Multiuser MIMO Techniques with Feedback

Jee Hyun Kim

ILMENAU UNIVERSITY OF TECHNOLOGY

Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität Ilmenau

Multiuser MIMO Techniques with Feedback

Jee Hyun Kim

Dissertation zur Erlangung des
akademischen Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing)

Anfertigung im:	Fachgebiet Nachrichtentechnik Institut für Informationstechnik Fakultät für Elektrotechnik und Informationstechnik
Gutachter:	Univ.-Prof. Dr.-Ing. Martin Haardt
	Univ.-Prof. Dr.-Ing. Anja Klein
	Univ.-Prof. Dr.-Ing. Giovanni Del Galdo
Vorgelegt am:	11.02.2019
Verteidigt am:	28.10.2020

urn:nbn:de:gbv:ilm1-2020000484

ACKNOWLEDGEMENTS

People say that life is full of surprises and some say that God is in the detail, and my journey to completion of this dissertation is one good example. This dissertation would not have been possible without many people to whom I am indebted.

First of all, I would like to express my sincere gratitude to my supervisor Prof. Martin Haardt. His consistent supervision encouraged me to enhance my work. His immense knowledge and deep understanding on the fundamental theory as well as their practical implication to the real world scenario never cease to amaze me. His intensive attention to the details helped me to take this dissertation to the next level. There were times when I thought it would not be possible to meet his high standard, I should admit, but I am very glad that I finally made it. Without his support and supervision this would not have been possible at all.

I would like to thank Wolfgang Zirwas at (now) Nokia Bell Labs in Munich. I first got to know him when I started working on a master thesis work under his supervision at (then) Siemens AG. When I decided to come back to Germany, after having briefly worked for two years in Korea, he generously offered me a doctoral thesis work position. During my time at Siemens and later Nokia Siemens Networks, he gave me an opportunity to explore my own ideas with full degree of freedom, which I believe is rare in the industry-funded program. He has been always open to share his expert knowledge and to provide his unique point of view. It was truly great pleasure working with him. I also would like to thank Prof. Anja Klein and Prof. Giovanni Del Galdo for kindly accepting the role as an examiner of this dissertation and providing me with valuable review comments.

I was lucky to have an opportunity to work under or with many creative and brilliant researchers, Dr. Egon Schulz, Dr. Thomas Haustein, Dr. Jesse (Jijun) Luo, Dr. Hanguang Wu, Dr. Ying Zhang, and Mr. Ruediger Halfmann, to name a few. I would like to thank Kwanjeong Educational Foundation and Siemens AG (which became Nokia Siemens Networks GmbH & Co. KG during the course of my study) for their financial support.

Last but not least, I would like to thank my parents for their unconditioned love. I also would

like to appreciate my loving wife Soo Hyun having supported me throughout this journey. I dedicate this dissertation to my children, i.e., Dong-Ha, Sua, Mina and Ina. They are reasons for my life.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF ABBREVIATIONS

2G 2nd Generation

3G 3rd Generation

3GPP 3rd Generation Partnership Project

4G 4th Generation

5G 5th Generation

aCB adaptive codebook

ACF autocorrelation function

ADM Adaptive Delta Modulation

AMC Adaptive Modulation and Coding

ARQ automatic repeat request

AVQ Adaptive Vector Quantization

AWGN Additive White Gaussian Noise

BD Block Diagonalization

BER Bit Error Rate

BFPR Binary Floating Point number Representation

BS Base Station

CA Cooperative Area

cCA centralized CA

cCB combined codebook

cdf cumulative distribution function

CDMA Code Division Multiple Access

CEAC Complementary Envelope Autocorrelation Coefficient

CO Cyclic Overflow

CoMP coordinated multi-point transmission/reception

COOPA cooperative antenna

CP Cyclic Prefix

CPU Central Processing Unit

CQ Channel Quantization

CQI Channel Quality Information

CSI Channel State Information

CU Central Unit

DAS Distributed Antenna System

dCA distributed CA

D-Givens differential Givens

DL downlink

DwPTS downlink pilot timeslot

EAC Envelope Autocorrelation Coefficient

EVD eigenvalue decomposition

FDD Frequency Division Duplex

FDMA Frequency Division Multiple Access

FFT Fast Fourier Transform

FPU Floating Point Unit

HOSVD Higher Order Singular Value Decomposition

HSPA High Speed Packet Access

ICI inter-cell interference

IEEE Institute of Electrical and Electronics Engineers

IFFT Inverse Fast Fourier Transform

i.i.d. independent and identically distributed

IoT Internet of Things

IP Internet Protocol

IPR Intellectual Property Right

IR Interference Removal

ISI inter-symbol interference

ITU International Telecommunication Union

JD Joint Detection

JT Joint Transmission

KPI Key Performance Indicator

LBG Linde-Buzo-Gray

LOS Line Of Sight

LTE Long Term Evolution

MAC Medium Access Control

MCS Modulation and Coding Scheme

MIMO Multiple-Input Multiple-Output

MISO Multiple-Input Single-Output

mLBG modified LBG

MMSE Minimum Mean Squared Error

MS Mobile Station

MSE mean squared error

MUI multi-user interference

MU-MIMO Multi-User Multiple-Input Multiple-Output

MUS MU-MIMO User Service

nACvF normalized autocovariance function

NLOS Non Line Of Sight

NPF non-prefix-free

OFDM Orthogonal Frequency Division Multiplexing

OFDMA Orthogonal Frequency Division Multiple Access

pCh perfect channel

PHY physical layer

PMI Precoding Matrix Indicator

PRB Physical Resource Block

QoS Quality of Service

rCB recursive codebook

RBD Regularized Block Diagonalization

RE Resource Element

RNC Radio Network Controller

RVQ Recursive Vector Quantization

SCM Spatial Channel Model

SCMe extended Spatial Channel Model

SDMA Space Division Multiple Access

SF Signal Factor

SINR Signal to Interference and Noise Ratio

SMMSE Successive Minimum Mean Squared Error

SNR Signal to Noise Ratio

SU-MIMO Single-User Multiple-Input Multiple-Output

SVD Singular Value Decomposition

TDD Time Division Duplex

TDMA Time Division Multiple Access

THP Tomlinson-Harashima Precoding

TM Tracking Mode

TS Transmission Sequence

TTI Transmit Time Interval

TxMF transmit matched filter

TxWF transmit Wiener filter

TxZF transmit zero-forcing filter

UL uplink

VQ Vector Quantization

WiMAX Worldwide Interoperability for Microwave Access

WINNER Wireless World Initiative New Radio

WWRF Wireless World Research Forum

ZF zero-forcing

ABSTRACT

Multiuser MIMO Techniques with Feedback

by

Jee Hyun Kim

Cooperative antenna systems have recently become a hot research topic, as they promise significantly higher spectral efficiency than conventional cellular systems. The gain is acquired by eliminating inter-cell interference (ICI) through coordination of the base antenna transmissions. Recently, distributed organization methods have been suggested. One of the main challenges of the distributed cooperative antenna system is channel estimation for the downlink channel especially when FDD is used. All of the associated base stations in the cooperative area need to know the full channel state information to calculate the corresponding precoding weight matrix. This information has to be transferred from mobile stations to base stations by using uplink resources. As several base stations and several mobile stations are involved in cooperative antenna systems and each base station and mobile station may be equipped with multiple antennas, the number of channel state parameters to be fed back is expected to be big. In this thesis, efficient feedback techniques of the downlink channel state information are proposed for the multi-user multiple-input multiple-output case, targeting distributed cooperative antenna systems in particular.

