Project Number:	**CELTIC / CP5-026**
Project Title:	Wireless World Initiative New Radio – WINNER+
Document Type:	PU (Public)

Document Identifier:	D1.7
Document Title:	**D1.7 Intermediate Report on Advanced Antenna Schemes**
Source Activity:	WP1
Editor:	Petri Komulainen
Authors:	Mats Bengtsson, Emil Björnson, Loïc Brunel, Petri Komulainen, Yang Liu, Afif Osseiran, Lars Rasmussen, Florian Roemer, Malte Schellmann, Serdar Sezginer, Bhavani Shankar, Bin Song, Lars Thiele, Antti Tölli, Guillaume Vivier, Ming Xiao
Status / Version:	Final / 1.0
Date Last changes:	09.11.09
File Name:	D1.7.doc

Abstract:	This deliverable captures the second set of best innovative concepts identified in the field of Advanced Antenna Schemes for potential inclusion into the WINNER+ system concept. The concepts consist of promising principles or ideas as well as detailed innovative techniques. For each concept, the associated benefits as well as the corresponding requirements on the system architecture and protocols, measurements and signalling, are considered.

Keywords:	Channel state information, MIMO, network coding, transmit precoding, relaying

Disclaimer:

Table of Contents

Authors

Partner	Name	Phone / Fax / e-mail
Ericsson AB	Afif Osseiran	Fax: +46 10 7172092 e-mail: afif.osseiran@ericsson.com
	Jawad Manssour	Fax: +46 10 7172092 e-mail: jawad.manssour@ericsson.com
Fraunhofer HHI	Malte Schellmann	Phone: +49 30 31002 770 Fax: +49 30 31002 647 e-mail: malte.schellmann@hhi.fraunhofer.de
	Lars Thiele	Phone: +49 30 31002 429 Fax: +49 30 31002 647 e-mail: lars.thiele@hhi.fraunhofer.de
Kungliga Tekniska Högskolan	Lars Rasmussen	Phone: +46(0)87908419 Fax: e-mail: lkra@kth.se
	Ming Xiao	Phone: +46(0)87906577 Fax: e-mail: ming.xiao@ee.kth.se
	Emil Björnson	Phone:+46(0)87908470 Fax: +46(0)87907260 e-mail: emil.bjornson@ee.kth.se
	Bhavani Shankar	Phone:+46(0)87908435 Fax: +46(0)87907260 e-mail: bhavani.shankar@ee.kth.se
	Mats Bengtsson	Phone:+46(0)87908463 Fax: +46(0)87907260 e-mail: mats.bengtsson@ee.kth.se
Mitsubishi Electric R&D Centre Europe	Loïc Brunel	Phone: +33 (0)2 23 45 58 21 Fax: +33 (0)2 23 45 58 59 e-mail: l.brunel@fr.merce.mee.com
	Yang Liu	Phone: +33 (0)2 23 45 58 58 Fax: +33 (0)2 23 45 58 59 e-mail: y.liu@fr.merce.mee.com

SEQUANS Communications	Serdar Sezginer	Phone:+33(0)170721682 Fax: +33(0)170721609 e-mail: serdar@sequans.com
	Guillaume Vivier	Phone : +33(0)170721600 Fax : +33(0)170721609 e-mail : gvivier@sequans.com

Technical University Ilmenau	Florian Roemer	Phone:+49(0)3677691269 Fax: +49(0)3677691195 e-mail: florian.roemer@tu-ilmenau.de
	Bin Song	Phone:+49(0)3677691269 Fax: +49(0)3677691195 e-mail: bin.song@tu-ilmenau.de

University of Oulu	Petri Komulainen	Phone: +358 8 553 2971 Fax: +358 8 553 2845 e-mail: petri.komulainen@ee.oulu.fi
	Antti Tölli	Phone: +358 8 553 2986 Fax: +358 8 553 2845 e-mail: antti.tolli@ee.oulu.fi

Executive Summary

This deliverable captures the second set of best innovative concepts identified in the field of Advanced Antenna Schemes for potential inclusion into the WINNER+ system concept. The concepts consist of promising principles or ideas as well as detailed innovative techniques. For each concept, the associated benefits and the corresponding requirements on the system architecture and protocols, measurements and signalling, are considered.

The scope of this document is in various communication concepts related to antenna processing. One of the most interesting and promising novel techniques, proposed for cellular systems, is the Coordinated MultiPoint (CoMP) transmission and reception. The framework of CoMP and the related innovative concepts are presented in another WINNER+ deliverable, D1.8 "Intermediate Report on CoMP and Relaying in the Framework of CoMP".

The innovative concepts are described in Chapters 2, 3 and 4. The first set of proposals includes signal processing solutions for multiuser MIMO systems. Here, the context is the downlink of a cellular network, where a base station employing an antenna array communicates with user terminals, each equipped with one or more antenna elements. The framework of the presented solutions consists of spatial user multiplexing or scheduling, and beamforming by means of linear transmit precoding. Since both the precoding and the scheduling depend heavily on the CSI knowledge in the transmitter (CSIT), the proposals focus on how to make the CSIT available. The problem of acquiring the CSIT consists of multiple tasks, such as pilot signal design, channel state and quality estimation, as well as feedback signal design. All these aspects are addressed in order to enhance the system performance. These proposals form a realistic and promising set of improvements for accommodating precoded MIMO transmission in multiuser systems.

The second set of innovative concepts focuses on communication strategies and network topologies involving relay nodes. The main aim of relaying is to increase the cell coverage, and to provide more uniform service quality over the whole geographical area comprising the cell. The first proposal shows how multiuser relaying by network coding can utilize multiple relay nodes, by employing different, linearly independent codes in the relays. The second concept proposes a new two-way MIMO amplify-and-forward relaying strategy for terminal-to-terminal communication in TDD mode.

Finally, the third set of innovations is related to the coding and decoding in the point-to-point communication context. One concept explores how to employ network coding techniques in MIMO transmission and reception. In another proposal, the general receiver processing problem of joint channel estimation, equalization and decoding is addressed by employing the iterative belief propagation algorithm.

Chapter 5 is dedicated for an overview of the MIMO schemes of WiMAX systems, i.e., in the IEEE 802.16e standard and its enhancements in IEEE 802.16m system description document. In particular, the chapter focuses on the diversity-rate trade-off from a receiver complexity point of view and highlights various precoding schemes, their performance and resulting complexity. Similarly to the LTE track, the main emphasis is on codebook-based precoding, but in the TDD mode sounding based precoding will be supported as well.

List of acronyms and abbreviations

3G	3rd Generation
3GPP	3rd Generation Partnership Project
AF	Amplify-and-Forward
AMC	Adaptive Modulation and Coding
AP	Access Point
APP	A Posteriori Probability
ARQ	Automatic Repeat reQuest
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BLER	Block Error Rate
BP	Belief Propagation
BP-DUGA	Belief Propagation with Downward and Upward Gaussian Approximation
BPSK	Binary Phase Shift Keying
BS	Base Station
CAZAC	Constant Amplitude Zero Auto-Correlation
CDF	Cumulative Distribution Function
CCDF	Complementary Cumulative Distribution Function
CoMP	Coordinated MultiPoint
CQI	Channel Quality Indicator
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
DF	Decode-and-Forward
DFT	Discrete Fourier Transform
DL	Downlink
DPC	Dirty Paper Coding
EM	Expectation-Maximisation
eNB	Evolved Node B
FDD	Frequency Division Duplex
FFT	Fast Fourier Transform
HARQ	Hybrid Automatic Repeat Request
IEEE	Institute of Electrical and Electronics Engineers
IMT	International Mobile Telecommunications
IMT-A	IMT Advanced
ISI	Inter-Symbol Interference
LA	Local Area
LDPC	Low-density Parity-check
LI	Linearly Independent
LOS	Line Of Sight
LTE	Long Term Evolution of 3GPP mobile system
LTE-A	LTE-Advanced
MA	Metropolitan Area
MAC	Multiple Access Channel
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
MSE	Mean Squared Error
MU	Multi-User
MUMR	Multiple-User Multiple-Relay
NC	Network Coding
OFDM	Orthogonal Frequency Division Multiplexing

OFDMA	Orthogonal Frequency Division Multiple Access
pdf	Probability Density function
PHY	Physical Layer
PMI	Precoding Matrix Index
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
RN	Relay Node
RRM	Radio Resource Management
Rx	Receive
SDMA	Spatial Division Multiple Access
SM	Spatial Multiplexing
SINR	Signal to Interference plus Noise Ratio
SISO	Single-Input Single-Output
SNR	Signal to Noise Ratio
STBC	Space Time Block Codes
STC	Space-Time Code
STNC	Space-Time Network Coding
SVD	Singular Value Decomposition
TC	Turbo Code
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
Tx	Transmit
UE	User Equipment
UL	Uplink
UT	User Terminal
WA	Wide Area
WiMAX	Worldwide Interoperability for Microwave Access
XOR	eXlusive OR
ZF	Zero-Forcing

1. Introduction

This deliverable captures the second set of best innovative concepts identified in the field of Advanced Antenna Schemes for potential inclusion into the WINNER+ system concept. The concepts consist of promising principles or ideas as well as detailed innovative techniques. Most of the proposals are completely new, while others continue the work presented in earlier WINNER+ deliverables [WIN+D13] and [WIN+D14]. For each concept, the associated benefits and the corresponding requirements on the system architecture and protocols, measurements and signalling, are considered. The baseline system, against which the benefits of the proposed innovations are to be evaluated, is the current LTE concept enhanced by selected features from WINNER II. The basic characteristics of the baseline are OFDMA-based Multiple-Input Multiple-Output (MIMO) transmission, fast radio resource allocation, link adaptation and retransmissions, as well as optimization of the transmission parameters according to the user terminal velocity. The innovative concepts are described in Chapters 2, 3 and 4.

One of the principal radio techniques to be considered when developing future radio systems is MIMO communication, based on multiple antennas both at the transmitters (TX) and the receivers (RX). The spectral efficiency of MIMO transmission can be significantly increased if channel state information (CSI) is available at the transmitter, allowing the system to effectively adapt to the radio channel and take full advantage of the available spectrum. The main challenge is to make the CSI available at the transmitter (CSIT). This can be achieved by conveying feedback information over the reverse link as in frequency division duplex (FDD) systems. However, providing full CSI via feedback may cause an excessive overhead, and hence quantized instantaneous and/or statistical CSI are preferable in practice. A time division duplex (TDD) system uses the same carrier frequency alternately for transmission and reception, and thus the CSI can be tracked at the transmitter during receive periods, provided that fading is sufficiently slow and the radio chains are well calibrated.

Chapter 2 discusses multiuser MIMO systems, and especially the problem of acquiring CSIT, in the context of a cellular network, comprising a base station that employs an antenna array and mobiles with possibly multiple antenna elements as well. Here, the role of multiantenna techniques is essentially to schedule and multiplex users and data streams, and to take advantage of all the degrees of freedom offered by multiantenna processing. The intelligence in the network lies in the base station that gathers CSIT towards each active mobile and performs scheduling or SDMA in a centralized manner.

One recent trend in research and standardization is to enhance the conventional cellular networks by relay nodes. The main aim is to increase the cell coverage, and to provide more uniform service quality over the whole geographical area comprising the cell. While the relaying entails the design challenge to accommodate multihop communications, it also introduces opportunities for novel innovative communication concepts such as network coding and terminal-to-terminal communication via a relay station. Chapter 3 presents innovative concepts in this field. Finally, Chapter 4 explores how to employ network coding techniques in point-to-point MIMO transmission. Furthermore, the general receiver processing problem of joint channel estimation, equalization and decoding is addressed in a concept employing an iterative belief propagation algorithm.

The scope of this document is in various communication concepts related to antenna processing. One of the most interesting and promising novel techniques, proposed for cellular systems, is the Coordinated MultiPoint (CoMP) transmission and reception. The framework of CoMP and the related innovative concepts are presented in another WINNER+ deliverable, D1.8 “Intermediate Report on CoMP and Relaying in the Framework of CoMP”.

IMT-Advanced is a long term endeavor envisioned to provide higher data rates under high mobility beyond what IMT-2000 can offer. It anticipates technologies capable of supporting upto 100 Mbps in high mobility scenarios and up to 1 Gbps in low mobility or nomadic services. Several technology proposals are being considered for IMT-Advanced, the prominent ones being the LTE-Advanced by 3GPP and the IEEE 802.16m from IEEE. While LTE-Advanced is based on LTE, IEEE 802.16m is based on enhancements developed to the IEEE 802.16 standard. Even though the WINNER+ system concept is targeted for LTE-Advanced, it is interesting to follow some of the developments taking place in the IEEE 802.16m track. To this end, Chapter 5 is devoted to an overview of the MIMO schemes present in IEEE 802.16e standard and its enhancements in IEEE 802.16m system description document. In particular, the chapter focuses on the diversity-rate trade-off from a receiver complexity point of view and highlights various precoding schemes, their performance and resulting complexity.

2. Innovative concepts in multiuser MIMO systems

This chapter presents four innovative concepts related to multiuser MIMO systems. The context is the downlink of a cellular network, where a base station employing an antenna array communicates with user terminals, each equipped with one or more antenna elements. The framework of the presented solutions consists of spatial user multiplexing or scheduling, and beamforming by means of linear transmit precoding. Since both the precoding and the scheduling depend on the CSI knowledge in the transmitter (CSIT), the proposals focus on how to make the CSIT available. The problem of acquiring the CSIT consists of multiple tasks, such as pilot signal design, channel state and quality estimation, as well as feedback signal design. All these aspects are addressed in order to enhance the system capacity.

The efficiency of spatial multiplexing by transmit precoding depends on the accuracy of CSIT. In frequency- and time-invariant channels, short-term CSI can be utilized. However, when short-term CSI is not available, transmit precoding can be based on long-term CSI. Here, long-term CSI means the second-order spatial channel statistics averaged over a time period or frequency bandwidth, within which the channel may change significantly.

Section 2.1 presents a method for low-rank modelling of the long-term CSI, estimated over a finite time and frequency bandwidth. In Section 2.2, a downlink pilot signal design technique, optimized for improving the CSI and channel quality estimation accuracy, is introduced. Section 2.3 describes a novel signalling concept for reducing the overhead caused by uplink CSI sounding, needed for multiuser precoding in TDD systems. Finally in 2.4, a predictive channel quality estimation method, utilizing the knowledge of the physical antenna array setup in vehicular receivers, is proposed.

2.1 Multi-user MIMO downlink precoding for time-variant correlated channels

2.1.1 Introduction

Multi-user multiple-input multiple-output (MU-MIMO) systems provide a significantly increased capacity and spectral efficiency by exploiting the benefits of space division multiple access (SDMA). Linear precoding, as a sub-optimal SDMA strategy, has attracted much attention due to its lower complexity compared to dirty paper coding (DPC). If perfect channel state information (CSI) is available at the base station (BS), the multi-user interference (MUI) can be effectively eliminated by performing linear precoding at the BS. If the channel varies too fast to obtain short-term CSI, long-term CSI can be used alternatively to improve the system performance. In this proposal, we propose a new approach to multi-user precoding based on long-term CSI, which can be applied to previously defined precoding techniques originally requiring perfect CSI at the BS. It is shown that a significant performance improvement is achieved by the new approach as compared to a state of the art approach [SH05] to multi-user precoding with long-term CSI, especially for the case when a user has a line of sight (LOS) channel.

Duplexing mode	TDD
Topology / links involved	Basic cellular / downlink
Network deployment	Local and wide area
Target system	LTE-A
History	New
Field of main contribution	DL transmit precoding, CSI estimation

2.1.2 Description of system model

We consider a multi-user MIMO OFDM downlink system. There are K users, each of them is equipped with M_{Ri} receive antennas. The BS has M_T transmit antennas. The total number of receive antennas of all users is denoted by M_R. We use $\boldsymbol{H}_i(N_f, N_t) \in C^{M_{R_i} \times M_T}$ to denote the propagation channel between the BS and the user i at the subcarrier N_f and OFDM symbol N_t. Then the combined MIMO channel matrix of all users can be defined as

$$\boldsymbol{H}(N_f, N_t) = [\boldsymbol{H}_1^T(N_f, N_t) \quad \boldsymbol{H}_2^T(N_f, N_t) \quad \ldots \quad \boldsymbol{H}_K^T(N_f, N_t)]^T. \tag{2.1}$$

We assume that it is not possible to track fast variations of users' channels but the information about spatial correlations of the channels can be obtained.

The downlink input output data model with linear precoding matrix $\boldsymbol{F}$ and decoding matrix $\boldsymbol{D}$ can be expressed as

$$\boldsymbol{y} = \boldsymbol{D}(\boldsymbol{H}(N_f, N_t)\boldsymbol{F}\boldsymbol{x} + \boldsymbol{n}), \tag{2.2}$$

where the vectors $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{n}$ represent the vectors of transmitted symbols, received signals at all users, and additive noise at the receive antennas, respectively. $\boldsymbol{F} = [\boldsymbol{F}_1 \quad \ldots \quad \boldsymbol{F}_K]$ denotes the joint precoding matrix used to mitigate MUI and $\boldsymbol{D} \in C^{r \times M_R}$ is a block-diagonal decoding matrix containing each user's receive filter, $\boldsymbol{D}_i \in C^{r_i \times M_{R_i}}$, which is designed to combine the signals of the user's antennas efficiently. The dimensions r and r_i denote the total number of data streams and the number of data streams at the i user terminal, respectively.

We define a chunk as the basic resource element. A chunk contains N_T consecutive OFDM symbols in the time direction and N_F subcarriers in the frequency direction. Therefore, the number of $N_{chunk} = N_F \cdot N_T$ symbols are available within each chunk. Chunk-wise precoding and decoding is assumed.

2.1.3 Previous long-term CSI method

The authors in [SH05][RFH08] introduce a method to exploit the long-term CSI for multi-user precoding. They define the spatial correlation matrix estimate $\hat{\boldsymbol{R}}_{i,b}$ for the user i and the chunk b as

$$\hat{\boldsymbol{R}}_{i,b} = \frac{1}{N_{\mathrm{chunk}}} \sum_{N_f=1}^{N_F} \sum_{N_t=1}^{N_T} \boldsymbol{H}_i^{\mathrm{H}}(N_f, N_t)\boldsymbol{H}_i(N_f, N_t) \ . \tag{2.3}$$

Its singular value decomposition (SVD) is

$$\hat{\boldsymbol{R}}_{i,b} = \boldsymbol{V}_{i,b}\boldsymbol{\Lambda}_{i,b}\boldsymbol{V}_{i,b}^{\mathrm{H}} \ . \tag{2.4}$$

The multi-user MIMO precoding is now performed on the equivalent channel defined as follows

$$\hat{\boldsymbol{H}}_{i,b} = \boldsymbol{\Lambda}_{i,b}^{1/2}\boldsymbol{V}_{i,b}^{\mathrm{H}} \ . \tag{2.5}$$

2.1.4 ROLT-CSI

When channel varies too rapidly to track, only the information relative to the geometry of the propagation paths is captured by a spatial correlation matrix. In order to effectively perform precoding based on the available CSI at the BS, we propose to exploit the knowledge of the spatial correlation with a new approach called rank-one approximated long-term CSI (ROLT-CSI).

The ROLT-CSI approach is designed to effectively represent the channel by using a rank one approximation of the estimated long-term channel spatial correlation matrix per receive antenna. We use $\boldsymbol{h}_{i,l}^{H}(N_f, N_t)$ to denote the lth row of the channel matrix $\boldsymbol{H}_i(N_f, N_t)$. The index l indicates the lth receive antenna of user i. In this work we estimate the spatial correlation matrix of the lth receive antenna of user i by averaging over one chunk. Let $\hat{\boldsymbol{R}}_{i,b,l}$ denote the estimated spatial correlation matrix of user i, chunk b, and receive antenna l. Then we have

$$\hat{\boldsymbol{R}}_{i,b,l} = \frac{1}{N_{\mathrm{chunk}}} \sum_{N_f=1}^{N_F} \sum_{N_t=1}^{N_T} \boldsymbol{h}_{i,l}(N_f, N_t)\boldsymbol{h}_{i,l}^{\mathrm{H}}(N_f, N_t) \tag{2.6}$$

and its singular value decomposition (SVD) as

$$\hat{\boldsymbol{R}}_{i,b,l} = \boldsymbol{V}_{i,b,l}\boldsymbol{\Lambda}_{i,b,l}\boldsymbol{V}_{i,b,l}^{\mathrm{H}} \ , l = 1, \ldots, M_{\mathrm{R}_i} \ . \tag{2.7}$$

According to [BO02], when only the second-order channel statistics are available at transmitter, the optimum strategy is to transmit along the dominant eigenmode of the matrix $\hat{\boldsymbol{R}}_{i,b,l}$. Therefore, we define the equivalent channel matrix of user i in chunk b as

$$\hat{\boldsymbol{H}}_{i,b} = \boldsymbol{A}_{i,b}\boldsymbol{B}_{i,b} \in \mathbb{C}^{M_{\mathrm{R}_i} \times M_{\mathrm{T}}} \tag{2.8}$$

where

$$\boldsymbol{A}_{i,b} = \begin{bmatrix} \sqrt{\boldsymbol{\Lambda}_{i,b,1}(1,1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sqrt{\boldsymbol{\Lambda}_{i,b,2}(1,1)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sqrt{\boldsymbol{\Lambda}_{i,b,M_{\mathrm{R}_i}}(1,1)} \end{bmatrix},$$

and

$$\boldsymbol{B}_{i,b} = \begin{bmatrix} \boldsymbol{V}^{\mathrm{H}}_{i,b,1}(:,1) \\ \boldsymbol{V}^{\mathrm{H}}_{i,b,2}(:,1) \\ \vdots \\ \boldsymbol{V}^{\mathrm{H}}_{i,b,M_{\mathrm{R}_i}}(:,1) \end{bmatrix}.$$

Here $\boldsymbol{\Lambda}_{i,b,l}(1,1)$ indicates the largest eigenvalue of $\hat{\boldsymbol{R}}_{i,b,l}$ and $\boldsymbol{V}^{H}_{i,b,l}(:,1)$ denotes the corresponding eigenvector of $\hat{\boldsymbol{R}}_{i,b,l}$.

The multi-user MIMO precoding can now be performed on the equivalent channel as defined in equation (2.8). Clearly, the rank-one approximation in equation (2.8) can effectively represent the channel if its spatial correlation matrix in equation (2.6) has a low rank.

2.1.5 Expected performance or benefits

Based on ROLT-CSI, any linear precoding technique designed for perfect CSI at the BS, can be modified for long-term CSI. We just use the equivalent channel in equation (2.8) from the ROLT-CSI approach instead of the exact channel in linear precoding matrix derivation. Here, we use ***uplink dedicated pilots*** to estimate the channel between the user terminal and all BS antennas. For each chunk, there are several pilots available. We compute one channel estimate per pilot and then interpolate between these estimates for every symbol in the chunk. Then we calculate the equivalent channel of the chunk with equation (2.8) for the ROLT-CSI approach and with equation (2.5) for the previous long-term CSI method, respectively. Then the BS can compute the precoding matrix for each chunk. The linear precoding schemes used in the simulations are block diagonalization (BD) [SSH04] and regularized block diagonalization (RBD) [SH08]. We evaluate the throughput performance of BD and RBD precodings, when only long-term CSI is available. The simulation scenario is presented in Appendix A.1.

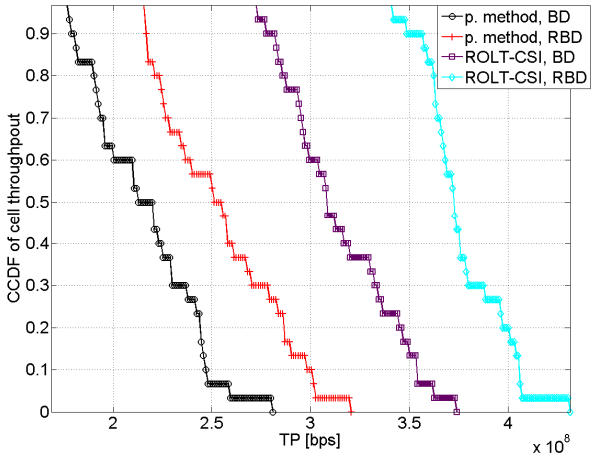


Figure 2.1: Complementary CDF (CCDF) of the sum rates with BD and RBD precoding based on long-term CSI at the transmitter, respectively. p. method indicates the previous long-term CSI method.

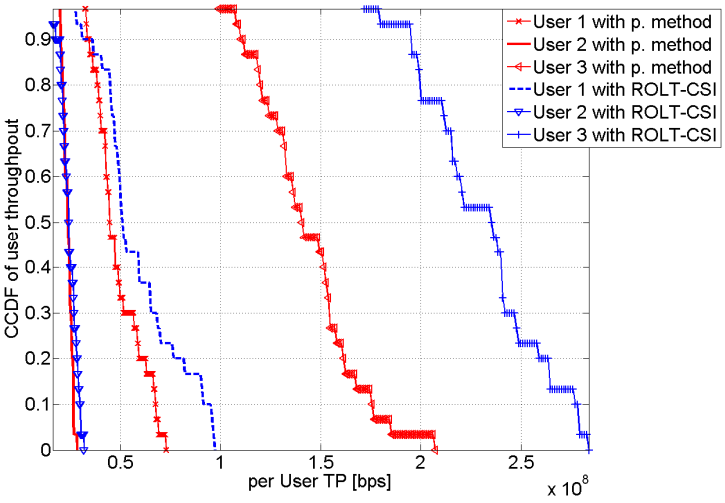


Figure 2.2: CCDF of the individual user throughput with BD precoding based on long-term CSI at the transmitter, p. method indicates the previous long-term CSI method.

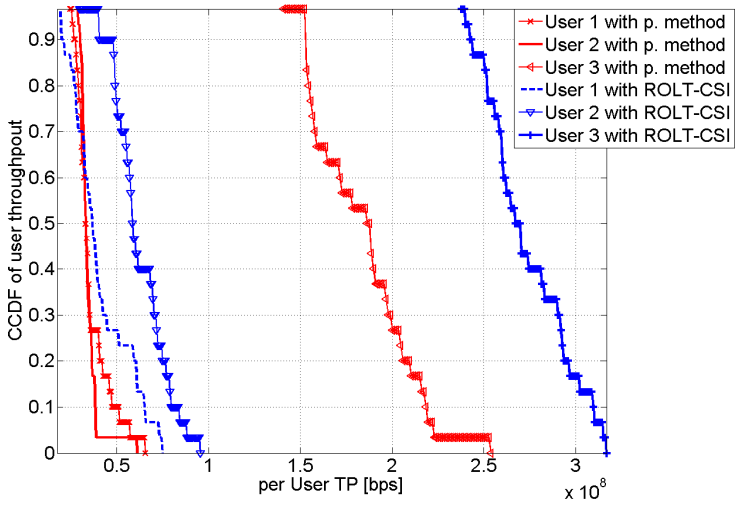


Figure 2.3: CCDF of the individual user throughput with RBD precoding based on long-term CSI at the transmitter, p. method indicates the previous long-term CSI method.

From Figure 2.1 to Figure 2.3, we assume that the channel estimate per pilot of each chunk is perfectly performed. In Figure 2.1, we compare the throughput of the system with precoding based on ROLT-CSI proposed in this paper to the throughput based on the state of the art long-term CSI method in [SH05, RFH08]. We can see that RBD precoding can achieve a higher data rate than BD precoding. When linear precoding is performed based on long-term CSI, a significant performance gain can be achieved by our new approach relative to the previous long-term CSI method.

In Figure 2.2 and Figure 2.3, the individual user throughputs based on ROLT-CSI and the previous long-term CSI approach are compared. It is shown that the ROLT-CSI approach is particularly efficient for the user whose spatial correlation matrix of the channel has low rank. Even for the users who only have NLOS channels, which means that the spatial correlation matrix of these user channels have a high rank, relative to the previous long-term CSI method there are still some performance gains available for the presented ROLT-CSI approach.

2.1.6 Expected requirements on signalling and measurements

Uplink dedicated pilots are needed to estimate the channel between the user terminal and all BS antennas. For each chunk, there are several pilots available. We compute one channel estimate per pilot and then interpolate between these estimates for every symbol in the chunk.

