded modulation (BICM) [7]. We will not be dealing with BICM here, nor do we tackle the problem of diversity maximization and coding. As of the year 2007, the combination of multiple-input multiple output (MIMO) systems and OFDM technology, i.e. MIMO-OFDM, seems to be the most promising proposal for fourth-generation mobile cellular system air interface. Inspired by this movement, in this chapter, we will extend the non-uniform pilot allocation scheme introduced in the previous chapters to a simple MIMO-OFDM system employing two transmit and one receive antennas. To exploit space-time diversity we opt to employ the Alamouti space-time block code (STBC) first introduced in [1]. The Alamouti diversity

[1] Mathematically, the diversity order determines the slope of the probability of error versus SNR curve in a log-log scale at asymptotic SNR.

scheme is well known and fully analyzed in literature. Omitting trivial observations, we immediately begin by deriving the channel estimation process for Alamouti STBC-OFDM with general pilot locations. We finish with simulated results and comments.

6.1 Alamouti STBC-OFDM

An Alamouti space-time block coded OFDM (ASTBC-OFDM) system is shown in Fig. 6.1. Consider two successive OFDM symbols $\mathbf{S}_k$ and $\mathbf{S}_{k+1}$ in time. Following the notation of [19], denote the first symbol as the odd symbol $\mathbf{S}_o$ and the second as the even symbol $\mathbf{S}_e$. For the first antenna, $\mathbf{S}_o$ is transmitted during the first time slot followed by $-\mathbf{S}_e^*$ in the second time slot. For the second antenna, $\mathbf{S}_e$ is transmitted first followed by $\mathbf{S}_o^*$. Since at each time instance two OFDM symbol are transmitted into the channel, to make future comparisons with single antenna systems fair in terms of power, we reduce the power of each symbol by half. Equivalently these symbols are encoded in space and in time to form the transmission matrix:

$$\mathcal{G}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{S}_o & -\mathbf{S}_e^* \\ \mathbf{S}_e & \mathbf{S}_o^* \end{pmatrix} \tag{6.1}$$

where the columns and rows of $\mathcal{G}_2$ represent temporal and spatial (antenna) dimensions respectively. Assume that $\mathbf{H}_1$ and $\mathbf{H}_2$ are the CFRs between the single receiver antenna and the first and second transmitting antennas respectively. Also assume that these channels are constant and independent of each other during the two consecutive OFDM time slots. The received signals in the corresponding time slots are

$$\mathbf{R}_1 = \frac{1}{\sqrt{2}} (\mathbf{S}_o \mathbf{H}_1 + \mathbf{S}_e \mathbf{H}_2) + \mathbf{N}_1 \tag{6.2}$$

$$\mathbf{R}_2 = \frac{1}{\sqrt{2}} (-\mathbf{S}_e^* \mathbf{H}_1 + \mathbf{S}_o^* \mathbf{H}_2) + \mathbf{N}_2 \tag{6.3}$$

which can conveniently be shown in matrix form:

$$\begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{S}_o & \mathbf{S}_e \\ -\mathbf{S}_e^* & \mathbf{S}_o^* \end{pmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{bmatrix} \tag{6.4}$$

where $\mathbf{N}_1$ and $\mathbf{N}_2$ are uncorrelated AWGN with identical correlation matrices of $\sigma_n^2 \mathbf{I}_N$.