First, a subspace based channel quantization method is proposed which employs a predefined codebook. An iterative codebook design algorithm is proposed which converges to a local optimum codebook. Furthermore, feedback overhead reduction methods are devised exploiting temporal correlation of the channel. It is shown that the proposed adaptive codebook method in conjunction with a data compression scheme achieves a performance close to the perfect channel case, requiring much less feedback overhead compared with other techniques.

The subspace based channel quantization method is extended by introducing multiple antennas at the transmitter side and/or at the receiver side and the performance of a regularized block diagonalization (RBD) precoding(/decoding) scheme has been investigated as well as a zero-forcing (ZF) precoding scheme. A cost-efficient decoding matrix quantization method is proposed which can avoid a complex computation at the mobile station while showing only a slight degradation.

The thesis is concluded by comparing the proposed feedback methods in terms of their performance, their required feedback overhead, and their computational complexity. The techniques that are developed in this thesis can be useful and applicable for 5G, which is envisioned to support the high granularity/resolution codebook and its efficient deployment schemes.

keywords: MU-MIMO, COOPA, limited feedback, CSI, CQ, feedback overhead reduction, Givens rotation

ZUSAMMENFASSUNG

Multiuser MIMO Techniques with Feedback

von

Jee Hyun Kim

Kooperative Antennenanlagen haben vor kurzem einen heißen Forschungsthema geworden, da Sie deutlich höhere spektrale Effizienz als herkömmliche zelluläre Systeme versprechen. Der Gewinn wird durch die Eliminierung von Inter-Zelle Störungen (ICI) durch Koordinierung der-Antenne Übertragungen erworben. Vor kurzem, verteilte Organisation Methoden vorgeschlagen. Eine der größten Herausforderungen für das Dezentrale kooperative Antennensystem ist Kanalschätzung für den Downlink Kanal besonders wenn FDD verwendet wird. Alle zugehörigen Basisstationen im genossenschaftlichen Bereich müssen die vollständige Kanal Informationen zu Wissen, die entsprechenden precoding Gewicht Matrix zu berechnen. Diese Information ist von mobilen Stationen übertragen werden Stationen mit Uplink Ressourcen zu stützen. Wird als mehrere Basisstationen und mehreren mobilen Stationen in kooperativen Antennensysteme und jede Basisstation und Mobilstation beteiligt sind, können mit mehreren Antennen ausgestattet sein, die Anzahl der Kanal Parameter wieder gefüttert werden erwartet, groß zu sein. In dieser Arbeit wird ein effizientes Feedback Techniken der downlink Kanal Informationen sind für die Multi-user Multiple Input Multiple Output Fall vorgeschlagen, der insbesondere auf verteilte kooperative Antennensysteme zielt.

Zuerst wird ein Unterraum-basiertes Kanalquantisierungsverfahren vorgeschlagen, das ein vorbestimmtes Codebuch verwendet. Ein iterativer Codebuchentwurfsalgorithmus wird vorgeschlagen, der zu einem lokalen optimalen Codebuch konvergiert. Darüber hinaus werden Feedback-Overhead-Reduktionsverfahren entwickelt, die die zeitliche Korrelation des Kanals ausnutzen. Es wird gezeigt, dass das vorgeschlagene adaptive Codebuchverfahren in Verbindung mit einem Datenkomprimierungsschema eine Leistung nahe an dem perfekten Kanalfall erzielt, was viel weniger Rückkopplungsoverhead im Vergleich zu anderen Techniken erfordert.

Das auf dem Unterraum basierende Kanalquantisierungsverfahren wird erweitert, indem mehrere Antennen auf der Senderseite und/oder auf der Empfängerseite eingeführt werden, und die Leistung eines Vorcodierungs- (/Decodierungs-) Schemas mit regulierter Blockdiagonalisierung (RBD) wurde

untersucht. Es wird ein kosteneffizientes Decodierungsmatrixquantisierungsverfahren vorgeschlagen, das eine komplexe Berechnung an der Mobilstation vermeiden kann, während es nur eine leichte Verschlechterung zeigt.

Die Arbeit wird abgeschlossen, indem die vorgeschlagenen Feedback-Methoden hinsichtlich ihrer Leistung, ihres erforderlichen Feedback-Overheads und ihrer Rechenkomplexität verglichen werden.

CHAPTER I

Introduction

The growing demand for mobile broadband has been finally materialized. The volume of high speed mobile data traffic has increased explosively since the deployment of High Speed Packet Access (HSPA) networks. By mid-2007 mobile broadband traffic exceeded voice traffic in some HSPA coverage areas. Mobile broadband subscriptions[1] are growing by around 20 percent year-by-year, increasing by approximately 140 million in Q1 2016 alone. Data traffic grew around 10 percent quarter-by-quarter and 60 percent year-by-year between Q1 2015 and Q1 2016 (*Ericsson Mobility Report*, 2016). According to Ericsson, global total mobile data traffic reached around 33EB (exabyte) per month by the end of 2019, and is projected to grow by a factor close to 5 to reach 164EB per month in 2025 (*Ericsson Mobility Report*, 2020). Therefore, the spectral efficiency of future wireless network systems needs to be further improved to allow increased flexibility to serve a large number of simultaneous users and different services.

This thesis focuses on developing efficient feedback techniques for future cellular networks beyond the existing 3rd Generation (3G)/4th Generation (4G) systems with advanced antenna technologies. Special focus is put on the design of Channel State Information (CSI) feedback schemes in cooperative antenna systems where signal processing can be performed across distributed base stations which form a common Cooperative Area (CA). In Section 1.1, general requirements for future mobile communication systems are presented. Section 1.2 introduces the primary radio techniques that should be considered when designing future radio systems. Section 1.4 describes the concept of a cooperative antenna system which is also known as Distributed Antenna System (DAS) or coordinated network systems. Section 1.5 briefly explains the limited feedback techniques developed for multiple antenna systems. Section 1.6 presents the scope and objectives of the thesis. Finally, Section 1.7 describes the contributions of the author and provides the outline of the thesis.

[1]Mobile broadband is defined as HSPA, LTE, CDMA2000 EV-DO, TD-SCDMA and Mobile WiMAX.

