# TRANSMIT BEAMFORMING FOR FREQUENCY–SELECTIVE CHANNELS WITH SUBOPTIMUM EQUALIZATION

by

LIANG, YANG WEN

B.ENG., McMaster University, 2004

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

August 2006

# Abstract

In this thesis, we propose beamforming schemes for frequency–selective channels with decision–feedback equalization (DFE) or linear equalization (LE) at the receiver and with, respectively, perfect and quantized channel state information (CSI) at the transmitter. For beamforming with perfect CSI and infinite impulse response (IIR) beamforming filters (BFFs) we derive a closed–form expression for the optimum BFFs. We also provide two efficient numerical methods for recursive calculation of the optimum finite impulse response (FIR) BFFs with perfect CSI. For beamforming with quantized CSI and finite–rate feedback channel, we propose a global vector quantization (GVQ) algorithm for codebook design. This algorithm is deterministic and independent of initial conditions and does not impose any constraints on the number of transmit and receive antennas, the antenna correlation, or the fading statistics. Our simulation results for typical GSM[1] [1] and EDGE[2] [2, 3, 4] channels show that in general short FIR BFFs are sufficient to closely approach the performance of IIR BFFs even in severely frequency–selective channels. Furthermore, finite–rate feedback beamforming with only a few feedback bits achieves significant performance gains over single–antenna transmission, transmit antenna selection, and optimized delay diversity [5] in frequency–selective fading.

[1]GSM: Global System for Mobile Communication.
[2]EDGE: Enhanced Data Rates for GSM Evolution.

# Table of Contents

# List of Figures

# List of Abbreviations and Symbols

## Acronyms

| | |
|---|---|
| 4G | The fourth generation wireless communications technology |
| AWGN | Additive white Gaussian noise |
| BER | Bit error rate |
| BFFs | Beamforming filters |
| BT | Time-bandwidth |
| CC | Centroid condition |
| CIR | Channel impulse response |
| CSI | Channel state information |
| DD | Delay diversity |
| DFE | Decision-feedback equalization |
| EDGE | Enhanced data rates for GSM evolution |
| EQ | Equalizer test |
| FBF | Feedback filter |
| FFF | Feed-forward filter |
| FIR | Finite impulse response |
| GMSK | Gaussian minimum shift keying |
| GSM | Global system for mobile communication |
| GVQ | Global Vector Quantization |
| HT | Hilly terrain |

| | |
|---|---|
| IIR | Infinite impulse response |
| ISI | Intersymbol interference |
| LBG | Linde-Buzo-Gray algorithm (or generalized Lloyd algorithm) |
| LE | Linear equalization |
| LOS | Line-of-sight |
| LSSE | Least sum of squared errors |
| MIMO | Multiple–input multiple–output |
| MLSE | Maximum-likelihood sequence estimation |
| MMSE | Minimum mean-square error |
| MQE | Mean quantization error |
| MSE | Mean–square error |
| NNC | Nearest neighborhood condition |
| ODD | Optimized delay diversity |
| OFDM | Orthogonal frequency division multiplexing |
| pdf | Probability density function |
| PEP | Pairwise error probability |
| PSK | Phase shift keying |
| QAM | Quadrature amplitude modulation |
| RA | Rural area |
| SDD | Standard delay diversity |
| SIMO | Single–input multiple–output |
| SISO | Single–input single–output |
| SNR | Signal–to–noise ratio |
| SRC | Square–root raised cosine |
| TU | Typical urban area |
| WiMax | Worldwide Interoperability for Microwave Access |
| WLAN | Wireless Local Area Network |
| WMF | Whitened matched filter |
| ZF | Zero forcing |

# Operators and Notation

| | |
|---|---|
| $\lvert\cdot\rvert$ | Absolute value of a complex number |
| $\cdot \otimes \cdot$ | Convolution |
| $\delta(\cdot)$ | Dirac delta function |
| $\mathcal{E}\{\cdot\}$ | Expectation |
| $\det(\cdot)$ | Determinant |
| $\mathrm{tr}(\cdot)$ | Trace |
| $[\cdot]^*$ | Complex conjugate |
| $[\cdot]^T$ | Matrix or vector transposition |
| $[\cdot]^H$ | Matrix or vector hermitian transposition |
| $[x]^+$ | $\max(x, 0)$ |
| $\mathbf{0}_m$ | All–zero column vector of length $m$ |
| $\boldsymbol{I}_m$ | Identity matrix with dimension $m \times m$ |

# Acknowledgments

Undoubtedly, the completion of this thesis came from the support of many individuals. First and foremost, I would like to extend my sincere thanks to my supervisor Professor Robert Schober for his invaluable guidance and for giving me immense insight into my research. Professor Schober provided much of the initial motivation for pursuing this investigation and also provided priceless feedback that has improved this work in nearly every aspect. Without him, this would never have been completed.

I would also like to thank my family who have always encouraged me in my quest for higher education and especially my wife Yifan for her persistent support.

Finally, I would like to extend my thanks to the colleagues at the Department of Electrical and Computer Engineering, UBC, for creating a stimulating and a friendly environment at work, and especially Simon Yiu for his detailed explanation on the simulation program that I use for this work.

YANG WEN LIANG

*The University of British Columbia*
*Vancouver, Canada*
*August 2006*

# Chapter 1

# Introduction

The following section provides an overview of the background information and motivation for this work in detail. We review the related work that has been proposed by other researchers in this field in this section as well. The contributions of this work are briefly summarized in the second section of this chapter, and the concluding section outlines the organization of this thesis.

## 1.1 Background and Motivation

In recent years, the application of multiple antennas in wireless communication systems has received considerable interest from academia and industry [6, 7, 8]. In particular, beamforming was shown to be a simple yet efficient technique for exploiting the benefits of multiple transmit antennas, cf. e.g. [9] and references therein.

Beamforming generally requires channel state information (CSI) at the transmitter. Since perfect CSI may not be available at the transmitter in practical systems, recent research in this field has focused on beamforming with imperfect CSI, cf. e.g. [10, 11, 12]. From a practical point of view the assumption of quantized CSI is particularly interesting. In this case, the transmitter selects the beamforming vector from a pre–designed codebook based on information received from the receiver over a finite–rate feedback channel [11, 12].

Depending on the adopted performance measure and the statistical properties of the underlying channel, a closed–form solution for the optimum beamforming vector codebook may or may not exist, cf. [11, 12, 13]. For the latter case vector quantization algorithms have been proposed for codebook construction. In particular, Linde–Buzo–Gray (LBG) type algorithms [14] (also referred to as generalized Lloyd–type algorithms) have been adopted in [13] and [15], respectively. However, the LBG algorithm is a local search procedure and its performance heavily depends on the starting conditions [16].

Most of the existing literature on transmit beamforming with perfect or imperfect CSI has assumed frequency–nonselective fading. A notable exception is [17] where it was shown that beamforming with infinite impulse response (IIR) filters is asymptotically capacity achieving in strongly correlated frequency–selective multiple–input multiple–output (MIMO) fading channels. In addition, in [17] jointly optimum IIR beamforming and equalization filters were derived for various optimization criteria. For systems employing orthogonal frequency–division multiplexing (OFDM) to cope with the frequency selectivity of the channel, effective beamforming techniques for the imperfect CSI case were proposed in [18, 19].

## 1.2 Contributions

In this thesis, we consider transmit beamforming with, respectively, perfect and quantized CSI for frequency–selective fading channels which are typically encountered in high–rate transmission. We focus on single–carrier transmission and the developed beamforming schemes may be used to e.g. upgrade existing communication systems such as the Global System for Mobile Communications (GSM) and the Enhanced Data Rates for GSM Evolution (EDGE) system. Note that the multi–carrier based techniques in [18, 19] are not applicable in this case. Due to the intersymbol interference (ISI) caused by the frequency selectivity of the channel, equalization is necessary at the receiver and the

optimum beamformer depends on the equalizer used. Here, we adopt decision–feedback equalization (DFE) and linear equalization (LE) because of their low complexity, good performance, and practical relevance [20, 21, 22]. Although the main emphasis of this paper is on the practically relevant case of finite impulse response (FIR) beamforming, for the sake of completeness and since there are many interesting parallels and differences between the FIR and the IIR cases, we also consider IIR beamforming.

The main contributions of the present research work are as follows:

- For perfect CSI we derive a closed–form solution for the optimum IIR beamforming filters (BFFs) maximizing the signal–to–noise ratio (SNR) for receivers with DFE or LE. Interestingly, although our derivation of the optimum IIR BFFs is much simpler than that presented in [17] as we do not perform a *joint* optimization of the BFFs and the DFE or LE filters, our final result is identical to that given in [17]. More importantly, our approach can be readily extended to FIR BFFs which does not seem to be easily possible for the approach taken in [17].
- We show that, similar to the optimum IIR BFFs, the optimum FIR BFFs with perfect CSI are the solution to a nonlinear eigenvalue problem. However, in the FIR case a closed–form solution to this problem does not seem to exist, and we provide two efficient numerical methods for calculation of the optimum FIR BFFs.
- We propose a practical finite–rate feedback beamforming scheme for frequency–selective channels. The beamforming vector codebook design is based on the optimum FIR BFFs for perfect CSI. In particular, exploiting the findings in [23], we propose a global vector quantization (GVQ) algorithm for codebook design which performs a *deterministic* global search. The GVQ algorithm does not depend on starting conditions and employs the LBG algorithm as a local search procedure.
- Our simulation results show that short FIR BFFs can closely approach the performance of the optimum IIR BFFs. In fact, for quantized CSI

with small codebook sizes BFFs of length one (i.e., beamforming weights) are preferable. If the channel is severely frequency selective, longer BFFs become beneficial as the codebook size increases.

- For typical GSM/EDGE channel profiles beamforming with finite–rate feedback enables large performance gains compared to single–antenna transmission, transmit antenna selection, and optimized delay diversity [5].

The results of our work are summarized in the following papers:

- Y. Liang, R. Schober, and W. Gerstacker. Transmit beamforming for frequency–selective channels with decision–feedback equalization. *Submitted to IEEE Transactions on Wireless Communications*, May 2006.

- Y. Liang, R. Schober, and W. Gerstacker. FIR beamforming for frequency–selective channels with linear equalization. *Submitted to IEEE Communication Letters*, May 2006.

- Y. Liang, R. Schober, and W. Gerstacker. Transmit beamforming with finite–rate feedback for frequency–selective channels. *Accepted for presentation at the IEEE Global Telecommunications Conference (GLOBECOM), San Francisco, USA*, December 2006.

- Y. Liang, R. Schober, and W. Gerstacker. Transmit beamforming for frequency–selective channels. *Accepted for presentation at the IEEE Vehicular Technology Conference (VTC), Montreal, Canada*, September 2006.

- Y. Liang. Transmit beamforming with linear equalization. *Accepted for presentation at the First Canadian Summer School on Communications and Information Theory, Banff, Canada*, August 2006.

## 1.3 Thesis Organization

To explain the above findings in detail, the thesis is organized as follows. In Chapter 2, we describe the adopted correlated MIMO frequency–selective Rayleigh fading model, and the GSM and EDGE power delay profiles. The adopted overall communications system model is also presented in this chapter. The optimization of IIR and FIR BFFs with perfect CSI is discussed in Chapters 3 and 4, respectively. In Chapter 5, basic vector quantization concepts are introduced, finite–rate feedback beamforming is discussed, and the proposed GVQ algorithm is presented. Simulation results are provided in Chapter 6, and some conclusions and recommendations for future work are given in Chapter 7.

# Chapter 2

# Transmission System

In this chapter, the overall transmission system consisting of signal mapper, beamforming filters (BFFs), pulse shaping filters, correlated MIMO channel, receiver input filters, and equalizer will be discussed. It will be first shown that the correlated MIMO channel with $N_T$ transmit antennas and $N_R$ receive antennas can be modeled by matrices with dimension $N_R \times N_T$. We will then show that the overall channel model, continuous in time, can be obtained by convolving the correlated MIMO channel with the pulse shaping filters, and the receiver input filters. Furthermore, an overall discrete–time channel model is obtained by sampling and truncating the continuous–time channel impulse response (CIR). Finally, an equivalent channel model containing the combined effect of the overall discrete–time channel and BFFs is derived.

## 2.1 Channel Model

In order to design transmit beamforming schemes, an established knowledge of the transmission channel and its properties are crucial. The correlated MIMO frequency–selective fading channel model is adopted in this work. In a MIMO wireless link, the data stream is broken up into separate signals and sent over different transmit antennas. To get a proper understanding of this MIMO frequency–selective channel, we will first explain MIMO frequency–

nonselective channel. In the later section, we will extend our discussion from frequency–nonselective channel to frequency–selective channel.

### 2.1.1 Frequency–nonselective Channel

The complex baseband frequency–nonselective MIMO channel can be modeled by the following channel matrix [24]:

$$\boldsymbol{H}_C(t) = \begin{bmatrix} h_C^{11}(t) & h_C^{12}(t) & \dots & h_C^{1N_T}(t) \\ h_C^{21}(t) & h_C^{22}(t) & \dots & h_C^{2N_T}(t) \\ \vdots & & & \vdots \\ h_C^{N_R1}(t) & h_C^{N_R2}(t) & \dots & h_C^{N_RN_T}(t) \end{bmatrix}. \tag{2.1}$$

$h_C^{n_rn_t}(t)$ is the continuous–time channel gain between transmit antenna $n_t$, $1 \leq n_t \leq N_T$, and receive antenna $n_r$, $1 \leq n_r \leq N_R$, where $N_T$ and $N_R$ are the total number of transmit and receive antennas, respectively. Furthermore, $h_C^{n_rn_t}(t)$ can be modeled as a Rayleigh fading channel or a Rician fading channel, if there is a line–of–sight (LOS) path.

In case of Rayleigh fading channels, the complex gain $h_C^{n_rn_t}(t)$ can be described as a continuous–time zero mean Gaussian random process

$$h_C^{n_rn_t}(t) = h_I^{n_rn_t}(t) + jh_Q^{n_rn_t}(t), \tag{2.2}$$

where $h_I^{n_rn_t}(t)$ and $h_Q^{n_rn_t}(t)$ are the real and imaginary parts of $h_C^{n_rn_t}(t)$, respectively [25]. The envelope of the process, $\zeta^{n_rn_t}(t) = |h_C^{n_rn_t}(t)|$, is Rayleigh distributed with probability density function (pdf)

$$p_\zeta(x) = \begin{cases} \frac{x}{\sigma_0^2} \exp\left(-\frac{x^2}{2\sigma_0^2}\right), & \text{for} \quad x \geq 0 \\ 0, & \text{for} \quad x < 0 \end{cases}, \tag{2.3}$$

where $\sigma_0^2$ is the variance of the two quadrature channels. Since $h_I^{n_rn_t}(t)$ and $h_Q^{n_rn_t}(t)$ are assumed to be independent and identically distributed, the variance of $h_C^{n_rn_t}(t)$ is equal to $2\sigma_0^2$.

In case of Rician fading channels, $h_C^{n_r n_t}(t)$ has a time dependent complex mean value $\hat{m}(t)$, and can be modeled as

$$\begin{aligned} h_C^{n_r n_t}(t) &= h_I^{n_r n_t}(t) + j h_Q^{n_r n_t}(t) \\ &= \tilde{h}_I^{n_r n_t}(t) + j \tilde{h}_Q^{n_r n_t}(t) + \hat{m}(t), \end{aligned} \tag{2.4}$$

where $h_I^{n_r n_t}(t)$ and $h_Q^{n_r n_t}(t)$ are the real and imaginary parts of $h_C^{n_r n_t}(t)$, respectively [25]. Furthermore, $\tilde{h}_I^{n_r n_t}(t)$ and $\tilde{h}_Q^{n_r n_t}(t)$ are statistically independent zero mean Gaussian random processes with common variance $\sigma_0^2$. The envelope of the process, $\zeta^{n_r n_t}(t) = |h_C^{n_r n_t}(t)|$, is now Rice distributed with probability density function (pdf)

$$p_\zeta(x) = \begin{cases} \frac{x}{\sigma_0^2} \exp\left(-\frac{x^2+\rho^2}{2\sigma_0^2}\right) I_0\left(\frac{x\rho}{\sigma_0^2}\right), & \text{for} \quad x \geq 0 \\ 0, & \text{for} \quad x < 0 \end{cases}, \tag{2.5}$$

where $I_0(x)$ is the modified zero order Bessel function of the first kind, $\sigma_0^2$ is the variance of the two quadrature channels, and $\rho^2 = \mathcal{E}^2\{h_I^{n_r n_t}(t)\} + \mathcal{E}^2\{h_Q^{n_r n_t}(t)\}$, where $\mathcal{E}\{\cdot\}$ denotes statistical expectation. Again, since $h_I^{n_r n_t}(t)$ and $h_Q^{n_r n_t}(t)$ are assumed to be independent, the variance of $h_C^{n_r n_t}(t)$ is equal to $2\sigma_0^2$. From Eq. (2.5), the Rice pdf converges to the Rayleigh pdf for $\rho \to 0$.

### 2.1.2 Frequency–selective Channel

The frequency–nonselective model described by Eq. (2.1) is only valid when the signal bandwidth is much smaller than the coherence bandwidth of the channel. If the signal has a bandwidth greater than the coherence bandwidth, the transmitted signal is subjected to different gains and phase shifts across the band. In such a case, the channel is said to be frequency–selective [25]. A frequency–selective channel causes intersymbol interference (ISI). The received signal will be the superposition of several transmitted signals.

A frequency–selective MIMO model with $L$ multipath components is shown in Figure 2.1. $x_{n_t}(t)$ represents the signal transmitted by transmit antenna $n_t$, while $y_{n_r}(t)$ represents the signal received by receive antenna $n_r$. $\tau_l$, $l =$

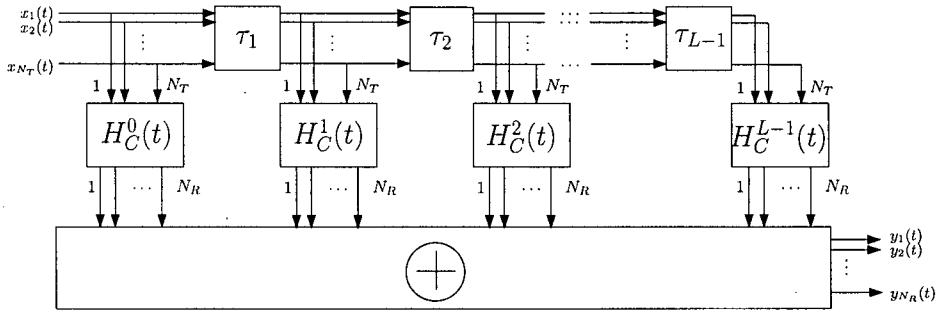


Figure 2.1: Frequency selective MIMO channel.

$1, \ldots, L-1$, represents the delay of the multipath component $l$. Each matrix, $\boldsymbol{H}_C^l(t)$, $l = 0, \ldots, L-1$, has dimension $N_R \times N_T$ and its elements can be written as [24, 25]

$$\boldsymbol{H}_C^l(t) = \begin{bmatrix} h_C^{11,l}(t) & h_C^{12,l}(t) & \ldots & h_C^{1N_T,l}(t) \\ h_C^{21,l}(t) & h_C^{22,l}(t) & \ldots & h_C^{2N_T,l}(t) \\ \vdots & \vdots & \ddots & \vdots \\ h_C^{N_R1,l}(t) & h_C^{N_R2,l}(t) & \ldots & h_C^{N_RN_T,l}(t) \end{bmatrix}. \tag{2.6}$$

The matrices $\boldsymbol{H}_C^l(t)$ are independent for different $l$, $l = 0, \ldots, L-1$, and their elements, $h_C^{n_rn_t,l}(t)$, are continuous–time random processes as defined in Eq. (2.2) or Eq. (2.4).

The overall MIMO weight function $\boldsymbol{H}_C(\tau, t)$ is also a matrix with dimension $N_R \times N_T$. It relates to the matrices $\boldsymbol{H}_C^l(t)$ in the following way

$$\boldsymbol{H}_C(\tau, t) = \sum_{l=0}^{L-1} \boldsymbol{H}_C^l(t)\delta(\tau - \tau_l), \tag{2.7}$$

where $\delta(\cdot)$ is the Dirac delta function [25] and $\tau_0$ is equal to zero.

Therefore, the matrix elements in Eqs. (2.6) and (2.7) are related by

$$h_C^{n_rn_t}(\tau, t) = \sum_{l=0}^{L-1} h_C^{n_rn_t,l}(t)\delta(\tau - \tau_l). \tag{2.8}$$

Adopting the power delay profile given in [26], we define the equivalent power delay profile of the channel for each receive–transmit antenna pair as

$$p(\tau) = \sum_{l=0}^{L-1} \left(\sigma_C^{n_rn_t,l}\right)^2 \delta(\tau - \tau_l), \tag{2.9}$$

where $\left(\sigma_C^{n_r n_t, l}\right)^2$ is the variance of $h_C^{n_r n_t, l}(t)$ defined as

$$\left(\sigma_C^{n_r n_t, l}\right)^2 = \mathcal{E}\left\{\left|h_C^{n_r n_t, l}(t)\right|^2\right\}. \tag{2.10}$$

In practice, the summation of $\left(\sigma_C^{n_r n_t, l}\right)^2$ with respect to $l$ is normalized to 1, i.e.,

$$\sum_{l=0}^{L-1}\left(\sigma_C^{n_r n_t, l}\right)^2 = 1. \tag{2.11}$$

For GSM and EDGE system, four different power delay profiles are specified [1]: rural area (RA), hilly terrain (HT), typical urban area (TU) and equalizer test (EQ). For EQ, HT, and TU it is assumed that the amplitudes of all propagation paths, $h_C^{n_r n_t, l}(t)$, are continuous–time zero mean Gaussian random processes as described by Eq. (2.2). Their envelopes are Rayleigh distributed with pdf as defined in Eq. (2.3). For RA, it is assumed that the amplitudes of all propagation paths are continuous–time non–zero mean Gaussian random processes as defined in Eqs. (2.4) and (2.5). The mean value is due to the LOS path between a transmit antenna and a receive antenna. This results in a Rician fading channel. In this work, the EQ, HT, and TU profiles are considered.