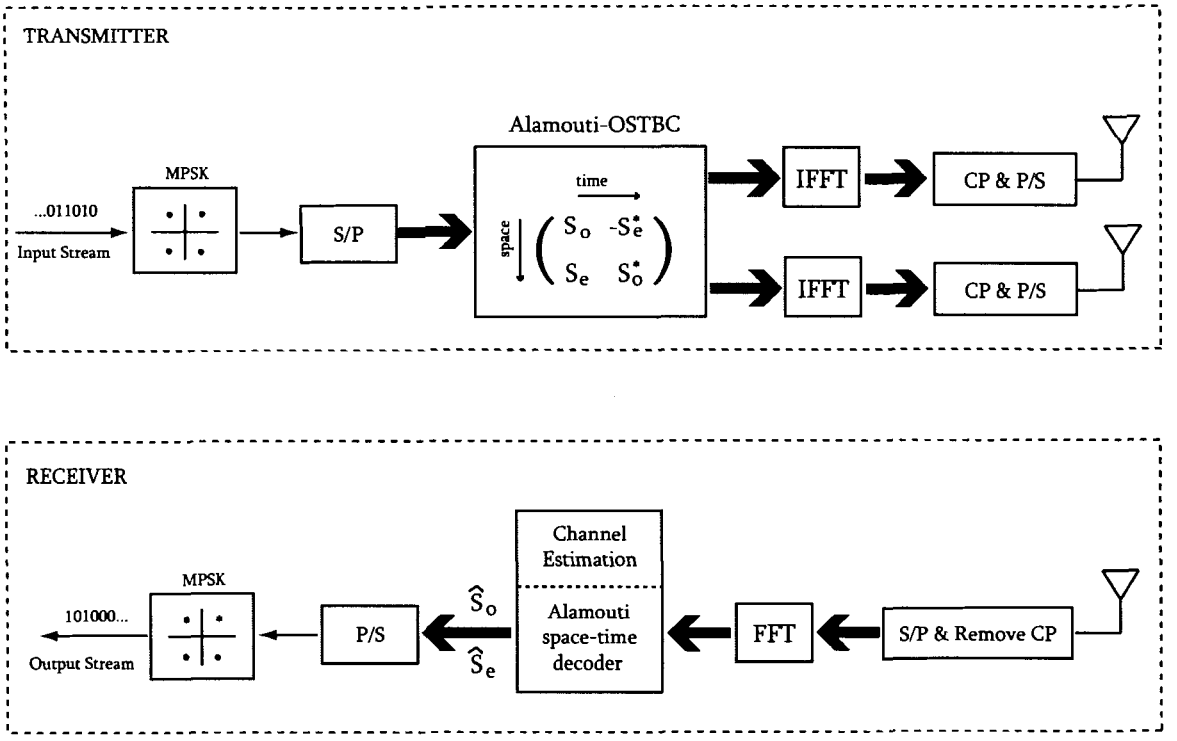

Figure 6.1: Alamouti 2×1 STBC employing OFDM, transmitter and receiver models.

Channel Estimation

Assume as in Chapter 4 that the pilots locations can be represented by a (diagonal) matrix $\mathbf{X} = diag\{x(1), x(2), \ldots, x(N)\}$, where $x(j) \in \{0, 1\}$ for $1 \leq j \leq N$. If the j^{th} sub-carrier is a pilot symbol then $x(j) = 1$ and $x(j) = 0$ otherwise. Therefore the received signals at the pilot locations only are

$$\begin{aligned}
\begin{pmatrix} \mathbf{X} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \mathbf{X} \end{pmatrix} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix} &= \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{S}_o\mathbf{X} & \mathbf{S}_e\mathbf{X} \\ -\mathbf{S}_e^*\mathbf{X} & \mathbf{S}_o^*\mathbf{X} \end{pmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{N}_1\mathbf{X} \\ \mathbf{N}_2\mathbf{X} \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{pmatrix} s_p\mathbf{X} & s_p\mathbf{X} \\ -(s_p\mathbf{X})^* & (s_p\mathbf{X})^* \end{pmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} \\
&= \frac{1}{\sqrt{2}} \mathbf{S}_p \begin{pmatrix} \mathbf{X} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \mathbf{X} \end{pmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix}
\end{aligned} \tag{6.5}$$

where $\begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{N}_1\mathbf{X} \\ \mathbf{N}_2\mathbf{X} \end{bmatrix}$ and s_p is a known pilot symbol and we have defined

$$\mathbf{S}_p = \begin{pmatrix} s_p\mathbf{I}_N & s_p\mathbf{I}_N \\ -s_p^*\mathbf{I}_N & s_p^*\mathbf{I}_N \end{pmatrix}, \quad \mathbf{S}_p^H\mathbf{S}_p = 2\mathbf{I}_{2N}$$