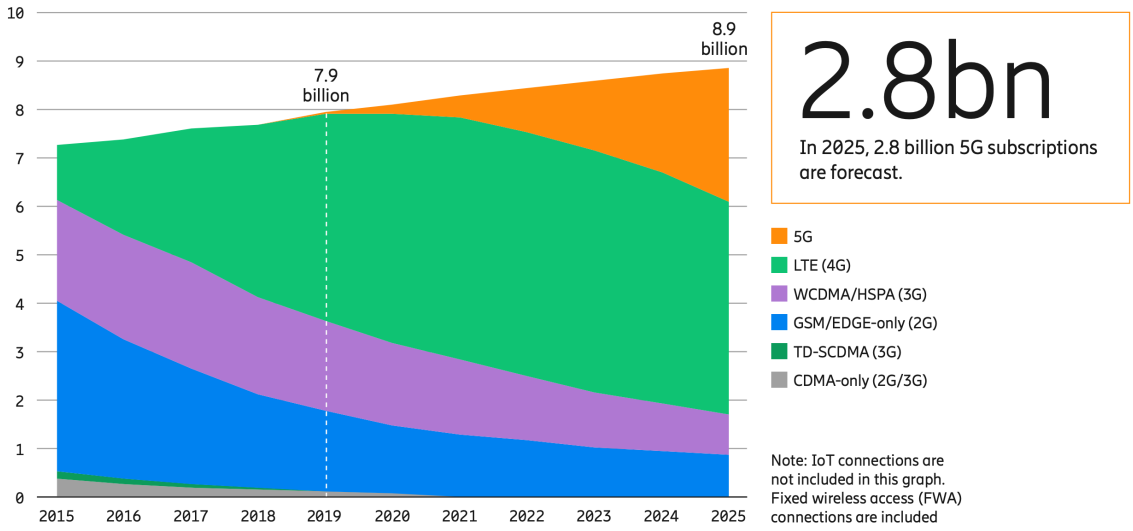


Figure 1.1: Mobile service subscription forecast (Source: Ericsson)

1.1 Requirements of Future Mobile Communication Systems

The 3rd Generation Partnership Project (3GPP) has taken a step towards higher data rates through the introduction of HSPA for 3G systems (*Holma and Toskala*, 2006). The 3GPP has been also carrying out a study that focuses on the long term evolution (LTE) of 3G (*3GPP*, 2008), which aims at an evolved radio access technology that is capable of providing service performance comparable to that of current fixed line accesses (*Dahlman et al.*, 2006), e.g., downlink (DL) peak data rate of 150 Mbps, uplink (UL) peak data rate of 50 Mbps, both in 20 MHz bandwidth. Alternative evolution paths towards higher data rates may involve wireless local- or metropolitan-area network (WLAN, WMAN) (*ANSI/IEEE*, 2003; *IEEE*, 2004) type of solutions. They are supposed to provide increased capacity and coverage as well as improved support for mobility and Quality of Service (QoS). IEEE 802.16e Worldwide Interoperability for Microwave Access (WiMAX) standard (*IEEE*, 2005) is one of such solutions. Major features of the requirements of these standards are high data transmission rate in both DL and UL, short latency time, flat network hierarchy, etc[2].

In parallel with the activities related to 3G evolution, there is also an increased research effort on future radio access beyond 3G, generally referred to as 4G radio access. The beyond 3G project is aimed at providing a ubiquitous system, fully based on the Internet Protocol (IP), where a wide range of services can be offered at a reasonable cost with QoS comparable to wireline technologies.

[2]As of this writing, 3GPP-based LTE seems to be the winner of this race for the next generation mobile broadband networks. Please take a look at Fig. 1.1 (*Ericsson Mobility Report*, 2020). The first commercial deployment of LTE took place in late 2009, which has been followed by a rapid and worldwide network deployment.

A 4G system is anticipated to provide peak data rates of up to 100 Mbps with wide-area coverage and up to 1 Gbps with local-area coverage, thus fulfilling the requirements set by the International Telecommunication Union (ITU) for beyond 3G systems (*Dahlman et al.*, 2006; *Astely et al.*, 2006). A large part of European research activities on 4G have been put into the Wireless World Initiative New Radio (WINNER) project (*Mohr*, 2007), while other independent research activities have been carried out, e.g., in the Wireless World Research Forum (WWRF)[3].

Taking one step ahead, the industry is already well on the road towards beyond 4G mobile communication systems, i.e., the 5th Generation (5G). A 5G system not only envisions enhancing the traditional mobile broadband use case, but also aims to support a significantly wider range of use cases (*Dahlman et al.*, 2016). A 5G system should be considered as a platform enabling wireless connectivity to all kinds of services. Connectivity will be provided essentially anywhere, anytime to anyone and anything. One of the major non-traditional use cases is a massive machine-type communication, i.e., in the Internet of Things (IoT), which motivates the need for new design concepts for the radio interface and the network architecture (*Schulz et al.*, 2017). The wide range of requirements for 5G system calls for a high degree of network flexibility/automation[4] as well (*Mwanje et al.*, 2016a). Another interesting use case for 5G is a device-to-device (D2D) communication (*M. Soleymani et al.*, 2016; *Mueckenheim et al.*, 2016).

1.2 System Design: Radio Access Technologies

Wireless cellular communication systems suffer from performance degradation which is mainly caused by the multi-user interference (*Sklar*, 1997). Apart from external interference and noise, signals are corrupted by time-varying fading, which can be further divided into large- and small-scale fading (*Sklar*, 1997; *Goldsmith*, 2005; *Proakis*, 2000; *Tse and Viswanath*, 2005). Large-scale fading represents the change of average path loss due to mobility over large geographical areas and it is usually modeled with an experimental model (*Goldsmith*, 2005; *Hata*, 1980). Small-scale fading, on the other hand, refers to the dramatic changes in signal amplitude and phase caused by small changes (an order of half a wavelength) in the distance between the transmitter and the receiver. The transmitted signal may propagate over multiple reflective paths which cause fluctuations in the amplitude, phase, and angle of arrival of the received signal. This phenomenon is generally referred

[3]Interested readers can find further information at http://www.wireless-world-research.org/.

[4]As a fully autonomous approach to self-organizing network, Prof. Mitschele-Thiel's group has proposed to introduce a machine learning solution to make the system adaptable to different network and operation scenarios and states (*Mwanje et al.*, 2016b; *Mwanje and Mitschele-Thiel*, 2014; *Sallakh et al.*, 2014; *Mwanje and Mitschele-Thiel*, 2015).

to as "multipath" fading causing frequency selective fading (*Sklar*, 1997; *Proakis*, 2000; *Biglieri et al.*, 1998).

The inter-symbol interference (ISI) imposed by multipath fading can be efficiently mitigated, e.g., via the use of an equalizer, as is done in the existing 2nd Generation (2G) and 3G cellular systems (*Juntti et al.*, 2007; *Redl et al.*, 1998). Orthogonal Frequency Division Multiplexing (OFDM) is an ISI resilient transmission method in which the transmitted data is modulated via an Inverse Fast Fourier Transform (IFFT) module and demodulated at the receiver with an Fast Fourier Transform (FFT) module. Together with Cyclic Prefix (CP) insertion, the OFDM modem transforms the frequency selective fading channel into a set of multiple flat fading orthogonal subchannels and as a result eliminates the ISI between subsequent data blocks as long as the CP is chosen to be longer than the maximum delay spread of the channel (*Goldsmith*, 2005; *Bingham*, 1990; *Wang and Giannakis*, 2000). This significantly simplifies the channel equalization at the receiver in comparison with conventional single-carrier modulation, as the channel at one subcarrier can be characterized by one tapped delay line model. Thus, OFDM has developed into a popular scheme for both existing and future broadband communication systems including WiMAX and Long Term Evolution (LTE) (*3GPP*, 2008; *Dahlman et al.*, 2006; *ANSI/IEEE*, 2003; *IEEE*, 2004; *Mohr*, 2007).