Finally, it should be noted that if $N_T$ and $N_R$ are both equal to one, the MIMO channel in Eq. (2.6) reduces to a single–input single–output (SISO) frequency–selective fading channel. Furthermore, if $L = 1$, the channel reduces to a frequency–nonselective channel resulting in only scalar multiplicative distortion of the transmitted signals.

## 2.2 Correlation of CIR Coefficients

In general, an independent and identically distributed (i.i.d.) channel model assuming rich uniform scattering will not be an accurate description of real–world multi–antenna channels [27], since in practice, insufficient antenna spacing and a lack of scattering cause the antennas to be correlated. Therefore,

spatial correlation is assumed to occur at both the transmit and receive antennas in this work. Under this assumption, the matrix taps in Eq. (2.6) can be written as [27]

$$\boldsymbol{H}_C^l(t) = \boldsymbol{R}^{1/2}\boldsymbol{H}^l(t)(\boldsymbol{S}^{1/2})^H, \tag{2.12}$$

where $\boldsymbol{H}^l(t)$, $\boldsymbol{R} = \boldsymbol{R}^{1/2}(\boldsymbol{R}^{1/2})^H$, and $\boldsymbol{S} = \boldsymbol{S}^{1/2}(\boldsymbol{S}^{1/2})^H$ are the channel matrix taps with i.i.d. entries, the receive correlation matrix, and the transmit correlation matrix, respectively. The superscripts $1/2$ and $H$ denote the matrix square–root and Hermitian transposition, respectively. Although not completely general, this simple correlation model has been validated through recent field measurements as a sufficiently accurate representation of the fade correlations seen in actual cellular systems [27, 28]. $\boldsymbol{S}$ and $\boldsymbol{R}$ are positive definite matrices with dimensions $N_T \times N_T$ and $N_R \times N_R$, respectively.

From now on, we assume the MIMO model defined in Eq. (2.7) to be a spatially correlated frequency–selective MIMO channel with matrix taps described by Eq. (2.12). For simplicity, we assume that the spatial correlation is identical for all matrix taps. Setups with up to three transmit and two receive antennas are considered in this work. Since matrices $\boldsymbol{S}$ and $\boldsymbol{R}$ have the same form, we will concentrate on the transmit correlation matrix in the following discussion.

There is only one correlation factor for the two antennas case. The correlation matrix $\boldsymbol{S}$ can be written as

$$\boldsymbol{S} = \begin{bmatrix} 1 & \rho_{12}^t \\ \rho_{12}^t & 1 \end{bmatrix}, \tag{2.13}$$

where $\rho_{12}^t$ is the correlation factor between transmit antenna one and transmit antenna two and it is defined by

$$\rho_{12}^t = \frac{\mathcal{E}\left\{h_C^{n_r1,l}(t)\left(h_C^{n_r2,l}(t)\right)^*\right\}}{\sqrt{(\sigma_C^{n_r1,l})^2(\sigma_C^{n_r2,l})^2}}. \tag{2.14}$$

There are three correlation factors for the three antennas case, $\rho_{12}^t$, $\rho_{23}^t$, and $\rho_{13}^t$. They represent the correlation between transmit antenna one and transmit

antenna two, between transmit antenna two and transmit antenna three, and between transmit antenna one and transmit antenna three, respectively. The resulting correlation matrix is a $3 \times 3$ matrix with elements similarly defined as in Eq. (2.14)

$$\boldsymbol{S} = \begin{bmatrix} 1 & \rho_{12}^t & \rho_{13}^t \\ \rho_{12}^t & 1 & \rho_{23}^t \\ \rho_{13}^t & \rho_{23}^t & 1 \end{bmatrix}. \tag{2.15}$$

The square root of the correlation matrix can be calculated by using Cholesky decomposition such that $\boldsymbol{S}^{1/2}$ and $\boldsymbol{R}^{1/2}$ are lower triangular matrices whereas $(\boldsymbol{S}^{1/2})^H$ and $(\boldsymbol{R}^{1/2})^H$ are upper triangular matrices [29]. $\boldsymbol{S}^{1/2}$ for the two and three transmit antennas case can be written as

$$\boldsymbol{S}^{1/2} = \begin{bmatrix} 1 & 0 \\ \rho_{12}^t & \sqrt{1-(\rho_{12}^t)^2} \end{bmatrix}, \tag{2.16}$$

and

$$\boldsymbol{S}^{1/2} = \begin{bmatrix} 1 & 0 & 0 \\ \rho_{12}^t & \sqrt{1-(\rho_{12}^t)^2} & 0 \\ \rho_{13}^t & \frac{\rho_{23}^t - \rho_{12}^t \rho_{13}^t}{\sqrt{1-(\rho_{12}^t)^2}} & \sqrt{1-(\rho_{13})^2 - \frac{(\rho_{23}^t - \rho_{12}^t \rho_{13}^t)^2}{1-(\rho_{12}^t)^2}} \end{bmatrix}, \tag{2.17}$$

respectively. Similar results can be obtained for the receive antennas by replacing the correlation factors in the above matrices with the respective receive correlation factors. For future convenience,

$$\boldsymbol{\rho}_t = \begin{bmatrix} \rho_{12}^t & \rho_{13}^t & \rho_{23}^t \end{bmatrix} \tag{2.18}$$

and

$$\boldsymbol{\rho}_r = \begin{bmatrix} \rho_{12}^r & \rho_{13}^r & \rho_{23}^r \end{bmatrix} \tag{2.19}$$

are defined, where $\boldsymbol{\rho}_t$ and $\boldsymbol{\rho}_r$ are the correlation vectors for the transmit and receive antennas, respectively.

## 2.3 Equivalent Baseband System Model

The channel model presented in the previous section is only a part of the overall mobile communications transmission model. The other parts of the

model are discussed in this section. The block diagram of the continuous–time overall transmission model in complex baseband representation is shown in Figure 2.2.

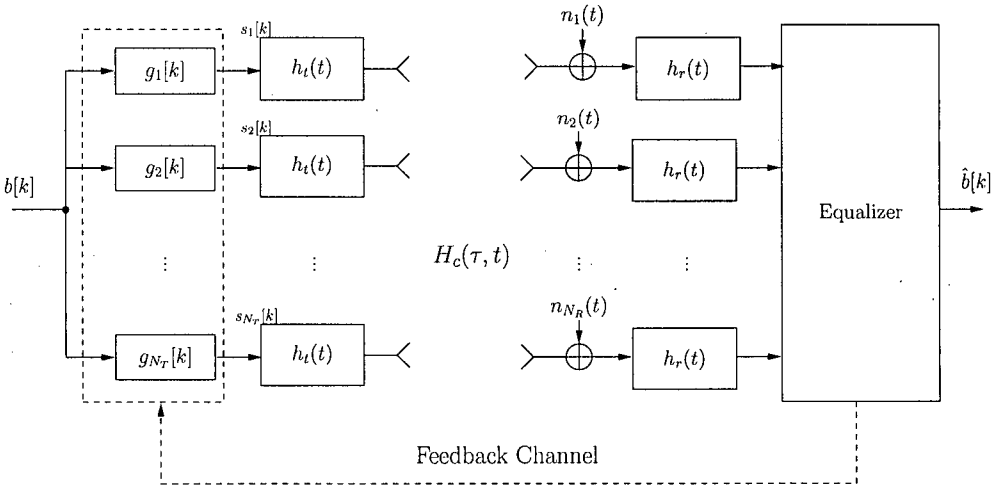


Figure 2.2: Block diagram of the continuous–time overall transmission system. $\hat{b}[k]$ are estimated symbols at the receiver.

The i.i.d. symbols $b[k]$ are taken from a scalar symbol alphabet $\mathcal{A}$ such as phase–shift keying (PSK) or quadrature amplitude modulation (QAM), and have variance $\sigma_b^2 \triangleq \mathcal{E}\{|b[k]|^2\} = 1$. Since we consider GSM and EDGE in this work, the mapped symbol is either a 2–PSK or an 8–PSK symbol depending on whether GSM or EDGE is used. Note that GSM uses binary Gaussian Minimum Shift Keying (GMSK), which can be approximated as filtered 2–PSK. EDGE improves spectral efficiency by employing 8–PSK modulation instead. However, other system parameters such as symbol rate and burst duration remain unchanged in order to enable a smooth transition from GSM to EDGE [2].

Before transmit pulse shaping, the modulated symbols, $b[k]$, are first filtered by the discrete–time beamforming filters (BFFs), $g_{n_t}[k]$. The BFFs depicted in Figure 2.2 are discrete–time filters, which can be realized as tapped delay lines. The transmit BFF impulse response coefficients of antenna $n_t$,

$1 \leq n_t \leq N_T$, are denoted by $g_{n_t}[k]$, $-q_l \leq k \leq q_u$, and their energy is normalized to $\sum_{n_t=1}^{N_T} \sum_{k=-q_l}^{q_u} |g_{n_t}[k]|^2 = 1$. For infinite impulse response (IIR) BFFs $q_l \to \infty$ and $q_u \to \infty$ and for finite impulse response (FIR) BFFs $q_l = 0$ and $q_u = L_g - 1$, where $L_g$ is the FIR BFF length. For future reference we define the FIR BFF vector $\boldsymbol{g} \triangleq [g_1[0]\ g_1[1]\ \ldots\ g_1[L_g-1]\ g_2[0]\ \ldots\ g_{N_T}[L_g-1]]^T$ of length $N_T L_g$.

As a result, the filtered symbols $s_{n_t}[k]$ of antenna $n_t$ can be obtained for the modulated symbols $b[k]$ at time $k$ by

$$s_{n_t}[k] = g_{n_t}[k] \otimes b[k], \tag{2.20}$$

where $\otimes$ refers to convolution.

### 2.3.1 Transmit Pulse Shaping

For transmit pulse shaping, the linearized impulse $h_t(t)$ corresponding to GMSK with time–bandwidth (BT) product 0.3 is employed [25]. Therefore, the transmit filter impulse response is given by [2, 25]

$$h_t(t) = \begin{cases} \prod_{k=0}^{3} s(t+kT), & 0 \leq t \leq 5T; \\ 0, & \text{otherwise}, \end{cases} \tag{2.21}$$

with

$$s(t) = \begin{cases} \sin\left(\pi \int_0^t \dot{g}(\tau)d\tau\right), & 0 \leq t \leq 4T; \\ \sin\left(\frac{\pi}{2} - \pi \int_0^{t-4T} \dot{g}(\tau)d\tau\right), & 4T \leq t \leq 8T; \\ 0, & \text{otherwise}, \end{cases} \tag{2.22}$$

where $T = 3.69\mu s$ is the symbol duration. The impulse $\dot{g}(t)$ of duration $4T$ is given by

$$\begin{aligned} \dot{g}(t) = &\frac{1}{2T}\left[Q\left(2\pi \cdot 0.3 \frac{t - \frac{5T}{2}}{T\sqrt{\ln(2)}}\right)\right. \\ &\left. - Q\left(2\pi \cdot 0.3 \frac{t - \frac{3T}{2}}{T\sqrt{\ln(2)}}\right)\right], \quad 0 \leq t \leq 4T, \end{aligned} \tag{2.23}$$

where $Q(\cdot)$ denotes the complementary Gaussian error integral [25],

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_{t}^{+\infty} \mathrm{e}^{-\tau^2/2} d\tau. \tag{2.24}$$

By this choice of the transmit filter for 2–PSK and 8–PSK symbols, the transmit spectra of EDGE and GSM are approximately equal, and the spectral masks of GSM are fulfilled [2].

### 2.3.2 Receiver Processing

The continuous–time signals are transmitted over the correlated MIMO channel $\boldsymbol{H}_C(\tau, t)$ discussed in the last section. At the receiver, the continuous–time received signal at antenna $n_r$ is impaired by additive white Gaussian noise (AWGN) $n_{n_r}(t)$. The choice of the receiver input filter, $h_r(t)$, is up to the receiver designer. We assume a filter with square–root Nyquist frequency response. This allows us to model the channel noise after sampling as a spatially and temporally white discrete–time Gaussian random process, which will be discussed in Section 2.4.

Two filters which have a square–root Nyquist frequency response are the whitened matched filter (WMF) [25], which belongs to the class of optimum receiver input filters [30], and the square–root raised cosine (SRC) filter [25, 26]. We use a fixed filter in this work, namely the SRC receive filter with roll–off factor 0.3 [26]. This filter offers a similar performance as the optimum WMF. However, the implementation of the SRC filter is much simpler because, in contrast to the WMF, it does not have to be adapted to a particular channel impulse response [2]. The discrete–time received signals are obtained by sampling the output of the receiver input filters at times $t = kT$. Finally, the receiver, assumed to have perfect knowledge of the overall channel impulse response (CIR), performs equalization of the received signals and estimates the transmit symbols. In other words, ISI can be mitigated by employing an equalizer at the receiver side. Maximum–likelihood sequence estimation (MLSE),

decision–feedback equalization (DFE), and linear equalization (LE) are some of the equalization methods which are commonly used in practice. We adopt DFE and LE at the receiver in this work due to their low complexity, good performance, and practical relevance.

It should be noted that the BFFs $g_{n_t}[k]$, the pulse shaping filters $h_t(t)$, and the receiver input filters $h_r(t)$, introduce additional ISI to the MIMO channel. In addition, the pulse shaping filters and receiver input filters introduce temporal correlation to the channel.

## 2.4 Equivalent Discrete–Time System Model

The overall channel model discussed in the previous section is in continuous–time and contains different blocks including the pulse shaping filters $h_t(t)$, the physical channel $\boldsymbol{H}(\tau, t)$, and the receiver input filters $h_r(t)$. It is convenient to derive an equivalent discrete–time model containing the combined effects of all these blocks. In this section, we will show how the discrete–time model can be obtained. Let us consider again a MIMO system with $N_T$ transmit and $N_R$ receive antennas. The block diagram of the discrete–time overall transmission system in complex baseband representation is shown in Figure 2.3.

### 2.4.1 Discrete–time Channel Model

In this work, block fading is assumed. That is, the wireless channel coefficients $h_C^{n_r n_t, l}(t)$ defined in Eq. (2.6) are approximately constant during one burst but vary from burst to burst. In other words, the coefficients $h_C^{n_r n_t, l}(t)$ are time-invariant within each burst. This assumption is valid for small–to–moderate burst lengths and low vehicle speeds. With this assumption, the time depen-

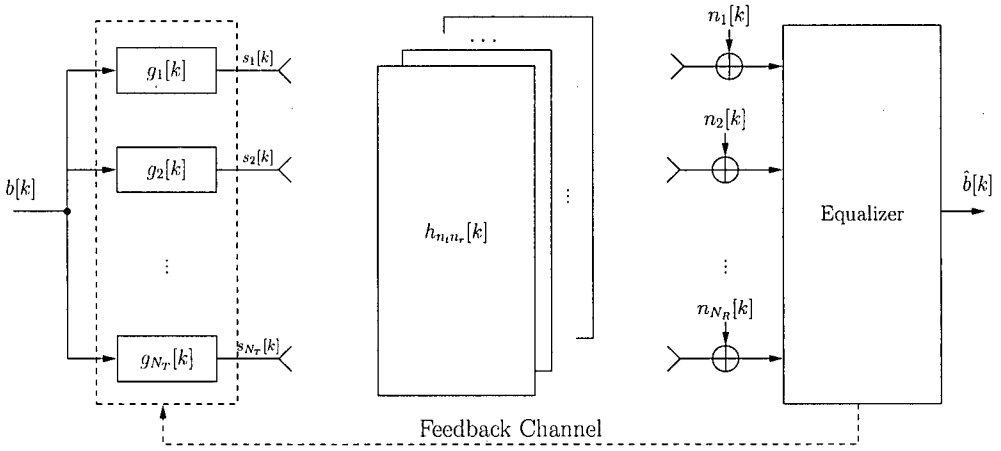


Figure 2.3: Block diagram of the discrete–time overall transmission system. $\hat{b}[k]$ are estimated symbols at the receiver.

dence of $h_C^{n_r n_t,\,l}(t)$ can be dropped and Eq. (2.6) reduces to

$$\boldsymbol{H}_C^l = \begin{bmatrix} h_C^{11,l} & h_C^{12,l} & \dots & h_C^{1N_T,l} \\ h_C^{21,l} & h_C^{22,l} & \dots & h_C^{2N_T,l} \\ \vdots & \vdots & \ddots & \vdots \\ h_C^{N_R 1,l} & h_C^{N_R 2,l} & \dots & h_C^{N_R N_T,l} \end{bmatrix}. \tag{2.25}$$

Now, the continuous–time overall CIR can be obtained from

$$h_{n_r n_t}(t) = h_t(t) \otimes h_C^{n_r n_t}(t) \otimes h_r(t), \tag{2.26}$$

where

$$h_C^{n_r n_t}(t) = \sum_{l=0}^{L-1} h_C^{n_r n_t,\,l}\,\delta(t - \tau_l). \tag{2.27}$$

One can also obtain the above equation from Eq. (2.8). Since the channel is time–invariant, $t$ in Eq. (2.8) is fixed and can be dropped from the equation. Therefore, the only variable left is $\tau$. Replacing $\tau$ with $t$ yields $h_C^{n_r n_t}(t)$.

In principle, the overall CIR is of infinite length. However, in practice, it can be sampled and truncated to $L$ consecutive taps which exhibit maximum energy [31]. Therefore, the sampled and truncated overall CIR can be written as

$$h_{n_r n_t}[l] \triangleq h_{n_r n_t}(lT + t_0), \qquad l = 0, \dots, L-1, \tag{2.28}$$

where $t_0$ is a small time delay, and $T$ is the symbol duration. $L$ and $t_0$ are chosen so that only a negligible amount of power is disregarded.

### 2.4.2 Equivalent SIMO Model

With this discrete–time channel model, the $T$–spaced, sampled version of the received signal at receive antenna $n_r$, $1 \leq n_r \leq N_R$, can be modeled as

$$\begin{aligned} r_{n_r}[k] &= \sum_{n_t=1}^{N_T} \sum_{l=0}^{L-1} h_{n_r n_t}[l] s_{n_t}[k-l] + n_{n_r}[k] \\ &= \sum_{n_t=1}^{N_T} h_{n_t n_r}[k] \otimes s_{n_t}[k] + n_{n_r}[k], \end{aligned} \tag{2.29}$$

where $s_{n_t}[k]$ is defined in Eq. (2.20) and $n_{n_r}[k]$ denotes additive (spatially and temporally) white Gaussian noise (AWGN) with variance $\sigma_n^2 \triangleq \mathcal{E}\{|n_{n_r}[k]|^2\} = N_0$, where $N_0$ denotes the single–sided power spectral density of the underlying continuous–time passband noise process. Note that $n_{n_r}[k] = n_{n_r}(kT + t_0)$ in Eq. (2.29) is spatially and temporally white because the SRC receive filter autocorrelation function fulfills the first Nyquist criterion [25].

Likewise, Eq. (2.29) can be written as

$$\begin{aligned} r_{n_r}[k] &= \sum_{n_t=1}^{N_T} h_{n_r n_t}[k] \otimes b[k] \otimes g_{n_t}[k] + n_{n_r}[k] \\ &= h_{n_r}^{\text{eq}}[k] \otimes b[k] + n_{n_r}[k] \\ &= \sum_{l=0}^{L+L_g-2} h_{n_r}^{\text{eq}}[l] b[k-l] + n_{n_r}[k], \end{aligned} \tag{2.30}$$

where $h_{n_r}^{\text{eq}}[k]$ is the equivalent CIR with length $L_{\text{eq}} = L + N - 1$ corresponding to receive antenna $n_r$ and is defined as

$$h_{n_r}^{\text{eq}}[k] \triangleq \sum_{n_t=1}^{N_T} h_{n_r n_t}[k] \otimes g_{n_t}[k]. \tag{2.31}$$

Eq. (2.30) shows that a MIMO system with beamforming can be modeled as an equivalent single–input multiple–output (SIMO) system. Therefore, at the receiver the same equalization, channel estimation, and channel tracking

techniques as for single–antenna transmission can be used. Here, we adopt receive–diversity zero–forcing (ZF) or minimum mean–squared error (MMSE) DFE [21] and LE [22] for detection and assume perfect channel estimation at the receiver.