2.1.7 Expected requirements on architecture and protocols

No additional requirements on architecture and protocols are expected.

2.2 Efficient feedback schemes combining long-term and short-term channel information

2.2.1 Introduction

Long term channel information, by virtue of easier acquisition, has been used to improve performance of multiple antenna systems when obtaining complete instantaneous Channel State Information (CSI) involves a premium. However in most systems, there is a possibility of providing a channel quality indicator –quantized function of instantaneous channel – to the transmitter. When available, such short term information can be augmented with long term information to further improve the performance. In the present work, correlation between pairs of elements of the channel matrix constitutes the long term information. The feedback supported eigen-beamforming presented in [WIN+D14] utilizes effective quantized instantaneous channel norm in addition to the channel covariance matrices for designing schedulers and beamformers. This scheme is further evaluated in [WIN+D41] and some of the results are presented below for completeness.

Table 2.1 Performance of Feedback supported Eigen-beamforming scheme evaluated in [WIN +D41]

Feedback Resolution of Channel Norm	Channel Estimation	10%/ 50% / 90% throughput (in Mbps)	Sector throughput (Mbps)
Infinite	Ideal	5.3 / 36.6/ 95	186
1 bit	Ideal	5/36.5 / 95	186
1 bit	Actual	1/ 14.7 / 39.3	84.2

Table 2.1 depicts the throughput achieved with feedback supported eigen-beamforming. In the set-up leading to Table 2.1, each base station has 4 antennas with each user terminal containing two antennas. Two streams are served per user and hence the number of active users per base station transmission is 2. A statistical zero-forcing beamformer is used for transmit precoding. The effective channel norm, obtained from the resulting eigen-beamformed channel, is fed back after quantization. A detailed description of the channel estimation technique is presented in [WIN2D341] and the precoding and quantization scheme is described in [BHO09]. It is clear from Table 2.1 that the proposed scheme performs below par when the channel is estimated at the receiver. This warrants further investigation into the estimation of channel and its norm. Such an investigation is detailed in the ensuing section.

Duplexing mode	FDD or TDD
Topology / links involved	Basic cellular / downlink and uplink
Network deployment	Local and wide area
Target system	LTE-A
History	Refinement from D1.4 and WINNER II
Field of main contribution	DL pilot signalling, feedback signalling, channel and channel norm estimation in terminals

2.2.2 Description

We consider the downlink of a multi-user system with multiple antennas at each end. The transmitter has covariance information about the channel and interference for each user. Each user, in addition to the covariance information, also estimates the instantaneous channel based on pilot transmissions. Many pilot transmission schemes have been proposed in literature providing varied performance. In the [WIN+D14] proposal, the channel norm is estimated based on the channel estimates. In this refinement, we provide a training based estimation scheme that allows for a direct estimation of the channel norm in addition to the instantaneous channel. The long term information is exploited in the design of the training scheme. For

the purpose of completeness, we consider estimation of true channel norm. The general results can be easily simplified for feedback of effective channel norm in the proposed eigen-beamforming scheme.

Long-term information based training scheme

In this scheme [BO09], [BO10], the pilot matrix is defined in terms of its SVD with the singular values depicting the power loading. An expression for the MSE of channel estimation is then obtained in terms of the long term channel and interference covariance matrices as well as the pilot matrix. The various parameters of the pilot matrix are obtained by minimizing the MSE of the channel estimation. For the case of Kronecker structured covariance matrices with no spatial receiver correlation, an explicit structure for the MSE minimizing pilot matrices has been obtained. This is extended to general receiver spatial correlation matrices in [BO10]. In [BO10], it is shown that the obtained structure assigns the j^{th} strongest channel eigen-direction with the j^{th} weakest interference eigen-direction and vice-versa. This can also be thought of as measuring the strongest channel mode when the interference is as weak as possible. At high training powers, the pilot matrix will allocate power for estimation of all channel eigenmodes, while only the strongest eigenmode is estimated at low powers. Furthermore, the channel MSE reduces with increasing transmitter correlation for a fixed receiver correlation. The spatial correlation is also shown to reduce the pilot length. While the optimal pilot matrix is derived for specific covariance structures, a heuristic based on this structure is developed for arbitrarily correlated channel and interference.

An MMSE estimator of the channel norm is obtained and can be employed for any pilot matrix structure. When optimizing the pilot matrix structure towards MSE minimization, it turns out that the optimal power loading is different for estimating the channel and its norm. The differences are not withstanding, the pilot matrices for the two problems tend to have similar ranks for very high and very low training powers. This motivates the use of channel optimized pilot matrix in obtaining MSE estimator of the channel norm. Further details are presented in Appendix A.2.

2.2.3 Expected performance or benefits

Preliminary results for a single cell downlink multi-user MIMO system using proportional fair scheduling have been presented in [BHO09] and the corresponding channel and norm estimation results in [BO09] and [BO10] are repeated here for completeness.

Figure 2.4 compares the normalized MSE of channel estimation for various schemes in an 8 transmit and 4 receive antenna system. The Weichselberger model [WHOB06] is chosen to generate channel realizations for this simulation due to its accurate representation of measured data. The channel mean is chosen as zero and the coupling matrices of Weichselberger model are chosen to be chi-squared distributed [WHOB06], to induce randomized spatial correlation. The standard uniform pilot matrix is only asymptotically optimal, while the heuristic scheme has a negligible loss compared to the optimal performance.

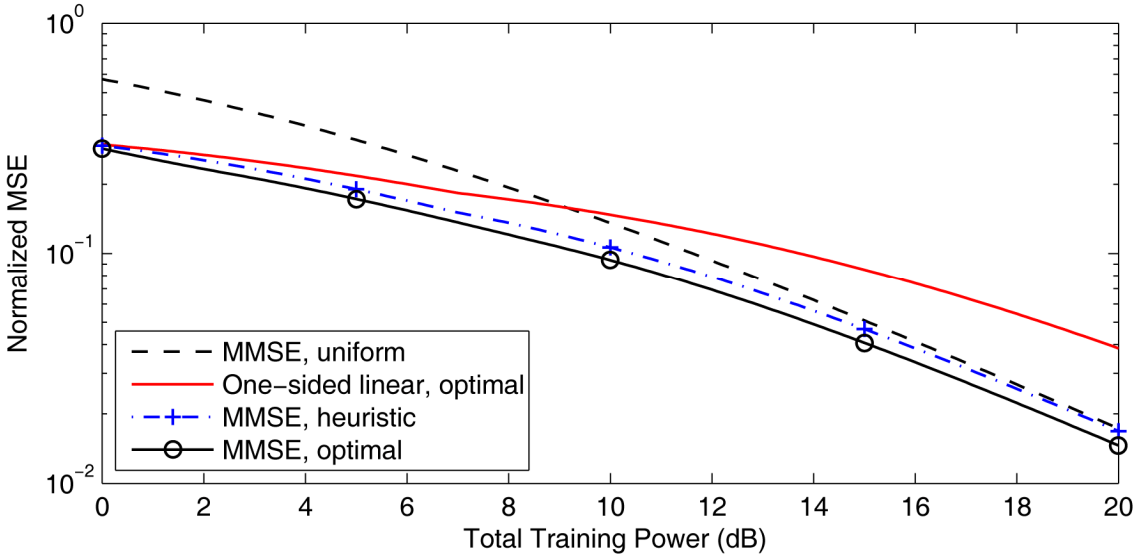


Figure 2.4: The average normalized MSEs of MMSE channel matrix estimation as a function of the training power for optimal and heuristic pilot matrices proposed in [BO09] and [BO10]. The performance of the MMSE estimator with simple uniform training matrix is also shown, along with the performance of the suboptimal linear scheme proposed in [BG06].

Figure 2.5 compares the normalized MSE of channel norm estimation for various schemes in an 8 transmit and 4 receive antenna system [BO09]. A Kronecker model is chosen for the channel with uncorrelated receivers and an exponentially correlated transmitter array. The factor of correlation is chosen to be 0.8 (high correlation). The schemes described as *channel based* involve estimating the norm using the channel estimates. The uniform as well as optimal pilot schemes for channel estimation are used. Further, the figure also depicts the MSE when the norm is obtained directly by minimizing its estimation error for various pilot matrices. It is clearly seen that obtaining norm from the channel estimates involves higher error and that the pilot matrix used for channel estimation yields a satisfactory performance. Based on this result, we propose to apply the heuristic pilot matrix optimized for channel matrix estimation for MMSE estimation of both the channel and its norm.

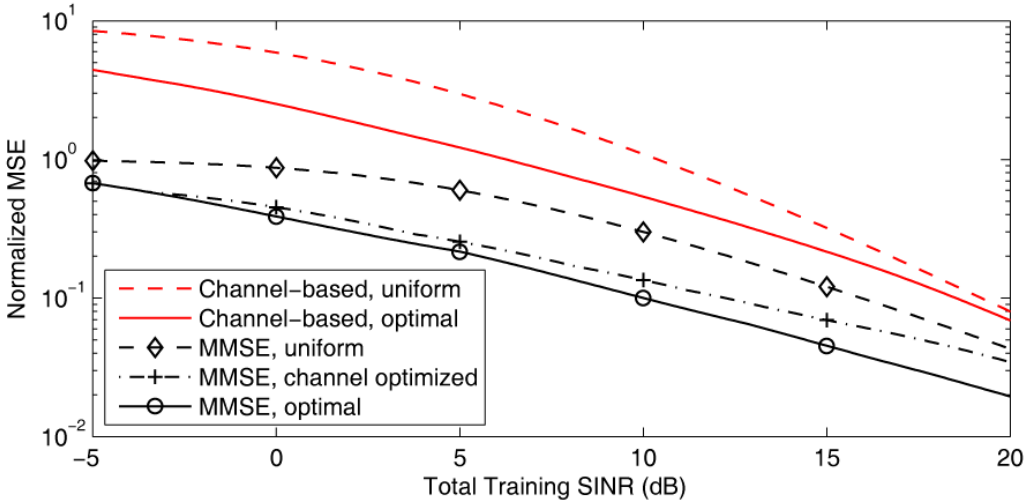


Figure 2.5: The average normalized MSEs of squared channel norm estimation as a function of the training power for direct and indirect estimation schemes.

2.2.4 Expected requirements on signalling and measurements

Since the pilots are precoded, there is a need for dedicated pilots to each user. Further, the statistics of interference at each user must be measured and be made available to the transmitter. However, this can be updated on a lower frequency. The feedback is compatible with LTE Release 8 as it can utilize the CQI reporting of LTE.

2.2.5 Expected requirements on architecture and protocols

The short term CQI fits into previously proposed protocols.

2.3 Pilot overhead reduction for multiuser MIMO systems in TDD mode

2.3.1 Introduction

Efficient transmit precoding or beamforming requires complex spatial channel state information in the transmitter (CSIT). Multiuser precoding, i.e., simultaneous precoding for multiple users, requires centralized CSIT of all the terminals. In the time division duplex (TDD) mode, CSIT for the BS is provided by means of uplink CSI sounding pilot signals. CSIT can be used as a reference for scheduling as well. However, antenna-specific uplink pilot streams cause an extensive overhead that restricts the size of the practical user group and the terminal antenna setup that can be handled within the same time-frequency slot.

Conventionally, the number of the required mutually orthogonal CSI sounding pilot streams corresponds to the aggregate number of terminal antennas that are simultaneously active. Let there be K user terminals in the spatial signal processing group, each with N_k antennas, $k = 1,\ldots,K$, and let the BS have N_B antennas. In practice, $\sum_k N_k$ mutually orthogonal – in time and/or frequency domain – pilot sequences are needed. Thus, system standards must set limits to the number of terminal antennas supported.

The goal of this concept is to reduce the required uplink CSI sounding overhead by letting the terminals form $J_k < N_k$ uplink pilot beams by transmit precoding instead of transmitting N_k antenna-specific pilots.

As a result, the number of the required orthogonal uplink pilot resources reduces from $\sum_k N_k$ to $\sum_k J_k$. Consequently, terminal k appears as a J_k -antenna device to the BS. In the simplest form of operation, J_k equals to unity so that all terminals may be treated as single-antenna devices. The number J_k can be imposed either statically by a standard or dynamically by the BS. This concept is evaluated in more detail in paper [KTL+09] that is based on the research done within WINNER+.

The CSI sounding beams are formed based on the knowledge of the user-specific MIMO channels, obtained via a downlink common pilot signal. This way part of the signalling overhead is moved to the downlink. The common pilot signal is resource efficient since only N_B orthogonal pilot sequences are needed. In cellular systems, downlink common pilots exist for facilitating the reception of common channels.

Duplexing mode	TDD
Topology / links involved	Basic cellular / downlink and uplink
Network deployment	Local area
Target system	LTE-A
History	New
Field of main contribution	Pilot signalling: uplink CSI sounding

2.3.2 Description

The proposed signaling stages are depicted in Figure 2.6, and the corresponding TDD frame structure in Figure 2.7. We assume that the terminals can estimate their individual N_k x N_B MIMO channels $\mathbf{H}_k$ by means of a transmit-antenna-specific downlink common pilot signal before performing CSI sounding. The best choice for the sounding beamformers is then based on the strongest spatial eigenmodes so that the precoding matrix $\mathbf{S}_k = [\mathbf{u}_{k,1} \ldots \mathbf{u}_{k,J}]$ contains the corresponding J_k left singular vectors of the estimated channel. As a result, the BS cannot explicitly estimate the channel matrices but only the J_k x N_B pilot responses $\mathbf{H}_k^T\mathbf{S}_k$. Thus, an estimate of the best signal subspace of each user's channel is declared to the BS. Note that in addition to the estimation noise, CSIT is affected by the channel time variations between uplink and downlink frames.

Many multiuser MIMO scheduling and precoding strategies for the downlink are based on the user-specific channel eigenmodes. From the system sum rate point of view, the optimal number L_k of data streams to be allocated per user is usually less than N_k, especially when either K or N_k is large. Therefore, the weak eigenmodes, neglected in the reduced overhead sounding concept, would rarely be utilized. In the simplest form of operation, the number of streams per user can be restricted to one. For large K, the strategy of allocating at most one beam per user, is asymptotically capacity optimal.

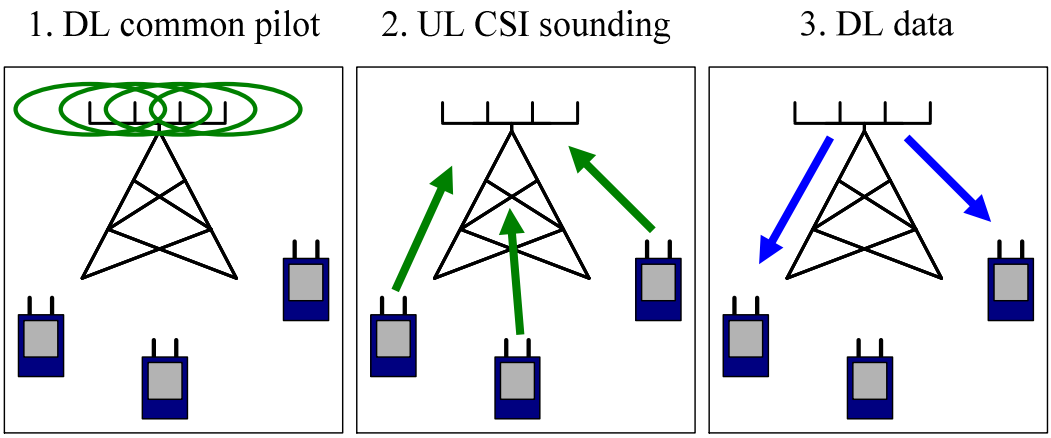


Figure 2.6: Pilot and data signaling stages.

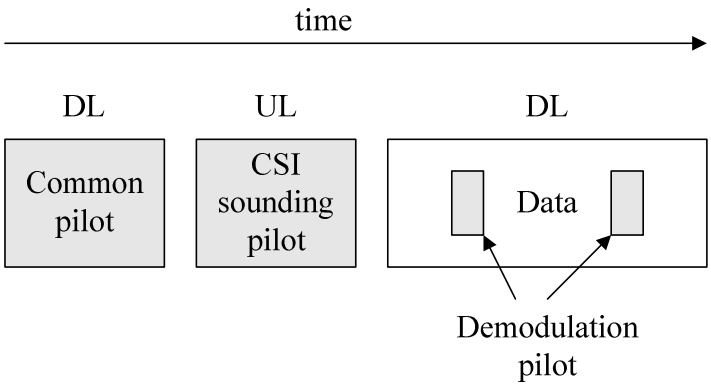


Figure 2.7: Simplified TDD frame structure.

2.3.3 Expected performance or benefits

The concept reduces the pilot overhead caused by CSI sounding. Alternatively, mobile terminals or relay stations that have a different or higher number of antenna elements than supported by system standards, can hide their true number of antennas from the base station. Thus, more advanced user equipment is allowed to be used in the system.

Since the reduced overhead results in reduced CSIT in the BS, the system capacity could be expected to decrease compared to the full overhead case. In Appendix A.3, the performance of the strategy in the context of beam selection and multiuser zero-forcing by coordinated transmit-receive processing is evaluated. According to the results, in the case of perfect CSI estimation, the performance loss induced by the incomplete sounding is minor, as the beamforming gain provided by multiple terminal antennas, and the multiuser diversity seen by the BS are retained. As shown in Figure 2.8, when taking into account the CSI estimation error in the BS, caused by limited pilot power, the overhead reduction turns out to improve robustness and even increase the average system capacity. This is due to the power efficiency of the CSI sounding concept: Uplink transmit power is not wasted on the weak eigenmodes that are unlikely to be utilized.

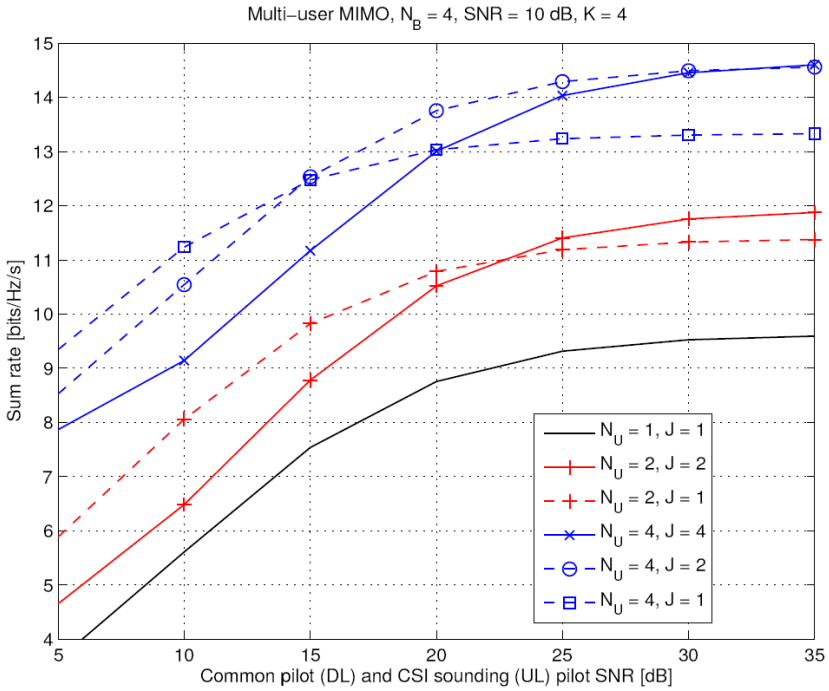


Figure 2.8: Average sum rate with noisy DL common pilot and CSI sounding, LMMSE receivers, $N_B = 4$, $K = 4$, $N_k = N_U$, $J_k = J$.

2.3.4 Expected requirements on signalling and measurements

The required uplink pilot overhead caused by CSI sounding can be reduced. On the other hand, in order to ensure system performance, it may be beneficial to increase the overhead in the downlink, or to distribute the downlink common pilot signals evenly over the frequency spectrum.

2.3.5 Expected requirements on architecture and protocols

The concept requires that the terminal is capable of transmit beamforming. Furthermore, the base station may inform the terminals of the number of allowed CSI sounding beams per user. Alternatively, an advanced terminal may choose a number independently so that the BS does not need to know how many antennas the terminal actually has.

2.4 Adaptive MIMO transmission in time-varying channels: Predicting future channels

2.4.1 Introduction

Channel adaptive transmission in multi-user MIMO systems is seen as a promising concept to achieve high spectral efficiencies in future radio networks. Facilitation of this concept requires some information on the current channel state at the transmitter. For the downlink in FDD systems, this information may be provided by the user terminals to the base station via a (usually limited) feedback channel. The base station may then allocate transmission resources to the user terminals where they can support high data rates. However, due to the delay between evaluation of the channels at the terminals and application of the resource allocation decision, the concept operates conveniently only under quasi-static channel conditions. In case user terminals are moving, the channel may vary and thus the adaptive concept may suffer strong performance degradations. For the channel-adaptive transmission concept based on evaluation of SINR conditions, which has been introduced in Section 2.1.2 in [WIN+D14], a solution to predict future SINR conditions based on channel interpolation techniques is proposed. For a limited prediction interval depending on the receive antenna spacing, the technique is capable of diminishing potential SINR losses significantly.

Duplexing mode	FDD
Topology / links involved	Basic cellular / downlink
Network deployment	Wide area, vehicular terminals
Target system	LTE-A
History	Refinement from D1.4
Field of main contribution	Signal processing: SINR prediction in terminals

2.4.2 Description

In a practical system, evaluation of the SINRs at the UT will be carried out based on the channel measured at time instant i_0, while the scheduling decision will be applied at time instant $i_k > i_0$, resulting in a delay Δt, see Figure 2.9. During this time, the channel may change, so that the SINR conditions determined from the channel measured at i_0 may no longer be valid. In case the SINR conditions in a scheduled resource drop down, the channel will be overloaded, and a bit rate that cannot be supported by the channel any longer will be assigned for a resource. Hence, detection errors are very likely to occur, resulting in severe performance degradations. In particular, variations of the SINR conditions will take effect if the delay Δt is in the order of the channel's coherence time T_c, which we define as $T_c = 1/ f_D$, with f_D being the maximum Doppler frequency. For a fixed system configuration, Δt is in general a constant system parameter, and hence only the Doppler frequency f_D, which is invoked by the speed of the UT, impacts the SINR variations.

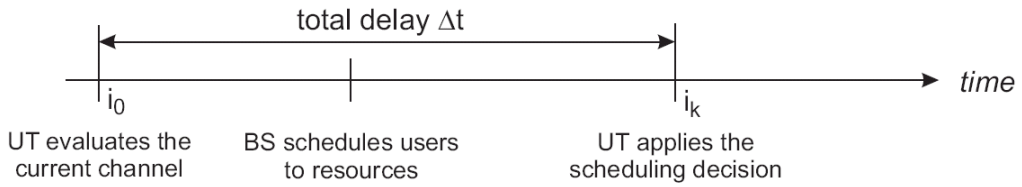


Figure 2.9: Influence of the delay caused by the feedback of UTs.

To alleviate the effect of SINR degradations, it would be desirable to predict the channel state at time instant $i_k = i_0 + \Delta t$ at the UT and carry out the SINR evaluation process based on that channel, $H(i_k)$. For channel prediction, the channels measurements gathered up to time instant i_0 as well as statistical information on the channel dynamics can beneficially be used.

As variations of the channel H(i) over the duration Δt become significant rather at vehicular speeds, it is reasonable to assume that the receive antennas of the mobile terminal can be mounted at a fixed position on the moving vehicle. We assume the antennas to be arranged as a uniform linear array (ULA) with fixed antenna spacing, with its broadside oriented in the moving direction (see Figure 2.10) – an idea which has already been presented in [KK07]. With this configuration, the channel seen at receive antenna 2 (Rx2) is a delayed version of the one seen at Rx1. The delay D, measured in integer numbers of OFDM symbols of length T_o, depends on the spacing *a* of the two antennas as well as on the vehicle's moving speed. Thus, an observation at the ULA at time instant i delivers 2 sampling points of the channel impulse response $h_n(i)$, which characterizes the channel from a single transmit antenna. By using channel interpolation techniques [HKR97], we can then determine $h_n(j)$ for an arbitrary j and thus obtain an estimate of the channel $H(i_k)$ at the future time instant $i_k > i_0$. A requirement for the interpolation technique to operate conveniently results from the sampling theorem, yielding

$$2 f_D D T_o \leq 1 \quad \Leftrightarrow \quad a \leq \lambda/2$$

For further details of the interpolation process, refer to Appendix A.4.

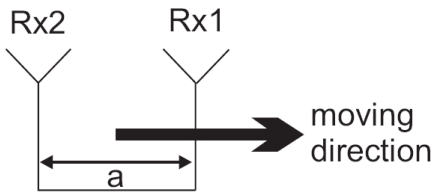


Figure 2.10: Configuration of the ULA at the UT for the prediction-based approach.

2.4.3 Expected performance or benefits

The achievable system performance is illustrated in Figure 2.11, where the antenna spacing at the ULA was set to λ/2. The performance measure is the SINR loss, ΔSINR, representing the difference between the SINR determined at the UT based on the predicted channel and the true SINR valid at time instant i_k. We focus on the 10-percentile of the overall CDF, which is plotted versus the channel dynamics, $\Delta t/T_c$. SINR conditions have been evaluated separately for the streams in multi-stream and single-stream MIMO mode (i.e. multiple or a single spatial beam simultaneously active, each used to transmit an independent data stream). Evaluation is based on a 2x2 MIMO configuration, where at most 2 spatial streams may be simultaneously active. The performance of the prediction-based approach is given by the dashed lines. As reference case (solid lines), the classical approach has been used, where channel evaluation at the UTs is carried out at time instant i_0. We observe that the prediction-based approach yields a significantly smoother degradation of the SINR conditions in the range $\Delta t < 0.5T_c$, suggesting that this technique may be seen as a promising solution to support instantaneous channel-adaptive MIMO transmission also in mobile environments. For further details on the proposed technique as well as the performance evaluation, refer to [STJ08].

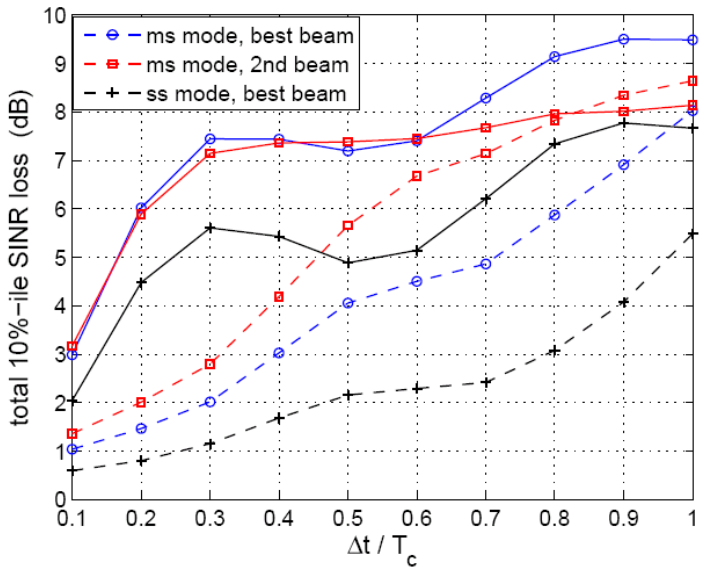


Figure 2.11: Total loss of the per-stream SINR (in dB) for the 10-percentile of the DSINR-CDF measured (solid) and predicted channels (dashed).

2.4.4 Expected requirements on signalling and measurements

- Feedback is given in terms of CQI and preferred beam indices for the favoured transmission mode per chunk, as in the channel-adaptive concept proposed already in [WIN+D14].
- Predicting the future channels based on past measured channels enables a chunk-wise evaluation of the future per-stream SINR achievable with the different spatial transmission modes (single-stream or multi-stream), resulting in the desired CQI.

2.4.5 Expected requirements on architecture and protocols

- Channel prediction requires tracking and storing the MIMO channel measurements per chunk for a set of successive transmission time slots.
- From past channel measurements, UTs derive statistical information on the channel dynamics (time-domain autocorrelation function).