Multiplying both sides of (6.5) by $\frac{1}{\sqrt{2}}\mathbf{S}_p^H$ and considering the CIR on the right hand side we get

$$\begin{aligned} \frac{1}{\sqrt{2}}\mathbf{S}_p^H(\mathbf{I}_2 \otimes \mathbf{X}) \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix} &= \begin{pmatrix} \mathbf{X} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \mathbf{X} \end{pmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} \\ &= \begin{pmatrix} \mathbf{XG} & \mathbf{0}_{N\times L} \\ \mathbf{0}_{N\times L} & \mathbf{XG} \end{pmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} \\ &= (\mathbf{I}_2 \otimes \mathbf{XG}) \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} \end{aligned} \tag{6.6}$$

The zero-forcing solution to (6.6) is

$$\begin{bmatrix} \widehat{\mathbf{H}}_1 \\ \widehat{\mathbf{H}}_2 \end{bmatrix} = \frac{1}{\sqrt{2}}(\mathbf{I}_2 \otimes \mathbf{G})(\mathbf{I}_2 \otimes \mathbf{XG})^\dagger \mathbf{S}_p^H(\mathbf{I}_2 \otimes \mathbf{X}) \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix} \tag{6.7}$$

Data Detection

Once the channel has been estimated, data detection can be performed by rewriting (6.2) as[2]

$$\begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2^* \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{H}_1 & \mathbf{H}_2 \\ \mathbf{H}_2^* & -\mathbf{H}_1^* \end{pmatrix} \begin{bmatrix} \mathbf{S}_o \\ \mathbf{S}_e \end{bmatrix} + \begin{bmatrix} \mathbf{N}_1 \\ \mathbf{N}_2^* \end{bmatrix} \tag{6.8}$$

multiplying both side by $\begin{pmatrix} \mathbf{H}_1 & \mathbf{H}_2 \\ \mathbf{H}_2^* & -\mathbf{H}_1^* \end{pmatrix}^H$ we have

$$\begin{bmatrix} \widehat{\mathbf{S}}_o \\ \widehat{\mathbf{S}}_e \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} (\|\mathbf{H}_1\|_2^2 + \|\mathbf{H}_2\|_2^2)\mathbf{I}_N & \mathbf{I}_N \\ \mathbf{I}_N & (\|\mathbf{H}_1\|_2^2 + \|\mathbf{H}_2\|_2^2)\mathbf{I}_N \end{pmatrix} \begin{bmatrix} \mathbf{S}_o \\ \mathbf{S}_e \end{bmatrix} + \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \end{bmatrix} \tag{6.9}$$

[2]We understand that $\mathbf{H}_1$ and $\mathbf{H}_2$ here are $N \times N$ diagonal matrices and $\mathbf{S}_o$ and $\mathbf{S}_e$ are length N column vectors.

where $\mathbf{W}_1$ and $\mathbf{W}_2$ are independent AWGN with covariance matrix of $2N\sigma_n^2\mathbf{I}_N$. The transmitted symbols are disjointly decoded using this final set of equations[3].