### 2.4.3 Feedback Channel

Furthermore, we assume that a feedback channel from the receiver to the transmitter is available, cf. Figure 2.3. In case of perfect CSI, the receiver sends the optimum BFFs (or equivalently the CIR) to the transmitter. If the feedback channel allows only the transmission of $B$ bits per channel update, the receiver and the transmitter have to agree on a pre–designed BFF vector codebook $\mathcal{G} \triangleq \{\hat{\boldsymbol{g}}_1, \hat{\boldsymbol{g}}_2, \ldots, \hat{\boldsymbol{g}}_N\}$ of size $N = 2^B$, where $\hat{\boldsymbol{g}}_n$ is a vector of length $N_T L_g$ and $\hat{\boldsymbol{g}}_n^H \hat{\boldsymbol{g}}_n = 1$, $1 \leq n \leq N$. For a given CIR, the receiver determines the address $n$ of the codeword (BFF vector) $\hat{\boldsymbol{g}}_n \in \mathcal{G}$, $1 \leq n \leq N$, which yields the minimum bit error rate (BER). Subsequently, address $n$ is sent to the transmitter which then utilizes $\boldsymbol{g} = \hat{\boldsymbol{g}}_n$ for beamforming. Since the primary goal of this work is to investigate the achievable performance of beamforming with, respectively, perfect and quantized CSI at the transmitter, similar to [11, 12, 13] we assume that the feedback channel is error–free and has zero delay.

The discussions in this chapter are valid for both GSM and EDGE systems because they use the same frequency bands, transmit pulse shaping filters, and receiver input filters [1]. It should also be noted that the model is not restricted to GSM and EDGE systems, but is applicable to any system that employs linear single–carrier modulation and has a feedback channel.

# Chapter 3

# Beamforming with Perfect CSI and IIR Filters

In this section, we assume perfect CSI and IIR BFFs. For this scenario jointly optimum BFFs and equalization filters were derived in [17, Section IV]. Thereby, the BFFs and equalization filters were optimized for a fixed overall CIR (including the BFFs, the channel, and the equalization filters), and subsequently this overall CIR was chosen or optimized. Here, we pursue a much simpler and more straightforward approach. In particular, we assume that the receiver employs the optimum DFE (ZF–DFE or MMSE–DFE) filters or LE (ZF–LE or MMSE–LE) filters for given MIMO channel and BFFs, and optimize the BFFs for maximization of the SNR under this assumption. The SNRs achievable with IIR BFFs given in this chapter will serve as upper bounds for the SNRs achievable with the FIR BFFs, which will be discussed in Chapter 4 and 5. For convenience the frequency responses of the IIR BFFs $G_{n_t}(f) \triangleq \mathcal{F}\{g_{n_t}[k]\}$ are collected in vector $\boldsymbol{G}(f) \triangleq [G_1(f)\ G_2(f)\ \ldots\ G_{N_T}(f)]^T$ in the following sections.

## 3.1 IIR Beamforming with Decision–Feedback Equalization

In this section, we consider IIR beamforming with ZF– and MMSE–DFE at the receiver. The optimization problem is carefully stated first, then the optimum BFFs for maximization of the SNR of a given channel realization are derived under these assumptions.

### 3.1.1 Optimization Problem

For a given channel realization and a given beamforming vector $\boldsymbol{G}(f)$, the unbiased SNR for DFE with optimum IIR feedforward and corresponding feedback filtering is given by [20, 21]

$$\mathrm{SNR}(\boldsymbol{G}(f)) = \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left[ \xi + \sum_{n_r=1}^{N_R} |H_{n_r}^{\mathrm{eq}}(f)|^2 \right] \mathrm{d}f \right\} - \chi, \tag{3.1}$$

where $\chi = 0$, $\xi = 0$ and $\chi = 1$, $\xi = \sigma_n^2/\sigma_b^2$ for ZF- and MMSE–DFE filter optimization, respectively. In Eq. (3.1), the equivalent channel frequency response $H_{n_r}^{\mathrm{eq}}(f) \triangleq \mathcal{F}\{h_{n_r}^{\mathrm{eq}}[k]\}$ is given by $H_{n_r}^{\mathrm{eq}}(f) = \sum_{n_t=1}^{N_T} G_{n_t}(f) H_{n_t n_r}(f)$ with $H_{n_t n_r}(f) \triangleq \mathcal{F}\{h_{n_t n_r}[k]\}$.

The optimum BFF vector $\bar{\boldsymbol{G}}(f) \triangleq [\bar{G}_1(f)\ \bar{G}_2(f),\ \ldots\ \bar{G}_{N_T}(f)]^T$ shall maximize $\mathrm{SNR}(\boldsymbol{G}(f))$ subject to the transmit power constraint

$$\int_{-1/2}^{1/2} \boldsymbol{G}^H(f)\boldsymbol{G}(f)\,\mathrm{d}f = 1. \tag{3.2}$$

A convenient approach for calculating $\bar{\boldsymbol{G}}(f)$ is the classical Calculus of Variations method [32]. Following this method, we model the BFF of antenna $n_t$ as $G_{n_t}(f) = \bar{G}_{n_t}(f) + \varepsilon_{n_t} B_{n_t}(f)$, where $B_{n_t}(f)$ and $\varepsilon_{n_t}$ denote an arbitrary function of $f$ and a complex–valued variable, respectively. The optimization

problem can now be formulated in terms of its Lagrangian

$$\begin{aligned} L(\boldsymbol{\varepsilon}) =& \mathrm{SNR}(\boldsymbol{G}(f)) - \mu \int_{-1/2}^{1/2} \boldsymbol{G}^H(f)\boldsymbol{G}(f)\,\mathrm{d}f \\ =& \int_{-1/2}^{1/2} \ln \left[ \frac{\sigma_b^2 \sum_{n_r=1}^{N_R} \left| \sum_{n_t=1}^{N_T} \left( \bar{G}_{nt}(f) + \varepsilon_{nt} B_{nt}(f) \right) H_{n_t n_r}(f) \right|^2}{\sigma_n^2} + \chi \right] \mathrm{d}f \\ & - \mu \int_{-1/2}^{1/2} \sum_{n_t=1}^{N_T} \left| \bar{G}_{n_t}(f) + \varepsilon_{nt} B_{nt}(f) \right|^2 \mathrm{d}f \end{aligned} \tag{3.3}$$

where $\boldsymbol{\varepsilon} \triangleq [\varepsilon_1 \ \varepsilon_2 \ \ldots \ \varepsilon_{N_T}]^T$ and $\mu$ is the Lagrange multiplier. The optimum BFF vector $\bar{\boldsymbol{G}}(f)$ has to fulfill the condition [32]

$$\left. \frac{\partial L(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}^*} \right|_{\boldsymbol{\varepsilon}=\mathbf{0}_{N_T}} = \mathbf{0}_{N_T}, \tag{3.4}$$

for arbitrary $B_{n_t}(f)$, $1 \le n_t \le N_T$. Therefore, for all $s = 1, 2, \ldots, N_T$, Eq. (3.4) can be written as

$$\begin{aligned} & \left. \frac{\partial L(\boldsymbol{\varepsilon})}{\partial \varepsilon_s^*} \right|_{\boldsymbol{\varepsilon}=\mathbf{0}_{N_T}} \\ =& \int_{-1/2}^{1/2} \frac{\sum_{n_r=1}^{N_R} B_s^*(f) H_{sn_r}^*(f) \left[ \sum_{n_t=1}^{N_T} \left( \bar{G}_{nt}(f) + \varepsilon_{nt} B_{nt}(f) \right) H_{n_t n_r}(f) \right]}{\sum_{n_r=1}^{N_R} \left| \sum_{n_t=1}^{N_T} \left( \bar{G}_{nt}(f) + \varepsilon_{nt} B_{nt}(f) \right) H_{n_t n_r}(f) \right|^2 + \xi} \mathrm{d}f \Bigg|_{\boldsymbol{\varepsilon}=\mathbf{0}_{N_T}} \\ & - \mu \int_{-1/2}^{1/2} (\bar{G}_s(f) + \varepsilon_s B_s(f)) B_s^*(f)\,\mathrm{d}f \Big|_{\boldsymbol{\varepsilon}=\mathbf{0}_{N_T}} \\ =& \int_{-1/2}^{1/2} \frac{\sum_{n_r=1}^{N_R} B_s^*(f) H_{sn_r}^*(f) \left[ \sum_{n_t=1}^{N_T} \bar{G}_{nt}(f) H_{n_t n_r}(f) \right]}{\sum_{n_r=1}^{N_R} \left| \sum_{n_t=1}^{N_T} \bar{G}_{nt}(f) H_{n_t n_r}(f) \right|^2 + \xi} \mathrm{d}f \\ & - \mu \int_{-1/2}^{1/2} \bar{G}_s(f) B_s^*(f) \mathrm{d}f \\ =& 0, \end{aligned} \tag{3.5}$$

where $[\cdot]^*$ denotes complex conjugate. In order to achieve Eq. (3.5), the fol-

lowing condition must be satisfied for all $s = 1, 2, \ldots, N_T$,

$$\mu \bar{G}_s(f) \left( \sum_{n_r=1}^{N_R} \left| \sum_{n_t'=1}^{N_T} \bar{G}_{n_t'}(f) H_{n_t' n_r}(f) \right|^2 + \xi \right) = \sum_{n_r=1}^{N_R} H_{s n_r}^*(f) \left( \sum_{n_t=1}^{N_T} \bar{G}_{n_t}(f) H_{n_t n_r}(f) \right), \tag{3.6}$$

or equivalently,

$$\mu \bar{G}_s(f) \left( \sum_{n_r=1}^{N_R} \left| \sum_{n_t'=1}^{N_T} \bar{G}_{n_t'}(f) H_{n_t' n_r}(f) \right|^2 + \xi \right) = \sum_{n_t=1}^{N_T} \bar{G}_{n_t}(f) \left( \sum_{n_r=1}^{N_R} H_{s n_r}^*(f) H_{n_t n_r}(f) \right). \tag{3.7}$$

### 3.1.2 Optimum IIR BFFs

Combining Eqs. (3.4) and (3.7), it can be shown that vector $\bar{\boldsymbol{G}}(f)$ has to fulfill

$$\boldsymbol{S}(f)\, \bar{\boldsymbol{G}}(f) = \tilde{\mu} \left[ \bar{\boldsymbol{G}}^H(f) \boldsymbol{S}(f) \bar{\boldsymbol{G}}(f) + \xi \right] \bar{\boldsymbol{G}}(f), \tag{3.8}$$

where $\tilde{\mu}$ is a constant and $\boldsymbol{S}(f)$ is an $N_T \times N_T$ matrix given by

$$\begin{aligned} \boldsymbol{S}(f) &\triangleq \sum_{n_r=1}^{N_R} \boldsymbol{H}_{n_r}(f) \boldsymbol{H}_{n_r}^H(f) \\ &= \begin{bmatrix} \sum_{n_r=1}^{N_R} H_{1 n_r}^*(f) H_{1 n_r}(f) & \sum_{n_r=1}^{N_R} H_{1 n_r}^*(f) H_{2 n_r}(f) & \ldots & \sum_{n_r=1}^{N_R} H_{1 n_r}^*(f) H_{N_T n_r}(f) \\ \sum_{n_r=1}^{N_R} H_{2 n_r}^*(f) H_{1 n_r}(f) & \sum_{n_r=1}^{N_R} H_{2 n_r}^*(f) H_{2 n_r}(f) & \ldots & \sum_{n_r=1}^{N_R} H_{2 n_r}^*(f) H_{N_T n_r}(f) \\ \ldots & \ldots & \ldots & \ldots \\ \sum_{n_r=1}^{N_R} H_{N_T n_r}^*(f) H_{1 n_r}(f) & \sum_{n_r=1}^{N_R} H_{N_T n_r}^*(f) H_{2 n_r}(f) & \ldots & \sum_{n_r=1}^{N_R} H_{N_T n_r}^*(f) H_{N_T n_r}(f) \end{bmatrix}, \end{aligned} \tag{3.9}$$

and

$$\boldsymbol{H}_{n_r}(f) \triangleq [H_{1 n_r}(f) \;\; H_{2 n_r}(f) \;\; \ldots \;\; H_{N_T n_r}(f)]^T. \tag{3.10}$$

Eq. (3.8) is a nonlinear eigenvalue problem and $\bar{\boldsymbol{G}}(f)$ can be expressed as

$$\bar{\boldsymbol{G}}(f) = X(f)\, \boldsymbol{E}(f), \tag{3.11}$$

where $\boldsymbol{E}(f) \triangleq [E_1(f)\ E_2(f)\ \ldots\ E_{N_T}(f)]^T$ is that unit-norm eigenvector of $\boldsymbol{S}(f)$ which corresponds to its largest eigenvalue $\lambda_{\max}(f)$, and $X(f)$ is a scalar factor. For example, for $N_R = 1$ we obtain $\boldsymbol{E}(f) = \boldsymbol{H}_1(f)/\sqrt{\boldsymbol{H}_1^H(f)\boldsymbol{H}_1(f)}$. Eq. (3.11) shows that in general the optimum IIR BFFs can be viewed as concatenation of two filters: A filter $X(f)$ which is common to all transmit antennas and a filter $E_{n_t}(f)$ which is transmit antenna dependent. In the following, we derive filter $X(f)$ for ZF–DFE and MMSE–DFE.

*1) ZF–DFE:* In this case, $\xi = 0$ is valid. Combining Eqs. (3.8) and (3.11), we can get

$$\begin{aligned} \boldsymbol{S}(f)\, X(f)\, \boldsymbol{E}(f) &= \tilde{\mu}\, \boldsymbol{E}^H(f)\, X^*(f)\, \boldsymbol{S}(f) X(f)\, \boldsymbol{E}(f)\, X(f)\, \boldsymbol{E}(f) \\ \lambda_{\max}(f)\, X(f) &= \tilde{\mu} \lambda_{\max}(f)\, |X(f)|^2\, X(f) \\ |X(f)|^2 &= \frac{1}{\tilde{\mu}}, \end{aligned} \tag{3.12}$$

or equivalently $|X(f)| = 1/\sqrt{\tilde{\mu}}$. Furthermore, from the power constraint in Eq. (3.2), we obtain

$$\begin{aligned} \int_{-1/2}^{1/2} \boldsymbol{G}^H(f)\boldsymbol{G}(f)\, \mathrm{d}f &= 1 \\ \int_{-1/2}^{1/2} \left(X^*(f)\, \boldsymbol{E}^H(f)\right)\, X(f)\, \boldsymbol{E}(f)\, \mathrm{d}f &= 1 \\ \int_{-1/2}^{1/2} \frac{1}{\tilde{\mu}}\, \mathrm{d}f &= 1 \\ \tilde{\mu} &= 1. \end{aligned} \tag{3.13}$$

Therefore, the optimum IIR BFFs for ZF–DFE are given by

$$\bar{\boldsymbol{G}}(f) = \boldsymbol{E}(f)\, e^{j\varphi(f)}, \tag{3.14}$$

where $\varphi(f)$ is the phase which can be chosen arbitrarily. Replacing $\boldsymbol{G}(f)$ in Eq. (3.1) by $\bar{\boldsymbol{G}}(f)$ from Eq. (3.14) yields the maximum SNR for ZF–DFE

$$\begin{aligned}
\mathrm{SNR}(\bar{\boldsymbol{G}}(f)) &= \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left[ \sum_{n_r=1}^{N_R} \left| \sum_{n_t=1}^{N_T} \bar{G}_{nt}(f) H_{n_t n_r}(f) \right|^2 \right] \mathrm{d}f \right\} \\
&= \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left[ \bar{\boldsymbol{G}}^H(f) \boldsymbol{S}(f) \bar{\boldsymbol{G}}(f) \right] \mathrm{d}f \right\} \\
&= \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left[ |X(f)|^2 \lambda_{\max}(f) \right] \mathrm{d}f \right\} \\
&= \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left( \lambda_{\max}(f) \right) \mathrm{d}f \right\}. \qquad (3.15)
\end{aligned}$$

*2) MMSE–DFE:* For MMSE–DFE $\xi = \sigma_n^2/\sigma_b^2$ holds. As in the derivation for ZF–DFE case, combining Eqs. (3.8) and (3.11), we can get

$$\begin{aligned}
\boldsymbol{S}(f)\, X(f)\, \boldsymbol{E}(f) &= \tilde{\mu} \left[ \boldsymbol{E}^H(f)\, X^*(f)\, \boldsymbol{S}(f) X(f)\, \boldsymbol{E}(f) + \xi \right] X(f)\, \boldsymbol{E}(f) \\
\lambda_{\max}(f)\, X(f) &= \tilde{\mu} \left[ \lambda_{\max}(f)\, |X(f)|^2 + \xi \right] X(f) \\
|X(f)|^2 &= \hat{\mu} - \frac{\xi}{\lambda_{\max}(f)}. \qquad (3.16)
\end{aligned}$$

where $\hat{\mu} \triangleq 1/\tilde{\mu}$. Therefore, it can be shown that in this case the magnitude of the optimum $X(f)$ is given by

$$|X(f)| = \sqrt{\left[ \hat{\mu} - \frac{\xi}{\lambda_{\max}(f)} \right]^+}, \qquad (3.17)$$

where we took into account that $|X(f)|$ has to be non–negative. Furthermore, from the power constraint in Eq. (3.2), we obtain

$$\int_{-1/2}^{1/2} \boldsymbol{G}^H(f)\boldsymbol{G}(f)\,\mathrm{d}f = 1$$

$$\int_{-1/2}^{1/2} |X(f)|^2\,\mathrm{d}f = 1$$

$$\int_{-1/2}^{1/2} \left[\hat{\mu} - \frac{\xi}{\lambda_{\max}(f)}\right]^+ \mathrm{d}f = 1. \tag{3.18}$$

As Eq. (3.18) suggests, finding the optimum $\hat{\mu}$ is a typical Water Filling problem [29]. In particular, motivated by the power constraint in Eq. (3.2) we define

$$P(\tilde{\mu}) \triangleq \int_{-1/2}^{1/2} \left[\hat{\mu} - \frac{\xi}{\lambda_{\max}(f)}\right]^+ \mathrm{d}f. \tag{3.19}$$

The optimum $\hat{\mu}_{\text{opt}}$ can be found by slowly increasing a very small starting value $\hat{\mu} = \hat{\mu}_0$ which yields $P(\hat{\mu}_0) > 1$ until $P(\hat{\mu} = \hat{\mu}_{\text{opt}}) = 1$ is achieved.

Figures 3.1 and 3.2 illustrate results for the Water Filling algorithm for the same channel realization of the EQ channel with channel length $L = 7$ at $E_s/N_0 = 10\,\text{dB}$ and $E_s/N_0 = 20\,\text{dB}$, respectively, where $E_s$ is the average received energy per symbol. The communications system consists of 3 transmit antennas, 1 receive antenna and MMSE–DFE filter at the receiver. From Figures 3.1 and 3.2, the optimum "water" level, $\hat{\mu}_{\text{opt}}$, is a constant within the normalized frequency range, $-1/2 \leq f \leq 1/2$. The optimum IIR BFFs utilize the channel realization and allocate more power to frequencies with large eigenvalues, $\lambda_{\max}(f)$, of matrix $\mathbf{S}(f)$. As the $E_s/N_0$ increases from 10 dB to 20 dB, the "water" level $\hat{\mu}_{\text{opt}}$ drops from 1.15 to 1.01, respectively. Also, the applied power, $[\hat{\mu}_{\text{opt}} - \xi/\lambda_{\text{opt}}(f)]^+$, becomes flatter in Figure 3.2 compared to Figure 3.1 as $E_s/N_0$ increases.

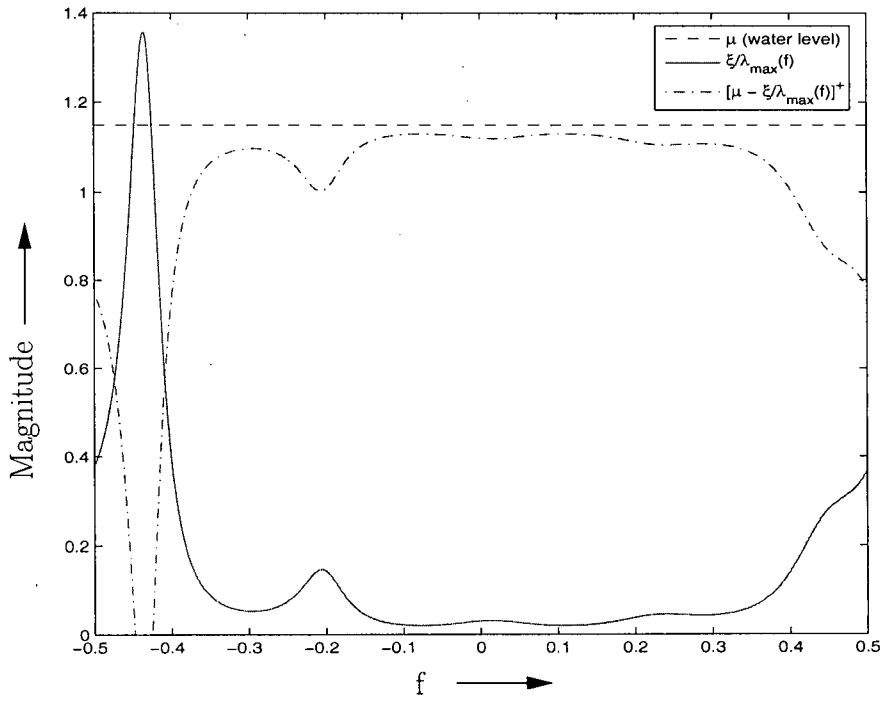


Figure 3.1: Water Filling for one realization of the EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, equal antenna correlation $\rho = 0.5$, and $10\log_{10}(E_s/N_0) = 10$ dB.