3. Innovative concepts in relaying and network coding

This chapter deals with concepts related to communication strategies and network topologies involving relay nodes. The main aim is to increase the cell coverage, and to provide more uniform service quality over the whole geographical area comprising the cell. While the relaying entails the design challenge to accommodate multihop communication, it also introduces opportunities for innovative communication concepts.

Section 3.1 proposes the use of multiple network codes in the context of uplink comprising multiple users and relay stations. In Section 3.2, a novel two-way MIMO amplify-and-forward (AF) relaying strategy for terminal-to-terminal communication via a relay station is presented.

3.1 Network coding for multiple-user multiple-relay systems

3.1.1 Introduction

As an efficient method for combating fading effects over wireless channels, relay networks/channels ([CG79][LTW04][KGG05]) have been subject to significant research efforts. For example, the three-node relaying approach has been adapted to LTE for increasing coverage. In relay networks, one or more intermediate nodes assist the source node in transmitting information, according to a specific protocol. Various transmission protocols have been proposed for relay networks, e.g., amplify-forward and (selective) decode-forward [CG79][LTW04][KGG05], each with different complexity and performance.

In addition, network coding ([LYC03][KM03]) was proposed to reduce bit error rates in multiple-source relay networks ([CKL6]). By combining information from different sources (or to different sinks) performance improvements are obtained based on joint information processing at intermediate nodes. The results show that relay networks with network coding have better performance (in the energy-efficiency, bit error rate etc) than those without network coding.

Here, we investigate the design of network codes for multiple-user multiple-relay (MUMR) wireless networks. In these networks, M ($M \geq 2$) users have independent information to be transmitted to a common base station (BS), with the help of N ($N \geq 2$) relays. The transmission links within the networks are modeled as independent quasi-static fading channels. This setting provides a general framework, encompassing previously investigated cases with only one relay. We propose to use linearly independent network codes for such scenario. Network codes are normally described by its encoding kernel, which denotes the linear relations between the source and the network codeword. For instance, two sources I1 and I2, two codeword C1 and C2 with encoding kernel K1 and K2, respectively. Then, C1 = K1 [I1, I2]; and C2 = K2[I1, I2]. Then, C1 and C2 are linearly independent, if K1 and K2 are linearly independent.

In particular, we investigate the performance of a class of deterministic network codes in such networks in terms of outage probabilities (to measure asymptotic performance with respective SNR: signal-to-noise ratio), and frame error rates. The former leads to theoretical performance limits constituting targets for practical schemes, while the latter provides for a performance measure for more practical settings.

Duplexing mode	Half-Duplex
Topology / links involved	Relay enhanced cellular / uplink
Network deployment	Wide or local area
Target system	LTE-A
History	Refinement from D1.3
Field of main contribution	Network coding

3.1.2 State of the art

Although the application of network coding to general relay networks is natural and clearly beneficial, most of previous schemes (e.g., [LJS06][YK07]) only consider two-source one-relay networks settings as shown in Figure 3.1. In [YK07], coding for an effective two-user one-relay network was considered where the sources also act as relays. However, to increase system performance, multiple relays can be used within a cell. Thus, future cellular wireless networks can be modeled as multiple-user multiple-relay

systems, as shown in Figure 3.2. Surprisingly, design principles for using network coding in multiple-user multiple-relay networks are mostly unexplored. It is therefore valuable to study network coding schemes for such networks. So far only binary network coding schemes have been considered. However, as we shall show, binary network coding is generally suboptimal for MUMR wireless networks, at least for high SNRs (and quasi-static fading channels).

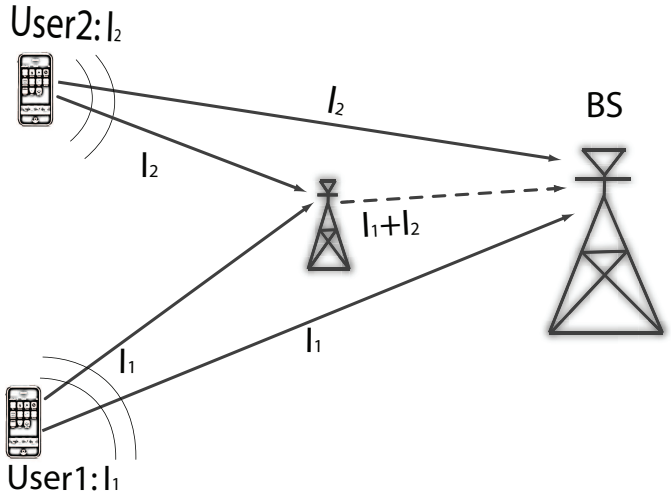


Figure 3.1: Two-source one relay with network coding.

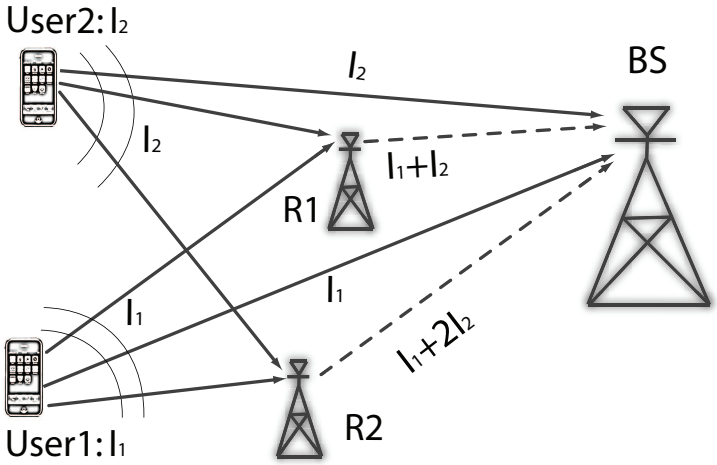


Figure 3.2: Two-user two-relay networks with network coding.

3.1.3 Description

Consider the MUMR network in Figure 3.2. When user nodes (user 1 and user 2) transmit to the BS, both relay nodes also receive the respective messages due to the broadcast property of the wireless medium. The relay nodes will attempt to decode, and if decoding is successful, each relay will forward the decoded messages to the BS following suitable channel and network encoding. Here successful decoding means that information is received error-free. Note that "+" operation means the operations in finite fields in GF(4). Thus, it will not cause any bandwidth expansion or extra power consumption.

To increase asymptotic performance, we propose to use linearly independent (LI) network codes in the two relays, as shown in Figure 3.2 for a two-user, two-relay network example. We showed ([XS09]) that such LI network codes are asymptotically optimal in terms of diversity (diversity order 3). Furthermore, the network codes are deterministic at each relay node. It follows that the BS receives four codeword transmissions with four different message combinations: I_1, I_2, I_1+I_2 and I_1+2I_2, constituting a resulting nonbinary LI network code. If the relay can only successfully decode one source message (outage in one SR channel), it transmits the message with the same channel codeword as the source.

3.1.4 Expected performance or benefits

Significant improvements in outage probability (and FER) are expected in the high SNR region for the two-users, two-relay network, as well as for networks with a higher number of users and relays. We are in fact expecting asymptotically optimal performance in the high SNR region. One example on the FER gain is shown in Figure 3.3. The rate is 0.5 bits/Second/Hz for each user. The channel is Rayleigh fading with unit mean. The channel codes are regular LDPC codes with [200, 400, 3] matrix. All nodes have only one antenna. Note that the channel codes have not been optimized. Otherwise, the outage probabilities and FER shall be much closer. The relaying and BS use CSC to check if the decoding is successful or not. The channels are orthogonal, either in frequency (FDMA) or time (TDMA). We assume BPSK modulation scheme. However, our network coding scheme is transparent for any modulation scheme.

We expect further that the coverage of multi-user networks is improved as compared to three-node relaying networks.

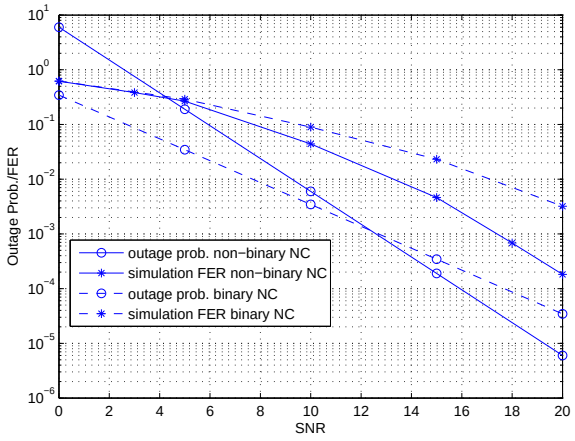


Figure 3.3: Outage probability and frame error rates of two-user two-relay networks with network coding. The binary NC means both relays use binary NC, and the non-binary is LI NC.

3.1.5 Expected requirements on signalling and measurements

The approach is based on simple common signalling with no strict constraints on modulation or channel coding scheme. Thus, we believe our proposed MUMR system can be relatively easy introduced into the current OFDM-based wireless networks.

The protocols are transparent to the users. Thus, user terminals do not need to change anything. If feedback is available, we may gain more in performance. This should be our future topic/deliverables.

3.1.6 Expected requirements on architecture and protocols

So far, orthogonal channel allocation in time or frequency domain has been assumed. However, this assumption is not necessary for our scheme. In the future, we will seek to extend the scheme to include MAC and BC channel scenarios as well.

3.2 Two-way relaying with MIMO-AF-relays

3.2.1 Introduction

Relaying is one of the key candidate technologies to achieve the ubiquitous demand of high data rate traffic which is expected for next generation mobile radio systems. Relays can be used in many different ways, e.g., to enhance the coverage of a radio cell (as in relay enhanced cells), to extract spatial diversity (using cooperative relaying), or to enhance the traffic density in the cell by playing the role of a direct communication partner (as in one-way or two-way relaying).

The two-way relaying scheme is a very promising candidate among the relaying protocols, since it uses the radio resources particularly efficiently [RW05]. In two-way relaying, two communication partners (which can be terminals or access points and will be referred to as “nodes” in the sequel) that need to exchange data are supported by a single relay in a two-step procedure: In the first step both nodes

transmit their data to the relay where their transmissions interfere, in the second step the relay transmits back to both nodes.

It is desirable to maximize the density of the network nodes in order to achieve a satisfactory quality of service of the radio access. Therefore it is of crucial importance to lower the cost of the relay stations as far as possible since network installation and maintenance cost is known as the prime barrier to achieving a large number of network nodes. We demonstrate in this contribution that two-way relaying is feasible with very simple amplify and forward (AF) relays which do not decode the users signals but merely forward the amplified version of the received superposition back to both nodes.

Therefore, the task to separate the desired data stream from the received self-interference must be performed by the nodes themselves. As we demonstrate this is feasible, provided they have sufficient channel state information about the channel between both nodes and the relay.

We therefore propose a channel estimation scheme and a corresponding pilot strategy to provide both nodes with all relevant channel parameters. We study the performance of this two-way relaying scheme under real-world constraints such as imperfect estimates due to noise, interference, and limited number of pilots, time-varying channels, and also asymmetric traffic rate requirements.

Duplexing mode	TDD
Topology / links involved	Relay enhanced / two-way communication of two terminals via relay station
Network deployment	Local area
Target system	LTE-A
History	New
Field of main contribution	Pilot signalling and communication protocol

3.2.2 Description

The scenario is depicted in Figure 3.4. Two nodes that are equipped with M_1 and M_2 antennas, respectively, exchange data with the help of one AF relay having M_R antennas not requiring any interaction from the base station. The two-way relaying scheme consists of two phases: In the first phase both nodes transmit to the relay where their transmissions interfere. Assuming frequency-flat fading, we can express the received signals at the relay as

$$\boldsymbol{r} = \boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \boldsymbol{n}_R \in \mathrm{C}^{M_R}, \tag{3.1}$$

where $\boldsymbol{H}_1 \in \mathrm{C}^{M_R \times M_1}, \boldsymbol{H}_2 \in \mathrm{C}^{M_R \times M_2}$ represent the quasi-static block fading MIMO channel matrices between the nodes and the relay, $\boldsymbol{x}_1 \in \mathrm{C}^{M_1}, \boldsymbol{x}_2 \in \mathrm{C}^{M_2}$ are the transmitted vectors, and $\boldsymbol{n}_R \in \mathrm{C}^{M_R}$ is the thermal noise at the relay. The AF relay amplifies the received signal by multiplying it with a complex amplification matrix $\boldsymbol{G} \in \mathrm{C}^{M_R \times M_R}$ and transmits the amplified signal in the second transmission phase. The second transmission takes place in a subsequent time slot in a TDD fashion. We assume that reciprocity is valid. We can therefore express the received signal in the second time slot in the following manner

$$\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \left(\boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \boldsymbol{n}_R \right) + \boldsymbol{n}_1 \in \mathrm{C}^{M_1} \\ \boldsymbol{y}_2 &= \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \left(\boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \boldsymbol{n}_R \right) + \boldsymbol{n}_2 \in \mathrm{C}^{M_2}. \end{aligned} \tag{3.2}$$

Expanding these equations we find the following alternative representation

$$\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \tilde{\boldsymbol{n}}_1 \\ \boldsymbol{y}_2 &= \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_2 \cdot \boldsymbol{x}_2 + \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_1 \cdot \boldsymbol{x}_1 + \tilde{\boldsymbol{n}}_2. \end{aligned} \tag{3.3}$$

We observe that the received signals for both nodes comprise three terms: The first term represents the self-interference the node receives from its own transmissions. The second term is the desired information from the other node. The third term represents the effective noise contribution which consists of the forwarded relay noise and the node's own thermal noise contribution. We can conclude that the bidirectional two-way relaying transmission is feasible if both nodes have sufficient knowledge of the channel matrices $\boldsymbol{H}_1, \boldsymbol{H}_2$ since the self-interference term can then be cancelled and the data transmissions can be decoded.

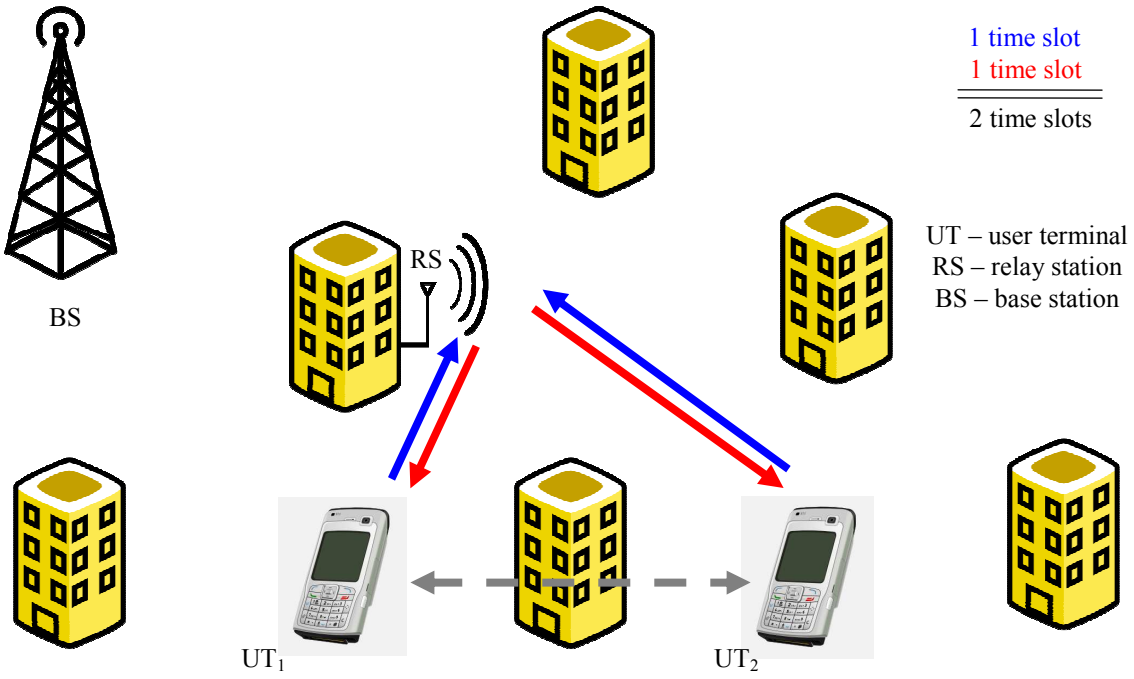


Figure 3.4: Two-Way Relaying Scenario.

While acquiring its own channel is fairly straight forward for each terminal in a system where reciprocity holds, the main difficulty is that each terminal also needs knowledge of the channel between the other terminal and the relay. Thus the key issue is to devise a feedback-free channel estimation scheme that provides the relevant channel parameters to both nodes. The salient features of this algorithm are presented below and the reader is referred to Appendix A.5.1 for details.

- During the training phase, a particular training signalling and the complex relay amplification matrices are devised.
- Crucial to the derivation of the estimator is the expression of equation (3.3) in tensor form.
- Application of the TENCE algorithm to estimate all relevant channel parameters [RH09a] .
- Use of Structured Least Squares based iterative refinement of the estimate obtained via TENCE [RH09b].

Data transmission

Once the channels are estimated in the training phase, the nodes can transmit their data streams. To this end introduce the effective channel matrices

$$\begin{aligned} \boldsymbol{H}^{(e)}{}_{1,1} &= \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_1 \\ \boldsymbol{H}^{(e)}{}_{1,2} &= \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_2 \\ \boldsymbol{H}^{(e)}{}_{2,1} &= \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_1 \\ \boldsymbol{H}^{(e)}{}_{2,2} &= \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_2 . \end{aligned} \tag{3.4}$$

If the channel estimates are perfect, the receivers can cancel their self-interference in the following way, which is often referred to as Analogue Network Coding (ANC)

$$\begin{aligned} \boldsymbol{z}_1 &= \boldsymbol{y}_1 - \boldsymbol{H}^{(e)}{}_{1,1} \cdot \boldsymbol{x}_1 = \boldsymbol{H}^{(e)}{}_{1,2} \cdot \boldsymbol{x}_2 + \tilde{\boldsymbol{n}}_1 \\ \boldsymbol{z}_2 &= \boldsymbol{y}_2 - \boldsymbol{H}^{(e)}{}_{2,2} \cdot \boldsymbol{x}_2 = \boldsymbol{H}^{(e)}{}_{2,1} \cdot \boldsymbol{x}_1 + \tilde{\boldsymbol{n}}_2 . \end{aligned} \tag{3.5}$$

We observe that if we ignore the channel estimation errors the two-way relaying channel is decoupled into two single-user MIMO channels on which any single-user transmission technique can be used. For example, dominant eigenmode transmission (DET) can be incorporated by choosing proper precoding and decoding vectors from the SVD of the effective channels. If we let s_1 and s_2 be the data symbols that the nodes would like to transmit, we can construct the transmitted vectors in the following way

$$\boldsymbol{x}_1 = \boldsymbol{w}_{\mathrm{Tx},1} \cdot s_1, \quad \boldsymbol{x}_2 = \boldsymbol{w}_{\mathrm{Tx},2} \cdot s_2$$
$$s_2^{\mathrm{est}} = \boldsymbol{w}_{\mathrm{Rx},1}{}^{\mathrm{H}} \cdot \boldsymbol{z}_1, \quad s_1^{\mathrm{est}} = \boldsymbol{w}_{\mathrm{Rx},2}{}^{\mathrm{H}} \cdot \boldsymbol{z}_2,$$

where

$$\boldsymbol{H}^{(e)}{}_{1,2} = \boldsymbol{U}_{1,2} \cdot \boldsymbol{\Sigma}_{1,2} \cdot \boldsymbol{V}_{1,2}{}^{\mathrm{H}}, \quad \boldsymbol{H}^{(e)}{}_{2,1} = \boldsymbol{U}_{2,1} \cdot \boldsymbol{\Sigma}_{2,1} \cdot \boldsymbol{V}_{2,1}{}^{\mathrm{H}}, \tag{3.6}$$
$$\boldsymbol{w}_{\mathrm{Tx},1} = \left[\boldsymbol{V}_{2,1}\right]_{:,1}, \quad \boldsymbol{w}_{\mathrm{Tx},2} = \left[\boldsymbol{V}_{1,2}\right]_{:,1},$$
$$\boldsymbol{w}_{\mathrm{Rx},1} = \left[\boldsymbol{U}_{1,2}\right]_{:,1}, \quad \boldsymbol{w}_{\mathrm{Rx},2} = \left[\boldsymbol{U}_{2,1}\right]_{:,1}.$$

Note that if the relay amplification matrix $\boldsymbol{G}$ is chosen symmetric, i.e., $\boldsymbol{G} = \boldsymbol{G}^{\mathrm{T}}$ then the effective channels are also symmetric, i.e., $\boldsymbol{H}^{(e)}{}_{1,2} = \boldsymbol{H}^{(e)}{}_{2,1}{}^{\mathrm{T}}$. Therefore, only one SVD has to be computed since $\boldsymbol{w}_{\mathrm{Rx},1} = \boldsymbol{w}_{\mathrm{Tx},1}{}^{*}$ and $\boldsymbol{w}_{\mathrm{Rx},2} = \boldsymbol{w}_{\mathrm{Tx},2}{}^{*}$.

The choice of the relay amplification matrix $\boldsymbol{G}$ depends on the availability of CSI at the relay as well as the requirements of the current transmission. If no CSI is available, one can for instance choose a properly scaled DFT matrix. This simple choice provides the full spatial multiplexing gain at high SNRs.

If channel knowledge is present, one can improve the system performance further by choosing $\boldsymbol{G}$ such that a suitable optimization criterion is maximized, e.g., the sum rate or the signal to noise ratio. For the maximization of the sum rate, this leads to a rather complicated nonlinear non-convex optimization problem. A simple algebraic solution is obtained if we maximize the squared Frobenius norms of the effective channel matrices $\boldsymbol{H}^{(e)}{}_{1,2}$ and $\boldsymbol{H}^{(e)}{}_{2,1}$, which gives rise to the Algebraic Norm-Maximizing (ANOMAX) Transmit Strategy [RH09c] (see Appendix A.5.2 for details). Via numerical simulations we found that ANOMAX is almost optimal in terms of the SNR.

Other proposals for the choice of $\boldsymbol{G}$ include the Dual Channel Matching scheme [VH08], which was proposed to simplify the optimization for the rate-optimal relay strategy or using ZF/MMSE transceivers to mitigate the interference at the relay instead of subtracting them at the nodes [UK09].

3.2.3 Expected performance or benefits

This two-way relaying scheme will improve the spectral efficiency of the system. A bidirectional transmission is achieved in only two transmission phases, which can be time slots in a TDD fashion or frequency slots in an interleaved FDD fashion. This also promises a particularly low latency which may be interesting for certain real-time services. Also the infrastructure requirements are low since no base station interaction is required and an amplify and forward relay station is sufficient.

3.2.4 Issues to be investigated

There are a number of open issues that still need to be investigated. Most importantly, channels have been assumed to be constant up to now. Since in any wireless radio access the channels are time-varying, a channel tracking scheme should be devised which can track the changes in the evolution of the propagation conditions. It is expected that this can be achieved on-line using the received data as additional training information to perform constant updates to the current channel knowledge.

Another aspect of time-varying channels is that the power levels at the relay station may also change which can force the relay to alter its amplification of the signals. This information must be signaled to the nodes since they require precise knowledge of any interaction the relay station performs.

It would also be desirable to extend this scheme for the case where the nodes have asymmetric QoS requirements, such as different data rates or different target bit error rates. These requirements can be supported by a clever design of the relay amplification matrix used during the data transmission phase.

Finally, from a system perspective it would be interesting to extend this scheme to a joint processing of several users (multi-user two-way relaying) or the interaction of several relays (cooperative relaying).

3.2.5 Expected requirements on signalling and measurements

Our proposed Two-Way Relaying transmission system requires the incorporation of the support for relaying-specific signalling and measurements into the system. First of all, to find suitable communication partners, a link quality indicator for the links between users and relay stations in their vicinity should be obtained (the simplest one could be their geometrical distance). This information can be used to assign a

relay station to each pair of communication partners. Next, we require a training phase for each of these pairs, in which the channels between nodes and the relays are estimated. As described above, this requires a total of $\left(M_1 + M_2\right) \cdot M_{\mathrm{R}}$ pilot slots. Note that no feedback of CSI is required. After this training phase, the data transmission phase can immediately be initiated.

3.2.6 Expected requirements on architecture and protocols

This transmission scheme is applicable only with a number of additional protocols. In the beginning, the presence of user terminals and relay stations must be detected by all communication partners. If a larger number of communicating network nodes are present, a resource allocation mechanism may be required to find communication partners for which this form of bidirectional transmission is feasible. In the next step, the training phase must take place, in which the nodes repeat their pilot sequences in subsequent frames and the relay station changes its amplification matrix for each frame. A protocol is required to initiate and control the transmission of these pilots. After finishing the training phase, a protocol mechanism should initiate the data transmission phase, where the nodes can transmit following the single-user precoding scheme described above. In this phase, the relay station only amplifies the received signal using a suitable relay amplification matrix.

To apply two-way relaying in a practical system where a bandwidth relevant to WINNER+ is present, it can directly be combined with an OFDM-based transmitter and receiver. As in chunk-wise adaptive precoding, the frequency axis is divided into chunks where the channel is approximately constant and the pre- and postprocessing is performed jointly for all subcarriers within the chunk. Consequently, our requirements on synchronization are the same as for every OFDM-based transmission: The nodes and the relay need to stay synchronized within the cyclic prefix.

4. Innovative concepts in coding and decoding

This section presents innovative concepts related to coding and decoding. Section 4.1 proposes the use of network coding methods for point-to-point MIMO communication, with the aim to provide diversity-multiplexing trade-off. In Section 4.2, an improved generic iterative joint channel estimation and decoding architecture employing belief propagation is presented.

4.1 Space-time network coding

4.1.1 Introduction

One way to introduce diversity in the received signal is to exploit the spatial diversity offered when multiple antennas are used at the transmitter with the possibility of using one or more antennas at the receiver. The use of multiple antennas offers significant diversity and multiplexing gains relative to single antenna systems. Multiple-Input Multiple-Output (MIMO) wireless systems can thus improve the link reliability and the spectral efficiency relative to Single-Input Single-Output (SISO) systems. MIMO schemes as such Space-Time Block Codes (STBC) are used at the transmitter in order to introduce diversity, but the data rate is reduced when the number of the transmit antennas is greater than two. Even for two antennas the space-time encoder rate is at most equal to one (in case of Alamouti).

A general trend in current space-time code design is that different symbols are transmitted independently of each other. However, by performing some linear combining (e.g. similar to the network coding operation) on the packets present at the input of the space-time encoder, the number of packets to be transmitted is effectively decreased, thereby increasing the rate of the encoder. Further, network coding is generally performed at intermediate nodes that combine data from two (or more) different sources. However, in the case of a wireless system where a transmitter has multiple antennas, the NC operation can be imitated and consequently resulting in a further exploitation of the spatial and temporal dimensions.

Duplexing mode	FDD or TDD
Topology / links involved	Applies to any point-to-point communication
Network deployment	Wide and local area
Target system	LTE-A
History	New
Field of main contribution	Space-time coding methods

4.1.2 Description

The proposed method consists of imitating network coding at the space-time encoder of a multi-antenna transmitting node in order to combine the symbols of various data streams and consequently increase the space-time encoder data rate. In the following we will present several codes that allow increasing the space-time code (STC) rate. An illustration of the transmitter structure with network coding at a source node is shown in Figure 4.1. Let s_1, s_2, s_3 and s_4 be the modulated data symbols. The modulated data streams are subject to space-time encoding where few or all of the encoded symbols are simply obtained by combining (e.g. by using the XOR operation on) two or more of the modulated symbols. In the following we will assume that the combining operation was done at the base band prior to the modulation operation. The coded symbols are then mapped to the physical transmitting antennas. The antenna mapping is dictated by the STC encoder.