6.2 Numerical Results

We simulated an A-STBC-OFDM system[4] with 2 transmit and 1 receive antenna (2×1). The channel length is $L = 4$ and the number of sub-carriers was set to $N = 64$. A total of $N_p = 8$ pilot tones are distributed within the sub-carriers for channel estimation purposes. In one scenario the pilot are allocated uniformly and in another scenario they are allocated using the DEPO method of Section. 3.1.2. Fig. 6.2 shows the average SER performance of both scenarios including the single antenna case of the previous chapters in a wide range of SNRs. In the low SNR regime ($\leq$ 10 dB), the performances of both uniform and non-uniform schemes are essentially equivalent. In the high SNR regime where fading is the dominant factor, there is considerable improvements using non-uniform pilots both for the 1×1 and the 2×1 case. This is mainly due to the fact that the non-uniform technique(s) *masks* or fills these fade locations with pilots. For example, at an error rate of 10^{-5}, the non-uniform 2×1 allocation scheme can save on average 5 dB in transmitter power compared to uniform 2×1 allocation. Interestingly, a comparison of the slope of the SER curves is an indication that the non-uniform pilot allocation is also enjoying frequency diversity in addition to space-time diversity from the Alamouti STBC.

6.3 Chapter Summary

We extended the framework of non-uniform pilot symbol to MIMO-OFDM systems. In particular we looked at a 2×1 Alamouti STBC and derived an expression for the channel estimation using arbitrary pilot locations. Finally we simulated the SER performance of the non-uniform pilot allocation in conjunction with this multiple antenna configuration. The results show considerable improvements in SER, especially at high SNR where diversity is the primary factor in performance.

[3]SER calculation is of course carried out only on the data location given by the $(\mathbf{I} - \mathbf{X})$ matrix.

[4]The IEEE 802.16 or WiMAX protocol has a transceiver mode similar to this configuration.

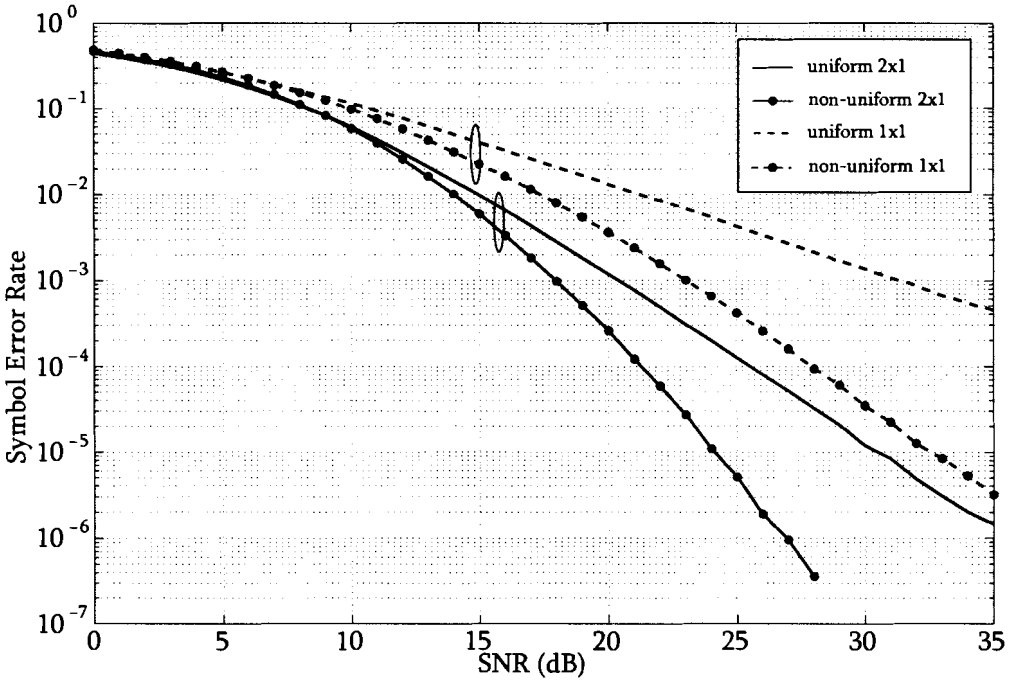

Figure 6.2: SER performance of Alamouti STBC-OFDM with non-uniform pilot allocation.