The optimum IIR BFFs for MMSE–DFE are given by

$$\bar{\boldsymbol{G}}(f) = \sqrt{\left[\hat{\mu}_{\text{opt}} - \frac{\xi}{\lambda_{\max}(f)}\right]^+} \boldsymbol{E}(f)\, e^{j\varphi(f)}, \tag{3.20}$$

where $\varphi(f)$ is again the phase which can be chosen arbitrarily. The corre-

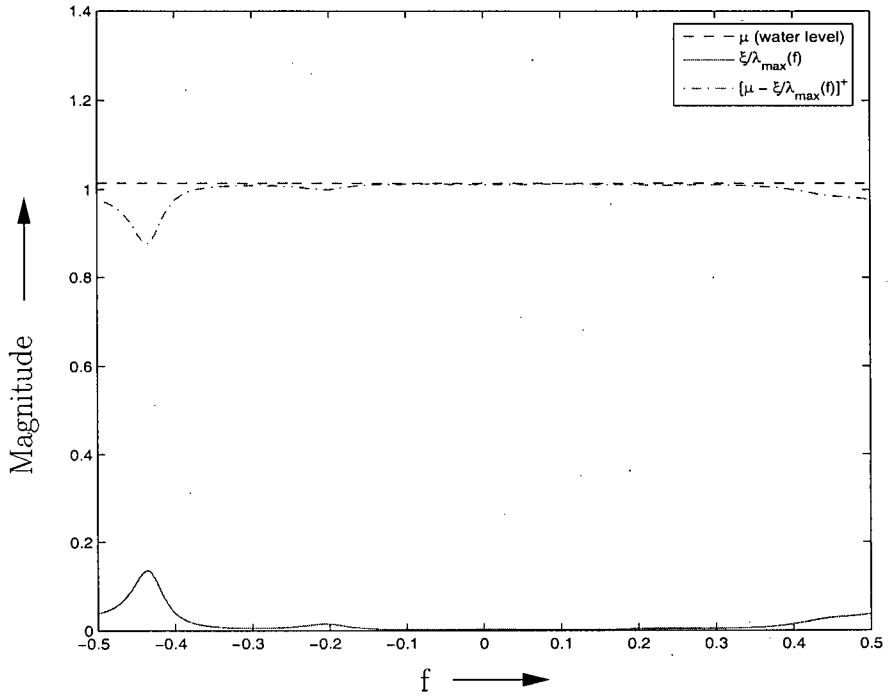


Figure 3.2: Water Filling for one realization of the EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, equal antenna correlation $\rho = 0.5$, and $10\log_{10}(E_s/N_0) = 20$ dB.

sponding maximum SNR is

$$\begin{aligned}
\mathrm{SNR}(\bar{\boldsymbol{G}}(f)) &= \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left[ \xi + \sum_{n_r=1}^{N_R} \left| \sum_{n_t=1}^{N_T} \bar{G}_{nt}(f) H_{n_t n_r}(f) \right|^2 \right] \mathrm{d}f \right\} - 1 \\
&= \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left[ \bar{\boldsymbol{G}}^H(f) \boldsymbol{S}(f) \bar{\boldsymbol{G}}(f) + \xi \right] \mathrm{d}f \right\} - 1 \\
&= \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left[ |X(f)|^2 \lambda_{\max}(f) + \xi \right] \mathrm{d}f \right\} - 1 \\
&= \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln\left( [\lambda_{\max}(f) \hat{\mu}_{\mathrm{opt}} - \xi]^+ + \xi \right) \mathrm{d}f \right\} - 1. \qquad (3.21)
\end{aligned}$$

Regarding the BFFs and the maximum SNRs for ZF–DFE and MMSE–DFE, we make the following interesting observations.

a) Although we have optimized the BFFs for fixed DFE filters, the optimum IIR BFF vectors given by Eqs. (3.14) and (3.20) are identical to the jointly optimum solution given in [17].

b) As expected, for $\sigma_n^2 \to 0$ (i.e., $\xi \to 0$) the optimum IIR BFFs for MMSE–DFE approach those for ZF–DFE. Furthermore, for arbitrary $\sigma_n^2$ and $L = 1$ (i.e., frequency–nonselective channel) the optimum BFFs for ZF–DFE and MMSE–DFE, respectively, have only one non–zero tap and are identical to the well–known beamforming weights, cf. e.g. [9].

c) In case of ZF–DFE, the optimum BFF frequency response vector $\bar{\boldsymbol{G}}(f)$ at frequency $f = f_0$ is just the dominant eigenvector of matrix $\boldsymbol{S}(f)$ at frequency $f = f_0$. Since $\boldsymbol{S}(f_0)$ only depends on the channel frequency responses $H_{n_t n_r}(f)$ at frequency $f = f_0$, $\bar{\boldsymbol{G}}(f_0)$ is independent of the $H_{n_t n_r}(f)$, $f \neq f_0$. This is not true for MMSE–DFE, where the optimum frequency response vector $\bar{\boldsymbol{G}}(f)$ at frequency $f = f_0$ also depends on the channel frequency responses $H_{n_t n_r}(f)$ at frequencies $f \neq f_0$, because of the constraint $P(\tilde{\mu}) = 1$, cf. Eq. (3.19). In fact, for MMSE–DFE $X(f)$ may be interpreted as a power allocation filter which allocates more power to frequencies with large eigenvalues $\lambda_{\max}(f)$.

## 3.2 IIR Beamforming with Linear Equalization

In this subsection, we consider IIR beamforming with ZF– and MMSE–LE at the receiver. Similar to the previous subsection's approach, the optimization problem is addressed followed by the derivation of the optimum BFFs for maximization of the SNR of a given channel realization under these assumptions.

### 3.2.1 Optimization Problem

From [20], the unbiased SNR for LE with optimum IIR feedforward filtering is

$$\mathrm{SNR}(\mathbf{G}(f)) = \frac{\sigma_b^2}{\sigma_e^2} - \chi, \tag{3.22}$$

where $\chi = 0$ and $\chi = 1$ for zero-forcing (ZF) and minimum mean–squared error (MMSE) LE filter optimization, respectively. The LE error variance is

$$\sigma_e^2 = \sigma_n^2 \int_{-1/2}^{1/2} \frac{1}{\sum_{n_r=1}^{N_R} |H_{n_r}^{\mathrm{eq}}(f)|^2 + \xi} \,\mathrm{d}f, \tag{3.23}$$

where $\xi = 0$ and $\xi = \sigma_n^2/\sigma_b^2$ are valid for ZF-LE and MMSE-LE, respectively [22]. Equivalently, we can rewrite Eq. (3.22) as

$$\mathrm{SNR}(\mathbf{G}(f)) = \frac{\sigma_b^2}{\sigma_n^2} \left( \int_{-1/2}^{1/2} \frac{1}{\sum_{n_r=1}^{N_R} |H_{n_r}^{\mathrm{eq}}(f)|^2 + \xi} \,\mathrm{d}f \right)^{-1} - \chi. \tag{3.24}$$

The equivalent channel frequency response $H_{n_r}^{\mathrm{eq}}(f) \triangleq \mathcal{F}\{h_{n_r}^{\mathrm{eq}}[k]\}$ is given by $H_{n_r}^{\mathrm{eq}}(f) = \sum_{n_t=1}^{N_T} G_{n_t}(f) H_{n_t n_r}(f)$ with $H_{n_t n_r}(f) \triangleq \mathcal{F}\{h_{n_t n_r}^{\mathrm{eq}}[k]\}$.

Again, the optimum BFF vector $\bar{\boldsymbol{G}}(f) \triangleq [\bar{G}_1(f)\ \bar{G}_2(f),\ \ldots\ \bar{G}_{N_T}(f)]^T$ shall maximize $\mathrm{SNR}(\boldsymbol{G}(f))$ subject to the transmit power constraint

$$\int_{-1/2}^{1/2} \boldsymbol{G}^H(f) \boldsymbol{G}(f) \,\mathrm{d}f = 1. \tag{3.25}$$

As in the DFE case, a convenient approach for calculating $\bar{\boldsymbol{G}}(f)$ is the classical Calculus of Variations method [32]. Following this method, we model the BFF of antenna $n_t$ as $G_{n_t}(f) = \bar{G}_{n_t}(f) + \varepsilon_{n_t} B_{n_t}(f)$, where $B_{n_t}(f)$ and $\varepsilon_{n_t}$ denote an arbitrary function of $f$ and a complex–valued variable, respectively. The

optimization problem can now be formulated in terms of its Lagrangian

$$\begin{aligned} L(\boldsymbol{\varepsilon}) =& \mathrm{SNR}(\boldsymbol{G}(f)) - \mu \int_{-1/2}^{1/2} \boldsymbol{G}^H(f)\boldsymbol{G}(f)\,\mathrm{d}f \\ =& \int_{-1/2}^{1/2} \frac{1}{\sum_{n_r=1}^{N_R} \left| \sum_{n_t=1}^{N_T} \left(G_{nt}(f) + \varepsilon_{nt} B_{nt}(f)\right) H_{n_t n_r}(f) \right|^2 + \xi} \mathrm{d}f \\ & - \mu \int_{-1/2}^{1/2} \sum_{n_t=1}^{N_T} \left| \bar{G}_{n_t}(f) + \varepsilon_{nt} B_{nt}(f) \right|^2 \mathrm{d}f \end{aligned} \tag{3.26}$$

where $\boldsymbol{\varepsilon} \triangleq [\varepsilon_1 \ \varepsilon_2 \ \ldots \ \varepsilon_{N_T}]^T$ and $\mu$ is the Lagrange multiplier. The optimum BFF vector $\bar{\boldsymbol{G}}(f)$ has to fulfill the condition [32]

$$\frac{\partial L(\boldsymbol{\varepsilon})}{\partial \boldsymbol{\varepsilon}^*}\Big|_{\boldsymbol{\varepsilon}=\mathbf{0}_{N_T}} = \mathbf{0}_{N_T}\,, \tag{3.27}$$

for arbitrary $B_{n_t}(f)$, $1 \le n_t \le N_T$. Therefore, for all $s = 1, 2, \ldots, N_T$, Eq. (3.27) can be written as

$$\begin{aligned} & \frac{\partial L(\boldsymbol{\varepsilon})}{\partial \varepsilon_s^*}\Big|_{\boldsymbol{\varepsilon}=\mathbf{0}_{N_T}} \\ =& \int_{-1/2}^{1/2} \frac{\sum_{n_r=1}^{N_R} B_s^*(f) H_{sn_r}^*(f) \left[ \sum_{n_t=1}^{N_T} \left(\bar{G}_{nt}(f) + \varepsilon_{nt} B_{nt}(f)\right) H_{n_t n_r}(f) \right]}{\left[ \sum_{n_r=1}^{N_R} \left| \sum_{n_t=1}^{N_T} \left(\bar{G}_{nt}(f) + \varepsilon_{nt} B_{nt}(f)\right) H_{n_t n_r}(f) \right|^2 + \xi \right]^2} \mathrm{d}f \Big|_{\boldsymbol{\varepsilon}=\mathbf{0}_{N_T}} \\ & - \mu \int_{-1/2}^{1/2} (\bar{G}_s(f) + \varepsilon_s B_s(f)) B_s^*(f)\,\mathrm{d}f \Big|_{\boldsymbol{\varepsilon}=\mathbf{0}_{N_T}} \\ =& \int_{-1/2}^{1/2} \frac{\sum_{n_r=1}^{N_R} B_s^*(f) H_{sn_r}^*(f) \left[ \sum_{n_t=1}^{N_T} \bar{G}_{nt}(f) H_{n_t n_r}(f) \right]}{\left[ \sum_{n_r=1}^{N_R} \left| \sum_{n_t=1}^{N_T} \bar{G}_{nt}(f) H_{n_t n_r}(f) \right|^2 + \xi \right]^2} \mathrm{d}f \\ & - \mu \int_{-1/2}^{1/2} \bar{G}_s(f) B_s^*(f) \mathrm{d}f \\ =& 0, \end{aligned} \tag{3.28}$$

In order to achieve Eq. (3.28), the following condition must be satisfied for all

$s = 1, 2, \ldots, N_T$,

$$\mu \bar{G}_s(f) \left( \sum_{n_r}^{N_R} \left| \sum_{n_t'=1}^{N_T} \bar{G}_{n_t'}(f) H_{n_t' n_r}(f) \right|^2 + \xi \right)^2 = \sum_{n_r=1}^{N_R} H_{s n_r}^*(f) \left( \sum_{n_t=1}^{N_T} \bar{G}_{n_t}(f) H_{n_t n_r}(f) \right), \tag{3.29}$$

or equivalently,

$$\mu \bar{G}_s(f) \left( \sum_{n_r}^{N_R} \left| \sum_{n_t'=1}^{N_T} \bar{G}_{n_t'}(f) H_{n_t' n_r}(f) \right|^2 + \xi \right)^2 = \sum_{n_t=1}^{N_T} \bar{G}_{n_t}(f) \left( \sum_{n_r=1}^{N_R} H_{s n_r}^*(f) H_{n_t n_r}(f) \right). \tag{3.30}$$

### 3.2.2 Optimum IIR BFFs

Combining Eqs. (3.27) and (3.30), it can be shown that vector $\bar{\boldsymbol{G}}(f)$ has to fulfill

$$\boldsymbol{S}(f)\, \bar{\boldsymbol{G}}(f) = \tilde{\mu} \left[ \bar{\boldsymbol{G}}^H(f) \boldsymbol{S}(f) \bar{\boldsymbol{G}}(f) + \xi \right]^2 \bar{\boldsymbol{G}}(f), \tag{3.31}$$

where $\tilde{\mu}$ is a constant and $\boldsymbol{S}(f)$ is an $N_T \times N_T$ matrix given by Eq. (3.9). The only difference between Eq. (3.31) and Eq. (3.8) of the DFE case is that there is a square term in Eq. (3.31) compared to Eq. (3.8). This difference is resulted from changes of the objective function for optimization. Similarly, Eq. (3.31) is a non–linear eigenvalue problem and $\bar{\boldsymbol{G}}(f)$ can be expressed as

$$\bar{\boldsymbol{G}}(f) = X(f)\, \boldsymbol{E}(f), \tag{3.32}$$

where $\boldsymbol{E}(f) \triangleq [E_1(f)\ E_2(f)\ \ldots\ E_{N_T}(f)]^T$ is that unit–norm eigenvector of $\boldsymbol{S}(f)$ which corresponds to its largest eigenvalue $\lambda_{\max}(f)$, and $X(f)$ is a scalar factor. Consequently, as in the DFE case, in general the optimum IIR BFFs can be viewed as concatenation of two filters: A filter $X(f)$ which is common to all transmit antennas and a filter $E_{n_t}(f)$ which is transmit antenna dependent.

*1) ZF–LE:* In this case, $\xi = 0$ is valid. Combining Eqs. (3.31) and (3.32), we

get

$$\begin{aligned}
\boldsymbol{S}(f)\,X(f)\,\boldsymbol{E}(f) &= \tilde{\mu}\,\left[\boldsymbol{E}^H(f)\,X^*(f)\,\boldsymbol{S}(f)X(f)\,\boldsymbol{E}(f)\,X(f)\right]^2\,X(f)\,\boldsymbol{E}(f) \\
\lambda_{\max}(f)\,X(f) &= \tilde{\mu}\lambda_{\max}^2(f)\,|X(f)|^4\,X(f) \\
|X(f)|^4 &= \frac{1}{\tilde{\mu}\lambda_{\max}(f)}\,, \qquad (3.33)
\end{aligned}$$

or equivalently $|X(f)| = 1/\sqrt[4]{\tilde{\mu}\lambda_{\max}(f)}$. Furthermore, from the power constraint in Eq. (3.2), we obtain

$$\int_{-1/2}^{1/2} \boldsymbol{G}^H(f)\boldsymbol{G}(f)\,\mathrm{d}f = 1$$

$$\int_{-1/2}^{1/2} \left(X^*(f)\,\boldsymbol{E}^H(f)\right)\,X(f)\,\boldsymbol{E}(f)\,\mathrm{d}f = 1$$

$$\int_{-1/2}^{1/2} \frac{1}{\sqrt{\tilde{\mu}\lambda_{\max}(f)}}\,\mathrm{d}f = 1$$

$$\tilde{\mu} = \left(\int_{-1/2}^{1/2} \frac{1}{\sqrt{\lambda_{\max}(f)}}\,\mathrm{d}f\right)^2 . \qquad (3.34)$$

Thereby, the optimum IIR BFFs for ZF-LE are given by

$$\bar{\boldsymbol{G}}(f) = \left(\sqrt{\lambda_{\max}(f)}\int_{-1/2}^{1/2}\sqrt{\frac{1}{\lambda_{\max}(f)}}\,\mathrm{d}f\right)^{-1/2} \boldsymbol{E}(f)\,e^{j\varphi(f)}\,, \qquad (3.35)$$

where $\varphi(f)$ is the phase which can be chosen arbitrarily. Unlike the result $X(f) = 1$ given by Eq. (3.14) for ZF–DFE, $X(f)$ is a frequency dependent term for ZF–LE. Interestingly, $X(f)$ allocates more power to frequencies with large eigenvalues $\lambda_{\max}(f)$ of $\mathbf{S}(f)$. Thereby, $X(f)$ for ZF–LE can be interpreted as a power allocation filter.

Replacing $\boldsymbol{G}(f)$ in Eq. (3.24) by $\bar{\boldsymbol{G}}(f)$ from Eq. (3.35) yields the corre-

sponding maximum SNR for ZF-LE

$$\begin{aligned}
\mathrm{SNR}(\bar{\boldsymbol{G}}(f)) &= \frac{\sigma_b^2}{\sigma_n^2}\left(\int_{-1/2}^{1/2} \frac{1}{\sum_{n_r=1}^{N_R}\left|\sum_{n_t=1}^{N_T}\bar{G}_{n_t}(f)H_{n_t n_r}(f)\right|^2}\,\mathrm{d}f\right)^{-1} \\
&= \frac{\sigma_b^2}{\sigma_n^2}\left(\int_{-1/2}^{1/2} \frac{1}{\bar{\boldsymbol{G}}^H(f)\boldsymbol{S}(f)\bar{\boldsymbol{G}}(f)}\,\mathrm{d}f\right)^{-1} \\
&= \frac{\sigma_b^2}{\sigma_n^2}\left(\int_{-1/2}^{1/2} \frac{\sqrt{\tilde{\mu}\lambda_{\max}(f)}}{\lambda_{\max}(f)}\,\mathrm{d}f\right)^{-1} \\
&= \frac{\sigma_b^2}{\sigma_n^2}\left(\sqrt{\tilde{\mu}}\int_{-1/2}^{1/2} \frac{1}{\sqrt{\lambda_{\max}(f)}}\,\mathrm{d}f\right)^{-1} \\
&= \frac{\sigma_b^2}{\sigma_n^2}\left(\int_{-1/2}^{1/2} \frac{1}{\sqrt{\lambda_{\max}(f)}}\,\mathrm{d}f\right)^{-2}. \qquad (3.36)
\end{aligned}$$

*2) MMSE–LE:* For MMSE-LE $\xi = \sigma_n^2/\sigma_b^2$ holds. Combining Eqs. (3.31) and (3.32), we can get

$$\begin{aligned}
\boldsymbol{S}(f)\,X(f)\,\boldsymbol{E}(f) &= \tilde{\mu}\left[\boldsymbol{E}^H(f)\,X^*(f)\,\boldsymbol{S}(f)X(f)\,\boldsymbol{E}(f)\,X(f)+\xi\right]^2 X(f)\,\boldsymbol{E}(f) \\
\lambda_{\max}(f)\,X(f) &= \tilde{\mu}\left[\lambda_{\max}(f)\,|X(f)|^2+\xi\right]^2 X(f) \\
|X(f)|^2 &= \frac{1}{\sqrt{\lambda_{\max}(f)}}\left[\hat{\mu}-\frac{\xi}{\sqrt{\lambda_{\max}(f)}}\right]^+, \qquad (3.37)
\end{aligned}$$

where $\hat{\mu} \triangleq 1/\sqrt{\tilde{\mu}}$ is a constant and we took into account that $|X(f)|^2$ has to be non–negative. Furthermore, from the power constraint in Eq. (3.26), we

obtain

$$\int_{-1/2}^{1/2} \boldsymbol{G}^H(f)\boldsymbol{G}(f)\,\mathrm{d}f = 1$$

$$\int_{-1/2}^{1/2} |X(f)|^2\,\mathrm{d}f = 1$$

$$\int_{-1/2}^{1/2} \frac{1}{\sqrt{\lambda_{\max}(f)}}\left[\hat{\mu} - \frac{\xi}{\sqrt{\lambda_{\max}(f)}}\right]^+ \mathrm{d}f = 1. \tag{3.38}$$

As Eq. (3.38) suggests, unlike the result derived in MMSE–DFE case, finding the optimum $\hat{\mu}$ is a *quasi* Water Filling problem [29] due to the fact that the "water" level, $\hat{\mu}/\sqrt{\lambda_{\max}(f)}$, is not a constant but depends on frequency $f$. In particular, motivated by the power constraint in Eq. (3.26) we define

$$P(\tilde{\mu}) \triangleq \int_{-1/2}^{1/2} \frac{1}{\sqrt{\lambda_{\max}(f)}}\left[\hat{\mu} - \frac{\xi}{\sqrt{\lambda_{\max}(f)}}\right]^+ \mathrm{d}f\,. \tag{3.39}$$

The optimum $\hat{\mu}_{\text{opt}}$ can be found by slowly increasing a very small starting value $\hat{\mu} = \hat{\mu}_0$ which yields $P(\hat{\mu}_0) < 1$ until $P(\hat{\mu} = \hat{\mu}_{\text{opt}}) = 1$ is achieved.