The main limitation for wireless communication systems is the fact that the limited resources, e.g., bandwidth, must be shared between multiple users (*Goldsmith*, 2005; *Tse and Viswanath*, 2005). Traditionally, the scarce radio resources have been shared among users in an orthogonal manner in time, frequency and/or code domains, using time, frequency or code division multiple access (TDMA, FDMA or CDMA), respectively (*Goldsmith*, 2005; *Tse and Viswanath*, 2005). The multicarrier transmission can be also combined with any type of multiple access to provide separation of multiple users. Orthogonal Frequency Division Multiple Access (OFDMA) combines Frequency Division Multiple Access (FDMA) with OFDM by assigning different numbers of orthogonal subchannels to different users, thus providing flexibility to allocate differentiated QoS to different users.

When developing future radio systems, we have to consider the other principal radio techniques like Multiple-Input Multiple-Output (MIMO) communication based on multiple antennas both at the transmitters and the receivers (*Foschini and Gans*, 1998; *Paulraj et al.*, 2004; *Raleigh and Cioffi*, 1998; *Sampath et al.*, 2002; *Telatar*, 1999) and Adaptive Modulation and Coding (AMC) (*Hanzo et al.*, 2002) as well. The spectral efficiency of MIMO transmission can be dramatically increased in the multiple user case if some level of CSI is available at the transmitter (*Tse and Viswanath*, 2005; *Raleigh and Cioffi*, 1998; *Telatar*, 1999), allowing the system to effectively adapt to and take

advantage of the available spectrum and the radio channel. The main challenge is to make the CSI available at the transmitter. Unlike Time Division Duplex (TDD) based systems in which the DL channel information can be estimated by looking into the UL channel, in case of Frequency Division Duplex (FDD) based systems the UL channel cannot play any significant role in DL CSI estimation due to the lack of channel reciprocity. In case of FDD, the task of delivering CSI to the transmitter can be achieved by conveying CSI as feedback information over the uplink. However, providing full CSI via feedback may cause an excessive overhead on the uplink, and hence it is difficult to implement in practice (*Tse and Viswanath*, 2005). Castañeda et al. have analyzed the achievable pairs of uplink and downlink rates considering the CSI overhead on the UL (*Castañeda et al.*, 2009).

A TDD system uses the same carrier frequency alternately for transmission and reception, and thus the CSI can be tracked at the transmitter provided that fading is sufficiently slow. In a TDD system with adaptive MIMO transmission, the modulation parameters in the communication link from Base Station (BS) to Mobile Station (MS) and vice versa can be adapted according to the channel conditions. Due to channel reciprocity, the DL channel can be estimated accurately during the previous UL frame. In order to provide channel estimates for DL transmission in a timely manner, the served users should have at least some minimum signalling (reference signals) in the UL direction. However the signalling consumes valuable uplink resources (capacity, battery power, etc.), especially if the user equipments have no data to transmit[5]. On the other hand, with CSI at transmitter, a large part of the signal processing load can be shifted from the MS to the BS both in UL and DL. This allows for a simple terminal design which is important as the terminal is limited in its computational capability and power due to its small form factor and portability nature.

Even though the CSI can be made available at the transmitter via time duplexing, the actual interference structure (interference level, frequency, time and space selectivity) observed by the multi-antenna terminal receiver in the DL can be very different from that measured by the BS receiver in the UL. In other words, the interference in the uplink and in the downlink is usually non-reciprocal, unlike the communication channel itself. In such a case, the obtained QoS at the receiver may differ significantly from the one desired if the transmission parameters are assigned on the basis of the uplink measurements only. In theoretical studies, the other-cell interference seen at the receiver is often conveniently assumed to be perfectly known at the transmitter in addition to the perfect channel knowledge both at the transmitter and the receiver (*Ye and Blum*, 2003). However, the amount of the required feedback would make the ideal feedback approach unpractical in most applications. For example, in practical adaptive MIMO OFDM cellular systems, the ideal

[5]The terminals, however, will send at least some acknowledgement packets in response to the downlink traffic.

feedback would require the interference covariance matrix to be reported to the transmitter for each subcarrier and for each transmitted data frame.

The particular challenge of the DL in case of Multi-User Multiple-Input Multiple-Output (MU-MIMO) systems is that while the BS has the ability to coordinate transmission from all of its antennas, the receiver antennas are grouped among different users that are typically unable to coordinate the transmission with each other (*Spencer et al.*, 2004a). While dirty paper coding is known to be a capacity achieving albeit very complex non-linear precoding technique in the DL (*Caire and Shamai*, 2003; *Vishwanath et al.*, 2003; *Vishwanath and Tse*, 2003; *Weingarten et al.*, 2006; *Yu and Cioffi*, 2004), linear precoding/beamforming (*Godara*, 1997b,a) is much simpler to implement to handle multi-user transmission. Hence, the linear beamforming gains popularity in practical system design (*Spencer et al.*, 2004a; *Alexiou and Haardt*, 2004). The resource allocation problem of the systems adopting a linear transmission scheme leads to a solution which can be divided into a two step approach, i.e., user selection or grouping for each orthogonal dimension (frequency, subcarrier, time, etc.) and the linear transceiver optimization for the selected set of users per orthogonal dimension subject to a power constraint. The optimal user selection per each orthogonal dimension is generally a difficult problem with integer constraints (*Sharif and Hassibi*, 2007, 2005; *Yoo and Goldsmith*, 2006). Therefore, sub-optimal allocation algorithms are commonly used in practice.

1.3 Multi-User Linear Precoding Methods

Several linear transmit processing techniques for MU-MIMO systems are summarized in this section (*Spencer et al.*, 2004a).

First, let $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ denote a MIMO channel with n_T transmitters and n_R receivers. x_j and $\mathbf{x}$ denote the signal or symbol transmitted from the jth transmitter and n_T dimensional vector which is a collection of x_j, respectively. The vector of received data $\mathbf{y}$ can be formulated as follows.

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

where $\mathbf{n}$ is an additive noise vector. In case of linear transmit processing, the transmitted signal $\mathbf{x}$ is generated by a linear combination of data symbols $\mathbf{s}$: $\mathbf{x} = \mathbf{W}\mathbf{s}$.

There are various approaches in designing $\mathbf{W}$. A transmit zero-forcing filter (TxZF) method imposes the constraint that all interference terms be zero. This can be accomplished by calculating

the pseudoinverse of the channel matrix: $\mathbf{W} = \mathbf{H}^{\dagger} = \mathbf{H}^{\mathcal{H}}(\mathbf{H}\mathbf{H}^{\mathcal{H}})^{-1}$ when $n_R \leq n_T$. TxZF is beneficial for low noise or high power situations, but it does not result in the linear capacity growth with respect to $\min(n_T, n_R)$ due to the stringent zero-interference requirements. An ill conditioned channel matrix will require a large normalization factor, which can lead to the dramatic reduction of SNR at the receivers.

A transmit Wiener filter (TxWF) method allows a limited amount of interference at each receiver and results in higher capacity for a given transmit power level. It adopts a Minimum Mean Squared Error (MMSE) approach: $\mathbf{W} = \mathbf{H}^{\mathcal{H}}(\mathbf{H}\mathbf{H}^{\mathcal{H}} + \alpha\mathbf{I})^{-1}$ where α is the loading factor. TxWF is shown to maximize the Signal to Interference and Noise Ratio (SINR) at the receiver, and it also results in the linear capacity growth with respect to $\min(n_T, n_R)$.