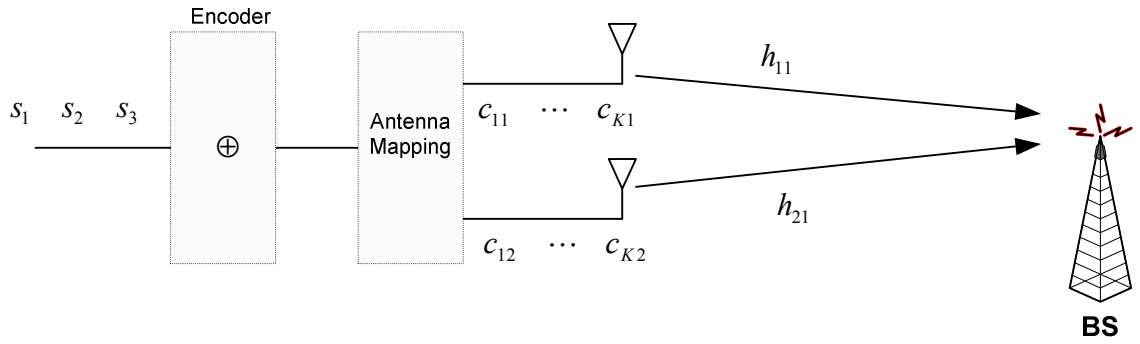


Figure 4.1: Network coding at a source node.

We may divide the code design into two categories. In the first category, network coding at the transmitter source is used to obtain multiplexing gain and low diversity. In the second category, a trade-off between diversity and multiplexing gain can be obtained. We assume in the following the case of a 2x2 MIMO system, where the channel matrix is given by:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \tag{4.1}$$

where h_{ij} refers to the channel between transmitting antenna i and receiving antenna j The generalization to an $m \times n$ MIMO system is analogous.

At the output of the encoder we obtain at the base band the following coded matrix:

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ \vdots & \vdots \\ c_{K1} & c_{K2} \end{bmatrix} \tag{4.2}$$

where K is the codeword length of the space-time encoder and the columns of $\mathbf{C}$ represent the transmitting antenna. The choice of c_{ij} depends on the desired spatial multiplexing and/or diversity gains. In particular the code matrix can be adapted to the radio environment and system service (e.g. user data rate, delay constraints, etc).

4.1.2.1 Spatial multiplexing & low diversity

In a dominant spatial multiplexing scenario the encoder matrix can be given as:

$$\mathbf{C} = \begin{bmatrix} s_1 & s_2 \\ s_2 \oplus s_4 & s_1 \oplus s_3 \end{bmatrix}. \tag{4.3}$$

In the first time slot, the transmission is identical to the case of horizontal spatial multiplexing. In the second transmission slot the signal transmitted from one antenna is simply the combination of the previous encoded symbol of the other antenna with a new symbol.

4.1.2.2 High diversity & spatial multiplexing

A 2 x 2 Alamouti scheme can provide a diversity order up to 4, but with a rate equal to 1. However, not all of this diversity gain might be useful (i.e., if the SINR from one link is already very high, not much can be gained by introducing more diversity) and it would be more beneficial to trade it off for some multiplexing gain. The following scheme allows, by using network coding at the transmitter, to increase the STC rate. The transmitted data is given by the following matrix:

$$\mathbf{C} = \begin{bmatrix} s_1 & s_1 \oplus s_3 \\ s_2 \oplus s_3 & s_2 \end{bmatrix} = \begin{bmatrix} s_1 & x_2 \\ x_1 & s_2 \end{bmatrix} \tag{4.4}$$

where $x_1 = s_2 \oplus s_3$, and $x_2 = s_1 \oplus s_3$. The streams are mapped to the antennas such that each of them (i.e. s_1, s_2 and s_3) is transmitted on the two different antennas, thus maximizing the transmit diversity. This scheme provides a transmission rate of 3/2.

4.1.3 Expected performance or benefits

The proposed method will obtain adaptively a tradeoff between diversity and multiplexing gains without requiring any channel state information at the transmitter (CSIT). As an example we expect that the proposed scheme outperforms the 2 × 2 Alamouti scheme through system level simulations.

4.1.3.1 SINR

In the following we will derive the signal to interference and noise ratio (SINR) equations based on successive interference cancellation and maximum ratio combining at the receiver. We assume the scenario of high diversity and spatial multiplexing as defined in Section 4.1.2.2. The transmission protocol consists of two transmission slots, T_1 and T_2, during which the channel remains constant. In the following SINR derivation, perfect interference cancellation was assumed. It shall be noted that the methodology used for the derivation of the pre-decoding SINR is similar to the one used in [WIN+D13].

Pre-decoding SINR

The pre-decoding SINRs can be easily shown to be given by:

$\Gamma_{s_1} = \Gamma_{x_1}$ and $\Gamma_{x_2} = \Gamma_{s_2}$

Depending on the received signal strength, we can distinguish between two cases.

Case 1

In the first case, the received power from the first transmit antenna is stronger than the received power from the second transmit antenna. In that case s_1 (resp. x_1) is detected first during T_1 (resp. T_2), then followed by x_2 (resp. s_2).

Case 2

In the second case, the received power from the second transmit antenna is stronger than the received power from the first transmit antenna.

Post-decoding SINR

Following the SINR evaluation of the transmitted coded symbols (i.e. pre-decoding SINR), the modulated symbols will be estimated by the space-time decoder. Once the symbols s_1, s_2, x_1 and x_2 have been detected, the decoding will be done depending on the pre-decoding SINR values of those symbols, resulting in the post-decoding SINRs that would directly determine the resulting capacity. We distinguish between two main decoding scenarios that offer different diversity-multiplexing trade-offs.

Scenario 1

In the first scenario, both s_1 and s_2 are decoded based on their direct transmissions such that we use x_1 and x_2 to obtain s_3. This is simply achieved by first decoding $s_1 \oplus s_3$ with s_1 (i.e. $s_1 \oplus (s_1 \oplus s_3)$) and obtaining the first estimate of s_3. The second estimate of s_3 is obtained by decoding $s_2 \oplus s_3$ with s_2. The two estimates of s_3 are then combined together. The equivalent (i.e. post-decoding) SINRs will then be given as follows:

$\Gamma'_{s_1} = \Gamma_{s_1}$; $\Gamma'_{s_2} = \Gamma_{s_2}$; $\Gamma'_{s_3} = \Gamma_{x_1} + \Gamma_{x_2}$.

Scenario 2

In the second scenario, one of s_1 and s_2 (the one with the higher pre-decoding SINR) is decoded based on its direct transmission such that we use the relevant x (i.e. x_1 or x_2) to obtain it, and use the other x to obtain s_3. The two possible cases are:

Case a: $\Gamma_{s_1} > \Gamma_{S_2}$

In this case, we use x_1 to increase the diversity gain of s_2, and x_2 to obtain multiplexing gain by decoding s_3. The equivalent SINRs will then be given by:

$$\Gamma'_{s_1} = \Gamma_{s_1} ; \Gamma'_{s_3} = \Gamma_{x_2} ; \Gamma'_{s_2} = \Gamma_{s_2} + \Gamma_{x_1} .$$

Case b: $\Gamma_{s_2} > \Gamma_{S_1}$

In this case, we use x_2 to increase the diversity gain of s_1, and x_1 to obtain multiplexing gain by decoding s_3. The equivalent SINRs will then be given by:

$$\Gamma'_{s_2} = \Gamma_{s_2} ; \Gamma'_{s_3} = \Gamma_{x_1} ; \Gamma'_{s_1} = \Gamma_{s_1} + \Gamma_{x_2} .$$

The sum-capacity for all different scenarios is given by:

$$C_{sum} = \log_2(1+\Gamma'_{s_1}) + \log_2(1+\Gamma'_{s_2}) + \log_2(1+\Gamma'_{s_3}) \quad (4.5)$$

where the post-decoding SINRs depend on the chosen decoding scenario. A main merit of the proposed scheme is that the diversity-multiplexing gains for the different transmitted streams can be adaptively controlled at the receiver based on desired performance measures by simply choosing the desired decoding scenario, consequently not requiring any CSIT. In fact the receiver may compare between various scenarios (scenario 1, 2a and 2b) then decide the suitable one in terms of diversity-spatial multiplexing trade-off (assuming more than one scenario can be adopted).

4.1.3.2 System-level performance

The proposed scheme in (4.4) is evaluated in a system level simulator [WIN+D41] and compared to the 2x2 Alamouti scheme with MRC combining at the receiver and the 2x2 channel capacity based on single value decomposition (SVD). Ideal symbol estimation and perfect interference cancellation were assumed. The cumulative distribution function (CDF) of the SINR performance is shown in Figure 4.2 . Although the SVD method achieves the highest sum-capacity, it has two major drawbacks: it requires full CSI at both the transmitter and the receiver, and half of the transmitted streams will have a very low performance (which might not yield any gains in a practical setup). On the other hand, the proposed method is able to exchange one of the transmitted streams into a better diversity performance so that the diversity-multiplexing behavior can be controlled by the receiver as opposed to the Alamouti scheme that provides a better diversity performance in general at the expense of a lower rate. It shall be noted that the SINR of both streams of the proposed method exhibits almost identical CDF as shown in Figure 4.2

Furthermore the CDF of the sum-capacity is shown in Figure 4.3. The average normalized sum-capacity of the Alamouti scheme is 2.4903 [b/s/Hz], whereas that of the SVD method is 4.4233 [b/s/Hz], and the proposed scheme is 3.5204 [b/s/Hz].

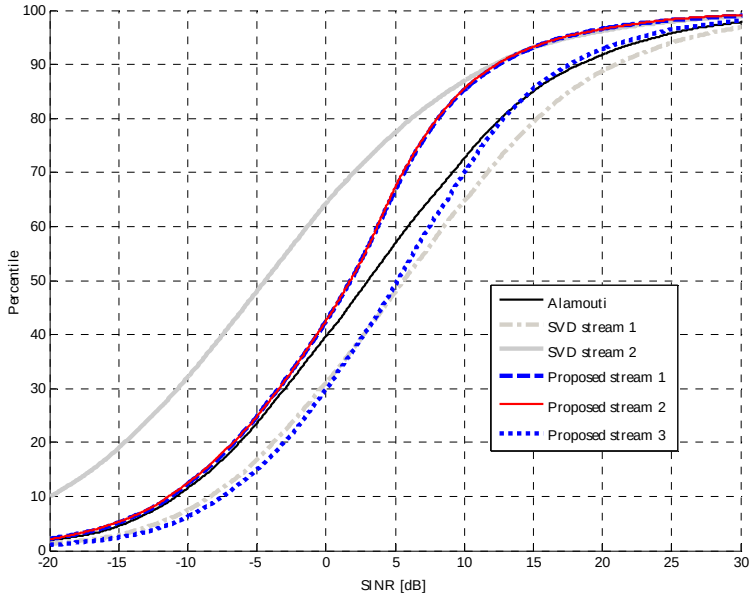


Figure 4.2: CDF of the SINR for the evaluated schemes.

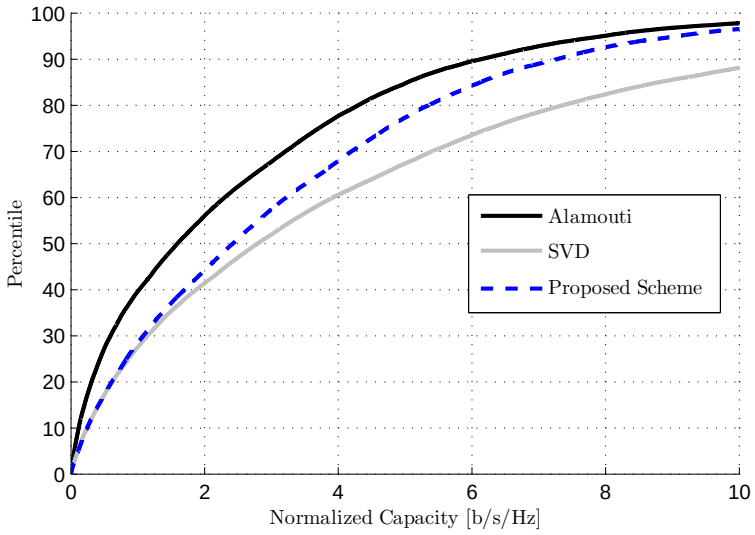


Figure 4.3: CDF of the normalized capacity of the evaluated schemes.

4.1.4 Expected requirements on signalling and measurements

The receiver needs to know that scheme used at the transmitter. In addition, depending of the transmitted code matrix the CSIT may be required at the transmitter.

4.1.5 Expected requirements on architecture and protocols

No requirements on architecture and protocols are expected.

4.2 Joint channel estimation and decoding using Gaussian approximation in a factor graph

4.2.1 Introduction

Propagating messages in a suitable factor graph [Loe04] is a systematic tool for deriving iterative algorithms. A message propagated over an edge connecting two nodes of a factor graph is a probability distribution for a discrete variable or a probability density function (pdf) for a continuous variable. Among various receiver issues solved using the belief propagation algorithm (BP), also called sum-product algorithm [KFL01], we can cite decoding, channel estimation, synchronization and detection [WSBM06]. [WS01] presents a BP handling continuous variables, in which canonical distributions are used for quantizing probability distributions, in order to propagate discrete probability distributions. However, the degree of quantization has a strong impact on estimation accuracy and performance. Even adapting the quantization in each iteration of BP, as proposed in [MV07] and [DKL07], does not fill the complexity gap between BP and other algorithms. Instead of relying on quantization, we introduce here a new proposal which is to model probability distributions as mixtures of Gaussian distributions [YBB09]. It allows for estimation improvement and complexity reduction simultaneously. We focus on BP with Gaussian approximation over a multipath channel.

A frequency-selective multi-path channel creates inter-symbol interference (ISI). Thus, received symbols must be equalized at the receiver, e.g., by an iterative a posteriori probability (APP) equalizer or a minimum mean square error (MMSE) equalizer. The equalizer has to work together with a channel estimator. A factor graph with BP can help defining an iterative receiver in a systematic way and implementing joint channel estimation and decoding. However, the quantization method will make BP unfeasible over a multipath channel, due to a prohibitive complexity, and the proposed Gaussian approximation becomes very attractive.

The section is structured as follows. Section 4.2.2 explains how the transmission system is modeled using a factor graph and how BP is applied. Section 4.2.3 presents the approximation of the distribution of channel estimates over a multipath channel in BP by a mixture of Gaussian distributions. In Section 4.2.4, APPs are computed from the approximated distribution. Continuous upward messages in the factor graph are presented in Section 4.2.5. Expected benefits and requirements are listed in Sections 4.2.6 to 4.2.8. The section ends with conclusions in Section 4.2.9.

In the sequel of Section 4.2, messages that are not based on quantized densities will be referred to as continuous messages.

Duplexing mode	FDD or TDD
Topology / links involved	Any point-to-point communication
Network deployment	Any
Target system	Any
History	New
Field of main contribution	Signal processing in the receiver

4.2.2 System model and factor graph

We consider a coded system with transmission over a multi-path channel as shown in Figure 4.4.

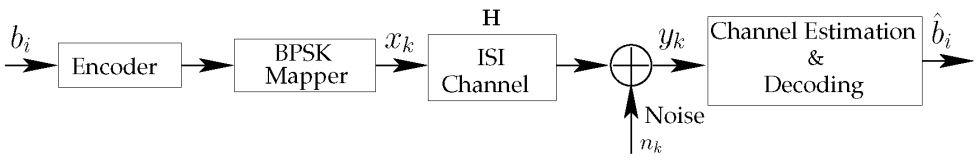


Figure 4.4: System model.

An information binary sequence b_i is encoded, modulated into N BPSK symbols x_k with unit energy and multiplexed with L_p pilot symbols with energy $\mathcal{E}_p$. After convolution with an impulse response made of L taps, i.e., L complex Gaussian coefficients h_l with zero mean, and addition of a complex Gaussian noise $n_k \sim \mathcal{CN}(0,2\sigma_n^2)$, the channel outputs y_k are processed by a receiver performing joint channel estimation

and decoding. Finally, the receiver outputs the estimated information sequence $\hat{b}_i$. The system model is described by

$$y_k = \sum_{l=0}^{L-1} h_l x_{k-l} + n_k \quad , \quad 0 \le k \le N-1. \tag{4.6}$$

We re-write (4.6) in a matrix form:

$$y_k = \mathbf{X}_k^{\mathrm{T}} \mathbf{H} + n_k \tag{4.7}$$

where $\mathbf{X}_k = (x_k, x_{k-1}, \ldots, x_{k-L+1})^{\mathrm{T}}$ represents the symbol vector at time instant k and $\mathbf{H} = (h_0, \ldots, h_{L-1})^{\mathrm{T}}$ represents the ISI channel.

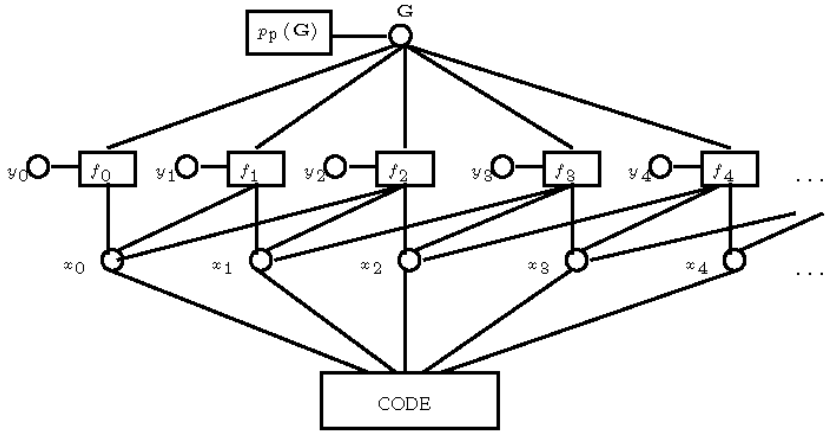


Figure 4.5: Factor graph for multipath channel.

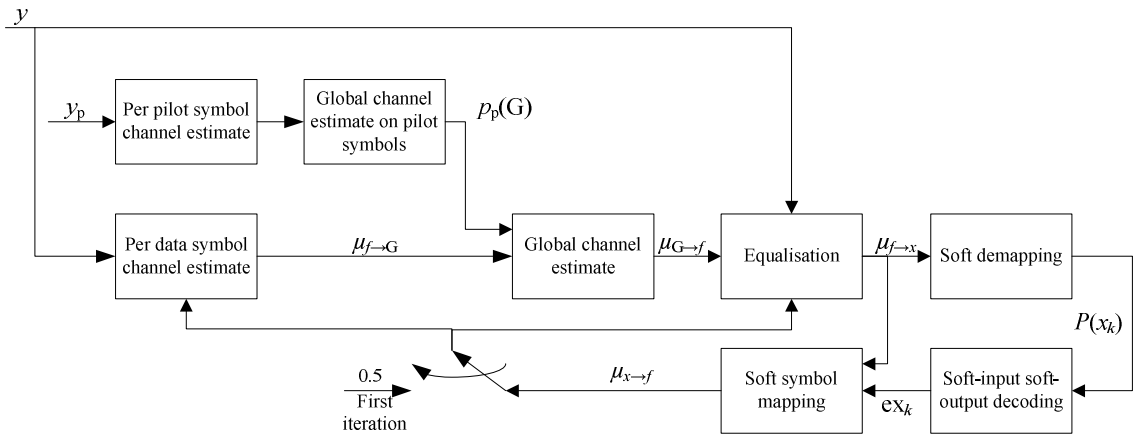


Figure 4.6: Iterative receiver.

The corresponding factor graph for 3 taps is depicted in Figure 4.5, following [WS01], where **G** is a quantized estimate of **H** and $p_{\mathrm{p}}(\mathbf{G})$ represents the quantized distribution of the known a priori of **G**. We use the estimate from pilots as the a priori of **G.** The corresponding block diagram with receiver functions is depicted in Figure 4.6. Based on channel estimation from pilots and data (computation of $\mu_{G \to f_k}$), the received signal is equalized and demapped (computation of $P(x_k)$). Then, decoding is performed (computation of ex_k) and ex_k is used in the iterative process to create a modulation symbol a priori probability (computation of $\mu_{x_k \to fk}$), which will be used to improve the channel estimation. According to the factor graph representation in Figure 4.5, we call *downward message* a message from node G (estimation) to node CODE (decoding) and *upward message* a message from node CODE to node G. The estimate from pilots is:

$$p_{\mathrm{p}}(\mathbf{G}) = \prod_{k=0}^{L_{\mathrm{p}}-1} \mu_{f_{\mathrm{p},k} \to \mathbf{G}} \tag{4.8}$$

as shown in Figure 4.7(a) (for simplicity the nodes $f_{\mathrm{p},k}$ are not shown in Figure 4.5). For upward messages, in node CODE, a forward-backward algorithm computes the extrinsic information for each

coded bit. Taking deinterleaving into account, the extrinsic information ex_k is propagated to nodes x_k. In node x_k, the message $\mu_{x_k \to f_{k+l}} = \xi_{k,l}$ is obtained by multiplying all messages into x_k:

$$\mu_{x_k \to f_{k+l}} = \xi_{k,l} = \mathrm{ex}_k \prod_{\substack{i=0 \\ i \neq l}}^{L-1} \mu_{f_{k+i} \to x_k} , \tag{4.9}$$

as shown in Figure 4.7(b). From each node f_k to node **G**, a discrete distribution $\mu_{f_k \to \mathbf{G}}$ of the quantized estimate of **H** is computed and propagated based on a marginalization of the likelihood $p(y_k|\mathbf{X}_k,\mathbf{G})$ with respect to the transmitted symbol $\mathbf{X}_k$, as shown in Figure 4.7 (c).

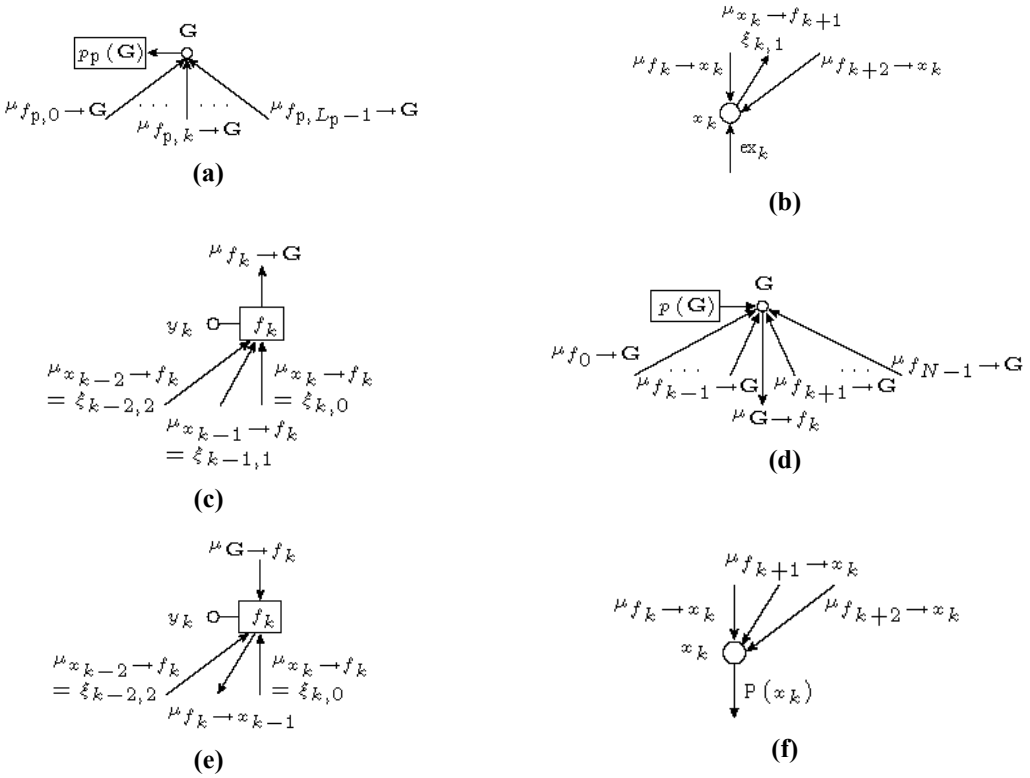


Figure 4.7: Message propagation in factor graph.

For downward messages, the message $\mu_{G \to f_k}$ is calculated as shown in Figure 4.7(d):

$$\mu_{\mathbf{G} \to f_k} = p_{\mathrm{p}}(\mathbf{G}) \prod_{\substack{i=0 \\ i \neq k}}^{N-1} \mu_{f_i \to \mathbf{G}} . \tag{4.10}$$

By multiplying message $\mu_{\mathbf{G} \to f_k}$ and all messages from $\mathbf{X}_{k,l} = (x_k, x_{k-1}, \ldots, x_{k-l+1}, x_{k-l-1}, \ldots, x_{k-L+1})$ into f_k (Figure 4.7(e)), the APP of each transmitted symbol x_{k-l} is computed, marginalizing the likelihood $p(y_k|\mathbf{X}_k,\mathbf{G})$ with respect to **G** and $\mathbf{X}_{k,l}$. The final APP of each coded bit $P(x_k)$ is obtained by multiplying all messages from node f_k to x_k (Figure 4.7(f)) and then propagated to node CODE. The whole process of propagating upward and downward messages is then iterated.

4.2.3 Distribution of the channel estimate

In the iterative receiver, the initial estimate is obtained from known pilots and subsequent estimates from data symbols. Thus, the distribution of the channel estimate will differ, depending on the iteration.

In Appendix A.6.1, we show that the distribution of the channel estimate when estimation is based on pilots can be approximated as one Gaussian distribution $\mathcal{CN}(\mathbf{H}, 2\sigma_n^2/L_{\mathrm{p}}/\mathcal{E}_{\mathrm{p}})$.

In Appendix A.6.2, we show that for each channel tap, the pdf $p_{\mathrm{d},k}(g_l)$ when estimation is based on data can be approximated as a mixture of two Gaussian distributions. The pdf $p_{\mathrm{d},k}(\mathbf{G})$ can be approximated as a mixture of multiple Gaussian distributions which are the product of all pdfs of each tap with variance $2\sigma_n^2/(N\text{-}1)$.

4.2.4 APP evaluation from downward messages

With the conclusions in Section 4.2.3, the known a priori discrete channel distribution $p_{\mathrm{p}}(\mathbf{G})$ can be approximated as one Gaussian distribution and the discrete distribution of the product

$$\prod_{i=0;\, i\neq k}^{N-1} \mu_{f_i \to \mathbf{G}} = p_{\mathrm{d},k}(\mathbf{G}) \tag{4.11}$$

can be approximated as a mixture of multiple Gaussian distributions, where $p_{\mathrm{p}}(\mathbf{G})$ is based on pilots and $p_{\mathrm{d},k}(\mathbf{G})$ is based on the messages in the current iteration. Furthermore, we can show that for each tap, there is always one dominant Gaussian distribution (with mean value h_l). Hence, when calculating APP, we consider only the dominant one ($\beta_l = 1$). Then, the discrete distributions of $p_{\mathrm{p}}(\mathbf{G})$ and $p_{\mathrm{d},k}(\mathbf{G})$ can both be reduced to L pairs of (mean, variance) parameters: $\left(\hat{h}_{\mathrm{p},l}, \hat{\sigma}^2_{h\mathrm{p}}\right)$ for $p_{\mathrm{p}}(\mathbf{G})$ and $\left(\hat{h}_{\mathrm{d},k,l}, \hat{\sigma}^2_{h\mathrm{d}}\right)$ for $p_{\mathrm{d},k}(\mathbf{G})$. Thus, $p_{\mathrm{p}}(\mathbf{G})$ times $p_{\mathrm{d},k}(\mathbf{G})$ can also be approximated by a mixture of Gaussian distributions, i.e., the discrete distribution of message $\mu_{\mathbf{G}\to f_k}$, denoted as $p_k(\mathbf{G})$, can be reduced to L pairs of parameters $\left(\hat{h}_{k,l}, \hat{\sigma}^2_{h}\right)$. $\left(\hat{h}_{k,l}, \hat{\sigma}^2_{h}\right)$ can be calculated from $\left(\hat{h}_{\mathrm{p},l}, \hat{\sigma}^2_{h\mathrm{p}}\right)$ and $\left(\hat{h}_{\mathrm{d},k,l}, \hat{\sigma}^2_{h\mathrm{d}}\right)$ as it will be shown in the following part. Thus, we can calculate each downward message $\mu_{f_k \to x_{k-l}}$ in a continuous way, instead of computing it for each codebook value $\mathbf{G}_c$, and then marginalizing with respect to $\mathbf{G}$. It reduces the computation complexity. Indeed, thanks to the computation in Appendix A.6.3 resulting in (A.29), a single APP computation instead of L^q computations (L^q is the quantization codebooke size) is performed for each symbol vector $\mathbf{X}_k$ with the Gaussian approximation. Thus, the global complexity is strongly reduced by the Gaussian approximation in the downward messages.