Figures 3.3 and 3.4 illustrate the quasi Water Filling results for the same channel realization of the EQ channel with channel length $L = 7$ at $E_s/N_0 = 10\,\text{dB}$ and $E_s/N_0 = 20\,\text{dB}$, respectively. The simulated system consists of 3 transmit antennas, 1 receive antenna and MMSE–LE filter at the receiver. Figures 3.3 and 3.4 confirm that, unlike the "water" level in the DFE case, the optimum "water" level, $\hat{\mu}_{\text{opt}}/\sqrt{\lambda_{\max}(f)}$, is *not* a constant within the frequency range but a variable with respect to $f$. As the $E_s/N_0$ increases from 10 dB to 20 dB, the "water" level varies more smoothly. Although it is hard to justify how the optimum IIR BFFs allocate power from these two figures, the resulting optimum SNR, i.e. Eq. (3.41) derived in the later section, at the equalizer confirms that the optimum IIR BFFs utilize the channel realization and allocate more power to frequencies with large eigenvalues, $\lambda_{\max}(f)$, of matrix $\mathbf{S}(f)$.

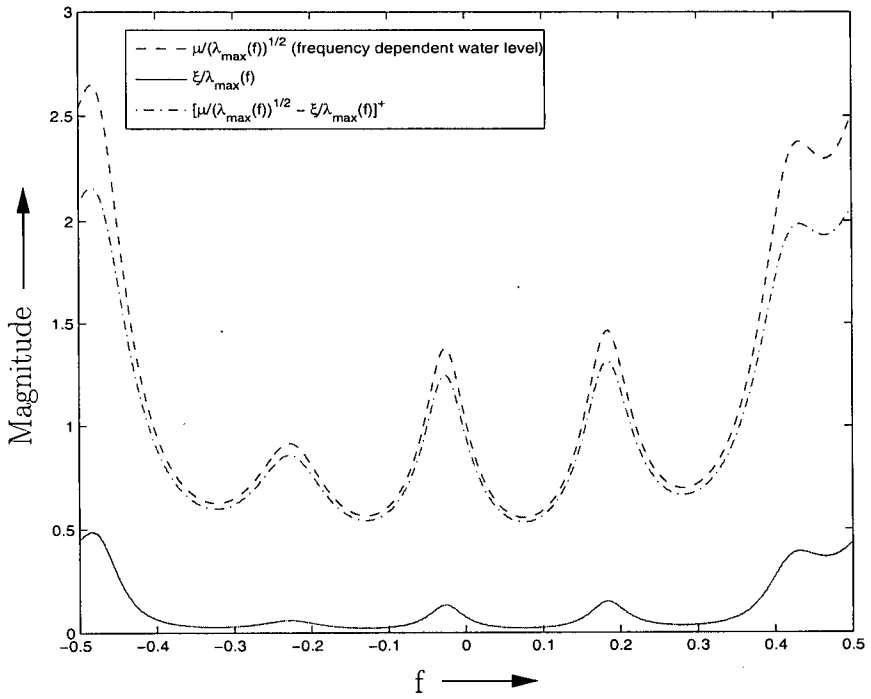


Figure 3.3: Quasi Water Filling for one realization of the EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, equal antenna correlation $\rho = 0.5$, and $10\log_{10}(E_s/N_0) = 10$ dB.

The optimum IIR BFFs for MMSE–LE are given by

$$\bar{\boldsymbol{G}}(f) = \frac{1}{\sqrt[4]{\lambda_{\max}(f)}} \sqrt{\left[\hat{\mu}_{\text{opt}} - \frac{\xi}{\sqrt{\lambda_{\max}(f)}}\right]^{+}} \boldsymbol{E}(f)\, e^{j\varphi(f)}, \qquad (3.40)$$

where $\varphi(f)$ is again the phase which can be chosen arbitrarily. Replacing $\boldsymbol{G}(f)$ in Eq. (3.24) by $\bar{\boldsymbol{G}}(f)$ from Eq. (3.40) yields the corresponding maximum SNR

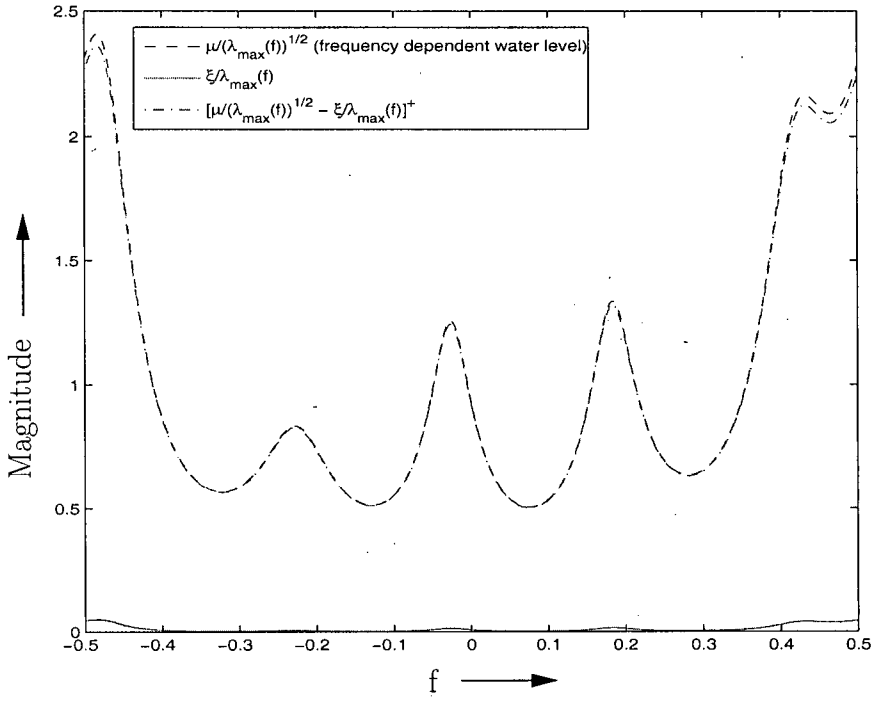


Figure 3.4: Quasi Water Filling for one realization of the EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, equal antenna correlation $\rho = 0.5$, and $10\log_{10}(E_s/N_0) = 20$ dB.

for MMSE–LE

$$\begin{aligned}
\mathrm{SNR}\left(\bar{\boldsymbol{G}}(f)\right) &= \frac{\sigma_b^2}{\sigma_n^2}\left(\int_{-1/2}^{1/2} \frac{1}{\sum_{n_r=1}^{N_R}\left|\sum_{n_t=1}^{N_T}\bar{G}_{n_t}(f)H_{n_t n_r}(f)\right|^2+\xi}\,\mathrm{d}f\right)^{-1} - 1 \\
&= \frac{\sigma_b^2}{\sigma_n^2}\left(\int_{-1/2}^{1/2} \frac{1}{\bar{\boldsymbol{G}}^H(f)\boldsymbol{S}(f)\bar{\boldsymbol{G}}(f)+\xi}\,\mathrm{d}f\right)^{-1} - 1 \\
&= \frac{\sigma_b^2}{\sigma_n^2}\left(\int_{-1/2}^{1/2} \frac{1}{\frac{1}{\sqrt{\lambda_{\max}(f)}}\left[\hat{\mu}_{\mathrm{opt}} - \frac{\xi}{\sqrt{\lambda_{\max}(f)}}\right]^+\lambda_{\max}(f)+\xi}\,\mathrm{d}f\right)^{-1} - 1 \\
&= \frac{\sigma_b^2}{\sigma_n^2}\left(\int_{-1/2}^{1/2} \frac{1}{\left[\sqrt{\lambda_{\max}(f)}\hat{\mu}_{\mathrm{opt}} - \xi\right]^+ + \xi}\,\mathrm{d}f\right)^{-1} - 1. \qquad (3.41)
\end{aligned}$$

Comparing the BFFs and the maximum SNRs for ZF–LE and MMSE–LE, we make the following interesting observations:

a) For $\sigma_n^2 \to 0$ (i.e., $\xi \to 0$) the optimum IIR BFFs for MMSE-LE approach those for ZF-LE. Although we have optimized the BFFs for fixed LE filters, the optimum IIR BFF vectors given by Eqs. (3.35) and (3.40) are identical to the jointly optimum solution given in [17].

b) In both ZF-LE and MMSE-LE cases, the optimum BFF frequency response vector $\bar{\boldsymbol{G}}(f)$ at frequency $f = f_0$ depends on the channel frequency responses $H_{n_t n_r}(f)$ at all frequencies $-1/2 \leq f < 1/2$. In fact, $X(f)$ may be interpreted as a power allocation filter which allocates more power to frequencies with large eigenvalues $\lambda_{\max}(f)$ in ZF-LE and MMSE-LE cases.

c) If the underlying channel is frequency-nonselective, $\mathbf{S}(f) = \mathbf{S}$ for all $f$. In this case, it is easy to see that the optimum BFFs have only one non-zero tap and are identical to the well-know beamforming weights for frequency-nonselective channels [9]. In this case, the LE structure collapses to a simple threshold detector, of course.

Although we derived promising IIR BFFs in this chapter, implementing IIR BFFs in practice is not realistic. There are two main reasons. Firstly, it is well know that designing stable IIR filters for any channel realization is difficult. Secondly, the transmitter requires to know perfect CSI in order to perform IIR beamforming. Due to the fact that the feedback channel is bandwidth limited, perfect CSI at the transmitter cannot be provided. However, the SNRs achievable with IIR BFFs given by Eqs. (3.15), (3.21), (3.36), and (3.41) will serve as upper bounds for the SNRs achievable with the FIR BFFs considered in Chapter 4 and 5.

# Chapter 4

# Beamforming with Perfect CSI and FIR Filters

The IIR BFFs derived in the previous chapter are not suitable for practical implementation. Therefore, in this chapter, we derive the optimum FIR BFFs for perfect CSI using the same approach as for the IIR case in Chapter 3. We note that although FIR BFFs are adopted in this section, the DFE and LE are still assumed to employ an IIR feedforward filter. This is not a major restriction as reasonably long FIR DFE and LE feedforward filters yield practically the same performance as IIR feedforward filters.

## 4.1 FIR Beamforming with Perfect CSI and DFE

### 4.1.1 Optimum FIR BFFs

Since the FIR BFFs have length $L_g$, the resulting equivalent overall CIR $h_{n_r}^{\text{eq}}[k]$ has length $L_{\text{eq}} = L + L_g - 1$. The frequency response of the equivalent channel can now be expressed as

$$H_{n_r}^{\text{eq}}(f) = \boldsymbol{d}^H(f)\boldsymbol{H}_{n_r}\boldsymbol{g}, \tag{4.1}$$

where $\boldsymbol{d}(f) \triangleq [1 \; e^{j2\pi f} \; \ldots \; e^{j2\pi f(L_{\text{eq}}-1)}]^T$, $\boldsymbol{H}_{n_r} \triangleq [\boldsymbol{H}_{1n_r} \; \boldsymbol{H}_{2n_r} \; \ldots \; \boldsymbol{H}_{N_T n_r}]$, and $\boldsymbol{H}_{n_t n_r}$ is an $L_{\text{eq}} \times L_g$ column–circulant matrix defined as

$$\mathbf{H}_{n_t n_r} \triangleq \begin{bmatrix} h_{n_t n_r}[0] & \cdots & 0 \\ \vdots & \ddots & \vdots \\ h_{n_t n_r}[L-1] & \ddots & 0 \\ 0 & \ddots & h_{n_t n_r}[0] \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_{n_t n_r}[L-1] \end{bmatrix}. \tag{4.2}$$

Combining Eqs. (3.1) and (4.1) the SNR of ZF- and MMSE–DFE with FIR BFFs is obtained as

$$\text{SNR}(\boldsymbol{g}) = \frac{\sigma_b^2}{\sigma_n^2} \exp\left\{ \int_{-1/2}^{1/2} \ln(\boldsymbol{g}^H \boldsymbol{B}(f) \boldsymbol{g} + \xi) \, \mathrm{d}f \right\} - \chi, \tag{4.3}$$

with $N_T L_g \times N_T L_g$ matrix $\boldsymbol{B}(f) \triangleq \sum_{n_r=1}^{N_R} \boldsymbol{H}_{n_r}^H \boldsymbol{d}(f) \boldsymbol{d}^H(f) \boldsymbol{H}_{n_r}$. The optimum BFF vector $\bar{\boldsymbol{g}}$ shall maximize $\text{SNR}(\boldsymbol{g})$ subject to the power constraint $\boldsymbol{g}^H \boldsymbol{g} = 1$. Thus, the Lagrangian of the optimization problem can be formulated as

$$L(\boldsymbol{g}) = \text{SNR}(\boldsymbol{g}) + \mu \, \boldsymbol{g}^H \boldsymbol{g}, \tag{4.4}$$

where $\mu$ denotes again the Lagrange multiplier. The optimum vector $\bar{\boldsymbol{g}}$ has to fulfill $\partial L(\boldsymbol{g}) / \partial \boldsymbol{g}^* = \mathbf{0}_{N_T L_g}$, which leads to the nonlinear eigenvalue problem

$$\left[ \int_{-1/2}^{1/2} \frac{\boldsymbol{B}(f) + \xi \boldsymbol{I}_{N_T L_g}}{\bar{\boldsymbol{g}}^H \left( \boldsymbol{B}(f) + \xi \boldsymbol{I}_{N_T L_g} \right) \bar{\boldsymbol{g}}} \, \mathrm{d}f \right] \bar{\boldsymbol{g}} = \bar{\boldsymbol{g}}, \tag{4.5}$$

where $\mu$ does not appear since Eq. (4.5) already includes the constraint $\bar{\boldsymbol{g}}^H \bar{\boldsymbol{g}} = 1$. However, in contrast to the IIR eigenvalue problem in Eq. (3.8), Eq. (4.5) does not seem to have a closed–form solution. To substantiate this claim, we discretize the integral in Eq. (4.3) and rewrite the optimization problem in Eq. (4.4) as

$$\tilde{L}(\boldsymbol{g}) = \prod_{i=1}^{S} \frac{\boldsymbol{g}^H (\boldsymbol{B}(f_i) + \xi \boldsymbol{I}_{N_T L_g}) \boldsymbol{g}}{\boldsymbol{g}^H \boldsymbol{g}}, \tag{4.6}$$

where $f_i \triangleq -1/2+(i-1)/(S-1)$ and $S$ is a large integer. Note that Eq. (4.6) is scale invariant (i.e., $\tilde{L}(\boldsymbol{g})$ does not depend on the magnitude of $\boldsymbol{g}$) and the resulting solution has to be scaled to meet $\boldsymbol{g}^H\boldsymbol{g} = 1$. Eq. (4.6) is a product of Rayleigh quotients. Unfortunately, it is well–known that the maximization of a product of Rayleigh quotients is a difficult mathematical problem which is not well understood and a closed–form solution is not known except for the trivial case $S = 1$, cf. e.g. [33, 34]. Therefore, we also do not expect to find a closed–form solution for the nonlinear eigenvalue problem in Eq. (4.5). Instead, we provide two efficient numerical methods for recursive calculation of the optimum FIR BFF vector $\bar{\boldsymbol{g}}$ in the next section.

### 4.1.2 Calculation of the Optimum FIR BFFs

For calculation of the optimum FIR BFFs we propose two different algorithms. Both algorithms recursively improve an initial BFF vector $\boldsymbol{g}_0$. Since the cost function in Eq. (4.3) is not concave, we cannot guarantee that the algorithms converge to the global maximum. However, adopting the initialization procedure explained below, the solutions found by both algorithms seem to be close–to–optimum, i.e., for $L_g \gg 1$ the obtained FIR BFFs approach the performance of the optimum IIR BFFs derived in Chapter 3.

*1) Gradient Algorithm (GA):* The first algorithm is a typical GA where in iteration $i+1$ the BFF vector $\boldsymbol{g}_i$ from iteration $i$ is improved in the direction of the steepest ascent [29]. The GA is summarized as follows:

1. Let $i = 0$ and initialize the BFF vector with some $\boldsymbol{g}_0$ fulfilling $\boldsymbol{g}_0^H\boldsymbol{g}_0 = 1$.

2. Update the BFF vector

$$\tilde{\boldsymbol{g}}_{i+1} = \boldsymbol{g}_i + \delta_i \left[ \int_{-1/2}^{1/2} \frac{\boldsymbol{B}(f) + \xi \boldsymbol{I}_{N_T L_g}}{\boldsymbol{g}_i^H \left(\boldsymbol{B}(f) + \xi \boldsymbol{I}_{N_T L_g}\right) \boldsymbol{g}_i} \, \mathrm{d}f \right] \boldsymbol{g}_i, \qquad (4.7)$$

   where $\delta_i$ is a suitable adaptation step size.

3. Normalize the BFF vector

$$\boldsymbol{g}_{i+1} = \frac{\tilde{\boldsymbol{g}}_{i+1}}{\sqrt{\tilde{\boldsymbol{g}}_{i+1}^H \tilde{\boldsymbol{g}}_{i+1}}}. \tag{4.8}$$

4. If $1 - |\boldsymbol{g}_{i+1}^H \boldsymbol{g}_i| < \epsilon$, goto Step 5, otherwise increment $i \rightarrow i+1$ and goto Step 2.

5. $\boldsymbol{g}_{i+1}$ is the desired BFF vector $\bar{\boldsymbol{g}}$.

The main drawback of the GA is that its speed of convergence critically depends on the adaptation step size $\delta_i$, which has to be empirically optimized. For the termination constant $\epsilon$ in Step 4 a suitably small value should be chosen, e.g. $\epsilon = 10^{-4}$. Ideally the adaptation step size $\delta_i$ should be optimized to maximize the speed of convergence. Here, we follow [29] and choose $\delta_i$ proportional to $1/\lambda_i$ where $\lambda_i$ is the maximum eigenvalue of matrix $\int_{-1/2}^{1/2} \boldsymbol{B}(f)\,\mathrm{d}f / \left(\boldsymbol{g}_i^H \boldsymbol{B}(f) \boldsymbol{g}_i + \xi\right)$ in iteration $i$. In particular, we found empirically that $\delta_i = 0.01/\lambda_i$ is a good choice.

*2) Modified Power Method (MPM):* It is interesting to observe that if the denominator under the integral in Eq. (4.5) was absent, $\bar{\boldsymbol{g}}$ would simply be the maximum eigenvalue of matrix $\int_{-1/2}^{1/2} (\boldsymbol{B}(f) + \xi \boldsymbol{I}_{N_T L_g})\,\mathrm{d}f$, which could be calculated efficiently using the so–called Power Method [29]. Motivated by this observation, we propose an MPM for recursive calculation of $\bar{\boldsymbol{g}}$. The corresponding algorithm does not involve an adaptation step size and is summarized as follows:

1. Let $i = 0$ and initialize the BFF vector with some $\boldsymbol{g}_0$ fulfilling $\boldsymbol{g}_0^H \boldsymbol{g}_0 = 1$.

2. Update the BFF vector

$$\tilde{\boldsymbol{g}}_{i+1} = \left[ \int_{-1/2}^{1/2} \frac{\boldsymbol{B}(f) + \xi \boldsymbol{I}_{N_T L_g}}{\boldsymbol{g}_i^H \left(\boldsymbol{B}(f) + \xi \boldsymbol{I}_{N_T L_g}\right) \boldsymbol{g}_i} \,\mathrm{d}f \right] \boldsymbol{g}_i. \tag{4.9}$$

3. Normalize the BFF vector

$$\boldsymbol{g}_{i+1} = \frac{\tilde{\boldsymbol{g}}_{i+1}}{\sqrt{\tilde{\boldsymbol{g}}_{i+1}^H \tilde{\boldsymbol{g}}_{i+1}}}. \tag{4.10}$$

4. If $1 - |\boldsymbol{g}_{i+1}^H \boldsymbol{g}_i| < \epsilon$, goto Step 5, otherwise increment $i \rightarrow i+1$ and goto Step 2.

5. $\boldsymbol{g}_{i+1}$ is the desired BFF vector $\bar{\boldsymbol{g}}$.

As mentioned earlier global convergence of the GA and the MPM to the maximum SNR solution cannot be guaranteed. However, we found empirically that convergence to the optimum or a close–to–optimum solution is achieved if the BFF length is gradually increased. If the desired BFF length is $L_g$, the GA or the MPM are executed $L_g$ times. For the $\nu$th, $2 \leq \nu \leq L_g$, execution of the algorithm the first $\nu - 1$ BFF coefficients of each antenna are initialized with the optimum BFF coefficients for that antenna obtained from the $(\nu - 1)$th execution and the $\nu$th coefficients are initialized with zero. For the first $(\nu = 1)$ execution, the BFF vector is initialized with the normalized all–ones vector of length $N_T$.

If this initialization procedure is used to calculate the optimum BFFs of length $L_g$, the optimum or close–to–optimum FIR BFFs of lengths 1, 2, ..., $L_g - 1$ are also obtained as a by–product. This property may be useful when comparing FIR BFFs of different length. Of course, this comes at the cost of increased computational complexity. However, computational complexity is not a major concern here, since in practice beamforming with perfect CSI is not possible anyway. Nevertheless, FIR beamforming with perfect CSI is of interest because it (a) constitutes a benchmark for beamforming with quantized CSI and (b) the optimum FIR BFFs are the input to the (off–line) codebook design for beamforming with quantized CSI, cf. Chapter 5.