A Block Diagonalization (BD) method is developed for the case in which there are multiple antennas at the receiver (*Spencer et al.*, 2004b). It is generalization of channel inversion which optimizes the power transfer to a group of antennas of the same user rather than an individual single antenna in an effort to eliminate multi-user interference (MUI). BD outperforms TxZF, but the zero MUI constraint can lead to a significant capacity loss when the users' subspaces largely overlap.

A Regularized Block Diagonalization (RBD) method is designed such that each user transmits on the eigenmodes of the combined channel of all other users but with the power which is inversely proportional to the regularized eigenvalues of this matrix (*Stankovic and Haardt*, 2008). The system performance at low SNRs is limited by noise whereas it is limited by MUI at high SNRs. Accordingly, at low SNRs the user transmits in the null and a part of the signal subspace of all other users, balancing the MUI in order to better use the available subspace. At high SNRs the user transmits only in the null subspace of all other users to minimize MUI. As a result, RBD has the same capacity as Successive Minimum Mean Squared Error (SMMSE) at low SNRs and as BD at high SNRs. For high data rates, RBD extracts the full diversity in the system.

1.4 Cooperative Antenna (COOPA) Systems with MIMO-OFDM

Cooperative antenna systems, which are based on cooperative processing across distributed base stations, have recently gained a lot of interest, as they promise significantly higher performance compared to conventional cellular mobile systems (*Yu et al.*, 2005; *Zhang et al.*, 2005; *Aktas et al.*, 2006; *Choi and Andrews*, 2007; *Dawod et al.*, 2006; *Karakayali et al.*, 2006; *Wyner*, 1994; *Zhang and Dai*, 2004; *Hu et al.*, 2007; *Zhang et al.*, 2009; *Somekh et al.*, 2006; *Furht and Ahson*, 2016).

Cooperative antenna systems are MU-MIMO systems, where several adjacent base stations are tightly coupled over the backbone network. The assumption is that cooperative signal processing can be performed in a centralized manner over adjacent BSs so that the MIMO antennas are distributed over a larger geographical area (e.g. hundreds of meters), as illustrated in Fig. 1.2. The new logical area in which the MSs are served by multiple associated BSs is named *Cooperative Area*[6].

In conventional cellular systems, each base station transmits signals intended for users within its cell coverage. Depending on the users' channel conditions, interference caused by the neighboring cell transmissions can sharply degrade the received signal quality. Thus, the downlink capacity of cellular wireless networks is limited by inter-cell interference (ICI). (Refer to the first figure in Fig. 1.3[7]. Each individual color represents a signal useful for a given MS). As depicted in the figure, each MS receives the signal (solid line in its own identification color) as well as the interferences intended for other MSs (dotted line in other colors). Fortunately, since the base stations can be connected via a high-speed backbone, there is an opportunity to coordinate the base antenna transmissions so as to minimize the inter-cell interference, and hence to increase the downlink system capacity. The objective of cooperative antenna or *network coordination* is to enable cooperation between the base stations so that useful signals, as opposed to the interference, can be received from the neighboring base antennas. See the second figure in Fig. 1.3. Now the interferences turn into the useful signals.

The great advantage of using multiple antennas is that, without increasing power or bandwidth, the capacity of a point-to-point link scales linearly with the minimum of the number of transmit or receive antennas deployed. The gain, in terms of the marginal increase in rate when an additional antenna is deployed, is especially large when the Signal to Noise Ratio (SNR) is high (*Foschini et al.*, 2003). It has been shown that, without coordination, the link qualities can be very poor because of ICI (*Karakayali et al.*, 2006). In this case, the network does not benefit significantly from multiple antennas. When the coordination is employed, the inter-cell interference is mitigated so that the links can operate in the high SNR ratio regime. This enables the cellular network to enjoy the great spectral efficiency improvement associated with using multiple antennas.

The benefits of cooperative antenna systems are summarized as follows:

- The system exploits the freely available spatial dimension as it is well known for MIMO systems. In case of spatial multiplexing the capacity can be enhanced by factors compared to that of a

[6] Fig. 1.2 depicts the example in which three sectors form a cooperative area, which results in the size of a cooperative area being equal to that of a cell (three sectors). In general, a cooperative area can be extended to cover multiple cells, depending on the number of cooperating BSs.

[7] The graphical representation of the concept of the cooperative antenna system in comparison with the conventional cellular network is inspired by the figure introduced in (*Karakayali et al.*, 2006).

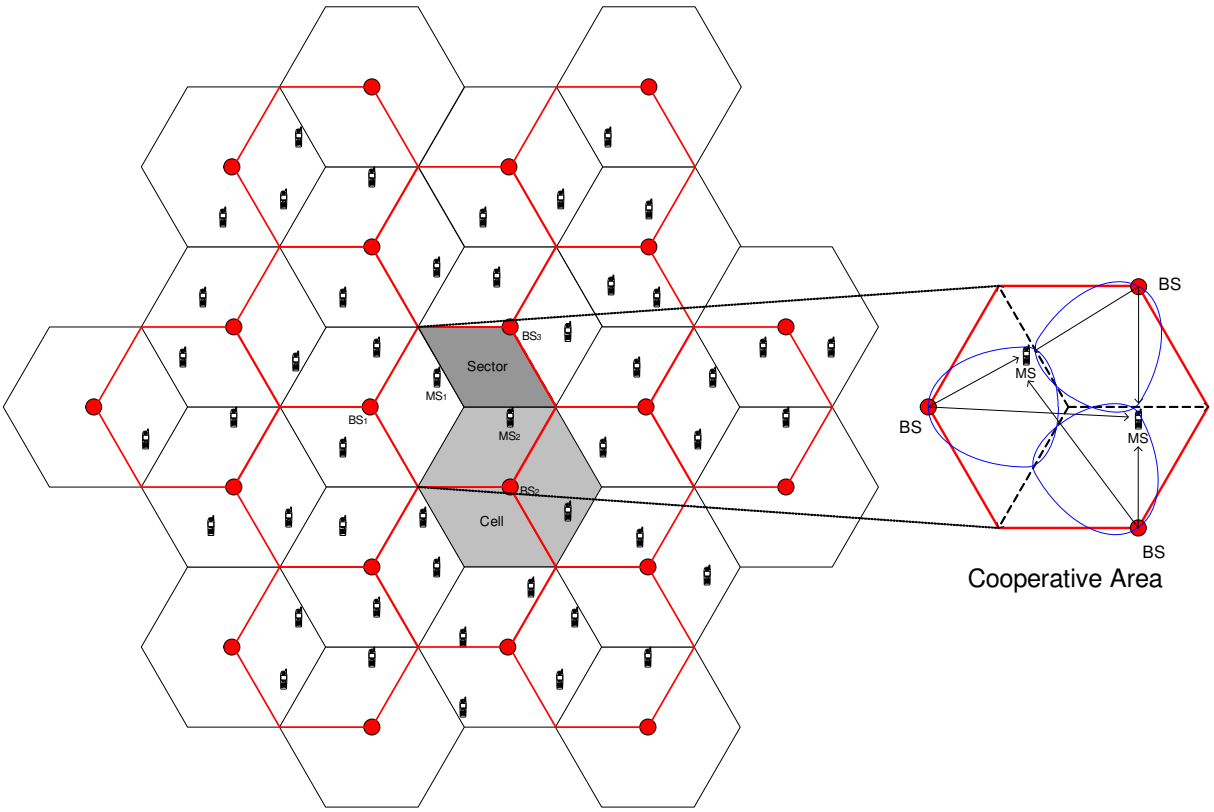
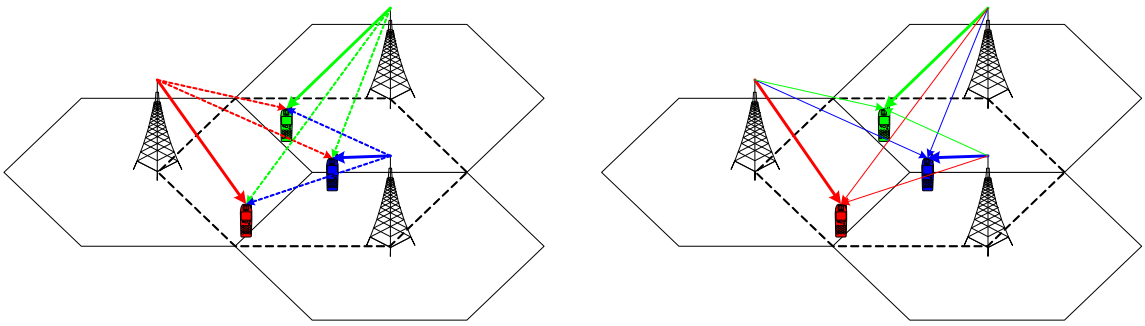