4.2.5 Estimation from upward messages

In order to improve the performance of the Gaussian approximation, we propose to increase the accuracy of $\hat{\mathbf{H}}_k$ using a continuous upward message. Derivation details are given in Appendix A.6.4.

With (A.36), we obtain $\hat{\mathbf{H}}_{\mathrm{d},k} = \left(\hat{h}_{\mathrm{d},k,0}, \cdots, \hat{h}_{\mathrm{d},k,L-1}\right)^{\mathrm{T}}$ for $p_{\mathrm{d},k}(\mathbf{G})$ by using the messages in current iteration. With (A.38), we obtain $\hat{\mathbf{H}}_{\mathrm{p}} = \left(\hat{h}_{\mathrm{p},0}, \cdots, \hat{h}_{\mathrm{p},L-1}\right)^{\mathrm{T}}$. Together with $\hat{\sigma}^2_{h\mathrm{d}}$ and $\hat{\sigma}^2_{h\mathrm{p}}$, we get

$$\hat{h}_{k,l} = \frac{\hat{\sigma}^2_{h\mathrm{p}} \hat{h}_{\mathrm{d},k,l} + \hat{\sigma}^2_{h\mathrm{d}} \hat{h}_{\mathrm{p},l}}{\hat{\sigma}^2_{h\mathrm{p}} + \hat{\sigma}^2_{h\mathrm{d}}} \tag{4.12}$$

and

$$\hat{\sigma}^2_{h} = \frac{\hat{\sigma}^2_{h\mathrm{p}} \hat{\sigma}^2_{h\mathrm{d}}}{\hat{\sigma}^2_{h\mathrm{p}} + \hat{\sigma}^2_{h\mathrm{d}}} \tag{4.13}$$

where the value of $\hat{\sigma}^2_{h\mathrm{p}}$ and $\hat{\sigma}^2_{h\mathrm{d}}$ are obtained from (A.21) and (A.23): $\sigma_n^2/(L_p \mathcal{E}_p)$ for the pilot case and $\sigma_n^2/(N\text{-}1)$ for the data case.

4.2.6 Expected performance or benefits

The performance is expected to be improved and complexity reduced compared to the BP using quantization.

4.2.7 Expected requirements on signalling and measurements

There are no requirements on signalling and measurements.

4.2.8 Expected requirements on architecture and protocols

There are no requirements on architecture and protocols.

4.2.9 Conclusion and perspectives

Thanks to an approximation of the distribution of the channel estimate as a mixture of Gaussian distributions, we improved the performance of BP and reduced its complexity by propagating continuous messages in the factor graph for multipath channel. As shown by simulation results in Appendix A.6.5, the proposed BP with continuous downward and upward messages (BP-DUGA) almost achieves the APP equalizer performance and outperforms the MMSE equalizer. Even though in this work BPSK

modulation was assumed, the extension of the Gaussian approximation principle to a higher level modulation scheme is straightforward.

Further studies are in progress in order to apply the BP algorithm to coded OFDM transmission and finally extend it to multiple antennas.

5. MIMO schemes in WiMAX systems

5.1 Introduction

MIMO techniques have been incorporated in all of the recently developed wireless communications standards including IEEE 802.11n, IEEE 802.16e-2005, and Long-Term Evolution (LTE). In this part of the report, we will focus on the schemes existing in current IEEE 802.16e specifications and the solutions proposed for IEEE 802.16m.

5.2 Existing schemes in IEEE 802.16e

5.2.1 Description

IEEE 802.16e specifications [IEEE16e05] include several MIMO profiles for 2, 3, and 4 transmit antennas. They provide transmit diversity, spatial multiplexing (SM) or combine the advantages of both. Most of the MIMO schemes included in the IEEE 802.16e specifications are based on two schemes which are defined for two transmit antennas. The first one, called Matrix A in the specifications, is based on the space–time block code (STBC) proposed by Alamouti for transmit diversity [Ala98]. This code achieves a diversity order that is equal to twice the number of antennas at the receiver, but it is only rate-1 code since it only transmits two symbols using two time slots. The other profile, defined as Matrix B, provides spatial multiplexing (SM) and uses two transmit antennas to transmit two independent data streams. This scheme is a rate-2 code, but it does not benefit from any diversity gain at the transmitter, and, at best, it provides a diversity order equal to the number of receive antennas. Furthermore, these two schemes are the only options defined for uplink (UL) transmission. These two schemes have also been included in the WiMAX Forum specifications as two mandatory profiles for use on the downlink.

It is believed that these schemes using 2 transmit antennas will be two basic profiles of most future standards, such as the IEEE 802.16m for mobile WiMAX evolutions and the LTE-Advanced of the 3GPP. However, there may be a need to include new codes combining the respective advantages of the Alamouti code and the SM while avoiding their drawbacks. Such a code actually exists in the IEEE 802.16e-2005 specifications as Matrix C. This code is a variant of the Golden code [BRV05] (see also [YW03] and [DV05] for other variants), which is known to be one of the best 2×2 STBCs achieving the diversity-multiplexing frontier [TV07].

Below we provide these three schemes of IEEE 802.16e while retaining the notations of [IEEE16e05]:

$$A=\begin{bmatrix} s_i & -s_{i+1}^* \\ s_{i+1} & s_i^* \end{bmatrix},\ B=\begin{bmatrix} s_i \\ s_{i+1} \end{bmatrix},\ C=\frac{1}{\sqrt{1+r^2}}\begin{bmatrix} s_i + jr\cdot s_{i+3} & r\cdot s_{i+1}+s_{i+2} \\ s_{i+1}-r\cdot s_{i+2} & jr\cdot s_i + s_{i+3} \end{bmatrix}, \tag{5.1}$$

where $r=(-1+\sqrt{5})/2$.

As the number of transmit antennas increase, the complexity of full-rate full-diversity codes increase exponentially and, therefore, for higher number of antennas only the combination of Alamouti and SM is preferred to improve the performance while keeping the detection complexity reasonable. This is, in fact, the case with IEEE 802.16e specifications and the existing schemes mainly use Alamouti code and SM given in equation (5.1) for three and four transmit antennas. Particularly, for three transmit antennas, the main schemes are defined as

$$A=\begin{bmatrix} \tilde{s}_1 & -\tilde{s}_2^* & 0 & 0 \\ \tilde{s}_2 & \tilde{s}_1^* & \tilde{s}_3 & -\tilde{s}_4^* \\ 0 & 0 & \tilde{s}_4 & \tilde{s}_3^* \end{bmatrix},\ B=\begin{bmatrix} \sqrt{\frac{3}{4}} & 0 & 0 \\ 0 & \sqrt{\frac{3}{4}} & 0 \\ 0 & 0 & \sqrt{\frac{3}{2}} \end{bmatrix}\begin{bmatrix} \tilde{s}_1 & -\tilde{s}_2^* & \tilde{s}_5 & -\tilde{s}_6^* \\ \tilde{s}_2 & \tilde{s}_1^* & \tilde{s}_6 & \tilde{s}_5^* \\ \tilde{s}_7 & -\tilde{s}_8^* & \tilde{s}_3 & -\tilde{s}_4^* \end{bmatrix},\ C=\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}. \tag{5.2}$$

Here the complex symbols to be transmitted are taken as x_1, x_2, x_3, x_4, which take values from a square QAM constellation, and we have $s_i = x_i e^{j\theta}$ for i=1,2,…,8, where $\theta=\tan^{-1}(1/3)$. Then, the matrix

elements are obtained as $\tilde{s}_1 = s_{1I} + js_{3Q}$; $\tilde{s}_2 = s_{2I} + js_{4Q}$; $\tilde{s}_3 = s_{3I} + js_{1Q}$; $\tilde{s}_4 = s_{4I} + js_{2Q}$ where $s_i = s_{iI} + js_{iQ}$.

The first two matrices, namely, Matrix A and Matrix B, benefit from transmit diversity exploited by means of Alamouti code. Moreover, in these two schemes, the coordinate interleaved notion [Jaf01] is also added over the phase-rotated symbols to increase the transmit diversity. Indeed, it can be easily seen that Matrix A of 3 transmit antenna has full diversity. Both matrices are defined as space-time-frequency codes (i.e., they are transmitted over 2 time slots and two subcarriers) and exploit the orthogonality of the Alamouti code for complexity reduction. Matrix C with 3 transmit antennas is the pure SM.

Similar to the 3 transmit antenna case, for 4 transmit antennas, we have the following matrices.

$$A = \begin{bmatrix} s_1 & -s_2^* & 0 & 0 \\ s_2 & s_1^* & 0 & 0 \\ 0 & 0 & s_3 & -s_4^* \\ 0 & 0 & s_4 & s_3^* \end{bmatrix}, \; B = \begin{bmatrix} s_1 & -s_2^* & s_5 & -s_7^* \\ s_2 & s_1^* & s_6 & -s_8^* \\ s_3 & -s_4^* & s_7 & s_5^* \\ s_4 & s_3^* & s_8 & s_6^* \end{bmatrix}, \; C = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix}. \tag{5.3}$$

Again, the first two matrices benefit from transmit diversity by means of Alamouti scheme, and defined as rate-1 and rate-2 options, respectively. We have also pure SM as a rate-4 option. In addition, all these codes with 3 and 4 transmit antennas are extended based on antenna grouping and antenna selection options depending on the feedback information. In addition to the mentioned open loop schemes, closed-loop schemes have also been included in [IEEE16e05]. However, none of these schemes have been included in the WiMAX profile and implemented in the existing products.

5.2.2 Simulated performance

Although IEEE 802.16e specifications include the above mentioned matrices and different permutations of them, WiMAX profile only includes the two simple ones: Matrix A and Matrix B with 2 transmit antennas. During the standardization period, there have been discussions to include Matrix C with 2 transmit antennas in order to benefit from transmit diversity while maximizing the transmission rate. However, because of its high decoding complexity it has not been included in the final WiMAX profile.

Indeed, the optimal Maximum-Likelihood (ML) detection complexity grows exponentially with the modulation and number of transmit antennas and this prevents the usage of Matrix C especially for high constellation sizes. However, because of the rapid change in wireless technologies, it is evident that implementation of more complex decoders will be possible. Therefore, such optimum codes are still thought to be strong candidates for future standards. In retrospect, it is meaningful to investigate the performance of these codes in real environments.

We now present the performance of Matrix C in a real WiMAX environment, where advanced WiMAX features such as frequency permutation and convolutional turbo codes (CTCs) are used. We particularly compare Matrix B and Matrix C in a 2×2 downlink MIMO WiMAX system for different coding rates, while using a soft-output sphere decoder based on the single tree search algorithm [SBB06]. FFT size is chosen as 1024 which corresponds to a system bandwidth of 10 MHz. Both uncoded and coded cases (with CTC having coding rates of 1/2 and 3/4) are treated in order to demonstrate the effect of channel encoding on space-time codes using QPSK modulation. At the receiver, we use the soft-output Schnorr-Euchner decoder, with a bit de-interleaver and a soft-input CTC decoder. In the simulations, Jakes' channel model is used in a Pedestrian B environment at a speed of 3 km/h. The Pedestrian B test environment parameters are given in Table 5.1.

Table 5.1: Pedestrian B Test Environment Tapped-Delay-Line Parameters

Tap	Relative delay (ns)	Average power (dB)	Doppler spectrum
1	0	0	Classic
2	200	-0.9	Classic
3	800	-4.9	Classic
4	1200	-8.0	Classic
5	2300	-7.8	Classic
6	3700	-23.9	Classic

Figure 5.1 shows the bit-error-rate (BER) performance of Matrix B and Matrix C for an uncoded QPSK signal constellation as a function of signal-to-noise ratio (SNR), where SNR is defined as the ratio of the received signal energy per antenna to noise spectral density. In the uncoded case, as expected, Matrix C outperforms Matrix B with a transmit diversity advantage which can be clearly observed above 10 dB SNR. On the other hand, as shown in Figure 5.2 (a), in the presence of channel coding with a rate of 3/4, Matrix C still performs better than Matrix B. However, the gap between the BER curves is remarkably reduced compared to the uncoded case and both schemes exploit essentially the same diversity order. When we further decrease the coding rate to 1/2 (see Figure 5.2 (b)), Matrix C only becomes closer to Matrix B below the BER value of 10^{-4} which requires an SNR value above 5 dB. These figures simply show that BER performance is dominated more by the diversity exploited by channel coding than the diversity exploited by STBC. In other words, channel codes may recover the diversity loss that Matrix B suffers from. Another interesting observation is that Matrix B outperforms Matrix C in the SNR range of interest with a much lower ML detection complexity. Similar results obtained also for higher modulation sizes and presented in [KSB09]. As a conclusion, these results along with the long and complex decoding of Matrix C compared to Matrix B, do not justify its use at least in the current WiMAX systems.

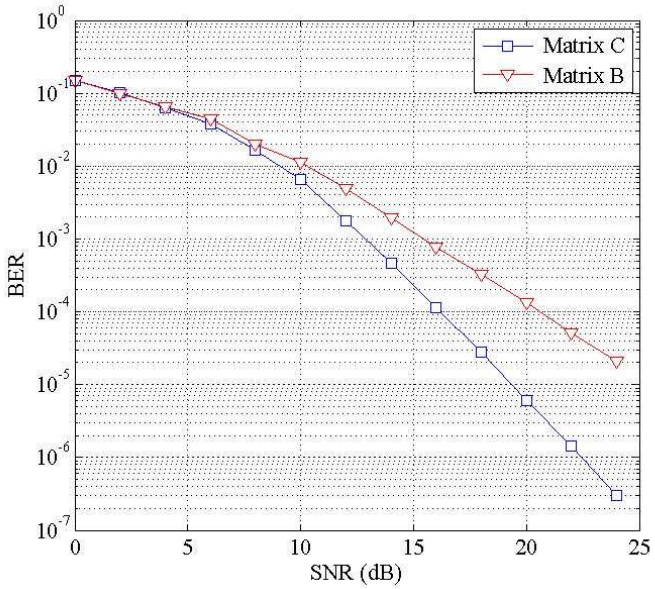
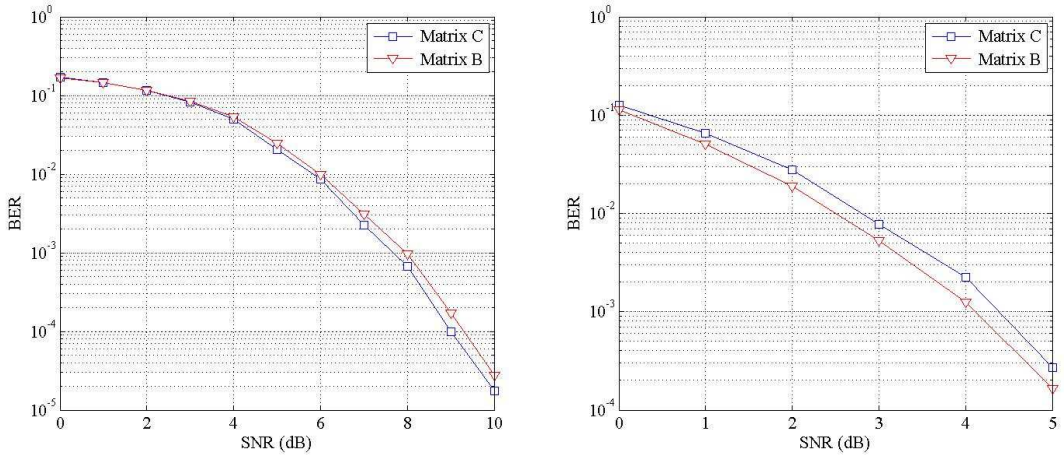


Figure 5.1: Uncoded BER, Matrix B and Matrix C, QPSK.



(a) QPSK CTC 3/4 (b) QPSK CTC 1/2

Figure 5.2: Coded BER, Matrix B and Matrix C.

5.2.3 Requirements on signalling and measurements

In the current WiMAX profile there is only one MIMO feedback option which allows a switch between Matrix A and Matrix B with two transmit antennas and requires a 6 bit feedback. Within this 6 bit we have both MIMO mode and the permutation information.

5.2.4 Requirements on architecture and protocols

Multiple transmit and receive antennas are needed in order to facilitate the MIMO transmission and reception of multiple spatial layers.

5.3 MIMO candidates for IEEE 802.16m

5.3.1 Description

In this section, we focus on the MIMO schemes included in the System Description Document (SDD) [IEEE16mSDD] of the IEEE 802.16m. Generally speaking, the SDD document includes the MIMO schemes (or at least the general descriptions of the schemes) which are decided to be included in the amendment document.

Concerning the IEEE 802.16m standardization period, the schemes considered up to now are modified versions of the ones existing in 802.16e specifications with 2 and 4 transmit antennas based on Alamouti and SM. The main concern is to find the codes which provide the best tradeoff between performance/rate/complexity. Recent attempts mainly focus on this issue and try to find different and better alternatives to the existing solutions. Despite the existence of numerous STBCs that provide interesting performance, they are not included in the SDD as they are all thought to be too complicated in terms of implementation.

In SDD, MIMO schemes are divided into two main groups, namely, open loop MIMO schemes and closed loop MIMO schemes. Currently, closed loop and open loop schemes use the same codebooks (or subset of these codebooks).

In open loop schemes, rate-1 schemes are collected in transmit diversity modes while the higher rate schemes have been put in spatial multiplexing modes. In particular, the transmit diversity modes are defined as

- 2Tx rate-1: For $M = 2$, SFBC with precoder, and for $M = 1$, a rank-1 precoder
- 4Tx rate-1: For $M = 2$, SFBC with precoder, and for $M = 1$, a rank-1 precoder
- 8Tx rate-1: For $M = 2$, SFBC with precoder, and for $M = 1$, a rank-1 precoder

where M denotes the number of symbols at a given time and SFBC refers to space-frequency block code. The precoding matrix will be based on the selected codebook and defined in the amendment document.

On the other hand, the spatial multiplexing modes include

- Rate-2 spatial multiplexing modes:
  - 2Tx rate-2: rate 2 SM with precoding
  - 4Tx rate-2: rate 2 SM with precoding
  - 8Tx rate-2: rate 2 SM with precoding
- Rate-3 spatial multiplexing modes:
  - 4Tx rate-3: rate 3 SM with precoding
  - 8Tx rate-3: rate 3 SM with precoding
- Rate-4 spatial multiplexing modes:
  - 4Tx rate-4: rate 4 SM with precoding
  - 8Tx rate-4: rate 4 SM with precoding

In closed-loop MIMO, unitary codebook based precoding is supported for both frequency-division duplex (FDD) and time-division duplex (TDD) systems. In TDD systems, sounding based precoding will also be supported. For codebook based precoding, two types of codebook are currently discussed in the SDD in TGm: 802.16e codebook and discrete Fourier transform (DFT) codebook.

5.3.2 Requirements on signalling and measurements

In FDD systems and TDD systems, a mobile station is required to feedback some of the following information in closed loop MIMO mode:

- Rank (Wideband or sub-band)
- Sub-band selection
- Channel Quality Indicator (CQI) : Wideband or sub-band, per layer
- Precoding Matrix Index (PMI) : Wideband or sub-band for serving cell and/or neighboring cell
- Long-term Channel State Information (CSI)

5.3.3 Requirements on architecture and protocols

Multiple transmit and receive antennas are needed in order to facilitate the MIMO transmission and reception of multiple spatial layers.

5.4 Concluding remarks

Although IEEE 802.16e specifications include many MIMO schemes, only the well-known Alamouti scheme and SM have been included in WiMAX profile for 2 transmit antennas. In order to have a 16m amendment more close to implementation, 16m working groups are trying to define minimum number of MIMO schemes (both mandatory and optional ones) which provide the best tradeoff between performance/rate/complexity. The optional ones will only be included only if they provide significant improvements compared to mandatory ones.

6. Conclusion

This deliverable captured the second set of best innovative concepts identified in the field of Advanced Antenna Schemes for potential inclusion into the WINNER+ system concept. The concepts consist of promising principles or ideas as well as detailed innovative techniques. For each concept, the associated benefits as well as the corresponding requirements on the system architecture and protocols, measurements and signalling, have been considered.

In Chapter 2, four innovative signal processing concepts for multiuser MIMO systems were presented. Here, the context is the downlink of a cellular network, where a base station employing an antenna array communicates with user terminals, each equipped with one or more antenna elements. The framework of the presented solutions consists of spatial user multiplexing or scheduling, and beamforming by means of linear transmit precoding. Since both the precoding and the scheduling depend heavily on the CSI knowledge in the transmitter (CSIT), the proposals focus on how to make the CSIT available. The problem of acquiring the CSIT consists of multiple tasks, such as pilot signal design, channel state and quality estimation, as well as feedback signal design. All these aspects were addressed in order to enhance the system performance.

The first concept in Chapter 2 presented a method for low-rank modelling of the long-term CSI, estimated over a finite time and frequency bandwidth. Compared to the conventional direct averaging, the low-rank modelling provides a more useful reference for precoding, especially when the directional components are dominating in the spatial channel. The second concept proposed a downlink pilot signal design technique, optimized for improving the CSI and channel quality estimation accuracy. The third proposal introduced a novel signalling concept for reducing the overhead caused by uplink CSI sounding, needed for multiuser precoding in TDD systems. According to the simulation results, the reduced pilot overhead actually improves the performance of the precoded transmission, due to the increased power efficiency of the sounding. Finally, the fourth concept described a predictive channel quality estimation method, utilizing the knowledge of the physical antenna array setup in vehicular receivers. The proposals in Chapter 2 form a realistic and promising set of improvements for accommodating precoded MIMO transmission in multiuser systems. Most of the concepts can be incorporated in the upcoming cellular systems, such as LTE-A, with minimal impact on the system specifications.

Chapter 3 proposed two innovative concepts related to communication strategies and network topologies involving relay nodes. The main aim of relaying is to increase the cell coverage, and to provide more uniform service quality over the whole geographical area comprising the cell. The first proposal showed how multiuser relaying by network coding can utilize multiple relay nodes, by employing different, linearly independent codes in the relays. The second concept proposed a new two-way MIMO amplify-and-forward relaying strategy for terminal-to-terminal communication. The proposal demonstrates that in TDD systems, relatively simple signal processing and pilot signaling techniques used in the relay node can accommodate spatial multiplexing of two terminals.

Finally, Chapter 4 presented two innovative concepts related to coding and decoding. The first one explored how to employ and benefit from network coding techniques in point-to-point MIMO transmission. In the second proposal, the general receiver processing problem of joint channel estimation, equalization and decoding was addressed in a concept employing an iterative belief propagation algorithm. The novel idea is to model the probability distributions as mixtures of Gaussian distributions. The approach allows for estimation improvement and complexity reduction simultaneously.

Chapter 5 was dedicated for an overview of the MIMO schemes of WiMAX systems, i.e., in the IEEE 802.16e standard and its enhancements in IEEE 802.16m system description document. In particular, the chapter focused on the diversity-rate trade-off from a receiver complexity point of view and highlighted various precoding schemes, their performance and resulting complexity. Similarly to the LTE track, the main emphasis is on codebook-based precoding, but in the TDD mode sounding based precoding will be supported as well.

7. References

[Ala98] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications,* vol. 16, no. 8, pp. 1451–1458, 1998.

[BG06] M. Biguesh and A.B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," IEEE Transactions on Signal Processing, vol. 54, pp. 884–893, 2006.

[BH07] F. Boccardi and H. Huang, "A near-optimum technique using linear precoding for the MIMO broadcast channel," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, vol. 3, Honolulu, Hawaii, Apr.15–20 2007, pp. III–17 – III–20.

[BHO09] E. Björnson, D. Hammarwall, B. Ottersten, "Exploiting Quantized Channel Norm Feedback Through Conditional Statistics in Arbitrarily Correlated MIMO Systems," IEEE Transactions on Signal Processing, October 2009.

[BO02] M. Bengtsson and B. Ottersten, "Optimum and suboptimum transmit beamforming," *in Handbook of antennas in wireless communication (L. C. Godara, eds.)*, CRC Press, 2002.

[BO09] E. Björnson and B. Ottersten, "Training-Based Bayesian MIMO Channel and Channel Norm Estimation", in Proc. IEEE ICASSP, April 2009.

[BO10] E. Björnson and B. Ottersten, "A Framework for Training-Based Estimation in Arbitrarily Correlated Rician MIMO Channels with Rician Disturbance," submitted to IEEE Transactions on Signal Processing, December 2008.

[BRV05] J. Belfiore, G. Rekaya, and E. Viterbo, "The Golden Code: A 2$$ 2 full rate Space-Time Code with non vanishing Determinants," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1432–1436, 2005.

[CG79] T. Cover, and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. On Inform. Theory*, pp. 572-584, Sep. 1979.

[CKL6] Y. Chen, S. Kishore and J. Li, "Wireless diversity through network coding,"*in Proceedings of IEEE WCNC 2006*, pp. 1681–1686.

[DHS03] G. Del Galdo, M. Haardt, and C. Schneider "Geometry-based channel modelling of MIMO channels in comparison with channel sounder measurements," *Advances in Radio Science- Kleinheubacher Berichte*, pp. 117-126, October 2003, more information on the model, as well as the source code and some exemplary scenarios can be found at http://tu- ilmenau.de/ilmprop.

[DJB95] C. Douillard, M. Jézéquel and C. Berrou, "Iterative correction of intersymbol interference: Turbo-equalization," *European Trans. On Telecommunications*, vol. 6, pp. 507-511, Sept./Oct. 1995.

[DKL07] J. Dauwels, S. Korl and H.-A. Loeliger, "Particle methods as message passing," *in Proceedings of IEEE ISIT'06*, pp. 2052-2056, July 2007.

[DV05] P. Dayal and M. Varanasi, "An optimal two transmit antenna space-time code and its stacked extensions," *IEEE Trans. on Information Theory,* , vol. 51, no. 12, pp. 4348–4355, 2005.

[GV97] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Information. Theory*, vol. 43, pp. 1986-1992, Nov 1997.

[Haa97] M. Haardt, "Structured least squares to improve the performance of ESPRIT-type algorithms," *IEEE Trans. Signal Processing*, vol. 45, pp. 792-799, Mar. 1997.

[HKR97] P. Hoeher, S. Kaiser, and P. Robertson, "Pilot-symbol-aided channel estimation in time and frequency," *in Proceedings of IEEE Global Telecommunications Conference (GLOBECOM'97)*, pp. 90–96, Nov 1997.

[HRG08] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems," *IEEE Trans. Signal Processing*, vol. 56, pp. 3198 - 3213, July 2008.

[IEEE16e05] IEEE 802.16-2005: IEEE Standard for Local and Metropolitan Area Networks – Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems – Amendment 2: Physical Layer and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands, Feb. 2006.

[IEEE16mSDD] IEEE 802.16 Broadband Wireless Access Working Group, "IEEE 802.16m system description document," IEEE 802.16m-08/003r7delta, Feb, 2009.

[Jaf01] H. Jafharkani, "A quasi-orthogonal space-time block code," *IEEE Trans. on Communications*, vol. 49, no. 1, pp. 1–4, 2001.

[KB08] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications", *SIAM Review*, vol. 51, no. 3, Sep. 2009, to appear.

[KFL01] F.R. Kschischang, B.J. Frey and H.-A. Loaliger, "Factor graphs and the sum-product algorithm," *IEEE Trans on Information Theory*, vol. 47, no. 2, pp. 498-519, Feb. 2001.

[KGG05] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. on Inform. Theory*, vol. 51, no. 9, pp. 3037-3063, Sep. 2005.

[KK07] P. Klenner and K.-D. Kammeyer, "Spatially interpolated OFDM with channel estimation for fast fading channels," *in Proceedings of Vehicular Technology Conference, 2007, VTC2007-Spring* pp. 2455–2459, April 2007.

[KM03] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. on Networking*, vol. 11, pp. 782-795, Oct. 2003.

[KSB09] R. Kobeissi, S. Sezginer, and F. Buda, "Downlink performance analysis of full-rate STCs in 2x2 MIMO WiMAX Systems," to appear *in Proceedings of VTC 2009 RAS Workshop*, Apr. 2009, Barcelona, Spain.