Extensive simulations have shown that the FIR BFFs obtained by respectively the GA and the MPM with the proposed initialization procedure approach the performance of the optimum IIR BFFs as the FIR filter length $L_g$ increases. Exemplary simulation results for both algorithms are shown and discussed in Chapter 6.

## 4.2 FIR Beamforming with Perfect CSI and LE

Similar to the approach in the previous section, we introduce a numerical method for calculation of the optimum FIR BFFs for LE in this section.

### 4.2.1 Optimum FIR BFFs

Combining Eqs. (3.24) and (4.1), the SNR for ZF–LE and MMSE–LE can be expressed as

$$\mathrm{SNR}(\boldsymbol{g}) = \frac{\sigma_b^2}{\sigma_n^2} \left[ \int_{-1/2}^{1/2} \frac{1}{\boldsymbol{g}^H \boldsymbol{B}(f) \boldsymbol{g} + \xi} \, \mathrm{d}f \right]^{-1} - \chi, \tag{4.11}$$

with $N_T L_g \times N_T L_g$ matrix $\boldsymbol{B}(f) \triangleq \sum_{n_r=1}^{N_R} \boldsymbol{H}_{n_r}^H \boldsymbol{d}(f) \boldsymbol{d}^H(f) \boldsymbol{H}_{n_r}$.

The optimum BFF vector $\boldsymbol{g}_{\mathrm{opt}}$ shall maximize $\mathrm{SNR}(\boldsymbol{g})$ subject to the power constraint $\boldsymbol{g}^H \boldsymbol{g} = 1$. Thus, the Lagrangian of the optimization problem can be formulated as

$$L(\boldsymbol{g}) = \mathrm{SNR}(\boldsymbol{g}) + \mu \, \boldsymbol{g}^H \boldsymbol{g}, \tag{4.12}$$

where $\mu$ denotes the Lagrange multiplier. The optimum vector $\bar{\boldsymbol{g}}$ has to fulfill $\partial L(\boldsymbol{g}) / \partial \boldsymbol{g}^* = \boldsymbol{0}_{N_T L_g}$, which leads to the non–linear eigenvalue problem

$$\left[ \int_{-1/2}^{1/2} \frac{\boldsymbol{B}(f)}{\left( \bar{\boldsymbol{g}}^H \boldsymbol{B}(f) \bar{\boldsymbol{g}} + \xi \right)^2} \, \mathrm{d}f \right] \bar{\boldsymbol{g}} = \tilde{\mu} \, \bar{\boldsymbol{g}}, \tag{4.13}$$

with eigenvalue $\tilde{\mu}$. The difference between Eq. (4.5) and Eq. (4.13) is that there is a square term in the denominator in Eq. (4.13). Unfortunately, Eq. (4.13) does not seem to have a closed–form solution. Instead, we proposed a gradient alorithm (GA) for recursive calculation of the optimum FIR BFF vector $\bar{\boldsymbol{g}}$ in the next section.

### 4.2.2 Calculation of the Optimum FIR BFFs

In this subsection, we propose a gradient algorithm (GA) for calculation of the optimum FIR BFFs, which recursively improves an initial BFF vector $\boldsymbol{g}_0$. Since the cost function in Eq. (4.12) is not concave, we cannot guarantee that the algorithm converges to a global maximum. However, adopting the initialization procedure similar to that for the DFE case, the solution found by the proposed GA seems to be close–to–optimum, cf. Chapter 3.

The *Gradient Algorithm* is summarized as follows:

1. Let $i = 0$ and initialize the BFF vector with a suitable $\boldsymbol{g}_0$ fulfilling $\boldsymbol{g}_0^H \boldsymbol{g}_0 = 1$.

2. Update the BFF vector

$$\tilde{\boldsymbol{g}}_{i+1} = \boldsymbol{g}_i + \delta_i \left[ \int_{-1/2}^{1/2} \frac{\boldsymbol{B}(f)\,\mathrm{d}f}{\left(\boldsymbol{g}_i^H \boldsymbol{B}(f) \boldsymbol{g}_i + \xi\right)^2} \right] \boldsymbol{g}_i, \tag{4.14}$$

   where $\delta_i$ is a suitable adaptation step size.

3. Normalize the BFF vector:

$$\boldsymbol{g}_{i+1} = \frac{\tilde{\boldsymbol{g}}_{i+1}}{\sqrt{\tilde{\boldsymbol{g}}_{i+1}^H \tilde{\boldsymbol{g}}_{i+1}}} \tag{4.15}$$

4. If $1 - |\boldsymbol{g}_{i+1}^H \boldsymbol{g}_i| < \epsilon$, goto Step 5, otherwise increment $i \rightarrow i+1$ and goto Step 2.

5. $\boldsymbol{g}_{i+1}$ is the desired BFF vector $\bar{\boldsymbol{g}}$.

For the termination constant $\epsilon$ in Step 4 a suitably small value should be chosen, e.g. $\epsilon = 10^{-4}$. Ideally the adaptation step size $\delta_i$ should be optimized to maximize the speed of convergence. Here, we follow [29] and choose $\delta_i$ proportional to $1/\lambda_i$ where $\lambda_i$ is the maximum eigenvalue of matrix $\int_{-1/2}^{1/2} \boldsymbol{B}(f)\,\mathrm{d}f / \left(\boldsymbol{g}_i^H \boldsymbol{B}(f) \boldsymbol{g}_i + \xi\right)^2$ in iteration $i$. In particular, we found empirically that $\delta_i = 0.01/\lambda_i$ is a good choice.

As the GA for the DFE case, the global convergence of the GA for LE cannot be guaranteed, but convergence to the optimum or a close–to–optimum solution is achieved if the BFF length is gradually increased. It is also worth mentioning that MPM is not suitable for calculating FIR BFFs for LE receivers. The reason for that is still unknown, and it may be the objective for future investigation.

Our numerical results in Chapter 6 show that the performance of optimum IIR beamforming can be closely approached with short FIR BFFs derived in this Chapter.

# Chapter 5

# FIR Beamforming with Quantized CSI

In this chapter, we consider FIR beamforming with quantized CSI. For this purpose, we first introduce vector quantization in the context of finite–rate feedback beamforming for frequency–selective channels. Subsequently, we discuss the mean quantization error and the distortion measure adopted for quantizer design. Then, we adapt the basic LBG algorithm to the problem at hand and present a global vector quantization (GVQ) algorithm for calculation of the BFF vector codebook $\mathcal{G}$ for DFE and LE.

## 5.1 Vector Quantization – Preliminaries

Clustering is an important and fundamental instrument in engineering and other scientific disciplines to solve problems such as pattern recognition, image processing, machine learning, and so on [16]. The simplest form of clustering is the partitioning clustering approach known as Vector Quantization [35], which aims at partitioning a given data set into disjoint subsets so that the objective function (distortion) representing the quantization error is minimized. In this manner, each element of the data in the same subset is represented by only one corresponding codeword. In this section, we apply general vector quantization

techniques to our specific FIR BFFs codebook searching problem.

For convenience we define the channel vector $\boldsymbol{h} \triangleq [h_{11}[0] \ \ldots \ h_{11}[L-1] \ h_{12}[0] \ \ldots \ h_{N_T N_R}[L-1]]^T$ of length $N_T N_R L$. We assume that a training set $\mathcal{H} \triangleq \{\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_T\}$ of $T$ channel vectors $\boldsymbol{h}_n$ is available. Thereby, channel vector $\boldsymbol{h}_n$ has length $N_T N_R L$ and contains the CIR coefficients of all $N_T N_R$ CIRs of the $n$th realization of the MIMO channel. In practice, the $\boldsymbol{h}_n$ may be obtained either from measurements or by simulating the MIMO channel. Using the GA or MPM summarized in Chapter 4 we calculate the optimum BFF vector $\bar{\boldsymbol{g}}_n$ for each channel $\boldsymbol{h}_n$, $1 \leq n \leq T$. The resulting BFF vector training set is denoted by $\mathcal{G}_T \triangleq \{\bar{\boldsymbol{g}}_1, \bar{\boldsymbol{g}}_2, \ldots, \bar{\boldsymbol{g}}_T\}$.

A vector quantizer $\mathcal{Q}$ is a mapping of the BFF vector training set $\mathcal{G}_T$ with $T$ entries to the BFF vector codebook $\mathcal{G} \triangleq \{\hat{\boldsymbol{g}}_1, \hat{\boldsymbol{g}}_2, \ldots, \hat{\boldsymbol{g}}_N\}$ with $N$ entries, where $N \ll T$ [14]. Therefore, the vector quantizer can be represented as a function $\mathcal{Q} : \mathcal{G}_T \rightarrow \mathcal{G}$. The elements $\hat{\boldsymbol{g}}_n$ of the codebook $\mathcal{G}$ are also referred to as codewords. Once $\mathcal{Q}$ is determined, we can define partition regions $\mathcal{R}_n$ constituted by subsets of the original training set

$$\mathcal{R}_n \triangleq \{\bar{\boldsymbol{g}} \in \mathcal{G}_T | \mathcal{Q}(\bar{\boldsymbol{g}}) = \hat{\boldsymbol{g}}_n\}, \quad 1 \leq n \leq N, \tag{5.1}$$

i.e., if $\bar{\boldsymbol{g}}$ falls into $\mathcal{R}_n$ it is quantized to $\hat{\boldsymbol{g}}_n$.

## 5.2 Mean Quantization Error (MQE) and Distortion Measure

In general, a vector quantizer is said to be optimum if it minimizes the *mean quantization error (MQE)* for a given codebook size $N$. The MQE is defined as

$$\text{MQE} \triangleq \frac{1}{T} \sum_{n=1}^{T} d(\mathcal{Q}(\bar{\boldsymbol{g}}_n), \bar{\boldsymbol{g}}_n), \tag{5.2}$$

where $d(\hat{\boldsymbol{g}}_m, \bar{\boldsymbol{g}}_n)$ is the so–called *distortion measure* and denotes the distortion caused by quantizing $\bar{\boldsymbol{g}}_n \in \mathcal{G}_T$ to $\hat{\boldsymbol{g}}_m \in \mathcal{G}$.

Here, our aim is to design a codebook $\mathcal{G}$ which minimizes the average BER, i.e., the average BER is the MQE. Therefore, the distortion measure $d(\hat{\boldsymbol{g}}_m, \bar{\boldsymbol{g}}_n)$ is the BER $P_e(\hat{\boldsymbol{g}}_m, \boldsymbol{h}_n)$ caused by $\hat{\boldsymbol{g}}_m \in \mathcal{G}$ for channel $\boldsymbol{h}_n \in \mathcal{H}$ with optimum BFF vector $\bar{\boldsymbol{g}}_n \in \mathcal{G}_T$, i.e.,

$$d(\hat{\boldsymbol{g}}_m, \bar{\boldsymbol{g}}_n) \triangleq P_e(\hat{\boldsymbol{g}}_m, \boldsymbol{h}_n). \tag{5.3}$$

In order to obtain a tractable expression for the BER, we assume Gray mapping, and that the residual error at the input of the equalizer, DFE or LE, decision device is a Gaussian random variable independent of the data symbol $b[k]$. The latter assumption is true for ZF–LE and ZF–DFE, and is a good approximation for MMSE–LE and MMSE–DFE. With these assumptions the BER of DFE (with error–free feedback) or LE can be approximated as

$$P_e(\hat{\boldsymbol{g}}_m, \boldsymbol{h}_n) \triangleq C \cdot Q\left(\sqrt{d_{\min}^2 \, \mathrm{SNR}(\hat{\boldsymbol{g}}_m, \boldsymbol{h}_n)/2}\right) \tag{5.4}$$

where $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} \, \mathrm{d}t$, and both the unimportant constant $C$ and the minimum Euclidean distance $d_{\min}$ depend on signal constellation $\mathcal{A}$. For example, $C = 2/\log_2 M$ and $d_{\min} = 2\sin(\pi/M)$ for $M$–ary PSK ($M$PSK). $\mathrm{SNR}(\hat{\boldsymbol{g}}_m, \boldsymbol{h}_n)$ can be obtained from Eq. (4.3) for DFE or Eq. (4.11) for LE.

## 5.3 LBG Algorithm

The LBG algorithm [14] can be used to improve a given initial codebook. This algorithm exploits two necessary conditions that an optimal vector quantizer satisfies (Lloyd–Max conditions):

1. *Nearest Neighborhood Condition (NNC)*: For a given codebook $\mathcal{G}$ the partition regions $\mathcal{R}_n$, $1 \leq n \leq N$, satisfy

$$\mathcal{R}_n = \{\bar{\boldsymbol{g}} \in \mathcal{G}_T | d(\hat{\boldsymbol{g}}_n, \bar{\boldsymbol{g}}) \leq d(\hat{\boldsymbol{g}}_m, \bar{\boldsymbol{g}}), \forall m \neq n\}, \tag{5.5}$$

   i.e., $\mathcal{R}_n$ is the Voronoi region of codeword $\hat{\boldsymbol{g}}_n$, $1 \leq n \leq N$.

2. *Centroid Condition (CC)*: For given partitions $\mathcal{R}_n$, $1 \leq n \leq N$, the optimum codewords satisfy

$$\hat{\bar{g}}_n = \underset{\bar{g} \in \mathcal{R}_n}{\operatorname{argmin}} \left\{ \mathrm{MQE}_n(\bar{g}) \right\}, \tag{5.6}$$

where the MQE for region $\mathcal{R}_n$ and candidate codeword $\bar{g}$ is defined as

$$\mathrm{MQE}_n(\bar{g}) \triangleq \frac{1}{|\mathcal{R}_n|} \sum_{\bar{g}' \in \mathcal{R}_n} d(\bar{g}, \bar{g}'), \tag{5.7}$$

where $|\mathcal{R}_n|$ denotes the number of training BFF vectors in region $\mathcal{R}_n$. The LBG algorithm applies the above two conditions in an iterative fashion. In the first part of each iteration, using the NNC the training set $\mathcal{G}_T$ is partitioned into $N$ regions $\mathcal{R}_n$, $1 \leq n \leq N$, based on the current codebook $\mathcal{G}$.. In the second part of the iteration, the CC is used to find a new codebook based on the partitions found in part one of the iteration. We note that, in contrast to the typically used Euclidean distance distortion measure [14, 16], for the distortion measure in Eq. (5.3) it is not possible to find a closed–form solution for the centroid in Eq. (5.6). Therefore, $\mathrm{MQE}_n(\bar{g})$ is computed for all training vectors $\bar{g} \in \mathcal{R}_n$, and that training vector which minimizes $\mathrm{MQE}_n(\bar{g})$ is chosen as the new codeword for region $\mathcal{R}_n$. The complexity of this operation can be considerably reduced by pre–computing and storing all $T^2$ possible $d(\bar{g}_m, \bar{g}_n)$, $1 \leq m, n \leq T$. The LBG algorithm terminates if the reduction in the global MQE given by Eq. (5.2) from one iteration to the next iteration becomes negligible.

Unfortunately, a codebook $\mathcal{G}$ satisfying the NCC and the CC may be a local optimum, and therefore, the final codebook obtained by the LBG algorithm may be a local optimum as well [35]. The most common approach to mitigate the effects of the local optimum problem is to run the LBG algorithm for many different random initial codebooks [16]. The final codebook which yields the lowest global MQE is then used for quantization. Recently, more systematic approaches to overcome the local optimum problem have been reported in the neural network and pattern recognition literature [16]. Two prominent examples are the *enhanced LBG algorithm* [36] and the *adaptive incremental LBG*

*algorithm* [37]. These algorithms try to optimize the codeword placement using the assumption that for the globally optimum vector quantizer the MQEs of all partition regions are approximately equal [38]. However, while this assumption is valid for large codebooks, it is questionable for codebooks with a small number of codewords [38]. Since for beamforming small codebooks are both desirable and sufficient to achieve close–to–perfect–CSI performance, we do not further pursue the enhanced and the adaptive incremental LBG algorithms here. Instead, we adapt the so–called global $k$–means clustering algorithm [23] to our problem, since it is particularly well suited for small codebooks [16].

## 5.4 Global Vector Quantization (GVQ) Algorithm

The global $k$–means clustering algorithm [23] is based on the assumption that the optimum codebook with $i$ codewords can be obtained by initializing the LBG algorithm with the optimal codebook with $i-1$ codewords. Although this assumption is difficult to prove theoretically, the excellent performance of the global $k$–means clustering algorithm has been shown experimentally in [23]. Therefore, we adapt the global $k$–means clustering algorithm to the finite–rate feedback beamforming problem and refer to the resulting algorithm as GVQ algorithm in the following. The GVQ algorithm can be used to either design the optimum codebook for a given number of codewords $N$ or to find the codebook with the minimum number of codewords for a given target MQE.

The proposed *GVQ algorithm* is summarized as follows:

1. Pre–define the total number of codewords as $N$ or pre–define the target MQE as $\mathrm{MQE}_{\mathrm{tar}}$.

2. Initialize the number of codewords with $i = 1$.

3. Calculate the optimum codeword $\hat{\boldsymbol{g}}_1[1]$ by searching the entire training

set $\mathcal{G}_T$ for that $\bar{\boldsymbol{g}}_n$ which minimizes the MQE in Eq. (5.7), where $\mathcal{R}_1 = \mathcal{G}_T$. Set $\mathcal{G}[1] = \hat{\boldsymbol{g}}_1[1]$ and record the corresponding MQE[1]. If $N = 1$ or $\text{MQE}[1] \leq \text{MQE}_{\text{tar}}$ goto Step 7, otherwise goto Step 4.

4. Increment the iteration number $i \rightarrow i + 1$.

5. Execute the LBG algorithm described in Section 5.3 for all $T - (i-1)$ initial codebooks given by $\{\hat{\boldsymbol{g}}_1[i-1], \hat{\boldsymbol{g}}_2[i-1], \ldots, \hat{\boldsymbol{g}}_{i-1}[i-1], \bar{\boldsymbol{g}}_n\}$, where $\bar{\boldsymbol{g}}_n \in \mathcal{G}_T$, $\bar{\boldsymbol{g}}_n \notin \mathcal{G}[i-1]$. Retain the final codebook delivered by the LBG algorithm with minimum MQE and record it as $\mathcal{G}[i]$. Record the corresponding MQE[$i$].

6. If $i < N$ or if $\text{MQE}[i] > \text{MQE}_{\text{tar}}$ goto Step 4, otherwise goto Step 7.

7. $\mathcal{G}[i]$ is the desired codebook.

It is interesting to note that in order to find the optimum codebook of size $N$, the GVQ algorithm computes all intermediate codebooks of size 1, 2, ..., $N-1$. This property is useful when comparing the performance of codebooks of different size as they can be obtained by executing the GVQ algorithm only once.

Note that the proposed GVQ algorithm is completely deterministic, which is a major advantage over related algorithms such as the enhanced and the adaptive incremental LBG algorithms [16]. On the other hand, for applications with large codebooks (e.g. $N > 500$) such as image compression the main drawback of the underlying $k$–means clustering algorithm is its high complexity. In particular, the $k$–means clustering algorithm (and therefore also the proposed GVQ algorithm) requires $\mathcal{O}(TN)$ executions of the LBG algorithm. However, for finite–rate feedback beamforming problems complexity is not a major concern as typically only a few thousand training BFF vectors are required to capture the statistical behavior of the channel and codebooks with $N < 200$ are usually sufficient to achieve close–to–perfect–CSI performance. Furthermore, it is important to note that codebooks for finite–rate feedback

beamforming are designed off-line. Therefore, the proposed GVQ algorithm is an attractive and feasible solution.

# Chapter 6

# Simulation and Numerical Results

In this chapter, we present simulation and numerical results for DFE and LE with transmit beamforming with perfect and quantized CSI, respectively. For all results shown we assume $N_R = 1$ receive antenna and $N_T = 3$ equally mutually correlated transmit antennas with correlation coefficient $\rho = 0.5$. As relevant practical examples we consider the severely frequency–selective equalizer test (EQ) and the moderately frequency–selective typical urban (TU) channel profiles of the GSM/EDGE system [1]. As the system model introduced in Chapter 2, for modulation we consider GMSK and 8–PSK used in GSM and EDGE, respectively.

In section 6.1, we will show that significant performance gains of beamforming with perfect CSI for both DFE and LE over single–antenna transmission can be achieved. We also find out that very short FIR BFFs closely approach the performance upper bound set by the IIR BFFs. In section 6.2, we only consider MMSE–DFE with finite–rate feedback beamforming due to the fact that MMSE–DFE gives better performance and GVQ, in Chapter 5, can be easily adopted with other equalizations.