Chapter 7

Conclusion

The primary focus of this thesis was to improve conventional OFDM channel estimation when feedback is made available to the transmitter. The objective function, i.e. cost function was the average SNR at the OFDM receiving antenna. Basic fundamentals of OFDM and the mathematical foundations of conventional PSACE-OFDM were the subject of the first two chapters. Chapter 3 showed how in the case of feedback, the optimum pilots are not necessarily equi-spaced. Subsequent novel contributions included:

- The utilization of a branch of evolutionary computational mathematics called genetic algorithms to solve for the optimum pilot locations (Chapter 4).
- Reduction in the feedback overhead using codebook clustering and the generalized Lloyd algorithm (Chapter 5).
- Extensions to MIMO-OFDM and Alamouti STBCs with mathematical derivations (Chapter 6).

Simulations were carried out to assess the performance of the new PSACE scheme in both the single antenna and in the MIMO-OFDM cases. Significant improvements were seen in both cases as a result of using optimum and even sub-optimum solutions to the pilot locations. When possible, these numerical results were validated with theoretical derivations.

Appendix A

Alternative Derivation of MSE

In this appendix we derive the MSE as a function of sub-carrier index and also the total MSE for each OFDM symbol.

From (2.9) the *per-carrier* channel mean square error (MSE) is

$$\begin{aligned}
\mathsf{MSE}(n) &\stackrel{\text{def}}{=} \mathbb{E}\left\{\left|\mathbf{H}(n)-\widehat{\mathbf{H}}(n)\right|^2\right\} \\
&= \mathbb{E}\left\{\left|\sum_{k=1}^{N_p}\mathbf{P}_{(n,k)}\mathbf{w}^{\langle p\rangle}(k)\right|^2\right\} \\
&= \mathbb{E}\left\{\left(\sum_{k=1}^{N_p}\mathbf{P}_{(n,k)}\mathbf{w}^{\langle p\rangle}(k)\right)\left(\sum_{k'=1}^{N_p}\mathbf{P}_{(n,k')}\mathbf{w}^{\langle p\rangle}(k')\right)^*\right\} \\
&= \sum_{k=1}^{N_p}\sum_{k'=1}^{N_p}\mathbf{P}_{(n,k)}\mathbf{P}^*_{(n,k')}\mathbb{E}\left\{\mathbf{w}^{\langle p\rangle}(k)(\mathbf{w}^{\langle p\rangle}(k'))^*\right\} \\
&= \sigma_n^2\sum_{k=1}^{N_p}\left|\mathbf{P}_{(n,k)}\right|^2 \\
&= \sigma_n^2\left(\mathbf{P}\mathbf{P}^H\right)_{(n,n)} \quad 1\le n\le N
\end{aligned} \tag{A.1}$$

For uniform allocation we have

$$\mathbf{P}\mathbf{P}^H = \mathbf{G}\mathbf{D}^{-1}\overbrace{\mathbf{B}^H\mathbf{B}}^{\mathbf{D}}\mathbf{D}^{-H}\mathbf{G}^H \tag{A.2}$$

$$= \mathbf{G}\mathbf{D}^{-1}\mathbf{G}^H$$

$$= \frac{1}{N_p}\mathbf{G}\mathbf{G}^H \tag{A.3}$$

Using (A.1), (A.2) and the fact that $\left(\mathbf{G}\mathbf{G}^H\right)_{(i,i)} = L$ we have

$$\mathsf{MSE}(n) = \sigma_n^2\left(\frac{L}{N_p}\right), \tag{A.4}$$

and the *total* MSE is

$$\overline{\mathsf{MSE}} = \frac{1}{N}\sum_{m=1}^{N}\mathsf{MSE}(n) \tag{A.5}$$

$$= \frac{1}{N}\sum_{m=1}^{N}\sigma_n^2\left(\frac{L}{N_p}\right)$$

$$= \sigma_n^2\left(\frac{L}{N_p}\right) \tag{A.6}$$

which is identical to (2.31).