[KTL+09] P. Komulainen, A. Tölli, M. Latva-aho, and M. Juntti, "Channel Sounding Pilot Overhead Reduction for TDD Multiuser MIMO Systems," 5th IEEE Workshop on Broadband Wireless Access, Honolulu, Hawaii, November 30, 2009.

[Loe04] H.-A. Loeliger, "An introduction to factor graphs, "*IEEE Signal Processing Mag.,* vol. 21, pp. 28-41, Jan. 2004.

[LJS06] P. Larsson, N. Johansson, and K. E. Sunell, "Coded bi-directed relaying," *IEEE Trans. on Vehicular Technology*, pp. 851-855, May 2006.

[LMV00] L. de Lathauwer, B. de Moor, and J. Vanderwalle, "A multilinear singular value decomposition", *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, 2000.

[LTW04] J. N. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. on Inform. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[LYC03] S. Li, R. W. Yeung and N. Cai, "Linear network coding," *IEEE Trans. on Inform. Theory*, vol. 49, pp. 371-381, Feb. 2003.

[MV07] F.Z. Merli and G.M. Vitetta, "A factor graph approach to the iterative detection of OFDM signals in the presence of carrier frequency offset and phase noise," *in Proceedings of IEEE ICC'07*, pp. 2865-2870, June 2007.

[NHH04] A. Nosratinia, T. E. Hunter and A. Hedayat, "Cooperative Communication in Wireless Networks," *IEEE Communications Magazine,* pp. 74-80, Oct 2004.

[RFH08] F. Roemer, M. Fuchs, and M. Haardt, "Distributed MIMO systems with spatial reuse for high-speed-indoor mobile radio access," *in Proceedings of the 20-th Meeting of the Wireless World Research Forum (WWRF), (Ottawa, ON, Canada)*, Apr. 2008.

[RH07] F. Römer and M. Haardt, "Tensor-structure structured least squares (TS-SLS) to improve the performance of multi-dimensional ESPRIT-type algorithms," *in Proceedings of IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, vol. II, Honolulu, HI, pp. 893 - 896, Apr. 2007.

[RH09a] F. Roemer and M. Haardt, "Tensor-Based channel estimation (TENCE) for Two-Way relaying with multiple antennas and spatial reuse," *in Proceedings of IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009, invited paper.

[RH09b] F. Roemer and M. Haardt, “Structured least squares (SLS) based enhancements of tensor-based channel estimation (TENCE) for two-way relaying with multiple antennas,” *in Proceedings of International ITG Workshop on Smart Antennas (WSA 2009)*, Berlin, Germany, Feb. 2009.

[RH09c] F. Roemer and M. Haardt, “Algebraic norm-maximizing (ANOMAX) transmit strategy for two-way relaying with MIMO amplify and forward relays,” *IEEE Signal Processing Letters*, vol. 16, issue 10, pp. 909-912, Oct 2009.

[RSB04] D. Rajan, A. Sabharwal and B. Aazhang, “Delay-Bounded Packet Scheduling of Bursty Traffic Over Wireless Channels,” *IEEE Trans. Information. Theory,* vol. 50, no. 1, pp. 125-144, Jan 2004.

[RW05] B. Rankov and A. Wittneben, “Spectral efficient signaling for halfduplex relay channels”, *in Proceedings of the 39th Annual Asilomar Conference on Signals, Systems and Computers*, pp. 1066–1071, Pacific Grove, CA, Oct. 2005.

[SBB06] C. Studer, A. Burg, and H. Bolcskei, “Soft-output sphere decoding: algorithms and VLSI implementation”, *in Proceedings of 40th Asilomar Conf. Signals, Systems and Computers*, pp. 2071-2076, Nov. 2006*,* Pacific Grove, CA, USA.

[SH05] V. Stankovic and M. Haardt, “Multi-user MIMO downlink beamforming over correlated MIMO channels,” *in Proceedings International ITG/IEEE Workshop on Smart Antennas (WSA’05)*, Apr. 2005.

[SH08] V. Stankovic and M. Haardt, “Generalized design of multi-user MIMO precoding matrices,” *IEEE Trans. on Wireless Communications*, 2008.

[SSH04] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multi-user MIMO channels,” *IEEE Trans. Signal Processing*, vol. 52, pp. 461-471, Feb. 2004.

[STJ08] M. Schellmann, L. Thiele, and V. Jungnickel, “Predicting SINR conditions in mobile MIMO-OFDM systems by interpolation techniques”, *in Proceedings of 42nd Asilomar Conference on Signals, Systems and Computers*, Oct. 2008.

[TJ05] A. Tölli and M. Juntti, “Scheduling for multiuser MIMO downlink with linear processing,” in Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun., vol. 1, Berlin, Germany, Sep. 2005, pp. 156–160.

[TV07] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[UBV01] B.S. Ünal, A. Berthet and R. Visoz, “Iterative channel estimation and coded symbol detection for dispersive channels,” *in Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 1, Sept./Oct. 2001, pp. C-100-C-106.

[UK09] T. Unger and A. Klein, "Duplex schemes in multiple antenna two-hop relaying", EURASIP Journal on Advances in Signal Processing, vol. 2008, 2008, doi: 10.1155/2008/128592.

[VH08] Rahul Vaze and Robert W. Heath Jr, “On the capacity and diversity-multiplexing tradeoff of the two-way relay channel”, 2008, arXiv: 0810.3900v2.

[WHOB06] W. Weichselberger, M. Herdin, H. Özcelik, E. Bonek, “A stochastic MIMO channel model with joint correlation of both link ends,” IEEE Transactions on Wireless Communications, vol. 5, Issue 1, Jan 2006, pages 90- 100.

[WIN+D13] WINNER+ D1.3 “Initial Report in Peer-to-Peer and Network Coding”

[WIN+D14] WINNER+ D1.4 “Initial Report on Advanced Multiple Antenna Systems”

[WIN+D41] WINNER+ D4.1 “D4.1 Results of Y1 proposed candidate proof-of-concept evaluation”

[WIN2D341] WINNERII D3.4.1 “The WINNER II Air Interface: Refined Spatial-Temporal Processing Solutions”

[WS01] A.P. Worthen and W.E. Stark, “Unified design of iterative receivers using factor graphs,” *IEEE Trans. On Information Theory*, vol. 47, no. 2, pp. 843-849, Feb. 2001.

[WSBM06] H. Wymeersch, H. Steendam, H. Bruneel and M. Moeneclaey, “Code-aided frame synchronization and phase ambiguity resolution,” *IEEE Trans. On Signal Processing*, vol. 54, no. 7, pp. 2747-2757, Jul. 2006.

[XS09] M. Xiao and M. Skoglund, “Design of network codes for multi-user multi-relay networks,” IEEE ISIT 2009.

[YBB09] Y. Liu, L. Brunel, J.J. Boutros, “Joint channel estimation and decosing using Gaussian approximation in a factor graph over multipath channel,” *in Proceedings of IEEE PIMRC’09*, 2009.

[YK07] S. Yang, and R. Koetter, “Network coding over a noisy relay: a belief propagation approach,” *in Proceedings of IEEE ISIT 2007*, pp. 801–804.

[YW03] H. Yao and G. Wornell, “Achieving the full MIMO diversity-multiplexing frontier with rotation-based space-time codes,” *in Proceedings Allerton Conf. Commun., Cont., and Computing,(Illinois)*, 2003.

A. Appendix

A.1 Multi-user MIMO downlink precoding for time-variant correlated channels

A.1.1 Simulation scenario

We consider a 3 users MIMO downlink system. The simulation scenario is illustrated in Figure A.1 . The channels between each user and the BS are generated by a geometry-based channel [DHS03] and is capable of dealing with time variant frequency selective scenarios.

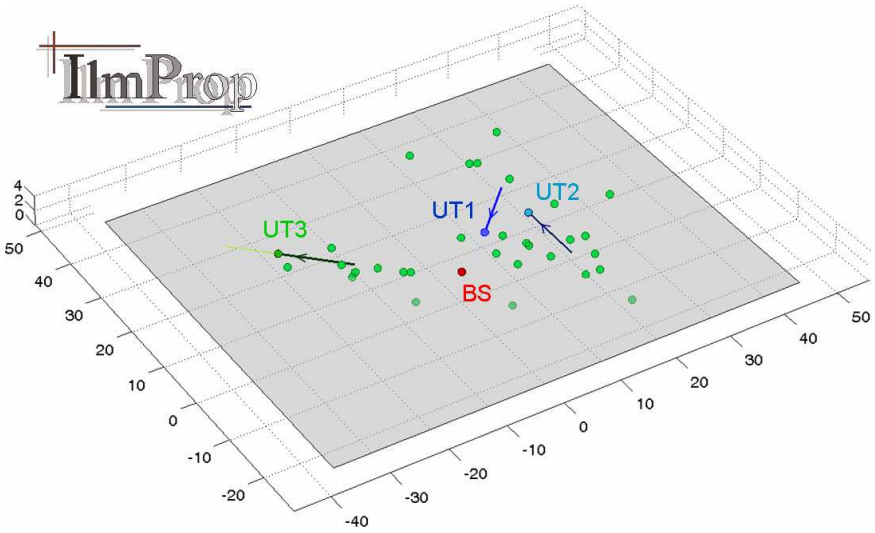


Figure A.1: The geometrical representation of the simulation scenario.

Each green point represents a fixed scatter. The channel impulse responses (CIR) are generated as a sum of propagation rays. The channel is computed from the superposition of the LOS component and a number of rays which represent the multi-path components. There are 8 transmit antennas at the BS and each user is equipped with 2 receive antennas. We simultaneously transmit two data streams to each user. User 1 and user 2 always have non-line of sight (NLOS) channels and user 3 always has a line of sight (LOS) channel. The velocities of the three users are 10 km/h. In Table A.1 the important OFDM parameters are listed.

Table A.1: OFDM Parameters

Parameters	Values
Carrier Frequency	5 GHz
Subcarrier Spacing	0.50196 MHz
Useful Symbol Duration	1.9922 µs
System Bandwidth	128.5 MHz
Used Subcarriers	[-128 : +128], 0 not used
Chunk Size	8 subcarriers, 15 OFDM symbols
Duplexing Mode	TDD

A.2 Efficient feedback schemes combining long term and short term information

A.2.1 System model

Since the scheme involves user specific precoded pilots, we focus on a single user. A communication system with n_T antennas at Base Station and n_R receive antennas is considered with the $n_R \times n_T$ channel matrix being denoted as $\boldsymbol{H}$. Entries of $\boldsymbol{H}$ are assumed to circularly symmetric complex Gaussian variables with vec($\boldsymbol{H}$) ~ $C(\mathbf{0}, \boldsymbol{R})$. Let $\boldsymbol{P}$ be the $n_T \times B$ training matrix with a power constraint trace ($\boldsymbol{P}^H\boldsymbol{P}$) = P. The input-output relation corresponding to the pilot transmission takes the form

$$\boldsymbol{Y} = \boldsymbol{H}\,\boldsymbol{P} + \boldsymbol{N}$$

where $\boldsymbol{N}$ includes interference and noise and is assumed to be spatially and temporally white with variance μ. While this model was used in [BO09] to treat white interference, investigation into the use of correlated interference is reported in [BO10].

A.2.2 MMSE channel estimation

The MMSE estimate of $\boldsymbol{H}$, denoted as $\boldsymbol{H}_{\boldsymbol{MMSE}}$, and the corresponding MSE can be shown to have the form,

$$vec(\boldsymbol{H}_{\boldsymbol{MMSE}}) = \boldsymbol{R}\,\boldsymbol{Q}\,(\boldsymbol{Q}\,\boldsymbol{R}\,\boldsymbol{Q}^H + \mu\,\boldsymbol{I})^{-1}\,vec(\boldsymbol{Y})$$
$$MSE(\boldsymbol{H}) = \mu\;Trace(\mu\,\boldsymbol{R}^{-1} + \boldsymbol{Q}^H\,\boldsymbol{Q})^{-1}$$

where $\boldsymbol{Q}$ is the Kronecker product of $\boldsymbol{P}^T$ and $\boldsymbol{I}$ (of dimension $n_R \times n_R$). The pilot matrix is then obtained by minimizing $MSE(\boldsymbol{H})$. While it is difficult to obtain the optimal pilot matrix for general correlations, further simplifications are possible when $\boldsymbol{R}$ can be written as a Kronecker product of $\boldsymbol{R}_{\boldsymbol{T}}^{\boldsymbol{T}}$ and $\boldsymbol{R}_{\boldsymbol{R}}$. Under this assumption, $\boldsymbol{P}$ has the form,

$$\boldsymbol{P} = \boldsymbol{U}_T\,\boldsymbol{\Sigma}\,\boldsymbol{V}^H$$

where $\boldsymbol{U}_T$ is the eigenvector matrix of $\boldsymbol{R}_{\boldsymbol{T}}$, $\boldsymbol{V}$ is an arbitrary orthonormal matrix and $\boldsymbol{\Sigma}$ is the diagonal power loading matrix. The system of equations leading to $\boldsymbol{\Sigma}$ is presented in [BO09]. Simplifications for special cases can also be found in [BO09].

Heuristic estimator:

While the optimal estimator is derived under the Kronecker structure assumption, it can be applied to arbitrarily correlated matrices by using $E\{\boldsymbol{H}^H\,\boldsymbol{H}\}$ instead of $\boldsymbol{R}_{\boldsymbol{T}}$.

A.2.3 MMSE channel norm estimation

It is tedious to obtain the pilot matrix minimizing the MSE of the channel (Frobenius) norm $\|\,\boldsymbol{H}\,\|^2$. Assuming that the pilot matrix has the form $\boldsymbol{P} = \boldsymbol{U}_T\,\boldsymbol{\Sigma}_1\,\boldsymbol{V}^H$, for some power loading matrix, the MMSE estimate of $\|\,\boldsymbol{H}\,\|^2$ and the corresponding MSE are given by,

$$\boldsymbol{\rho}_{mmse} = \mu\,\boldsymbol{1}^T\boldsymbol{B}\,\boldsymbol{1} + \boldsymbol{w}^T\,\boldsymbol{\Gamma}\,\boldsymbol{B}^2\,\boldsymbol{\Gamma}^H\,\boldsymbol{w}$$
$$MSE(\|\,\boldsymbol{H}\,\|^2) = \mu\,\boldsymbol{1}^T\boldsymbol{B}\,(2\mu\,\boldsymbol{\Gamma}\,\boldsymbol{\Phi}\,\boldsymbol{\Gamma}^H + \mu^2\,\boldsymbol{I})^{-1}\boldsymbol{B}\,\boldsymbol{1}$$

where $\boldsymbol{\Gamma}$ is the Kronecker product of $\boldsymbol{\Sigma}_1$ with $\boldsymbol{I}$, $\boldsymbol{B} = \boldsymbol{\Phi}\,(\boldsymbol{\Gamma}\,\boldsymbol{\Phi}\,\boldsymbol{\Gamma}^H + \mu\,\boldsymbol{I})^{-1}$, $\boldsymbol{w}$= $vec(\boldsymbol{Y}\,\boldsymbol{V})$, $\boldsymbol{\Phi}$ is the eigen-value matrix of $\boldsymbol{R}_T$ and $\boldsymbol{1}$ is a column vector of ones of appropriate dimension. Obtaining the optimal $\boldsymbol{\Sigma}_1$ is an optimization problem that only can be solved explicitly in a few special cases [BO9]. In general, the optimization problem is either convex or can be turned into convex by introducing a few rules on the power loading.

A.3 Pilot overhead reduction for multiuser MIMO systems in TDD mode

In this appendix, the performance of the pilot overhead reduction strategy described in Section 2.3 is evaluated with channel estimation, in the context of beam selection and multiuser zero-forcing by coordinated transmit-receive processing.

Coordinated linear transmitter-receiver processing by block diagonalization (BD) with greedy beam selection is a method to utilize all degrees of freedom available in multiuser MIMO networks [TJ05], [BH07]. By applying instantaneous channel state information in the transmitter (CSIT), the BD criterion offers zero-forcing between the downlink data streams of different users. The MIMO channels of different users are decoupled so that precoding based on singular value decomposition (SVD) can be carried out individually for each user. Any combination of the number of antennas in terminals and the base station can be supported.

Different multiuser MIMO scenarios were simulated in frequency flat fading and uncorrelated channels between antennas. In each simulation, constant antenna and pilot overhead reduction setups are applied in all terminals so that $N_k = N_U$ and $J_k = J$ for all k. The number of BS antennas is fixed to $N_B = 4$. The rates shown are averages over data fields only so that the fractional rate loss caused by pilot overhead is not taken into account.

The system signal-to-noise-ratio SNR for data was set to 10dB, and it is defined as P/N_0, where P is the total DL data transmit power. For power allocation, waterfilling with the sum power constraint is applied so that the system sum rate is maximized. While the objective of the simulations is to analyze the effect of imperfect or reduced CSIT, the terminal receivers for data demodulation are assumed to operate based on perfect knowledge of the DL responses.

Figure A.2 shows the average system sum rate performance of different terminal antenna and overhead setups versus the number of users K, in static channel and with ideal CSI estimation. As K grows, the loss from the incomplete sounding is reduced. When $K > 2$, two sounding beams per user ($J = 2$) is enough to achieve nearly optimal performance.

Figure 2.8 in Section 2.3.3 depicted the case where both the common pilot and the UL sounding are affected by estimation noise so that both are observed with equal sum SNR. As can be seen, the multiuser MIMO system is sensitive to the SNR of the sounding. However, the pilot overhead reduction actually decreases this sensitivity and improves the performance at low SNR.

According to the results, the performance loss induced by the incomplete sounding is minor, as the beamforming gain provided by multiple terminal antennas, and the multiuser diversity seen by the BS are retained. When taking into account the CSI estimation error in the BS, the overhead reduction turns out to improve robustness and even increase the average system capacity.

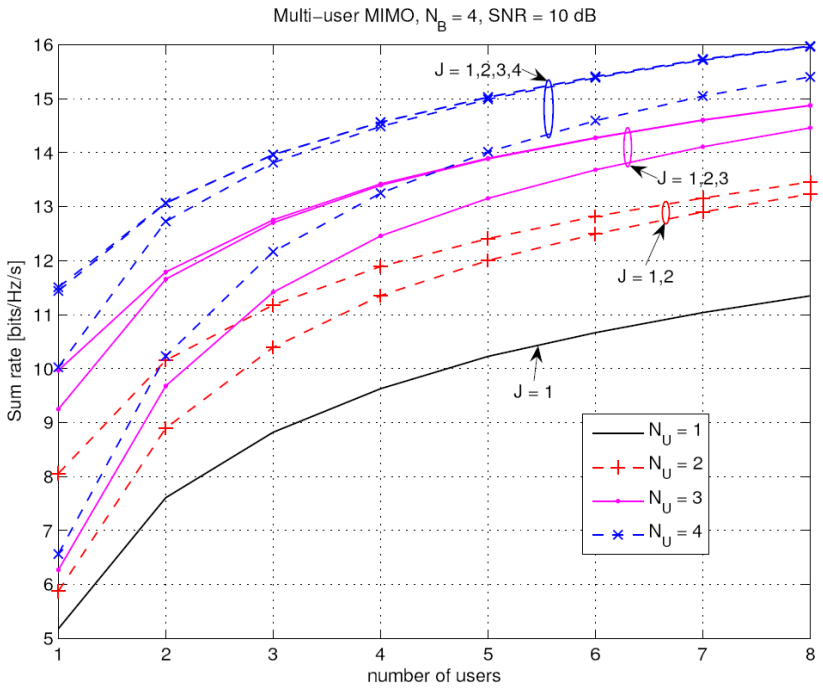


Figure A.2: Average sum rate vs. number of users, with ideal CSI estimation, $N_B = 4$.

A.4 Channel prediction based on linear interpolation techniques

According to Figure 2.10 in Section 2.4.2, the channel vector seen at the receive antenna array with N_r antennas can for n-th transmit antenna be given as

$$\mathbf{h}_n(i) = [h_n(i)\ h_n(i-D)\ \cdots\ h_n(i-(N_r-1)D)]^T \tag{A.1}$$

This notation highlights that the channel vector $h_n(i)$ is composed of N_r equi-spaced sampling points of the time-variant channel function $h_n(i)$.

We will first characterize the statistical properties of the channel vectors $h_n(i)$. We consider all channel functions $h_n(i)$, $n = 1,...,N_t$, constituting those vectors to have identical statistics. Hence, we focus on a channel for a single transmit antenna and omit the index n for notational convenience.

The channel's autocorrelation function (ACF) is given as $\varphi_{hh}(k) = E_i\{h(i)h^*(i+k)\}$. Assuming Jakes' model for the temporal evolution of the channel function h(i), the autocorrelation function yields

$$\varphi_{hh}(k) = E_i\{h(i)h^*(i+k)\} = J_0(2\pi k f_D \cdot T_o)$$

where $J_0(\cdot)$ is the the Bessel function of the first kind resulting from Jakes' Doppler power spectrum, and T_o is the OFDM symbol duration. The covariance matrix of the channel vector h(i) is defined as

$$\mathbf{R}_{hh} = E\{\mathbf{h}(i)\mathbf{h}^H(i)\}$$

Using (A.1), its elements can be related to the ACF φ_{hh} according to

$$[\mathbf{R}_{hh}]_{cd} = \varphi_{hh}((d-c)D) = J_0(2\pi(d-c)Df_D \cdot T_o) \tag{A.2}$$

where $[R_{hh}]_{cd}$ represents the element of matrix R_{hh} found in c-th row and d-th column. The delay D can be related to the antenna spacing a and the speed v of the mobile vehicle according to $DT_o = a/v$. Further, v relates to the Doppler frequency f_D via $v = f_D\lambda$, with λ being the wavelength of the carrier frequency. Thus, we obtain

$$D = \frac{a}{\lambda f_D T_o} \tag{A.3}$$

Inserting this expression into (A.2), we obtain for the elements in the correlation matrix R_{hh}

$$[\mathbf{R}_{hh}]_{cd} = J_0(2\pi(d-c)a/\lambda)$$

A.4.1 Channel prediction by linear interpolation techniques

The channel vector $h(i_0)$ according to (A.1) supplies N_r equi-spaced sampling points of the channel function h(i). Hence, we can use channel interpolation techniques to determine h(j) for an arbitrary j and thus obtain an estimate for the channel vector $h(i_k)$ for a future time instant $i_k > i_0$. However, for proper application of the interpolation techniques, it has to be ensured that the density of sampling points of h(i) obtained from the vector h(i) complies to the requirement of the sampling theorem, which yields [HKR97]

$$2f_D D T_o \le 1 \quad \Leftrightarrow \quad a \le \lambda/2$$

where we used equation (A.3). From this result, we can conclude that the channel prediction based on channel interpolation techniques requires an antenna spacing of at most $\lambda/2$. (Note that an antenna spacing $a < 0.5\lambda$ may result in modified radiation patterns due to mutual antenna coupling. However, this effect has not been taken further into account here.)

Next we turn our focus on the realization of the channel predictor. Note that the interpolation-based prediction gets the more reliable, the more information on the channel function h(i) can be taken into account. Hence, we use the past measured channels $h_n(i)$ gathered over an observation window of length N_o, i.e., $i = -N_o+1,...,0$, as input for the predictor. We assume here that $i_0 = 0$ is the index of the last OFDM symbol where measured channel information is available (see Figure A.3). Let

$$\mathbf{y}_m = [h(-(m-1)D - N_o + 1) \cdots h(-(m-1)D)]^T$$

be a vector comprising the N_o successive observations of the channel coefficient at m-th receive antenna. A compound observation vector is formed by stacking the single vectors y_m into one according to their temporal order, i.e., $y = [y_{N_r}^T \cdots y_1^T]$. The MMSE solution of the linear interpolator [HKR97] yields for the estimate of the future channel vector

$$\begin{aligned} \hat{\mathbf{h}}(i_k) &= \theta^H(i_k)\Phi^{-1}\mathbf{y} \\ \theta(i_k) &= E\{\mathbf{y}\mathbf{h}^H(i_k)\} \\ \Phi &= E\{\mathbf{h}\mathbf{h}^H\} + \gamma^{-1}\mathbf{I}_{N_r N_o} \end{aligned} \tag{A.4}$$

where γ is the SNR of the measured channels contained in y, characterizing the quality of the single measurements. The matrix $E\{yy^H\}$ constituting Φ can be structured into submatrices A_{mj} of dimension $N_o \times N_o$, which result from the outer products of the subvectors y_m in y, i.e., $A_{mj} = E\{y_m y_j^H\}$. Their elements relate to the channel's ACF $\varphi_{hh}(k)$ according to

$$[\mathbf{A}_{mj}]_{cd} = \varphi_{hh}((j-m)D + (d-c))$$

Correspondingly, the matrix $\theta(i_k)$ can be structured into submatrices $B_m = E\{y_m h^H(i_k)\}$ of dimension $N_o \times N_r$, whose elements relate to the ACF as

$$[\mathbf{B}_m]_{cd} = \varphi_{hh}(-(m-d)D - k + c)$$

Once we have obtained the predicted channel vectors $h_n(i_k)$, n = 1,...,N_t, from (A.4), we can construct the predicted channel matrix for the complete MIMO channel. Based on this MIMO channel matrix, the UT can then determine the future SINRs for the different spatial modes.

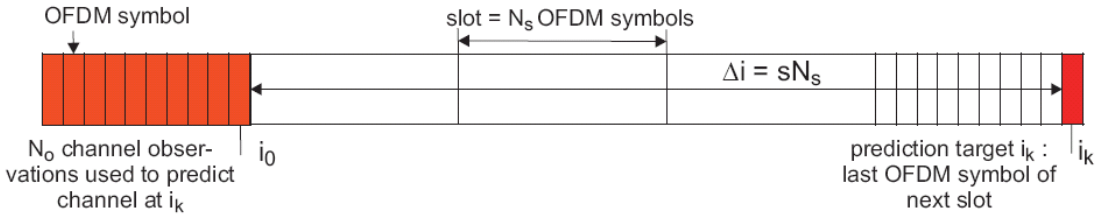


Figure A.3: Signal structure used for the prediction-based approach.

A.5 Two-way relaying with MIMO AF relays

A.5.1 Channel estimation

A.5.1.1 Description of channel estimation schemes

Training

In order to enable the channel estimation scheme we first require a training phase. We propose the following signalling scheme: The training phase is divided into M_{R} frames. In each frame, both nodes transmit sequences of N_{P} pilot symbols which we denote as $\boldsymbol{x}_{1,j} \in \mathrm{C}^{M_1}, \boldsymbol{x}_{2,j} \in \mathrm{C}^{M_2}$ for $j = 1,2,\ldots,N_{\mathrm{P}}$. The relay uses the relay amplification matrix $\boldsymbol{G}^{(i)} \in \mathrm{C}^{M_{\mathrm{R}} \times M_{\mathrm{R}}}$ in the *i*-th frame. All matrices $\boldsymbol{G}^{(i)}$ and all pilot symbols $\boldsymbol{x}_{1,j}, \boldsymbol{x}_{2,j}$ are designed beforehand and therefore known to both nodes. We conclude that the total number of pilot slots is equal to $N_{\mathrm{P}} \cdot M_{\mathrm{R}}$. As shown in [RH09a], we require $N_{\mathrm{P}} \geq M_1 + M_2$ and consequently at least $\left(M_1 + M_2\right) \cdot M_{\mathrm{R}}$ total number of pilots.

Tensor data model

We can conveniently express the data model in tensor form. The tensor representation can be seen as an alternative way of expressing the same data model. Due to its simplicity, the equations reveal significantly more of the structure inherent in the model, which helps to derive efficient solutions.