## 6.1 Beamforming with Perfect CSI

### 6.1.1 Decision–Feedback Equalization

In this subsection, we compare FIR and IIR beamforming assuming perfect CSI at the transmitter. However, first we illustrate the convergence behavior of the proposed GA and MPM for calculation of the FIR BFF vector $\boldsymbol{g}$ for DFE. We assume $L_g = 1$ and for both algorithms the BFF vector is initialized with $\boldsymbol{g}_0 = \frac{1}{\sqrt{3}}[1\ 1\ 1]^T$ as discussed in Section 4.2.2. Figure 6.1 shows the SNR of MMSE–DFE [calculated from Eq. (4.3)] vs. iteration number $i$ for one random realization of the EQ channel and $10\log_{10}(E_s/N_0) = 10$ dB, where $E_s$ is the average received energy per symbol. While both algorithms converge to

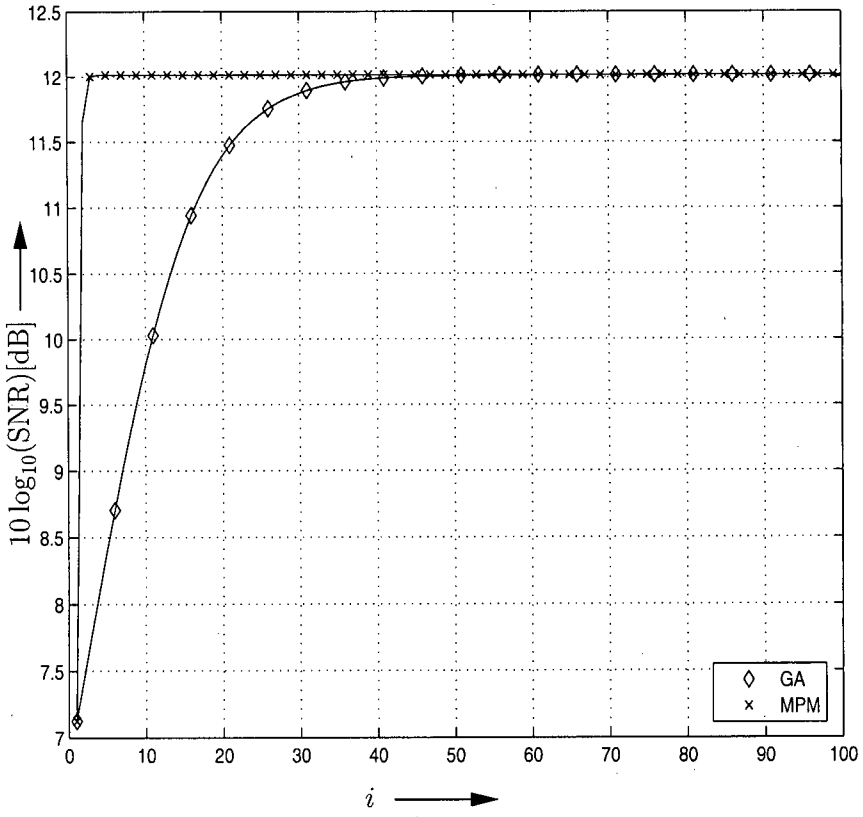


Figure 6.1: SNR of MMSE–DFE vs. iteration $i$ of the proposed GA and MPM for one realization of the EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, equal antenna correlation $\rho = 0.5$, $L_g = 1$, and $10\log_{10}(E_s/N_0) = 10$ dB.

the same SNR value, the convergence speed of the MPM is much higher than that of the GA.

Therefore, for all FIR beamforming results shown in the following for DFE, the BFFs were calculated for each channel realization using the MPM with $\epsilon = 10^{-4}$. The BFF vectors were initialized using the procedure proposed in Section 4.2.2.

In Figures 6.2 and 6.3 we show for, respectively, the EQ and the TU channel the average SNR ($\overline{\text{SNR}}$) of ZF–DFE with IIR BFFs and MMSE–DFE with FIR and IIR BFFs. Thereby, $\overline{\text{SNR}}$ was obtained by averaging the respective

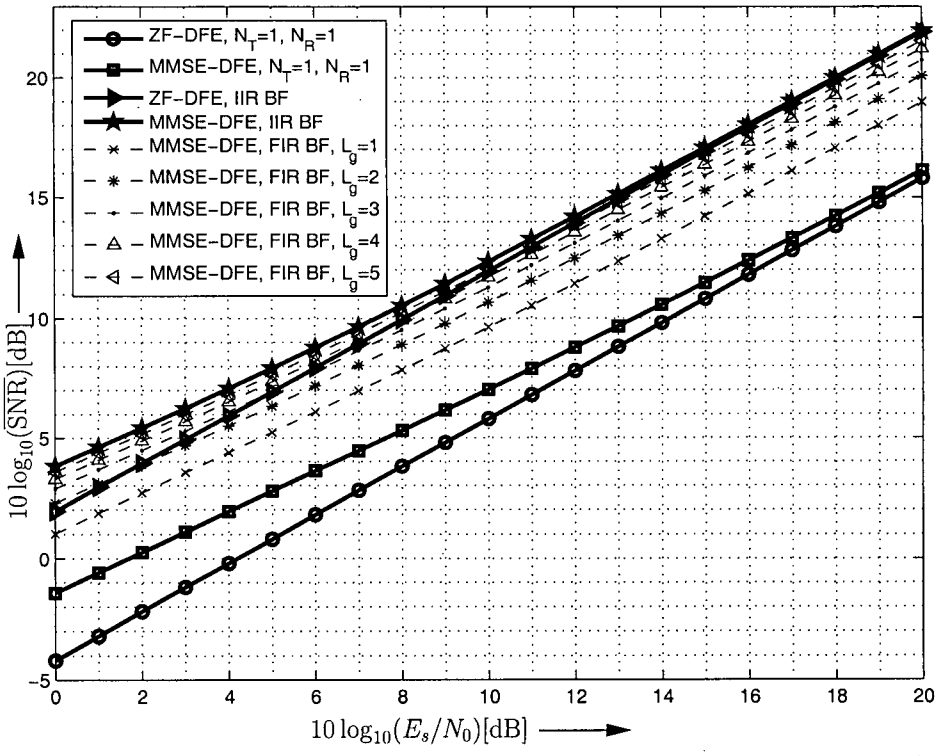


Figure 6.2: Average SNR of DFE for beamforming (BF) with perfect CSI and different BFFs. EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, and equal antenna correlation $\rho = 0.5$. Results for single–antenna transmission are also included.

SNRs in Eqs. (3.15), (3.21), and (4.3) over 500 independent realizations of the EQ and TU channels. For comparison we also show the average SNR of ZF–DFE and MMSE–DFE for single–antenna transmission ($N_T = 1$, $N_R = 1$) in

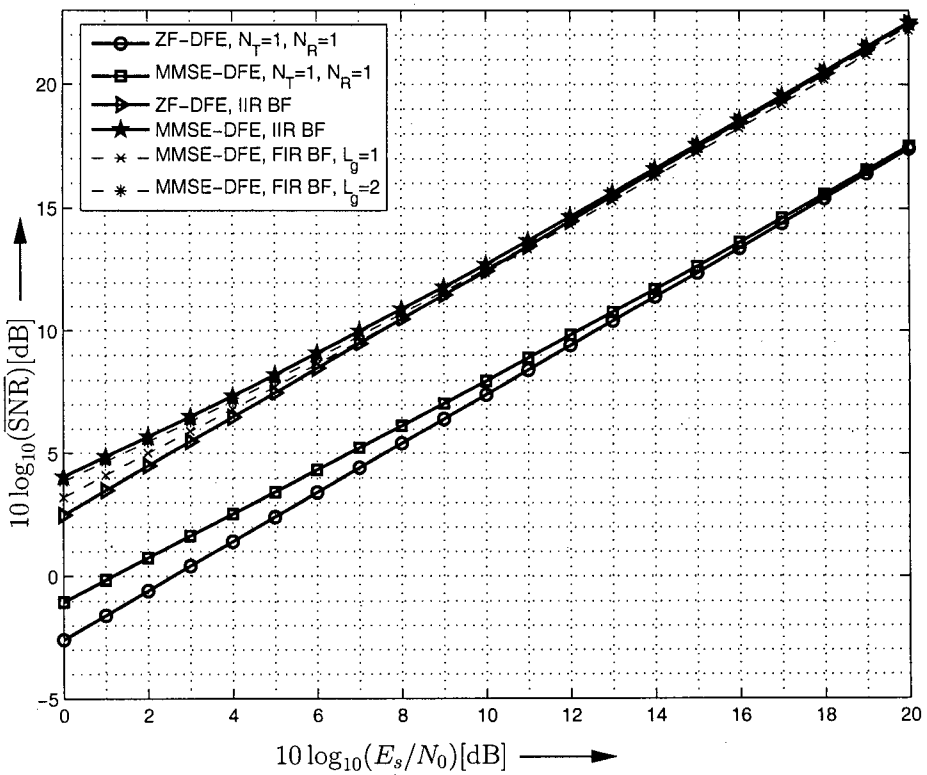


Figure 6.3: Average SNR of DFE for beamforming (BF) with perfect CSI and different BFFs. TU channel with $L = 5$, $N_T = 3$, $N_R = 1$, and equal antenna correlation $\rho = 0.5$. Results for single–antenna transmission are also included.

Figures 6.2 and 6.3. For both channel profiles transmit beamforming with IIR filters leads to performance gains of 5 dB or more compared to single–antenna transmission. However, while for the moderately frequency–selective TU channel an FIR BFF length of $L_g = 1$ achieves practically the same performance as IIR beamforming, especially at high $E_s/N_0$, for the severely frequency–selective EQ channel increasing the FIR BFF length beyond $L_g = 1$ is highly beneficial.

The fact that the FIR BFFs with large enough $L_g$ approach the performance of IIR beamforming in Figures 6.2 and 6.3 also confirms the effectiveness of the MPM for FIR BFF computation.

### 6.1.2 Linear Equalization

Figures 6.4 shows the average SNR ($\overline{\text{SNR}}$) vs. $E_s/N_0$ of MMSE–LE with FIR and IIR beamforming, respectively, where $E_s$ denotes the average received energy per symbol. $\overline{\text{SNR}}$ was obtained by averaging the respective SNRs over 500 independent realizations of the TU channel. For this purpose, in case of FIR beamforming, the SNR given in Eq. (4.11) was used and the corresponding BFFs were calculated with the GA introduced in Chapter 4. For IIR beamforming the results given by Eqs. (3.36) and (3.41) in Chapter 3 are used. As can be observed, IIR beamforming achieves a performance gain of more than 6.5 dB for ZF–LE and 4 dB for MMSE–LE compared to single–antenna transmission.

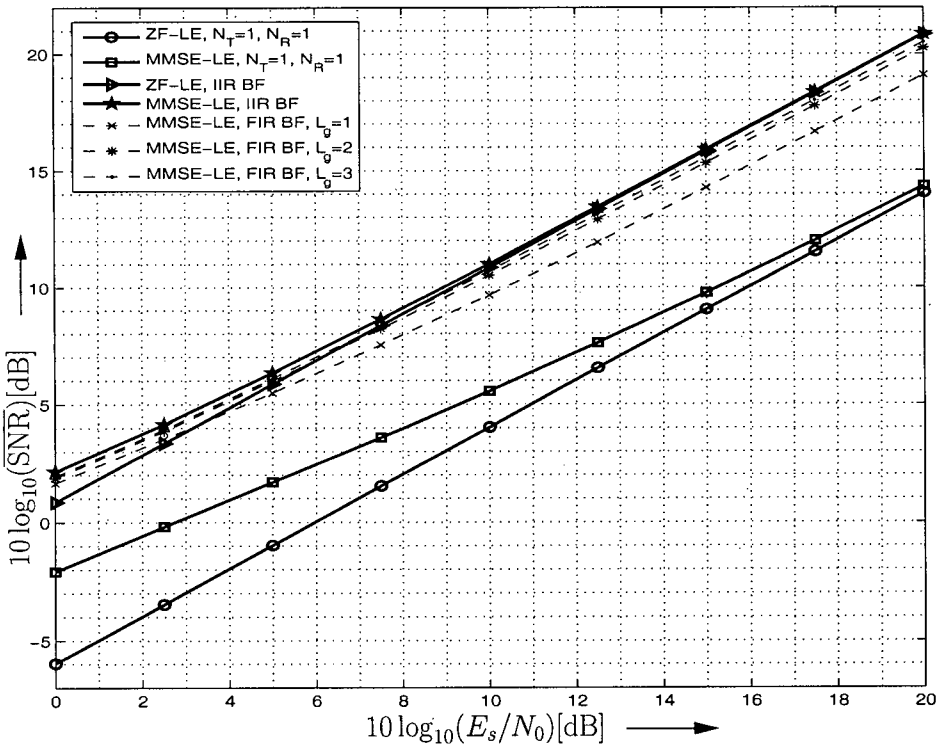


Figure 6.4: Average SNR of LE for beamforming (BF) with FIR and IIR filters. Transmission over TU channel with $L = 5$, $N_T = 3$, $N_R = 1$, and equal antenna correlation $\rho = 0.5$. The result for single–antenna transmission is also included.

As expected, beamforming with IIR BFFs constitutes a natural performance upper bound for beamforming with FIR filters. However, interestingly, for the TU channel an FIR filter length of $L_g = 3$ is sufficient to closely approach the performance of IIR beamforming. This also confirms the effectiveness of the proposed GA. We note that for high $E_s/N_0$ even an FIR BFF length of $L_g = 1$ achieves a performance gain of more than 4.5 dB compared to single–antenna transmission ($N_T = 1$, $N_R = 1$).

Additional simulations for other GSM/EDGE channel profiles have shown that in general FIR BFF lengths of $L_g \leq 6$ are sufficient to closely approach the performance of IIR beamforming. Thereby, the FIR filter length required to approach the performance of IIR beamforming seems to be shorter if the channel is less frequency selective.

Comparing Figures 6.3 and 6.4, they both achieve performances sufficiently close to IIR beamforming with short FIR BFFs. As expected, the average SNR achieved by IIR BFFs with DFE is about 1.5dB higher than that achieved by IIR BFFs with LE. However, about more than 1dB performance gain over single antenna transmission can be achieved at high $E_b/N_0$ by IIR beamforming with LE than that with DFE.

## 6.2 Finite–Rate Feedback Beamforming

In this section, we present simulation and numerical results for the bit error rate (BER) of MMSE–DFE with finite–rate feedback beamforming. For codebook design the proposed GVQ algorithm is applied to training sets of $T = 5000$ independent MIMO channel realizations generated for the EQ and the TU channel profiles, respectively, assuming an SNR of $10\log_{10}(E_b/N_0) = 10$ dB, where $E_b \triangleq E_s/\log_2 M$ denotes the received energy per bit.

Figures 6.5 and 6.6 show the BER of MMSE–DFE for finite–rate feedback beamforming (solid lines) as a function of the number of feedback bits $B$ for, respectively, BPSK transmission over the EQ channel and 8–PSK transmis-

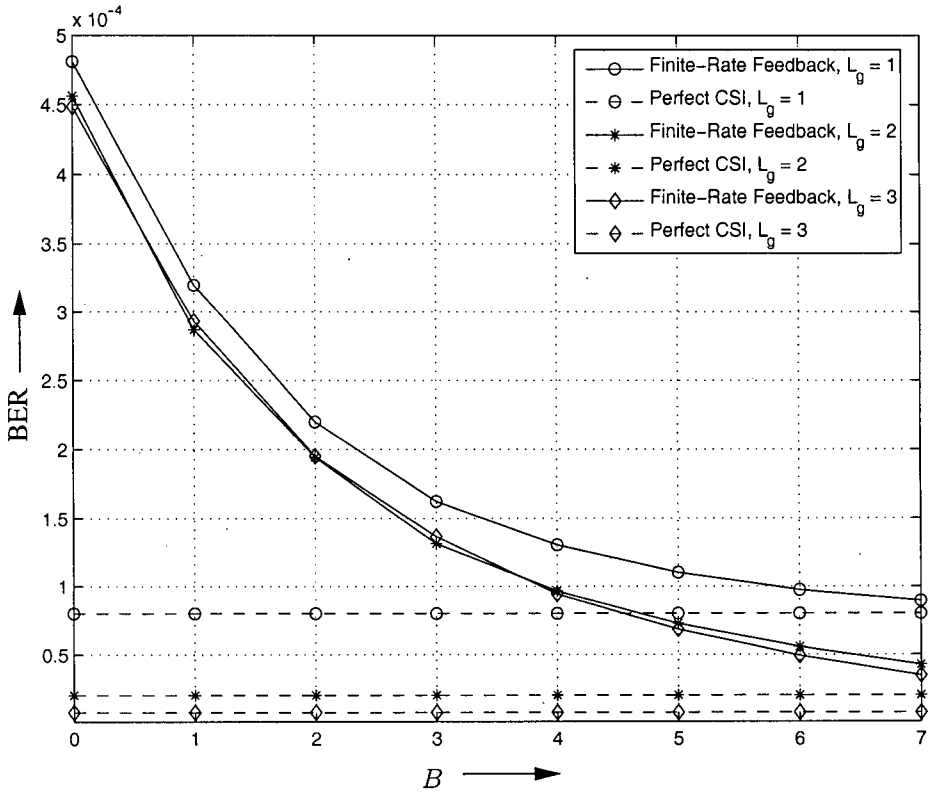


Figure 6.5: BER of MMSE–DFE vs. number of feedback bits $B$ per channel update. BPSK transmission over EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, equal antenna correlation $\rho = 0.5$, and $10\log_{10}(E_b/N_0) = 10$ dB. The BER is obtained from Eq. (5.2).

sion over the TU channel. The BER is identical to the global MQE and is obtained by evaluating Eq. (5.2) for the 5000 training channels. For comparison Figures 6.5 and 6.6 also contain the respective BERs of MMSE–DFE for beamforming with perfect CSI (dashed lines). For $B = 0$ the codebook has just one entry and no feedback is required. In this case, beamforming degenerates to delay diversity. As can be observed from Figures 6.5 and 6.6 finite–rate feedback beamforming approaches the performance of the perfect CSI case as $B$ increases. For the severely frequency–selective EQ channel increasing the BFF length from $L_g = 1$ to $L_g = 2$ or $L_g = 3$ results in a performance gain for both quantized and perfect CSI, cf. Figure 6.5. In contrast, as already

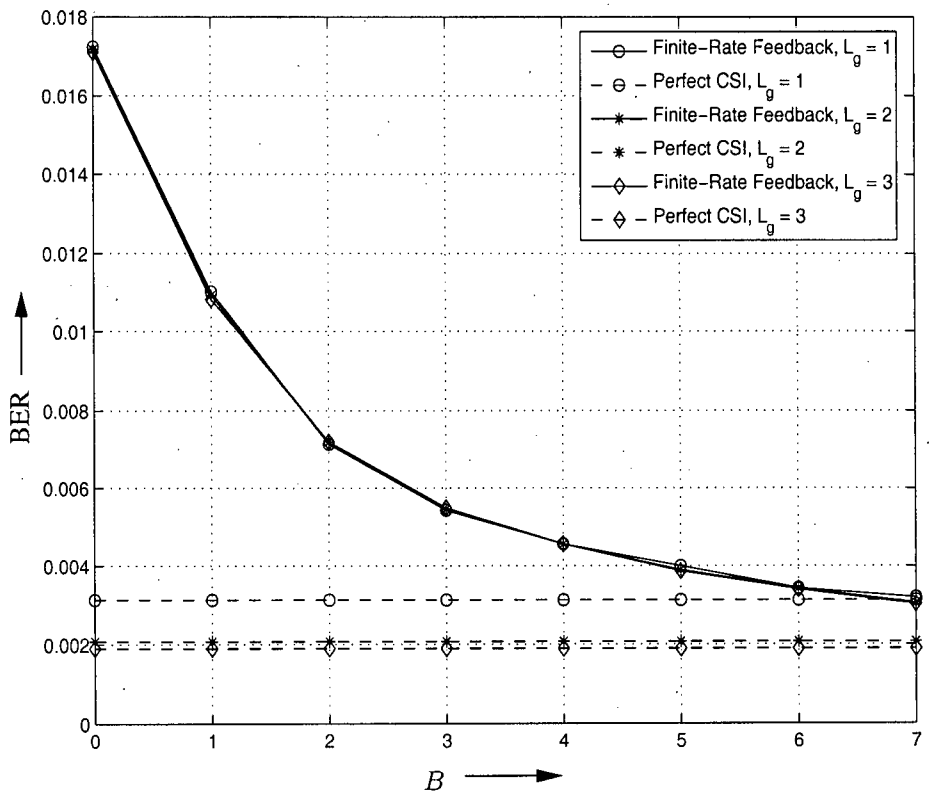


Figure 6.6: BER of MMSE–DFE vs. number of feedback bits $B$ per channel update. 8–PSK transmission over TU channel with $L = 5$, $N_T = 3$, $N_R = 1$, equal antenna correlation $\rho = 0.5$, and $10\log_{10}(E_b/N_0) = 10$ dB. The BER is obtained from Eq. (5.2).

expected from Figure 6.3, for the moderately frequency–selective TU channel $L_g = 1$ is near optimum and while small gains are possible with $L_g = 2, 3$ for perfect CSI, these gains cannot be realized with quantized CSI and $B \leq 7$ feedback bits, cf. Figure 6.6.

Figures 6.7 and 6.8 show the simulated BERs (averaged over 100,000 channel realizations) for BPSK modulation with MMSE–DFE and the EQ channel assuming finite–rate feedback beamforming with BFFs of lengths $L_g = 1$ and $L_g = 3$, respectively. For the simulations we implemented MMSE–DFE with FIR feedforward filters of length $4L_{\text{eq}}$ which caused negligible performance degradation compared to IIR feedforward filters. The DFE feedback filter had

optimum length $L_{\text{eq}} - 1 = L + L_g - 2$.

Figure 6.7 shows that even finite–rate feedback beamforming with BFFs of length $L_g = 1$ and a small number of feedback bits can achieve substantial performance gains over single–antenna transmission ($N_T = 1$, $N_R = 1$). Furthermore, finite–rate feedback beamforming with $B = 1$ feedback bit outperforms antenna selection which employs the codebook $\mathcal{G} \triangleq \{[1\ 0\ 0]^T, [0\ 1\ 0]^T, [0\ 0\ 1]^T\}$ and requires $B = 2$ feedback bits. Finite–rate feedback beamforming with $B = 7$ bits closely approaches the performance of beamforming with perfect CSI.