Appendix B

Average SNR derivation

We derive an expression for the average SNR at the OFDM receiver. Rewriting the average SNR from (3.2)

$$\overline{\mathsf{SNR}}_k = \frac{\left(\widehat{\mathbf{H}}_k^{\langle d \rangle}\right)^H \widehat{\mathbf{H}}_k^{\langle d \rangle}}{tr\left\{\mathbb{E}\left\{\mathbf{v}_k^{\langle d \rangle}\left(\mathbf{v}_k^{\langle d \rangle}\right)^H\right\}\right\} + tr\left\{\mathbb{E}\left\{\mathbf{n}_k^{\langle d \rangle}\left(\mathbf{n}_k^{\langle d \rangle}\right)^H\right\}\right\}} \tag{B.1}$$

The first noise power terms in (B.1) depends on the interpolation. Using (2.10)

$$
\begin{aligned}
P_{\mathbf{v}} &= tr\left\{\mathbb{E}\left\{\mathbf{v}_k^{\langle d\rangle}\left(\mathbf{v}_k^{\langle d\rangle}\right)^H\right\}\right\} \\
&= tr\left\{\mathbb{E}\left\{\left(\left(\mathbf{w}_k^{\langle p\rangle}\right)^H \mathbf{P}_k^H \mathbf{S}_k^H \mathbf{S}_k \mathbf{P}_k \mathbf{w}_k^{\langle p\rangle}\right)^{\langle d\rangle}\right\}\right\} \\
&= tr\left\{\mathbb{E}\left\{\left(\left(\mathbf{w}_k^{\langle p\rangle}\right)^H \mathbf{P}_k^H \mathbf{P}_k \mathbf{w}_k^{\langle p\rangle}\right)^{\langle d\rangle}\right\}\right\} \\
&= \mathbb{E}\left\{tr\left\{\left(\left(\mathbf{w}_k^{\langle p\rangle}\right)^H \mathbf{P}_k^H \mathbf{P}_k \mathbf{w}_k^{\langle p\rangle}\right)^{\langle d\rangle}\right\}\right\} \\
&= \mathbb{E}\left\{tr\left\{\left(\mathbf{w}_k^{\langle p\rangle}\left(\mathbf{w}_k^{\langle p\rangle}\right)^H \mathbf{P}_k^H \mathbf{P}_k\right)^{\langle d\rangle}\right\}\right\} \\
&= tr\left\{\mathbb{E}\left\{\left(\mathbf{w}_k^{\langle p\rangle}\left(\mathbf{w}_k^{\langle p\rangle}\right)^H \mathbf{P}_k^H \mathbf{P}_k\right)^{\langle d\rangle}\right\}\right\} \\
&= tr\left\{\left(\mathbb{E}\left\{\mathbf{w}_k^{\langle p\rangle}\left(\mathbf{w}_k^{\langle p\rangle}\right)^H\right\} \mathbf{P}_k^H \mathbf{P}_k\right)^{\langle d\rangle}\right\} \\
&= tr\left\{\left(\sigma_n^2 \mathbf{I}_{N_p} \mathbf{P}_k^H \mathbf{P}_k\right)^{\langle d\rangle}\right\} \\
&= \sigma_n^2 tr\left\{\left(\mathbf{P}_k^H \mathbf{P}_k\right)^{\langle d\rangle}\right\} \\
&= \sigma_n^2 ||\mathbf{P}_k^{\langle d\rangle}||_F^2
\end{aligned} \tag{B.2}
$$

where we have used the norm definition $\|A\|_F^2 = tr\left\{A^H A\right\}$ and the identity $tr\{AB\} = tr\{BA\}$ with repeated use of the interchangeability of the trace and expectation operators. We also used the assumption that $\widehat{\mathbf{H}}_k$ is deterministic to assume that the interpolating matrix $\mathbf{P}_k$ is also a deterministic (but unknown) function of $\widehat{\mathbf{H}}_k$, i.e. $\mathbb{E}\left\{\mathbf{P}_k^H \mathbf{P}_k\right\} = \mathbf{P}_k^H \mathbf{P}_k$. The second noise power in (B.1) is