The tensor operations we use are consistent with [LMV00]. For convenience we now summarize the required operations, for more details the reader is referred to [LMV00] or [HRG08]. A three-dimensional tensor $\boldsymbol{A} \in \mathrm{C}^{M_1 \times M_2 \times M_3}$ is a three-way array of size $M_1 \times M_2 \times M_3$. The *n*-mode vectors of $\boldsymbol{A}$ are obtained by collecting the elements of $\boldsymbol{A}$ into a vector where the *n*-th index is varied in its range and all other indices are held fixed. We can rearrange the elements of $\boldsymbol{A}$ into matrices in several ways. The particular matrix we obtain by aligning all *n*-mode vectors of $\boldsymbol{A}$ as columns of a matrix is referred to as the *n*-mode unfolding of $\boldsymbol{A}$ and written as $\left[\boldsymbol{A}\right]_{(n)} \in \mathrm{C}^{M_n \times M_1 \cdot M_2 \cdot M_3 / M_n}$ for $n = 1,2,3$. We can multiply a three-dimensional tensor and a matrix in each of its three modes. Formally, this is accomplished by the *n*-mode product operator. The *n*-mode product between a tensor $\boldsymbol{A} \in \mathrm{C}^{M_1 \times M_2 \times M_3}$ and a matrix $\boldsymbol{U}_n \in \mathrm{C}^{P_n \times M_n}$ is symbolized by $\boldsymbol{A} \times_n \boldsymbol{U}_n$ which means that all *n*-mode vectors of $\boldsymbol{A}$ are multiplied from the left-hand side by the matrix $\boldsymbol{U}_n$. Therefore, $\left[\boldsymbol{A} \times_n \boldsymbol{U}_n\right]_{(n)} = \boldsymbol{U}_n \cdot \left[\boldsymbol{A}\right]_{(n)}$. Finally, to concatenate two tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ along the *n*-th mode we use the concatenation operator $\left[\boldsymbol{A} \cup_{(n)} \boldsymbol{B}\right]$ [RH07].

With the help of these operators, the received data at the nodes can be written as

$$\begin{aligned} \boldsymbol{Y}_1 &= \boldsymbol{G} \times_1 \boldsymbol{H}_1^{\mathrm{T}} \times_2 \left(\boldsymbol{H} \cdot \boldsymbol{X}\right)^{\mathrm{T}} + \tilde{\boldsymbol{N}}_1 \in \mathrm{C}^{M_1 \times N_{\mathrm{P}} \times M_{\mathrm{R}}} \\ \boldsymbol{Y}_2 &= \boldsymbol{G} \times_1 \boldsymbol{H}_2^{\mathrm{T}} \times_2 \left(\boldsymbol{H} \cdot \boldsymbol{X}\right)^{\mathrm{T}} + \tilde{\boldsymbol{N}}_2 \in \mathrm{C}^{M_2 \times N_{\mathrm{P}} \times M_{\mathrm{R}}}. \end{aligned} \tag{A.5}$$

Here the tensors $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ collect all received samples during the training phase, where the first index references antennas, the second index the pilot symbols and the third index the frames. Moreover, the compound channel matrix $\boldsymbol{H}$ and the compound pilot matrix $\boldsymbol{X}$ are defined as

$$\begin{aligned} \boldsymbol{H} &= \left[\boldsymbol{H}_1, \boldsymbol{H}_2\right] \in \mathrm{C}^{M_{\mathrm{R}} \times (M_1 + M_2)} \\ \boldsymbol{X} &= \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_{1,1} & \cdots & \boldsymbol{x}_{1,N_{\mathrm{P}}} \\ \boldsymbol{x}_{2,1} & \cdots & \boldsymbol{x}_{2,N_{\mathrm{P}}} \end{bmatrix} \in \mathrm{C}^{(M_1 + M_2) \times N_{\mathrm{P}}}. \end{aligned} \tag{A.6}$$

The tensors $\tilde{\boldsymbol{N}}_1$ and $\tilde{\boldsymbol{N}}_2$ represent the effective noise contributions and the tensor $\boldsymbol{G} \in \mathrm{C}^{M_{\mathrm{R}} \times M_{\mathrm{R}} \times M_{\mathrm{R}}}$ is defined by aligning the relay amplification matrices $\boldsymbol{G}^{(i)}$ along the third mode of the tensor so that the *i*-th 3-mode "slice" of $\boldsymbol{G}$ is equal to $\boldsymbol{G}^{(i)}$, i.e., $\boldsymbol{G} = \left[\boldsymbol{G}^{(1)} \cup_{(3)} \boldsymbol{G}^{(2)} \cup_{(3)} \cdots \cup_{(3)} \boldsymbol{G}^{(M_{\mathrm{R}})}\right]$. Note that we

can also express the tensor $\boldsymbol{G}$ with the help of its PARAFAC decomposition [KB08] in the following manner

$$\boldsymbol{G} = \boldsymbol{I}_{3,M_{\mathrm{R}}} \times_1 \boldsymbol{G}_1 \times_2 \boldsymbol{G}_2 \times_3 \boldsymbol{G}_3, \tag{A.7}$$

where $\boldsymbol{I}_{3,M_{\mathrm{R}}}$ is the identity tensor of size $M_{\mathrm{R}} \times M_{\mathrm{R}} \times M_{\mathrm{R}}$ which is equal to one if all three indices are equal and zero otherwise and the matrices $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, and $\boldsymbol{G}_3$ are of size $M_{\mathrm{R}} \times M_{\mathrm{R}}$.

Since the models in $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are very similar we will focus on user terminal one in the following. The corresponding equations for UT2 are obtained by consistently exchanging $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$.

TENCE- an algebraic channel estimation scheme

In this section we briefly summarize the TENCE algorithm, which is a purely algebraic **Ten**sor-based **C**hannel **E**stimation scheme for two-way relaying scenarios with arbitrary antenna configurations [RH09a] . The first step of TENCE is to compute the matrix $\left(\boldsymbol{G}_3^{-1} \cdot \left[\boldsymbol{Y}_1\right]_{(3)}\right)^{\mathrm{T}}$. In the absence of noise, this matrix admits the following factorization

$$\left(\boldsymbol{G}_3^{-1} \cdot \left[\boldsymbol{Y}_1\right]_{(3)}\right)^{\mathrm{T}} = \left[\boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G}_1\right] \lozenge \left[\boldsymbol{X}^{\mathrm{T}} \cdot \boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{G}_2\right], \tag{A.8}$$

where $\lozenge$ represents the Khatri-Rao (column wise Kronecker) product between two matrices. In the presence of noise, this factorization represents an approximation. However, we can still find factors $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ such that $\boldsymbol{F}_1 \lozenge \boldsymbol{F}_2$ is the best approximation of $\left(\boldsymbol{G}_3^{-1} \cdot \left[\boldsymbol{Y}_1\right]_{(3)}\right)^{\mathrm{T}}$ in the least squares sense. This is accomplished via a least squares Khatri-Rao factorization algorithm explained in the Appendix A.5.1.2. Unfortunately, the Khatri-Rao factorization is only unique up to one scaling ambiguity per column. We therefore obtain factors which obey the following model

$$\begin{aligned} \boldsymbol{F}_1 &\approx \boldsymbol{H}_1 \cdot \boldsymbol{G}_1 \cdot \boldsymbol{\Lambda} \\ \boldsymbol{F}_2 &\approx \boldsymbol{X}^{\mathrm{T}} \cdot \boldsymbol{H}^{\mathrm{T}} \cdot \boldsymbol{G}_2 \cdot \boldsymbol{\Lambda}^{-1}, \end{aligned} \tag{A.9}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with the elements of the vector $\boldsymbol{\lambda} = \left[\lambda_1, \lambda_2, \ldots, \lambda_{M_{\mathrm{R}}}\right]$ on its main diagonal. Moreover, each $\lambda_m \in \mathrm{C}$ is an arbitrary complex number representing the scaling ambiguity.

In the second step of TENCE we resolve these scaling ambiguities by exploiting the structure of the data model even further. The details about how to obtain $\boldsymbol{\Lambda}$ can be found in the Appendix A.5.1.3. With the help of the estimated $\hat{\boldsymbol{\Lambda}}$ we can then obtain the final channel estimates

$$\begin{aligned} \boldsymbol{H}_1^{\mathrm{est}} &= \left(\boldsymbol{F}_1 \cdot \left(\boldsymbol{\Lambda}^{\mathrm{est}}\right)^{-1} \cdot \boldsymbol{G}_1^{-1}\right)^{\mathrm{T}} \\ \boldsymbol{H}_2^{\mathrm{est}} &= \left(\left[\boldsymbol{X}_2^{\mathrm{T}}\right]^{+} \cdot \boldsymbol{F}_2 \cdot \boldsymbol{\Lambda}^{\mathrm{est}} \cdot \boldsymbol{G}_2^{-1}\right)^{\mathrm{T}}, \end{aligned} \tag{A.10}$$

where the superscript $^{+}$ denotes the Moore-Penrose pseudo inverse. Note that all relevant scaling ambiguities have been resolved. What remains is one global sign ambiguity: Instead of finding an estimate for $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ we might as well find an estimate for $-\boldsymbol{H}_1$ and $-\boldsymbol{H}_2$. However, since in the data model in equation (3.3) this unknown sign cancels, this ambiguity can be considered irrelevant.

Structured Least Squares (SLS)-based iterative refinement

We can improve the channel estimate obtained via TENCE further by exploiting even more of the structure inherent in the received data in the training phase. What we cannot exploit in TENCE is that for each terminal its own channel is present in the 1-mode and the 2-mode of the received training data. While TENCE would generate independent estimates for these two channel matrices, we can improve the estimation if we force these two channel matrices to be equal. This leads to a non-linear least squares problem which can be solved iteratively by local linearization around the initial solution obtained via TENCE. In order to enhance the numerical stability we can additionally include regularization by adding penalty terms for solutions that deviate too much from the initial solution. The solution is inspired by similar ideas used in the Structured Least Squares (SLS) [Haa97] and the Tensor-Structure Structured

Least Squares (TS-SLS) [RH07] algorithms, which solve the highly structured shift invariance equations that appear in ESPRIT-based high-resolution direction of arrival estimation algorithms.

Let $\Delta\boldsymbol{H}_{1,k}$ and $\Delta\boldsymbol{H}_{2,k}$ be the correction terms for the channel estimates $\boldsymbol{H}_1^{\text{est}}$ and $\boldsymbol{H}_2^{\text{est}}$ so that the improved estimates after *k* iterations are given by $\boldsymbol{H}_1^{\text{est}} + \Delta\boldsymbol{H}_{1,k}$ and $\boldsymbol{H}_2^{\text{est}} + \Delta\boldsymbol{H}_{2,k}$. Then the overall cost function can be expressed in the following form [RH09b]

$$J\left(\Delta\boldsymbol{H}_k\right) = \left\| \boldsymbol{Y}_1 \times_2 \left(\boldsymbol{X}^{\mathrm{T}}\right)^{+} - \boldsymbol{G} \times_1 \left(\boldsymbol{H}_1^{\text{est}} + \Delta\boldsymbol{H}_{1,k}\right)^{\mathrm{T}} \times_2 \left(\boldsymbol{H}^{\text{est}} + \Delta\boldsymbol{H}_k\right)^{\mathrm{T}} \right\|_{\mathrm{H}}^{2} + \frac{M_1}{\alpha} \cdot \left\|\Delta\boldsymbol{H}_{1,k}\right\|_{\mathrm{F}}^{2} + \frac{M_2}{\alpha} \cdot \left\|\Delta\boldsymbol{H}_{2,k}\right\|_{\mathrm{F}}^{2}. \quad \text{(A.11)}$$

Here, $\boldsymbol{H}^{\text{est}} = \left[\boldsymbol{H}_1^{\text{est}}, \boldsymbol{H}_2^{\text{est}}\right]$ and $\Delta\boldsymbol{H}_k = \left[\Delta\boldsymbol{H}_{1,k}, \Delta\boldsymbol{H}_{2,k}\right]$ represent the compound channel estimate and the compound correction term, respectively. Moreover, $\|\boldsymbol{A}\|_{\mathrm{F}}$ and $\|\boldsymbol{A}\|_{\mathrm{H}}$ symbolize the matrix Frobenius norm and the higher-order Frobenius norm which are both defined as the square root of the sum of the squared magnitude of all elements. The second and the third term in the cost function contain the regularizations which penalize correction terms with a large norm. Note that these penalty terms are weighted by a scalar regularization parameter α which controls the amount of regularization to be used: The larger α is chosen the less regularization is included. Numerical evaluations have shown that a value of $\alpha \approx 100$ is a reasonable value which enhances the numerical stability at low signal to noise ratios. The iterative solution of the cost function in equation (A.11) is achieved by the following procedure

$$\begin{aligned} \Delta\boldsymbol{H}_{1,k+1} &= \Delta\boldsymbol{H}_{1,k} + \Delta\Delta\boldsymbol{H}_{1,k} \\ \Delta\boldsymbol{H}_{2,k+1} &= \Delta\boldsymbol{H}_{2,k} + \Delta\Delta\boldsymbol{H}_{2,k} \\ \Delta\boldsymbol{H}_{k+1} &= \left[\Delta\boldsymbol{H}_{1,k+1}, \Delta\boldsymbol{H}_{2,k+1}\right], \end{aligned} \quad \text{(A.12)}$$

where the update terms $\Delta\Delta\boldsymbol{H}_{1,k}$ and $\Delta\Delta\boldsymbol{H}_{2,k}$ are initialized with zero matrices for $k = 0$ and for $k > 0$ computed via

$$\begin{bmatrix} \mathrm{vec}\left\{\Delta\Delta\boldsymbol{H}_{1,k}\right\} \\ \mathrm{vec}\left\{\Delta\Delta\boldsymbol{H}_{2,k}\right\} \end{bmatrix} = -\begin{bmatrix} \tilde{\boldsymbol{F}}_k^{(1)} & \tilde{\boldsymbol{F}}_k^{(2)} \\ \sqrt{\frac{M_1}{\alpha}} \cdot \boldsymbol{I}_{M_1 \cdot M_{\mathrm{R}}} & \boldsymbol{0}_{M_1 \cdot M_{\mathrm{R}} \times M_2 \cdot M_{\mathrm{R}}} \\ \boldsymbol{0}_{M_2 \cdot M_{\mathrm{R}} \times M_1 \cdot M_{\mathrm{R}}} & \sqrt{\frac{M_2}{\alpha}} \cdot \boldsymbol{I}_{M_2 \cdot M_{\mathrm{R}}} \end{bmatrix}^{+} \cdot \begin{bmatrix} \mathrm{vec}\left\{\boldsymbol{R}_k\right\} \\ \sqrt{\frac{M_1}{\alpha}} \cdot \mathrm{vec}\left\{\Delta\boldsymbol{H}_{1,k}\right\} \\ \sqrt{\frac{M_2}{\alpha}} \cdot \mathrm{vec}\left\{\Delta\boldsymbol{H}_{2,k}\right\} \end{bmatrix} \quad \text{(A.13)}$$

$$\boldsymbol{R}_k = \boldsymbol{Y}_1 \times_2 \left(\boldsymbol{X}^{\mathrm{T}}\right)^{+} - \boldsymbol{G} \times_1 \left(\boldsymbol{H}_1^{\text{est}} + \Delta\boldsymbol{H}_{1,k}\right)^{\mathrm{T}} \times_2 \left(\boldsymbol{H}^{\text{est}} + \Delta\boldsymbol{H}_k\right)^{\mathrm{T}} \quad \text{(A.14)}$$

$$\begin{aligned} \tilde{\boldsymbol{F}}_k^{(1)} = {} & \boldsymbol{P}^{(3)}{}_{M_1, M_1+M_2, M_{\mathrm{R}}} \cdot \left(\boldsymbol{I}_{M_1} \otimes \left[\boldsymbol{G} \times_2 \left(\boldsymbol{H}^{\text{est}} + \Delta\boldsymbol{H}_k\right)^{\mathrm{T}} \right]_{(1)}^{\mathrm{T}} \right) \\ & + \boldsymbol{P}^{(1)}{}_{M_1, M_1+M_2, M_{\mathrm{R}}} \cdot \left(\boldsymbol{I}_{M_1+M_2} \otimes \left[\boldsymbol{G} \times_1 \left(\boldsymbol{H}_1^{\text{est}} + \Delta\boldsymbol{H}_{1,k}\right)^{\mathrm{T}} \right]_{(2)}^{\mathrm{T}} \right) \cdot \begin{bmatrix} \boldsymbol{I}_{M_1 \cdot M_{\mathrm{R}}} \\ \boldsymbol{0}_{M_2 \cdot M_{\mathrm{R}} \times M_1 \cdot M_{\mathrm{R}}} \end{bmatrix} \end{aligned} \quad \text{(A.15)}$$

$$\tilde{\boldsymbol{F}}_k^{(2)} = \boldsymbol{P}^{(1)}{}_{M_1, M_1+M_2, M_{\mathrm{R}}} \cdot \left(\boldsymbol{I}_{M_1+M_2} \otimes \left[\boldsymbol{G} \times_1 \left(\boldsymbol{H}_1^{\text{est}} + \Delta\boldsymbol{H}_{1,k}\right)^{\mathrm{T}} \right]_{(2)}^{\mathrm{T}} \right) \cdot \begin{bmatrix} \boldsymbol{0}_{M_1 \cdot M_{\mathrm{R}} \times M_2 \cdot M_{\mathrm{R}}} \\ \boldsymbol{I}_{M_2 \cdot M_{\mathrm{R}}} \end{bmatrix}. \quad \text{(A.16)}$$

Here, the operator $\mathrm{vec}\{.\}$ aligns all the elements of its operand into one vector and the notation $\boldsymbol{A} \otimes \boldsymbol{B}$ represents the Kronecker product. Moreover, the matrices $\boldsymbol{P}^{(n)}{}_{I,J,K}$ are the unique permutation matrices of size $I \cdot J \cdot K \times I \cdot J \cdot K$ defined via the relation $\boldsymbol{P}^{(n)}{}_{I,J,K} \cdot \mathrm{vec}\left\{\left[\boldsymbol{A}\right]_{(n)}\right\} = \mathrm{vec}\left\{\boldsymbol{A}\right\}$ for arbitrary tensors $\boldsymbol{A} \in \mathrm{C}^{I \times J \times K}$ [RH07].

The iteration is terminated when $\rho_k = \|\boldsymbol{R}_{k-1}\|_{\mathrm{H}} - \|\boldsymbol{R}_k\|_{\mathrm{H}} < \delta$, i.e., the norm of the residual tensor changes between iterations by less than a threshold parameter δ. Our simulations have shown that

$\delta = 10^{-3}$ represents a good trade-off between accuracy and computational complexity. We have also observed that between one and four iterations are usually sufficient.

A.5.1.2 Least Squares Khatri-Rao factorization scheme

Consider a matrix $\boldsymbol{C} \in \mathrm{C}^{M \cdot N \times P}$ which is an approximation of a Khatri-Rao product between a matrix $\boldsymbol{A} \in \mathrm{C}^{M \times P}$ and a matrix $\boldsymbol{B} \in \mathrm{C}^{N \times P}$, i.e., $\boldsymbol{C} \approx \boldsymbol{A} \lozenge \boldsymbol{B}$. Now we can find matrices $\boldsymbol{A}^{\mathrm{est}}$ and $\boldsymbol{B}^{\mathrm{est}}$ such that $\boldsymbol{A}^{\mathrm{est}} \lozenge \boldsymbol{B}^{\mathrm{est}}$ is the best approximation of $\boldsymbol{C}$ in the Frobenius norm sense using the following algorithm:

1. Step 1: Set $p = 1$.
2. Step 2: Let $\boldsymbol{c}_p$ be the p-th column of $\boldsymbol{C}$. Then, we can reshape $\boldsymbol{c}_p$ into a matrix $\tilde{\boldsymbol{C}}_p \in \mathrm{C}^{M \times N}$, such that $\mathrm{vec}\left\{\tilde{\boldsymbol{C}}_p\right\} = \boldsymbol{c}_p$.
3. Step 3: Compute the singular value decomposition of $\tilde{\boldsymbol{C}}_p$ as $\tilde{\boldsymbol{C}}_p = \boldsymbol{U}_p \cdot \boldsymbol{\Sigma}_p \cdot \boldsymbol{V}_p{}^{\mathrm{H}}$.
4. Step 4: Set $\boldsymbol{a}_p{}^{\mathrm{est}} = \sqrt{\sigma_1} \cdot \boldsymbol{v}_1^*$ and $\boldsymbol{b}_p{}^{\mathrm{est}} = \sqrt{\sigma_1} \cdot \boldsymbol{u}_1$, where $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ are the first columns of $\boldsymbol{U}_p$ and $\boldsymbol{V}_p$, respectively, and σ_1 is the largest singular value.
5. Step 5: If $p < P$ set $p = p + 1$ and go to step 2.
6. Step 6: The final estimates are given by $\boldsymbol{A}^{\mathrm{est}} = \left[\boldsymbol{a}_1{}^{\mathrm{est}}, \boldsymbol{a}_2{}^{\mathrm{est}}, \ldots, \boldsymbol{a}_P{}^{\mathrm{est}}\right]$ and $\boldsymbol{B}^{\mathrm{est}} = \left[\boldsymbol{b}_1{}^{\mathrm{est}}, \boldsymbol{b}_2{}^{\mathrm{est}}, \ldots, \boldsymbol{b}_P{}^{\mathrm{est}}\right]$.

A.5.1.3 Estimation of the scaling factors

Depending on the antenna configuration we have to treat two cases separately.

Case 1: $\min\{M_1, M_2\} \geq M_{\mathrm{R}}$

In this case, the vector $\boldsymbol{\lambda}$ can be estimated directly from the two LS Khatri-Rao factors $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ computed using the following steps:

1. Compute $\boldsymbol{L} = \left[\left(\left(\boldsymbol{X}_1{}^{\mathrm{T}}\right)^+ \cdot \boldsymbol{F}_2\right)^+ \cdot \boldsymbol{F}_1\right] \div \left(\boldsymbol{G}_2{}^{-1} \cdot \boldsymbol{G}_1\right)$, where $\div$ represents the inverse Schur (element wise division) operator.
2. Extract the symmetric (not Hermitian symmetric) part from $\boldsymbol{L}$ by defining $\tilde{\boldsymbol{L}} = \frac{1}{2}\left(\boldsymbol{L} + \boldsymbol{L}^{\mathrm{T}}\right)$.
3. Since $\tilde{\boldsymbol{L}}$ is symmetric, an SVD of $\tilde{\boldsymbol{L}}$ is given by $\tilde{\boldsymbol{L}} = \boldsymbol{U} \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{V}^{\mathrm{H}} = \boldsymbol{U} \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{U}^{\mathrm{T}}$.
4. Now, the estimate for $\boldsymbol{\lambda}$ is given by $\boldsymbol{\lambda}^{\mathrm{est}} = \sqrt{\sigma_1} \cdot \boldsymbol{u}_1$, where σ_1 is the largest singular value and $\boldsymbol{u}_1$ the corresponding singular vector (i.e. the first column of $\boldsymbol{U}$).

Case 2: $\min\{M_1, M_2\} < M_{\mathrm{R}}$

For this antenna configuration we require a different estimation scheme which takes advantage of the rank-one structure the matrix $\boldsymbol{L}$ should possess. Also, since we can design the relay amplification tensor, we choose the matrix $\tilde{\boldsymbol{G}} = \boldsymbol{G}_2{}^{-1} \cdot \boldsymbol{G}_1$ such that each column contains at most $\min\{M_1, M_2\}$ non-zero elements. We compute the matrix $\boldsymbol{L}$ in the following fashion:

1. Let $\boldsymbol{f}_{1,m}$ be the m-th column of $\boldsymbol{F}_1$ and $\tilde{\boldsymbol{g}}_m$ the m-th column of $\tilde{\boldsymbol{G}}$.
2. Then, compute $\boldsymbol{l}_m = \left[\left(\left(\boldsymbol{X}_1{}^{\mathrm{T}}\right)^+ \cdot \boldsymbol{F}_2\right)^+ \cdot \mathrm{diag}\{\tilde{\boldsymbol{g}}_m\}\right]^+ \cdot \boldsymbol{f}_{1,m}$ for $m = 1, 2, \ldots, M_{\mathrm{R}}$.

3. Collect the vectors $\boldsymbol{l}_m$ column wise into a matrix $\boldsymbol{L}$, i.e., $\boldsymbol{L} = \left[\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_{M_\mathrm{R}}\right]$.

4. This matrix contains zeros at the same position as $\tilde{\boldsymbol{G}}$. However, we know that it should be an estimate of the rank-one matrix $\boldsymbol{\lambda} \cdot \boldsymbol{\lambda}^\mathrm{T}$. Exploiting this structure, we can reconstruct the missing elements $l_{i,j}$ in $\boldsymbol{L}$ according to the following procedure:

- Set $l_{i,j} = l_{j,i}$ if the latter is known.
- If unknown elements are left, estimate the ratios $\rho_m = \lambda_m / \lambda_{m-1}$ for $m = 2, 3, \ldots, M_\mathrm{R}$ in the following manner:
  - For each row *m* find the indices *i* where $l_{m,i}$ and $l_{m-1,i}$ are known.
  - For each column *m* find the indices j where $l_{j,m}$ and $l_{j,m-1}$ are known.
  - Estimate ρ_m as the arithmetic average of $l_{m,i} / l_{m-1,i}$ $\forall i$ and $l_{j,m} / l_{j,m-1}$ $\forall j$.
- Apply the estimated ratios to fill the missing elements of the matrix:
  - If the element $l_{i,j-1}$ is known, an estimate of $l_{i,j}$ is given by $l_{i,j-1} \cdot \rho_m$.
  - If the element $l_{i-1,j}$ is known, an estimate of $l_{i,j}$ is given by $l_{i-1,j} \cdot \rho_m$.
  - If the element $l_{i,j+1}$ is known, an estimate of $l_{i,j}$ is given by $l_{i,j+1} / \rho_m$.
  - If the element $l_{i+1,j}$ is known, an estimate of $l_{i,j}$ is given by $l_{i+1,j} / \rho_m$.
  - If more than one estimate for $l_{i,j}$ is available, an arithmetic average of all estimates is computed.

Now that the missing elements in the matrix $\boldsymbol{L}$ are completed, we can follow the remaining steps of the previous case:

1. Extract the symmetric part from $\boldsymbol{L}$ by defining $\tilde{\boldsymbol{L}} = \frac{1}{2}\left(\boldsymbol{L} + \boldsymbol{L}^\mathrm{T}\right)$.
2. Since $\tilde{\boldsymbol{L}}$ is symmetric, an SVD of $\tilde{\boldsymbol{L}}$ is given by $\tilde{\boldsymbol{L}} = \boldsymbol{U} \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{U}^\mathrm{T}$, which can, for instance, be computed via a Takagi factorization.
3. Now, the estimate for $\boldsymbol{\lambda}$ is given by $\boldsymbol{\lambda}^{\mathrm{est}} = \sqrt{\sigma_1} \cdot \boldsymbol{u}_1$, where σ_1 is the largest singular value and $\boldsymbol{u}_1$ the corresponding singular vector.

A.5.1.4 Simulation results

In this paragraph we present two simulation results to depict the achievable channel estimation accuracy with the tensor-based channel estimation schemes. The results are shown in Figure A.4.

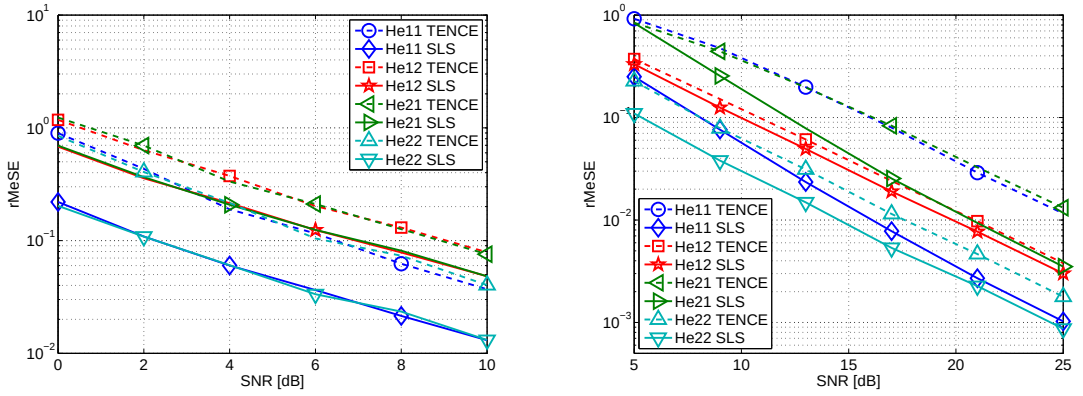


Figure A.4: Channel estimation error vs. the SNR for TENCE and the SLS-based refinement. Left: Uncorrelated Rayleigh fading scenario, right: mixed LOS/NLOS scenario.