Figure 1.2: Cooperative antenna system



(a) Conventional cellular networks

(b) Cooperative antenna systems

Figure 1.3: Cooperative antenna systems as inter-cell interference elimination scheme (solid line: signal, dotted line: interference)

single user served by a single link.

- MU-MIMO systems can be deployed with low cost MSs, which might be equipped with only one or two antenna elements.

- Due to the distributed nature, where several adjacent - but geographically distributed - associated BSs are used as transmit antennas, full macro diversity gains are available.

- Maybe the most important advantage for cellular radio systems is the elimination of ICI between those BSs, which are cooperating, due to the common processing in the Central Unit (CU).

In spite of these many advantages there is still some reluctance to adopt such system concepts. One of the main reasons is the necessity of a CU, which would impose a hierarchical network structure, prevailing in legacy 2G or 3G networks with the Radio Network Controller (RNC). But the vision of network planers, e.g., for 3GPP LTE, is a flat hierarchy, which allows an economy of scale of the BS hardware and fast and easy network deployment with at least simple network planning if any. To take care of these requirements, a decentralized and distributed cooperative radio concept organized over the air has been developed, minimizing the requirements for the backbone network (*Hu et al.*, 2007). Several issues arise when the CU is replaced by decentralized processing, like independently scheduling Medium Access Control (MAC) units at the BSs, sophisticated channel estimation, the calculation of the precoding weight matrix in a distributed manner, etc.

The issues, which should be taken into account for development of a distributed cooperation in the DL, are as follows:

- In real mobile radio systems the scheduling by the MAC entity in each BS is quite unpredictable as the MAC has to take into account the buffer length of different QoS classes, actual channel conditions, automatic repeat request (ARQ) messages, etc. For this reason in each radio frame the scheduling might change completely.

- Future backbone networks will use IP networks, so the delays over different routes in the backbone network are highly unpredictable, even in well managed networks (*Jungnickel et al.*, 2008).

- For highest benefit of cooperative transmission, all associated BSs shall have full channel knowledge and have access to all data for all associated MSs (*Thiele et al.*, 2009).

- For simultaneous transmission in an OFDM system all associated BSs and MSs have to be synchronized in frequency and time (*Jungnickel et al.*, 2008; *Manolakis et al.*, 2010).

This thesis tackles the third issue, i.e., developing an efficient mechanism to convey the downlink channel state information to the associated BSs. The cooperative antenna concept is considered

as one of the technology components enabling the LTE-Advanced, which is coined as coordinated multi-point transmission/reception (CoMP) (*3GPP*, 2009; *Parkvall et al.*, 2008; *Alcatel Lucent*, 2009; *Samsung Electronics Co. Ltd.*, 2009).

1.5 Limited Feedback for Multiple Antenna Systems

MIMO communication systems, especially in the single user case, can employ CSI feedback from the receiver to the transmitter so as to improve the data rates and link reliability on the downlink by a linear precoding (*Goldsmith*, 2005; *Andrews et al.*, 2007; *Love et al.*, 2008). With limited feedback, channel state information is quantized by a certain measure. The quantization is performed by selecting a predefined code from a codebook known to both the receiver and the transmitter. Quantized channel state information is used at the transmitter to design smart transmission strategies such as precoded spatial multiplexing and transmit beamforming (*Love et al.*, 2008; *Love and Heath*, 2005).

The linear precoding was utilized in (*Scaglione et al.*, 2002), where the optimal linear precoder that minimizes the symbol mean square error for linear receivers was derived under various constraints. The Bit Error Rate (BER) optimal precoder was introduced in (*Palomar et al.*, 2005), and the capacity optimal precoder was investigated in (*Telatar*, 1999). In (*Stankovic and Haardt*, 2008), Stankovic and Haardt proposed the general precoding matrix design principle for MU-MIMO systems which can target several optimization criteria like maximum information rate, maximum receive SNR, minimum BER, etc.

The use of partial CSI at the transmitter was presented in (*Narula et al.*, 2002), where the Lloyd algorithm is used to quantize the CSI. Other approaches focused on feeding back the mean of the channel (*Visotsky and Madhow*, 2001), or the covariance matrix of the channel (*Jafar et al.*, 2001). An overview of the achievable channel capacity with limited channel knowledge can be found in (*Goldsmith et al.*, 2003). Schemes that directly select a quantized precoder from a codebook at the receiver, and feed back the precoder index to the transmitter have been proposed in (*Mukkavilli et al.*, 2003). In this work, the authors proposed to design the precoder codebooks to maximize a subspace distance between two codebook entries, a problem which is known as the Grassmannian line packing problem. The advantage of directly quantizing the precoder is that the unitary precoder matrix (*Scaglione et al.*, 2002) has less degrees of freedom than the full CSI matrix, and is thus more efficient to quantize. Several subspace distances to design the codebooks were proposed in (*Love and Heath*, 2005), where the selected subspace distance depends on the function used to

quantize the precoding matrix. In (*Roh and Rao*, 2006a), a precoder quantization design criterion was presented that maximizes the capacity of the system and the corresponding codebook design was also presented. A quantization function that directly minimizes the uncoded BER was proposed in (*Zhou and Li*, 2006).

The references mentioned so far have been focusing on the single user case, assuming no external interferences, e.g., inter-user or inter-cell interference. Recently limited feedback concepts have been applied to more sophisticated system configurations such as MIMO-OFDM and multi-user MIMO and are proposed for current and next generation wireless systems (*Telatar*, 1999; *Zacarías et al.*, 2009). It is one of the major contributions of this thesis to develop an efficient feedback method for MU-MIMO systems. It has been shown that MU-MIMO systems are quantization error limited at high SNR, hence higher resolution feedback is required than in comparable single user systems (*Jindal*, 2006). More interestingly, the authors in (*Ravindran and Jindal*, 2008) have investigated question of whether low-rate feedback/many user systems or high-rate feedback/limited user systems are preferable in the context of MIMO downlink channels, and the surprising conclusion is that there is an extremely strong preference towards accurate channel information. Therefore, feedback compression becomes an important issue when it comes to the multi-user case.