$$
\begin{aligned}
P_{\mathbf{n}} &= tr\left\{\mathbb{E}\left\{\mathbf{n}_k^{\langle d\rangle}\left(\mathbf{n}_k^{\langle d\rangle}\right)^H\right\}\right\} \\
&= (N - N_p)\,\sigma_n^2.
\end{aligned} \tag{B.3}
$$

The expression in (B.1) therefore simplifies to

$$
\overline{\mathsf{SNR}}_k = \frac{\left(\widehat{\mathbf{H}}_k^{\langle d\rangle}\right)^H \widehat{\mathbf{H}}_k^{\langle d\rangle}}{\sigma_n^2 ||\mathbf{P}_k^{\langle d\rangle}||_F^2 + (N - N_p)\,\sigma_n^2}. \tag{B.4}
$$

This result can be further simplified by using the fact that

$$||\mathbf{P}_k^{\langle d \rangle}||_F^2 = ||\mathbf{P}_k||_F^2 - ||\mathbf{P}_k^{\langle p \rangle}||_F^2 = \|\mathbf{P}_k\|_F^2 - L \tag{B.5}$$

Substituting (B.5) into (B.4) we get the final result for the predicted average signal to noise ratio of the output at time k

$$\overline{\mathsf{SNR}}_k = \frac{\left(\widehat{\mathbf{H}}_k^{\langle d \rangle}\right)^H \widehat{\mathbf{H}}_k^{\langle d \rangle}}{\sigma_n^2 \left(\|\mathbf{P}_k\|_F^2 - L\right) + (N - N_p)\,\sigma_n^2}. \tag{B.6}$$

Bibliography

[1] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[2] T. Bäck and H. Schwefel, "An overview of evolutionary algorithms for parameter optimization," *Evolutionary Computation*, vol. 1, no. 1, pp. 1–23, 1993.

[3] J. Bagley, "The behavior of adaptive systems which employ genetic and correlative algorithms," Ph.D. dissertation, PhD thesis, University of Michigan, Ann Arbor, 1967.

[4] J. Baker, "Adaptive selection methods for genetic algorithms," in *Proc. 1st International Conference on Genetic Algorithms*. Lawrence Erlbaum Associates, Inc. Mahwah, NJ, USA, 1985, pp. 101–111.

[5] J. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, 1990.

[6] X. Cai and G. Giannakis, "Error probability minimizing pilots for OFDM with M-PSK modulation over Rayleigh-fading channels," in *Proc. IEEE Vehicular Technology Conference (VTC)*, vol. 53, no. 1, 2004, pp. 146–155.

[7] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, 1998.

[8] J. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, 1991.

[9] ——, "Pilot symbol assisted modulation and differential detection in fading and delay spread," *IEEE Transactions on Communications*, vol. 43, no. 7, pp. 2206–2212, 1995.

[10] M. Chang and Y. Su, "Performance analysis of equalized OFDM systems in Rayleigh fading," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 721–732, 2002.

[11] C. Chen, "An enhanced generalized Lloyd algorithm," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 167–170, 2004.

[12] H. Cheon and D. Hong, "Effect of channel estimation error in OFDM-based WLAN," *IEEE Communications Letters*, vol. 6, no. 5, 2002.

[13] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 48, no. 3, pp. 223–229, 2002.

[14] B. Crow, I. Widjaja, L. Kim, and P. Sakai, "IEEE 802.11 Wireless Local Area Networks," *IEEE Communications Magazine*, vol. 35, no. 9, pp. 116–126, 1997.

[15] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[16] A. Gersho and R. Gray, *Vector quantization and signal compression.* Kluwer Academic Pub, 1992.