We display the median of the channel estimation error for the channels $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ at terminal one and terminal two, where the median is computed with respect to random channel realizations. The curves labeled He*ij* depict the estimate of channel $\boldsymbol{H}_i$ at user terminal *j*. We depict both the initial solution obtained via TENCE with the dashed curves and the enhanced estimates based on SLS with the solid curves.

On the left-hand side we consider a scenario with uncorrelated Rayleigh fading channels in which both nodes and the relay are equipped with four antennas. We observe that especially the estimates of the own channels are particularly improved by the SLS-based technique whereas the improvement for the other channels is less pronounces.

The right-hand side displays a scenario where the first user has a Rician fading channel with a *K*-factor of 20 (corresponding to a relatively strong LOS component), whereas the second user has no line of sight (*K*=0). Also, both nodes and the relay are equipped with two antennas. The LOS component introduces a strong correlation into the first users' channel which deteriorates the performance of TENCE. The SLS-based technique can then significantly improve the estimation accuracy.

A.5.2 Algebraic Norm Maximizing (ANOMAX) transmit strategy

A.5.2.1 Description of ANOMAX

ANOMAX is a simple and yet efficient solution to choose the relay amplification matrix in a two-way relaying system with AF relays, where nodes and relay are equipped with multiple antennas. We start from the observations that by subtracting self-interference at the nodes, the two-way relaying system decouples into two parallel single-user MIMO channels with effective channel matrices $\boldsymbol{H}^{(e)}{}_{1,2} = \boldsymbol{H}_1^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_2$ and $\boldsymbol{H}^{(e)}{}_{2,1} = \boldsymbol{H}_2^{\mathrm{T}} \cdot \boldsymbol{G} \cdot \boldsymbol{H}_1$. ANOMAX tries to enhance the power of the received signal by maximizing the Frobenius norms of the effective channel matrices. The corresponding cost function is given by

$$J(\boldsymbol{G}) = \underset{\boldsymbol{G}, \|\boldsymbol{G}\|_{\mathrm{F}}=1}{\arg\max} \left[\beta^2 \cdot \left\| \boldsymbol{H}^{(e)}{}_{1,2} \right\|_2^2 + (1-\beta)^2 \cdot \left\| \boldsymbol{H}^{(e)}{}_{2,1} \right\|_2^2 \right], \tag{A.17}$$

where $\beta \in R_{[0,1]}$ is a weighting coefficient. This cost function can be solved algebraically. To this end, introduce the vector $\boldsymbol{g} = \mathrm{vec}\{\boldsymbol{G}\}$. Then, by a series of simple algebraic manipulations it can be shown that the cost function is identical to

$$J(\boldsymbol{g}) = \underset{\boldsymbol{g}, \|\boldsymbol{g}\|_2=1}{\arg\max} \left(\boldsymbol{K}_\beta{}^{\mathrm{T}} \cdot \boldsymbol{g} \right), \quad \boldsymbol{K}_\beta = \left[\beta \cdot (\boldsymbol{H}_2 \otimes \boldsymbol{H}_1), (1-\beta) \cdot (\boldsymbol{H}_1 \otimes \boldsymbol{H}_2) \right]. \tag{A.18}$$

Consequently, the maximizing $\boldsymbol{g}$ is given by $\boldsymbol{g} = \boldsymbol{u}_1^*$ where $\boldsymbol{u}_1$ is the dominant left singular vector of the matrix $\boldsymbol{K}_\beta$.

A.5.2.2 Simulation results

In this section we provide numerical simulation results to compare the performance of ANOMAX with other choices for the relay amplification matrix, namely the ZF/MMSE receivers from [UK09] and a fixed DFT matrix.

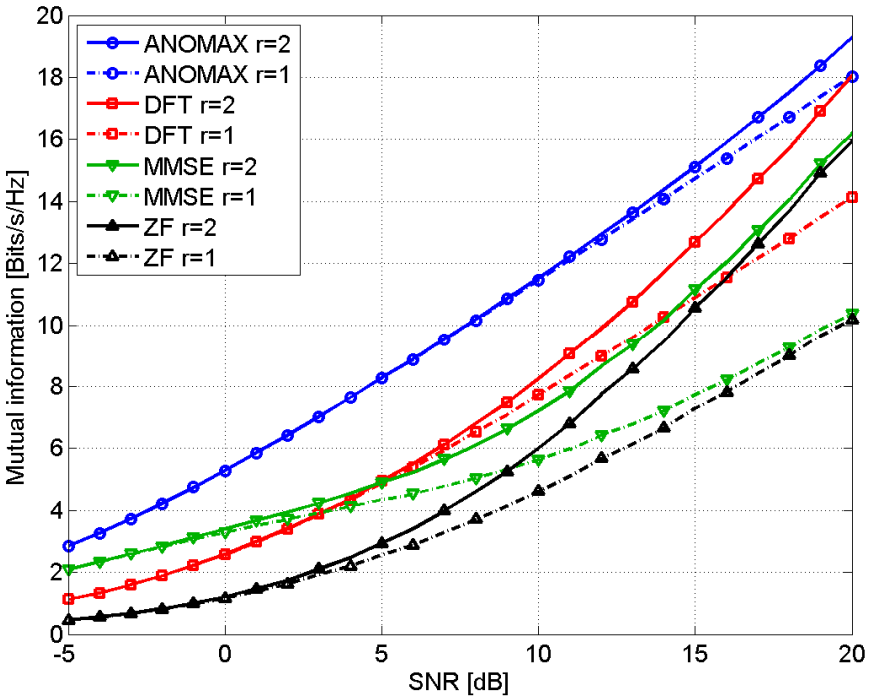


Figure A.5: Maximum mutual information for a scenario with $M_1 = M_2 = 2$ antennas at the UTs and $M_R = 5$ antennas at the relay. The dashed curves represent the case where only a single stream is used, the solid curves correspond to the case where both streams are active.

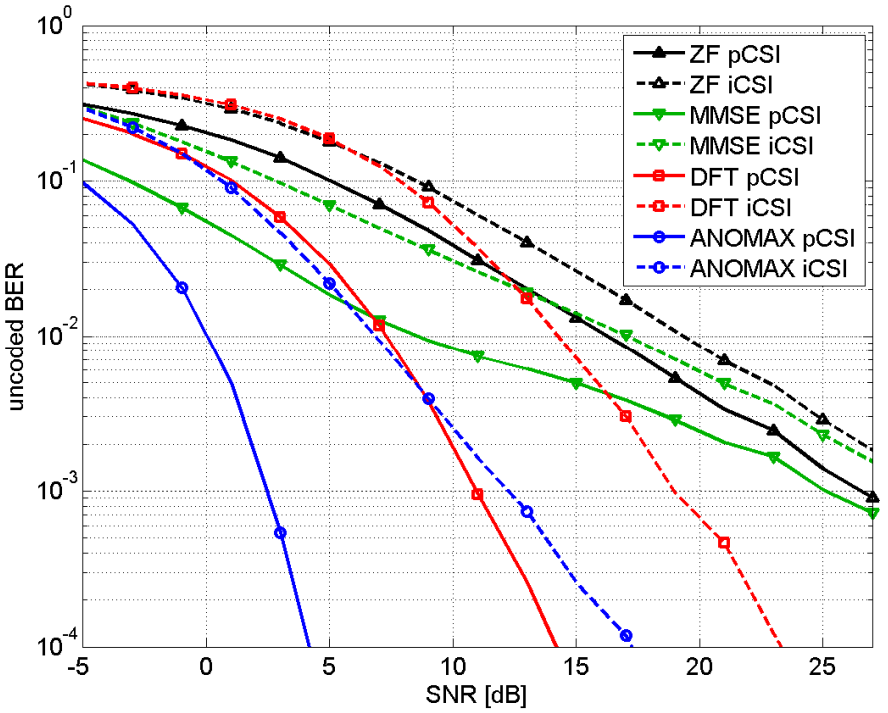


Figure A.6: Bit error rate of uncoded QPSK transmission in a scenario with $M_1 = M_2 = 2$ antennas at the UTs and $M_R = 5$ antennas at the relay. The solid curves represent the case where perfect CSI at the nodes and the relay is assumed, the dashed curves correspond to the case where all channels are estimated using TENCE and the SLS-based iterative refinement.

We consider uncorrelated Rayleigh fading with $M_1 = M_2 = 2$ antennas at the nodes and $M_R = 5$ antennas at the relay. In Figure A.5 we depict the maximum mutual information in Bits/s/Hz summed over the transmission from UT1 to UT2 and the transmission from UT2 to UT1. The dashed lines represent the

case where only a single stream is used, the solid curves correspond to the case where both streams can become active.

Figure A.6 displays the average bit error rate achieved via uncoded QPSK modulation and dominant eigenmode transmission. Here, the solid lines represent the case where the nodes and the relay have perfect CSI, whereas the dashed lines are based on channel estimates obtained via TENCE and the SLS-based iterative refinement of TENCE. In all cases, we observe that ANOMAX improves the SNR significantly and provides a good performance, even with estimates channels.

A.6 Joint channel estimation and decoding using Gaussian approximation in a factor graph

A.6.1 Distribution of the channel estimate for pilot-based estimation

L_p pilots $x_{p,k}$ ($0 \le k \le L_p$-1) are included in the transmitted sequence. From the L_p messages $\mu_{f_{p,k}\to\mathbf{G}}$ and (4.10) we get the discrete distribution of $\mathbf{G}$ [KFL01]:

$$p_p(\mathbf{G}) \propto \sum_{c=0}^{L^q-1} \delta(\mathbf{G}-\mathbf{G}_c) \prod_{k=0}^{L_p-1} \exp\left(-\frac{|y_{p,k}-\mathbf{X}_{p,k}^T\mathbf{G}_c|^2}{2\sigma_n^2}\right) \tag{A.19}$$

where:

$$y_{p,k} = \sum_{l=0}^{L-1} h_l\, x_{p,k-l} + n_{p,k} \tag{A.20}$$

$\mathbf{X}_{p,k} = (x_{p,k}, \ldots, x_{p,k-L+1}{}^T)$; $\mathbf{G}_c = (g_0{}^c, \ldots, g_{L-1}{}^c)^T$ is a quantization codebook of size L^q for channel estimate's probability density function (pdf) and $\delta(.)$ denotes the Dirac delta function. Using a constant amplitude zero autocorrelation (CAZAC) sequence as pilots [UBV01], (A.19) can be approximated as

$$p_p(\mathbf{G}) \propto \sum_{c=0}^{L^q-1} \delta(\mathbf{G}-\mathbf{G}_c) \exp\left(-\frac{L_p\mathcal{E}_p}{2\sigma_n^2}|\mathbf{G}_c-\mathbf{H}|^2\right) \tag{A.21}$$

where $\mathcal{E}_p$ represents the pilot energy. Hence, $p_p(\mathbf{G})$ can be approximated as one Gaussian distribution $\mathcal{CN}(\mathbf{H}, 2\sigma_n{}^2/L_p/\mathcal{E}_p)$.

A.6.2 Distribution of the channel estimate for data-based estimation

Let $\mathbf{S}_{i,m} = (s_{i,0,m}, s_{i,1,m}, \ldots, s_{i,L-1,m})^T$, where $0 \le m \le 2^L$-1, represent the m^{th} possible symbol vector and $\xi_{i-1,l}{}^m$ represent the probability that $x_{i-l} = s_{i,l,m}$. The product $\prod_{i=0;\, i\ne k}^{N-1} \mu_{f_i\to\mathbf{G}}$ can be expressed as $p_{d,k}(\mathbf{G})$ [KFL01]:

$$\begin{aligned} p_{d,k}(\mathbf{G}) &\propto \prod_{\substack{i=0\\ i\ne k}}^{N-1} \mu_{f_i\to\mathbf{G}} \propto \sum_{c=0}^{L^q-1} \delta(\mathbf{G}-\mathbf{G}_c) \times \prod_{\substack{i=0\\ i\ne k}}^{N-1} \left\{ \sum_{m=0}^{2^L-1} \exp\left(-\frac{|y_i-\mathbf{S}_{i,m}^T\mathbf{G}_c|^2}{2\sigma_n^2}\right) \prod_{l=0}^{L-1} \xi_{i-l,l}^m \right\} \\ &\propto \sum_{c=0}^{L^q-1} \delta(\mathbf{G}-\mathbf{G}_c) \times \sum_{j=0}^{2^{(N-1)L}-1} \exp\left(-\frac{1}{2\sigma_n^2}\sum_{\substack{i=0\\ i\ne k}}^{N-1}|y_i-\mathbf{S}_i^{jT}\mathbf{G}_c|^2\right) \prod_{\substack{i=0\\ i\ne k}}^{N-1}\prod_{l=0}^{L-1} \xi_{i-l,l}^j, \end{aligned} \tag{A.22}$$

where $\mathbf{S}_i{}^j = (s_{i,0}{}^j, s_{i,1}{}^j, \ldots, s_{i,L-1}{}^j)^T$ is the value of symbol $\mathbf{X}_i$ in sequence j and $\xi_{i-l,\,l}{}^j$ is the probability that x_{i-l} equals $s_{i,l}{}^j$. After some calculations, (A.22) can be approximated as

$$\begin{aligned} p_{d,k}(\mathbf{G}) \propto \sum_{c=0}^{L^q-1} \delta(\mathbf{G}-\mathbf{G}_c) \times \sum_{j=0}^{2^{(N-1)L}-1} &\left\{ \prod_{l=0}^{L-1} \exp\left\{ -\frac{(N-1)}{2\sigma_n^2}\left| g_l^c - \frac{\left(U_l^j - V_l^j\right)}{(N-1)} h_l \right|^2 \right\} \times \right. \\ &\left. \exp\left\{ -\frac{(N-1)|h_l|^2}{2\sigma_n^2}\left[1-\frac{\left(U_l^j-V_l^j\right)^2}{(N-1)^2}\right]\right\} \prod_{\substack{i=0\\ i\ne k}}^{N-1} \xi_{i-l,l}^j \right\}, \end{aligned} \tag{A.23}$$

where $U_l{}^j$ (resp. $V_l{}^j$) is the number of items with $s_{i,l}{}^j x_{i-l}{}^* = +1$ (resp. $s_{i,l}{}^j x_{i-l}{}^* = -1$) in sequence j.

1) With high SNR, the decoder almost provides perfect extrinsic information. Thus, for a single sequence j with $U_l{}^j = N$-1, all $\xi_{i-l,l}{}^j \to 1$ and other terms are null:

$$p_{\mathrm{d},k}(\mathbf{G}) \propto \sum_{c=0}^{L^q-1} \delta(\mathbf{G}-\mathbf{G}_c) \exp\left\{-\frac{(N-1)}{2\sigma_n^2}\left|\mathbf{G}_c-\mathbf{H}\right|^2\right\} \tag{A.24}$$

2) With low SNR, all probabilities $\xi_{i\text{-}l,l}{}^j$ are close to 1/2 for BPSK modulation. In (A.23), there are $2^{(N\text{-}1)L}$ items. However, when considering the first exponential item in (A.23), only the sequences with $(U_l{}^j - V_l{}^j)^2 \geq (N\text{-}1)^2$, i.e., with $U_l{}^j = N\text{-}1$ or $V_l{}^j = N\text{-}1$, are not close to zero. Therefore, there are only 2^L dominant terms:

$$p_{\mathrm{d},k}(\mathbf{G}) \propto \sum_{c=0}^{L^q-1} \delta(\mathbf{G}-\mathbf{G}_c) \prod_{l=0}^{L-1} \left\{\beta_l \exp\left\{-\frac{(N-1)}{2\sigma_n^2}\left|g_l^c-h_l\right|^2\right\} + (1-\beta_l)\exp\left\{-\frac{(N-1)}{2\sigma_n^2}\left|g_l^c+h_l\right|^2\right\}\right\} \tag{A.25}$$

where β_l denotes the normalized product of $\xi_{i\text{-}l,l}{}^j$; for a single channel tap, the distribution $p_{\mathrm{d},k}(g_l)$ is:

$$p_{\mathrm{d},k}(g_l) \propto \sum_{c=0}^{q-1} \delta(g_l-g_l^c) \left\{\beta_l \exp\left\{-\frac{(N-1)}{2\sigma_n^2}\left|g_l^c-h_l\right|^2\right\} + (1-\beta_l)\exp\left\{-\frac{(N-1)}{2\sigma_n^2}\left|g_l^c+h_l\right|^2\right\}\right\} \tag{A.26}$$

From (A.24), (A.25) and (A.26), for each channel tap, the pdf $p_{\mathrm{d},k}(g_l)$ can be approximated as a mixture of two Gaussian distributions. For the whole ISI channel, the pdf $p_{\mathrm{d},k}(\mathbf{G})$ can be approximated as a mixture of multiple Gaussian distributions which are the product of all pdfs of each tap with variance $2\sigma_n{}^2/(N\text{-}1)$.

A.6.3 APP computation

Using the discrete distribution, the probability of symbol vector $\mathbf{X}_k$ can be calculated as:

$$P(\mathbf{X}_k = \mathbf{S}_{k,m}) = \sum_{c=0}^{L^q-1} \exp\left\{-\frac{1}{2\sigma_n^2}\left|y_k - \mathbf{S}_{k,m}^{\mathrm{T}}\mathbf{G}_c\right|^2\right\} P_k(\mathbf{G}_c) \tag{A.27}$$

Equation (A.27) can be written in a continuous way as:

$$P(\mathbf{X}_k = \mathbf{S}_{k,m}) = \underbrace{\int \cdots \int}_{L} \exp\left\{-\frac{1}{2\sigma_n^2}\left|y_k - \sum_{l=0}^{L-1} s_{k,l,m} g_l\right|^2\right\} \prod_{l=0}^{L-1} p_k(g_l)\, dg_0 \cdots dg_{L-1} \tag{A.28}$$

After some calculations, we get

$$P(\mathbf{X}_k = \mathbf{S}_{k,m}) \propto \prod_{l=0}^{L-1} \frac{1}{\hat{\sigma}_h^2 \sum_{i=0}^{l} \left|s_{k,i,m}\right|^2 + \sigma_n^2} \exp\left\{-\frac{\left|y_k - \mathbf{S}_{k,m}^{\mathrm{T}}\hat{\mathbf{H}}_k\right|^2}{2\left(\hat{\sigma}_h^2 \sum_{l=0}^{L-1} \left|s_{k,l,m}\right|^2 + \sigma_n^2\right)}\right\} \tag{A.29}$$

where $\hat{\mathbf{H}}_k = \left(\hat{h}_{k,0}, \cdots, \hat{h}_{k,L-1}\right)^{\mathrm{T}}$. According to Figure 4.7 (e) and Figure 4.7 (f), we have

$$P(x_k = b_i) \propto \prod_{l=0}^{L-1} \sum_{s_{k+l,m,l}=b_i} P(\mathbf{X}_{k+l} = \mathbf{S}_{k+l,m}) \prod_{\substack{l'=0 \\ l' \neq l}}^{L-1} \mu_{x_{k+l-l'} \to f_{k+l}} \tag{A.30}$$

A.6.4 Computation of the mean and variance of the Gaussian distributions

Replacing the discrete distribution in (A.22) by an integral, we get a continuous distribution for $p_{\mathrm{d},k}(\mathbf{G})$:

$$p_{\mathrm{d},k}(\mathbf{G}) \propto \sum_{j=0}^{2^{(N-1)L}-1} \exp\left\{-\frac{1}{2\sigma_n^2} \sum_{\substack{i=0 \\ i \neq k}}^{N-1} \left|y_i - \mathbf{S}_i^{j\mathrm{T}}\mathbf{G}\right|^2\right\} \Delta_j \tag{A.31}$$

where $\Delta_j = \prod_{i=0;\, i\neq k}^{N-1} \prod_{l=0}^{L-1} \xi_{i-l,l}^{j}$. Thus, we can get the distribution of g_l from (A.31):

$$p_{\mathrm{d},k}(g_l) \propto \underbrace{\int \cdots \int}_{\mathbf{G}'} p_{\mathrm{d},k}(\mathbf{G})\, d\mathbf{G}' \tag{A.32}$$

where $\mathbf{G}' = (g_0,\ldots,g_{l-1},\ g_{l+1},\ldots,g_{L-1})$. After extensive calculation, the pdf $p_{\mathrm{d},k}(g_l)$ can be approximately written as

$$p_{\mathrm{d},k}(g_l) \propto \sum_{j=0}^{2^{(N-1)L}-1} \Delta_j \exp\left\{-\frac{1}{2\sigma_n^2}|g_l|^2\,\Omega_{k,l}^{j}\right\} \times \exp\left\{\frac{1}{2\sigma_n^2} 2\Re e\left\{g_l^{*}\Phi_{k,l}^{j} - g_l^{*}\sum_{\substack{l'=0\\ l'\neq l}}^{L-1}\frac{\Phi_{k,l'}^{j}}{\Omega_{k,l'}^{j}}\mathrm{R}_{k,l',l}^{j}\right\}\right\} \tag{A.33}$$

where

$$\Phi_{k,l}^{j} = \sum_{\substack{i=0\\ i\neq k}}^{N-1} y_i s_{i,l}^{j*},\ \Omega_{k,l}^{j} = \sum_{\substack{i=0\\ i\neq k}}^{N-1} \left|s_{i,l}^{j}\right|^2 \tag{A.34}$$

and

$$\mathrm{R}_{k,l',l}^{j} = \sum_{\substack{i=0\\ i\neq k}}^{N-1} s_{i,l'}^{j} s_{i,l}^{j*}\,. \tag{A.35}$$

Using (A.33) and considering normalization, we get

$$\hat{h}_{\mathrm{d},k,l} = \int_{g_l} g_l\, p_{\mathrm{d},k}(g_l)\, d\, g_l \approx \frac{1}{\Omega_{k,l}^{j_{\max}}}\left(\Phi_{k,l}^{j_{\max}} - \sum_{\substack{l'=0\\ l'\neq l}}^{L-1}\frac{\Phi_{k,l'}^{j_{\max}}}{\Omega_{k,l'}^{j_{\max}}}\mathrm{R}_{k,l',l}^{j_{\max}}\right) \tag{A.36}$$

where

$$j_{\max} = \arg\max_{j} \Delta_j \tag{A.37}$$

Following the same steps, the estimation with a CAZAC pilot sequence can be calculated as

$$\hat{h}_{\mathrm{p},l} \approx \frac{1}{L_{\mathrm{p}}\mathcal{E}_{\mathrm{p}}} \sum_{k=0}^{L_{\mathrm{p}}-1} y_{\mathrm{p},k} x_{\mathrm{p},k-l}^{*} \tag{A.38}$$

A.6.5 Simulation results

Some simulation results are shown in this section for an ISI Rayleigh channel with 3-tap rectangular impulse response. We use a half rate 64-state (133, 171) convolutional code and BPSK modulation. The pseudo-random interleaver size is 1000. There are 18 pilots with unit energy.

The proposed BP with continuous downward and upward messages (BP-DUGA) is compared with an iterative APP equalizer (APPEQ) and a MMSE equalizer (MMSEEQ), both using an expectation-maximisation (EM) channel estimation. From Figure A.7, we observe that the proposed BP-DUGA has a better BER performance with perfect channel state information (PerCSI) than the MMSE equalizer – about 1dB for 10^{-5} - where K=11 represents the number of complex-valued tap weight coefficients of the equalizer. With 5 iterations and actual estimation, it also outperforms MMSE+EM. From Figure A.8, we observe that BP-DUGA has a BER performance very close to that of the iterative APP equalizer with PerCSI and EM channel estimation. Using continuous downward and upward messages brings a complexity reduction compared to the quantization method.

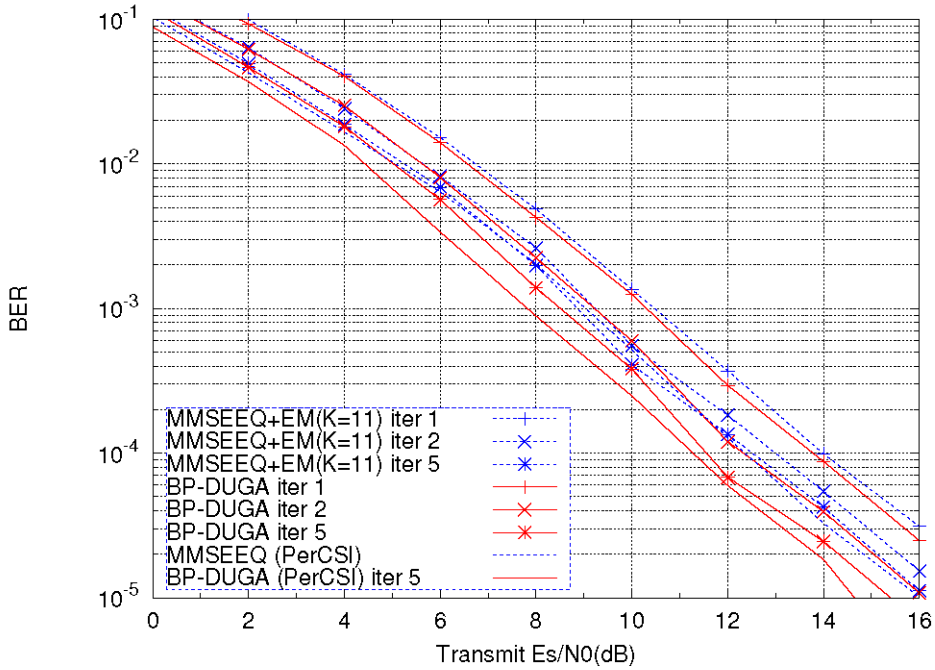


Figure A.7: Bit error rate performance comparison: BP-DUGA vs. MMSE Equalizer with EM channel estimation. Number of taps is 3. Rate 1/2 convolutional code (133, 171) with block length 1000. Number of pilots is 18.

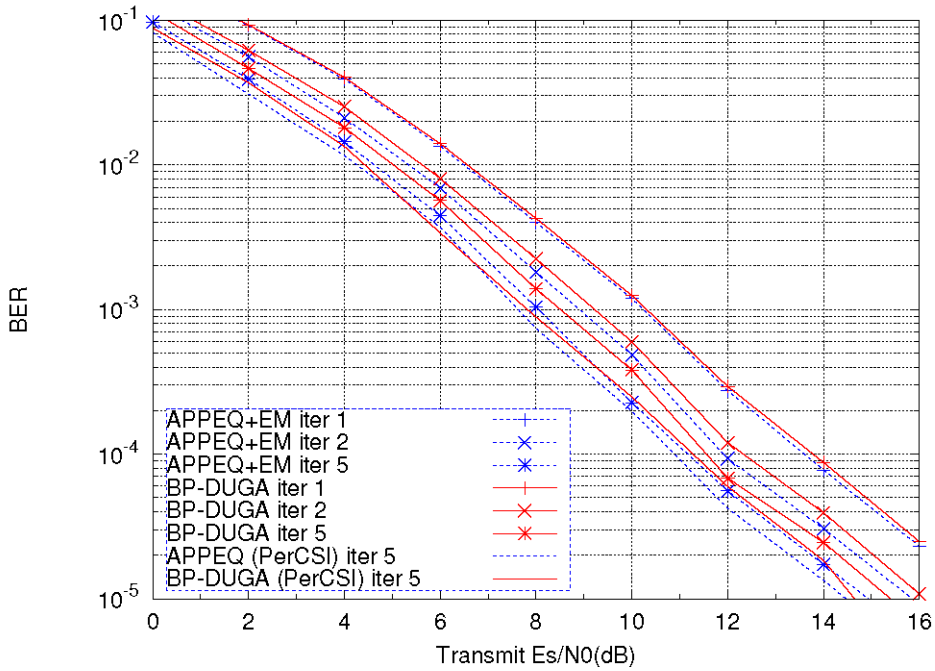


Figure A.8: Bit error rate performance comparison: BP-DUGA vs. APP Equalizer with EM channel estimation. Number of taps is 3. Rate 1/2 convolutional code (133, 171) with block length 1000. Number of pilots is 18.