Feedback overhead reduction methods have been investigated in (*Simon and Leus*, 2008b; *Heath et al.*, 2009; *Zacarías et al.*, 2009; *Bazzi et al.*, 2014). Most of the proposed methods are designed to exploit the time correlation of the channel to achieve a high resolution at the cost of a reasonable amount of feedback. Simon and Leus have proposed entropy encoding based feedback compression methods (*Simon and Leus*, 2008b), whereas Heath et al. have suggested a progressively scaled local codebook which enables high resolution quantization and reconstruction for multi-user MIMO with zero-forcing precoding (*Heath et al.*, 2009). The method proposed by Zacarías et al. is based on single-bit quantization tracking each channel element (*Zacarías et al.*, 2009), while the works in (*Simon and Leus*, 2008b; *Heath et al.*, 2009) focus on the codebook based channel vector quantization which better captures performance relevant channel features. Bazzi et al. have proposed an efficient two-stage precoding scheme for massive MIMO systems which can keep the required feedback amount of interfering links constant for a fixed number of receive antennas and transmitter/receiver pairs, regardless of the number of transmit antennas. In this thesis, novel feedback compression schemes exploiting time and frequency correlation have been investigated and the feedback frequency decision method is also proposed.

Most prior work on limited feedback MIMO assumes the error-free and delay-free feedback link

which would be unrealistic in actual wireless systems. To address this issue, Markov models to analyze the effect of the channel time evolution and consequently, the feedback delay are proposed in (*Wang and Moayeri*, 1995; *Huang et al.*, 2009). Other temporal correlation models and measurement results of the wireless channel are used in (*Roh and Rao*, 2004; *Daniels et al.*, 2008) to evaluate the effect of the feedback delay. In (*Roh and Rao*, 2004), the authors quantize the parameters of the channel to be fed back using adaptive delta modulation, taking into account the composite delay due to processing and propagation. The effect of feedback delay on the achievable rate and bit error rate performance of MIMO systems has been investigated in several scenarios (*Huang et al.*, 2006a; *Zhou and Giannakis*, 2004; *Isukapalli and Rao*, 2007). The feedback delay has been found to reduce the achievable throughput (*Huang et al.*, 2006a), and to cause interference between spatial data streams. Channel prediction was proposed in (*Zhou and Giannakis*, 2004; *Kobayashi et al.*, 2008) to remedy the effect of the feedback delay. Albeit not in the context of limited feedback, the authors use pilot symbol assisted modulation to predict the channel based on the Jakes model for temporal correlation of the channel. The authors in (*Huang et al.*, 2009, 2006a) derived expressions for the feedback bit rate, throughput gain and feedback compression rate as a function of the delay on the feedback channel.

Buzuverov et al. have investigated the achievable degrees of freedom of the MIMO systems with delayed instantaneous channel state information at the transmitters (*Buzuverov et al.*, 2015, 2016) and Akoum et al. have derived the impact of the delay on the achievable sum rate of limited feedback MIMO systems in the presence of other cell interference (*Akoum and Heath*, 2009). The authors showed that the feedback delay, coupled with other cell interference at the mobile station, causes the spectral efficiency of the system to decay exponentially. The decay rate almost doubles when the mobile station is at the edge of its cell, and hence this scenario is interference limited. As already mentioned in Section 1.4, cooperative antenna systems can mitigate the inter-cell interference and thus can alleviate a delay related performance degradation problem by reducing the interference from other cells.

The impact of imperfect channel state information (CSI) has been discussed for adaptive OFDM-based systems in (*Ye et al.*, 2006; *Su and Schwartz*, 2001; *Marqués et al.*, 2006; *Xia et al.*, 2004), whereas Kuehne and Klein have provided a comparison of adaptive and non-adaptive multi-user OFDMA schemes in the presence of imperfect channel knowledge in (*Kuehne and Klein*, 2008). According to (*Kuehne and Klein*, 2008), at a certain level of channel quality information (CQI) imperfectness it is beneficial to switch from adaptive to non-adaptive transmission, i.e., depending

on the quality of the channel knowledge, all users apply the adaptive or nonadaptive transmission scheme. In a realistic scenario however, the level of CQI imperfectness differs from user to user. As a next step, Kuehne et al. have extended their investigation in (*Kuehne and Klein*, 2012) and provided algorithms determining which user is to be served adaptively or non-adaptively subject to the BER and minimum data rate constraints as well as an analytical expression of the system performance considering imperfect CQI and different user demands for a hybrid multi-user OFDMA scheme. Simulation results have shown the superiority of the hybrid OFDMA scheme in terms of achievable data rate and user satisfaction compared to pure adaptive and non-adaptive OFDMA schemes in the presence of user-specific imperfect CQI.

Massive MIMO systems, i.e., MIMO systems with a large number of base station antennas, have received much attention in industry as well as academia as a means to improve the spectral and energy efficiency, and to reduce processing complexity of the next generation mobile communication systems (*Ji et al.*, 2017). Such a benefit of the massive MIMO systems, however, can be realized only when the perfect CSI is available at the base station. This is one of the major challenges for frequency division duplex (FDD) systems which unlike in time division duplex (TDD) systems cannot make use of the channel reciprocity. Hence, they should rely on the feedback link for the base station to acquire the downlink channel state information. Since the amount of feedback needs to grow linearly with the number of transmit antennas (and SNR) (*Jindal*, 2006), this provides a great challenge for FDD-based massive MIMO systems. Byungju Lee et al. have shown that the overall channel information can be reconstructed by making use of the spatial and temporal correlation among antenna elements, which provides a novel framework for FDD-based massive MIMO systems that achieves a reduction in the CSI feedback overhead by mapping multiple correlated antenna elements to a single representative value using properly designed grouping patterns (*Lee et al.*, 2015). The proposed scheme in (*Lee et al.*, 2015) can be considered as a good candidate feedback technique for Full-Dimension MIMO systems as well (*Ji et al.*, 2017). In (*Choi et al.*, 2015), on the other hand, a trellis code is employed to reduce both the codeword search complexity and the feedback overhead in the presence of a temporal correlation of the channel. In (*Han et al.*, 2014), an adaptive feedback compression scheme which exploits the spatial and temporal correlation jointly is proposed to reduce the dimension of CSI feedback by projecting it on a lower dimensional subspace. Another challenge with massive MIMO systems rises in Over-the-Air testing (*Kotterman et al.*, 2017; *Schirmer et al.*, 2014) as well as in 3D channel modeling (*Kotterman et al.*, 2014).

1.6 Scope and Objectives of the Thesis

The scope of this thesis is to develop efficient feedback techniques for the multi-user MIMO case, targeting distributed cooperative antenna systems in particular. As is briefly explained in Section 1.4, all associated base stations should be able to acquire full CSI from all associated mobile stations. As distributed cooperative antenna systems do not have a means to prompt coordination across associated base stations, CSI feedback from the MSs should provide the BSs with all the necessary channel information, over the air. In short, we assume no CSI sharing among associated BSs, let alone cooperation among MSs, as far as time critical operation is concerned. It is a challenging task to design economically viable feedback schemes, considering the necessity of high resolution CSI driven by the presence of inter-user interference and an increased size of the channel matrix for MIMO systems. In order to fulfill these requirements, feedback compression methods have been developed as well as the efficient CSI feedback schemes.