[17] L. Goldfeld, V. Lyandres, and D. Wulich, "Minimum BER power loading for OFDM in fading channel," *IEEE Transactions on Communications*, vol. 50, no. 11, pp. 1729–1733, 2002.

[18] I. Katsavounidis and C. Zhang, "A new initialization technique for generalized Lloyd iteration," *IEEE Signal Processing Letters*, vol. 1, no. 10, pp. 144–146, 1994.

[19] K. Lee and D. Williams, "A space-time coded transmitter diversity technique for frequency selective fading channels," *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 149–152, 2000.

[20] Y. Li, L. Cimini Jr, and N. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 902–915, 1998.

[21] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, vol. 28, no. 1, pp. 84–95, 1980.

[22] H. Lo, D. Lee, and J. Gansman, "A study of non-uniform pilot spacing for PSAM," in *Proc. IEEE International Conference on Communications (ICC)*, vol. 1, 2000.

[23] D. Love and R. Heath Jr, "OFDM power loading using limited feedback," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1773–1780, 2005.

[24] Z. Michalewicz, D. Dasgupta, R. Le Riche, and M. Schoenauer, "Evolutionary algorithms for constrained engineering problems," *Computers & Industrial Engineering Journal*, vol. 30, no. 2, pp. 851–870, 1996.

[25] M. Mitchell, *An introduction to genetic algorithms.* Mit Pr, 1996.

[26] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for ofdm systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3065–3073, 2001.

[27] Y. Mostofi and D. Cox, "Average error rate analysis for pilot-aided OFDM receivers with frequency-domain interpolation," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 3, 2004.

[28] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 3, pp. 1122–1128, 1998.

[29] S. Ohno and G. Giannakis, "Average-rate optimal PSAM transmissions over time-selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 712–720, 2002.

[30] ——, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 2138–2145, 2004.

[31] T. Onizawa, M. Mizoguchi, T. Sakata, and M. Morikura, "A simple adaptive channel estimation scheme for OFDM systems," in *Proc. IEEE Vehicular Technology Conference (VTC)*, vol. 1, 1999.

[32] J. Richardson, M. Palmer, G. Liepins, and M. Hilliard, "Some guidelines for genetic algorithms with penalty functions," in *Proc. Third international conference on Genetic algorithms.* Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 1989, pp. 191–197.

[33] J. Rinne and M. Renfors, "Pilot spacing in orthogonal frequency division multiplexing systems on practical channels," *IEEE Transactions on Consumer Electronics*, vol. 42, no. 4, pp. 959–962, 1996.

[34] M. Russell and G. Stüber, "Terrestrial digital video broadcasting for mobile reception using OFDM," *Wireless Personal Communications*, vol. 2, no. 1, pp. 45–66, 1995.

[35] L. Schuchman, J. Miller, and R. Bruno, "Digital audio broadcasting system," Feb. 1 1994, u.S Patent 5,283,780.

[36] W. Song and J. Lim, "Pilot-symbol aided channel estimation for OFDM with fast fading channels," *IEEE Transactions on Broadcasting*, vol. 49, no. 4, pp. 398–402, 2003.

[37] X. Tang, M. Alouini, and A. Goldsmith, "Effect of channel estimation error on M-QAM BER performance in Rayleigh fading," *IEEE Transactions on Communications*, vol. 47, no. 12, pp. 1856–1864, 1999.

[38] J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE Vehicular Technology Conference (VTC)*, vol. 2, 1995.

[39] M. Vose, *The simple genetic algorithm: foundations and theory.* Bradford Books, 1999.

[40] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Transactions on Communications [legacy, pre-1988]*, vol. 19, no. 5, pp. 628–634, 1971.

[41] D. Young and N. Beaulieu, "The generation of correlated Rayleigh random variates by inverse discrete Fourier transform," *IEEE Transactions on Communications*, vol. 48, no. 7, pp. 1114–1127, 2000.