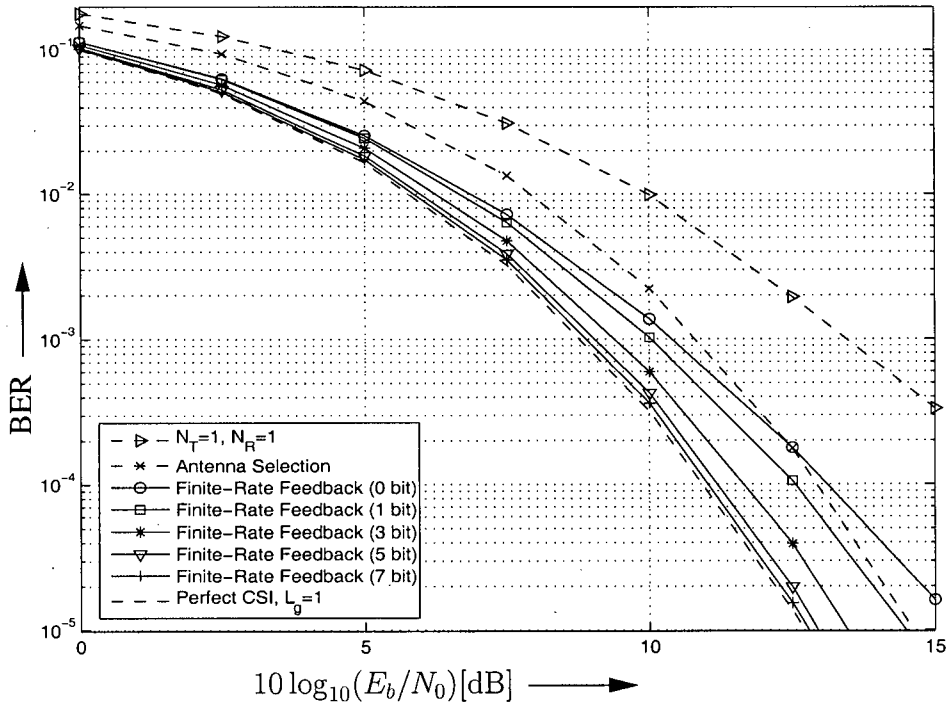


Figure 6.7: Simulated BER of MMSE–DFE for finite–rate feedback beamforming with BFFs of length $L_g = 1$. BPSK transmission over EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, and equal antenna correlation $\rho = 0.5$. Results for single–antenna transmission ($N_T = 1$, $N_R = 1$), antenna selection, and beamforming with perfect CSI are also included.

Figure 6.8 shows that for the degenerate case of $B = 0$ feedback bits the proposed method with $L_g = 3$ achieves a slightly better performance than

the optimized delay diversity (ODD) scheme in [5]. Note that both schemes

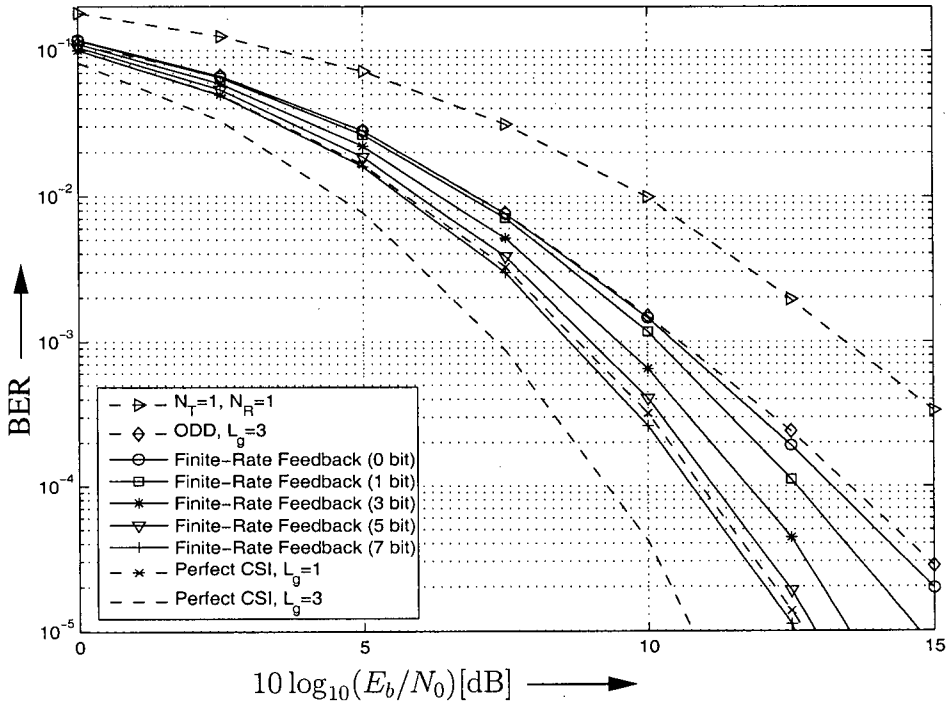


Figure 6.8: Simulated BER of MMSE–DFE for finite–rate feedback beamforming with BFFs of length $L_g = 3$. BPSK transmission over EQ channel with $L = 7$, $N_T = 3$, $N_R = 1$, and equal antenna correlation $\rho = 0.5$. Results for single–antenna transmission ($N_T = 1$, $N_R = 1$), ODD [5], and beamforming with perfect CSI are also included.

employ fixed transmit filters. However, the numerical methods used for filter calculation are completely different leading to small performance differences. Significant performance gains over ODD are possible even with few feedback bits. For example, for BER $= 10^{-4}$ finite–rate feedback beamforming with 1, 3, and 5 feedback bits yields a performance gain of 1.0 dB, 1.8 dB, and 2.4 dB over ODD, respectively. We also observe from Figure 6.8 that finite–rate feedback beamforming with $L_g = 3$ and $B = 7$ feedback bits outperforms beamforming with $L_g = 1$ and perfect CSI.

We note that the BERs at $10\log_{10}(E_b/N_0) = 10$ dB in Figures 6.7 and 6.8 are somewhat higher than the corresponding BERs in Figure 6.5. This

can be attributed to the fact that the BER in Figure 6.5 has been obtained by evaluating Eq. (5.2) which does not take into account the effect of error propagation in the DFE feedback filter, whereas the simulation results shown in Figures 6.7 and 6.8 include this effect, of course. Thereby, the performance for $L_g = 3$ is slightly more affected by error propagation than that for $L_g = 1$ since $L_g = 3$ results in a longer overall channel requiring a longer DFE feedback filter for equalization. This explains why the small performance gain promised by Figure 6.5 when increasing $L_g$ from 1 to 3 cannot be observed in Figures 6.7 and 6.8 for $B < 7$.

Figure 6.9 contains the same BER curves as Figure 6.7. However, now

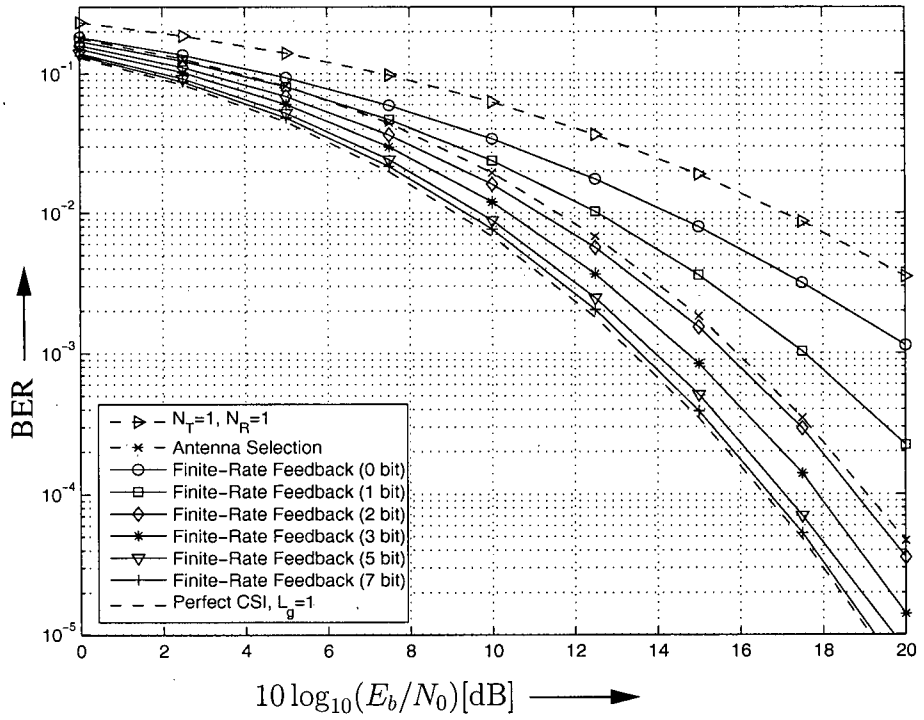


Figure 6.9: Simulated BER of MMSE–DFE for finite–rate feedback beamforming with BFFs of length $L_g = 1$. 8–PSK transmission over TU channel with $L = 5$, $N_T = 3$, $N_R = 1$, and equal antenna correlation $\rho = 0.5$. Results for single–antenna transmission ($N_T = 1$, $N_R = 1$), antenna selection, and beamforming with perfect CSI are also included.

8–PSK transmission over the TU channel is considered instead of BPSK trans-

mission over the EQ channel. Figure 6.9 shows that also for 8–PSK and the TU channel the performance loss incurred by quantized CSI becomes negligible for $B = 7$ feedback bits. Furthermore, for high $E_b/N_0$ finite–rate feedback beamforming with a sufficiently large number of feedback bits can achieve a performance gain of up to 2.2 dB compared to antenna selection.

# Chapter 7

# Conclusions and Future Work

This chapter concludes the thesis with some general comments on the beamforming system with feedback capability proposed in this work, followed by a discussion on possible future work for further investigation of general beamforming systems with finite–rate feedback channel.

## 7.1 Conclusions

In this work, we have considered beamforming with perfect and quantized CSI for single–carrier transmission over frequency–selective fading channels with DFE and LE at the receiver. For the case of perfect CSI we have provided a simple approach for derivation of closed–form expressions for the optimum IIR BFFs and we have developed two efficient numerical methods for calculation of the optimum FIR BFFs for DFE and one for LE. For beamforming with finite–rate feedback channel we have proposed a GVQ algorithm for codebook design. The GVQ algorithm performs a deterministic global search and is therefore independent of the starting conditions. This algorithm is applicable for any number of transmit and receive antennas, arbitrary antenna correlation, and arbitrary fading statistics. Simulation results for typical GSM/EDGE channels have shown that short FIR BFFs can approach the performance of IIR BFFs. Furthermore, for finite–rate feedback beamforming with BFFs of

length $L_g = 1$ few feedback bits are sufficient to approach the performance of beamforming with perfect CSI. In severely frequency–selective channels longer BFFs can further improve performance if a sufficient number of feedback bits can be afforded. Please refer to [39, 40, 41, 42, 43] for a summary of this work.

## 7.2 Recommendations for Future Work

We believe that the research work we initiated here on finite rate feedback beamforming for frequency–selective channels only scratch the tip of the iceberg and many important questions remain to be answered. We list some recommendations for future work as follows:

- In the system model, we assume that the feedback channel is error–free and has zero delay. However, in practice, a zero delay feedback channel cannot be achieved, and the channel realization is time variant within each channel usage. In this regard, researchers consider partial feedback channel with realistic short delay, using statistical models to model channel variations within one burst transmission [44]. It is interesting to perform investigation on imperfect feedback channel in the future work.

- Furthermore, since we only discuss transmit beamforming for uncoded systems in this work, it will be interesting to combined space–time coding and beamforming in the transmitter to cope with imperfect feedback channels and frequency–selective fading environments. However, current research in this field considers only the frequency non–selective case [45, 46].

- Also, in the BER simulations, the assumption that the receiver has perfect CSI is made. In practice, a least sum of squared errors (LSSE) channel estimation algorithm can be used to estimate the CSI from a known training sequences [47]. However, the impact of the channel estimation errors on the BER performance with feedback channel is un-

known. Therefore, channel estimation errors can be taken into account in the future work.

- Last but not least, OFDM has received considerable interests in recent years due to its easy implementation with current powerful digital signal processor (DSP). Due to its high data rate and frequency–selective resistance ability, OFDM is a promising technology for the fourth generation (4G) wireless communications, and has been adopted in the IEEE 802.11x Wireless Local Area Network (WLAN) standard and the IEEE 802.16e Worldwide Interoperability for Microwave Access (WiMax) standard. It is also interesting to note that WiMax and IEEE 802.11n supports partial feedback for beamforming purposes. Therefore, it will be interesting to extend the present work, which considers only single carrier transmission, to OFDM for single user and multiuser cases.

# Bibliography


[1] *GSM Recommendation 05.05: "Propagation Conditions", Vers. 5.3.0, Release 1996.*

[2] W. H. Gerstacker and R. Schober. Equalization concepts for edge. *IEEE Trans. Wireless Commun.*, 1:190–199, January 2002.

[3] A. Furuskar, S. Mazur, F. Muller, and H. Olofsson. Edge: Enhanced data rates for GSM and TDMA/136 evolution. *IEEE Pers. Commun.*, 6:56–66, June 1999.

[4] R. van Nobelen, N. Seshadri, J. Whitehead, and S. Timiri. An adaptive radio link protocol with enhanced data rates for GSM evolution. *IEEE Pers. Commun.*, 6:54–63, February 1999.

[5] S. Yiu, R. Schober, and W. Gerstacker. Optimization of delay diversity for decision–feedback equalization. In *Proceedings of IEEE Symposium on Personal, Indoor and Mobile Radio Communications*, pages 68–72, Barcelona, Spain, September 2004.

[6] J. Winters. On the capacity of radio communication systems with diversity in a Rayleigh fading environment. *IEEE J. Sel. Areas Commun.*, 5:871–878, June 1987.

[7] G. Foschini and M. Gans. On limits of wireless communication in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6:311–335, March 1998.

[8] I. E. Telatar. Capacity of multi–antenna gaussian channels. *European Trans. Telecom.*, 10:585–595, November-December 1999.

[9] P. Dighe, R. Mallik, and S. Jamuar. Analysis of transmit–receive diversity in Rayleigh fading. *IEEE Trans. Commun.*, 51:694–703, April 2003.

[10] G. Jongren, M. Skoglund, and B. Ottersten. Combining beamforming and orthogonal space–time block coding. *IEEE Trans. Inform. Theory*, 48:611–627, March 2002.

[11] D. Love, R. Heath, and T. Strohmer. Grassmannian beamforming for multiple–input multiple–output wireless systems. *IEEE Trans. Inform. Theory*, 49:2735–2747, October 2003.

[12] K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang. Beamforming with finite rate feedback in multiple antenna systems. *IEEE Trans. Inform. Theory*, 49:2562–2579, October 2003.

[13] J. Roh and B. Rao. Transmit beamforming in multiple antenna systems with finite rate feedback: A vq–based approach. *IEEE Trans. Inform. Theory*, IT-52:1101–1112, March 2006.

[14] Y. Linde, A. Buzo, and R. Gray. An algorithm for vector quantizer design. *IEEE Trans. Commun.*, 28:84–95, January 1980.

[15] P. Xia and G. Giannakis. Design and analysis of transmit–beamforming based on limited–rate feedback. *IEEE Trans. Signal Processing*, 54:1853–1863, May 2006.

[16] R. Xu and D. Wunsch II. Survey of clustering algorithms. *IEEE Trans. Neural Networks*, 16:645–678, May 2005.

[17] D.P. Palomar and M.A. Lagunas. Joint transmit–receive space–time equalization in spatially correlated MIMO channels: a beamforming approach. *IEEE J. Select. Areas Commun.*, SAC-21:730–743, 2003.

[18] J. Choi and R. Heath. Interpolation based transmit beamforming for MIMO–OFDM with limited feedback. *IEEE Trans. Signal Processing*, 53:4125–4135, November 2005.

[19] P. Xia, S. Zhou, and G. Giannakis. Adaptive MIMO–OFDM based on partial channel state information. *IEEE Trans. Signal Processing*, 52:202–213, January 2004.

[20] J. Cioffi, G. Dudevoir, M. Eyuboglu, and G. Forney Jr. MMSE decision–feedback equalizers and coding – part I: equalization results. *IEEE Trans. Commun.*, 43:2582–2594, October 1995.

[21] K.E. Baddour and P.J. McLane. Analysis of optimum diversity combining and decision feedback equalization in dispersive Rayleigh fading. In *Proceedings of IEEE International Communications Conference*, pages 21–26, June 1999.

[22] M.V. Clark, L.J. Greenstein, W.K. Kennedy, and M. Shafi. Optimum linear diversity receivers for mobile communications. *IEEE Trans. Veh. Technol.*, 43:47–56, February 1994.

[23] A. Likas, N. Vlassis, and J. Verbeek. The global K–means clustering algorithm. *Pattern Recognition*, 36:451–561, 2003.

[24] D. Gesbert, M. Shafi, Da-Shan Shiu, P. J. Smith, and A. Naguib. From theory to practice: an overview of MIMO space–time coded wireless systems. *IEEE J. Sel. Areas Commun.*, 21(3):281–302, April 2003.

[25] J. G. Proakis. *Digital Communications.* McGraw-Hill, New York, 4th edition, 2001.

[26] T. S. Rappaport. *Wireless Communications.* Prentice-Hall, Englewood Cliffs, NJ, 1996.

[27] H. Bolcskei and A. J. Paulraj. Performance of space–time codes in the presence of spatial fading correlation. In *Conf. Rec. of the 34th Asilomar*

*Conference on Signals, Systems and Computers*, volume 1, pages 686–693, 2000.

[28] D. Chizhik, J. Ling, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber. Multiple input multiple output measurements and modeling in Manhattan. In *Proceedings IEEE Vehicular Technology Conference (VTC)*, pages 107–110, Vancouver, BC, October 2002.

[29] T.K. Moon and W.C. Stirling. *Mathematical methods and algorithms for signal processing.* Prentice Hall, New York, 2000.

[30] G. D. Forney. Maximum–likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Transactions on Information Theory*, IT-18:363–378, May 1972.

[31] C. Luschi, M. Sandell, P. Strauch, and R.-H Yan. Adaptive channel memory truncation for TDMA digital mobile radio. In *Proceedings IEEE International Workshop Intelligent Signal Processing and Communication Systems*, pages 665–669, Melbourne, Australia, November 1998.

[32] R. Weinstock. *Calculus of Variations, with Applications to Physics and Engineering.* Dover, New York, 1974.

[33] R. Cameron. Minimizing the product of two Raleigh quotients. *Linear and Multilinear Algebra*, 13:177–178, 1983.

[34] R. K. Martin et al. Unification and evaluation of equalization structures and design algorithms for discrete multitone modulation systems. *IEEE Trans. Signal Processing*, 53:3880–3894, October 2005.

[35] A. Gersho and R. M. Gray. *Vector quantization and signal compression.* Kluwer Academic Publishers, Norwell, MA, USA, 1991.

[36] G. Patanè and M. Russo. The enhanced LBG algorithm. *Neural Networks*, 14:1219–1237, September 2001.

[37] F. Shen and O. Hasegawa. An adaptive incremental lbg for vector quantization. *To appear in Neural Networks*, 2006.

[38] A. Gersho. Asymptotically optimal block quantization. *IEEE Trans. Inform. Theory*, 25:373–380, July 1979.

[39] Y. Liang, R. Schober, and W. Gerstacker. Transmit beamforming for frequency–selective channels with decision–feedback equalization. *Submitted to IEEE Trans. Wireless Commun.*, May 2006.

[40] Y. Liang, R. Schober, and W. Gerstacker. Fir beamforming for frequency–selective channels with linear equalization.

[41] Y. Liang, R. Schober, and W. Gerstacker. Transmit beamforming for frequency–selective channels. *Accepted for presentation at the IEEE Vehicular Technology Conference, Montreal, Canada*, March 2006.

[42] Y. Liang, R. Schober, and W. Gerstacker. Transmit beamforming with finite–rate feedback for frequency–selective channels. *Accepted for presentation at the IEEE Global Telecommunications Conference (GLOBECOM), San Francisco, USA*, April 2006.

[43] Y. Liang. Transmit beamforming with linear equalization. *Accepted for presentation at the First Canadian Summer School on Communications and Information Theory*, Banff, Canada, June 2006.

[44] E. Kang and A. M. Sayeed. Versatile precoder codebook design method for orthogonal space time block codes. *Accepted by IEEE Symposium on Information Theory, Seattle, USA*, 2006.

[45] M. Skoglund G. Jongren and B. Ottersten. Combining beamforming and orthogonal space–time block coding. *IEEE Trans. Inform. Theory*, 48:611–627, Mar 2002.

[46] R. W. Heath Jr. D. J. Love. Limited feedback unitary precoding for orthogonal space–time block codes. *IEEE Trans. Signal Processing*, 53:64–73, Jan 2005.

[47] S. N. Crozier, D. D. Falconer, and S. A. Mahmoud. Least sum of squared errors (LSSE) channel estimation. *IEE Proceedings-F*, 138:371–378, issue: 4, August 1991.