The target system in consideration employs OFDM transmission technique with linear transceiver processing, i.e., zero-forcing (ZF) transmit filtering. The ZF methods are generally power inefficient because beamforming weights are not matched to the user channels. Despite its suboptimality, the ZF transmission can still perform well when some multi-user diversity is available, i.e., semi-orthogonal users can be grouped together for ZF processing from a larger pool of users. The ZF method is known to achieve the same asymptotic sum capacity as that of the dirty paper coding which is the optimal (channel capacity achieving) method, as the number of users goes to infinity (*Yoo and Goldsmith*, 2006). The main benefit of the ZF method is that the channel quantization design problems can be simplified into channel directional information quantization problems. This results from the fact that the ZF processing decouples the beamformer design and the power allocation. Regarding the physical layer parameters of the target system, the LTE frame structure, which is based on OFDMA for the DL transmission scheme, is adopted when applicable (Chapter V).

1.7 Contributions and Structure

The thesis is in part based on nine original publications, including one journal paper (*Kim et al.*, 2008c), three conference papers (*Kim et al.*, 2007, 2008d,e) and five patent applications (*Kim et al.*, 2008a,b; *Kim and Zirwas*, 2009; *Kim et al.*, 2009; *Kim*, 2016). The extended spatial channel model (SCMe) (*3GPP*, 2007; *Baum et al.*, 2005) used in the simulations is based on the code developed in the WINNER project (*Salo et al.*, 2005). In addition to the original publications (*Kim et al.*,

2007, 2008a,b,c,d,e; *Kim and Zirwas*, 2009; *Kim et al.*, 2009; *Kim*, 2016), the author coauthored a book chapter (*Zirwas et al.*, 2007) providing a part of the distributed system concepts of the DAS, in particular downlink channel estimation schemes and associated simulation results. The author has also contributed to two other works, as a coauthor of one conference paper (*Zirwas et al.*, 2006) performing the overall simulation campaign, and as a co-inventor of one patent application (*Haustein et al.*, 2009).

The author's research work has been performed taking into account the MU-MIMO systems based on OFDM transmission techniques, with possible application to the mobile communication systems like 3GPP LTE in mind. 3GPP released the first LTE specification (release 8) (*3GPP*, 2008) in 2008. It has turned out that there are two important differences between the author's proposal and the feedback scheme adopted in LTE (*Lim et al.*, 2013), which render the direct application of the author's schemes to the LTE standard not straightforward. First, LTE uses an implicit (precoding matrix) feedback whereas the author's proposal uses an explicit (channel matrix) feedback. Second, for LTE mobile stations the MIMO operation mode (Single-User Multiple-Input Multiple-Output (SU-MIMO) or MU-MIMO) is transparent, which means that the LTE release 8 compliant mobile stations use small size (4 bits) codebooks irrespective of the MIMO operation mode. The resolution of LTE release 8 codebooks might be sufficient for SU-MIMO, but it is not high enough to get the maximal benefit of deploying MU-MIMO. In the 3GPP standardization process, many approaches to improve the release 8 implicit CSI feedback mechanism have been proposed with focus on support of high resolution feedback for MU-MIMO (*Motorola*, 2010a,b,c; *Qualcomm Inc.*, 2010; *Samsung Electronics Co. Ltd.*, 2010a,b; *Huawei*, 2010). From the conceptual perspective, many similarities can be found between the author's proposal and the proposals in 3GPP , e.g., split of long-term/short-term feedback and exploiting temporal correlation of the channel for feedback overhead reduction. These feedback enhancement proposals could not be finalized in 3GPP release 10 (*3GPP*, 2012) due to a lack of time. However, it is envisioned that in future cellular systems, e.g., 5G, the high granularity/resolution codebook and its efficient deployment schemes should be indispensable for successful utilization of the MU-MIMO mode. For instance, according to the latest technical report on 5G New Radio Access Technology (*3GPP*, 2017), 3GPP has proposed a so called type II feedback for 5G. This new feedback type, which is an explicit channel feedback and/or codebook-based precoder feedback with higher spatial resolution, is introduced to support MU-MIMO more effectively, compared with the legacy type I feedback which is a codebook-based precoder feedback with normal spatial resolution. The techniques to cope with the high resolution

feedback that are developed in this thesis can be useful and applicable for 5G systems in this regard.

As mentioned in Section 1.6, the thesis focuses on developing efficient feedback techniques for MU-MIMO downlink transmission dealing with a practical application of the distributed cooperative antenna systems or multicell CoMP of LTE-Advanced into consideration. The structure of the thesis is as follows.

Chapter II proposes the codebook based limited feedback scheme which conveys the quantized channel state information rather than the beamformer index, in an effort to address the inter-user interference by allowing the BSs to calculate the precoding matrix based on the collected downlink CSI. Unlike other feedback methods developed so far, it is designed to address multi-user interference by directly delivering channel information to the BSs. The channel matrix estimated by the MS is first decomposed to reveal its directional structure, then it is quantized by comparing it with predefined codewords in terms of one of the subspace distance metrics, i.e., the chordal distance in our case. Chapter II also describes the novel codebook construction method based on the modified Linde-Buzo-Gray (LBG) Vector Quantization (VQ) algorithm. Numerical results show that the proposed subspace based channel quantization method outperforms the analog pilot retransmission method and the MU-MIMO case requires a higher resolution codebook than the SU-MIMO case to cope with the interferences. The results have been presented in (*Kim et al.*, 2007, 2008c).

In Chapter III we study Givens rotation based channel parameterization and quantization method, motivated by the previous works (*Roh and Rao*, 2007, 2006b). Simulation results using our cooperative antenna system set-up confirm the effectiveness of the Givens rotation method followed by an adaptive delta modulation (ADM), in the presence of slowly time-varying channels. The author's contribution to this issue is as follows. We investigate the tracking of the Givens parameters, and propose some novel guidelines to determine the frequency of the feedback for the Givens parameters for temporally-correlated channels. The simulation results have been in part shown in (*Kim et al.*, 2008c).

Chapter IV focuses on the feedback overhead reduction methods for temporally-correlated channels. The baseline of the chapter is Chapter II, i.e., the codebook based limited feedback scheme. As presented in Chapter II, MU-MIMO systems require higher resolution codebooks. This leads to an increase of the feedback overhead. In Chapter IV, we investigate some feedback compression methods to exploit the temporal correlation of the channel, which can render the proposed feedback scheme manageable in practical mobile communication systems. This research has been partially

published in (*Kim et al.*, 2008c,d,e). The adaptive codebook method, the most efficient of the proposed algorithms, has been granted as an European patent (*Kim et al.*, 2008a).

In Chapter V, we extend the adaptive codebook design method into two dimensions, i.e., time and frequency. This is a sensible extension of the method which relies on the correlation of the channel, as there exist correlations over adjacent subcarriers as well as subsequent time samples in multi-carrier transmission systems like OFDM. The chapter introduces the codeword interpolation method as well, which is motivated by the beamformer interpolation method proposed in (*Choi and Heath*, 2005) with modification on the phase parameter calculation. Note that the proposed codeword interpolation method is applicable to the single antenna receiver case only. Numerical results show that the two-dimensional adaptive codebook method can greatly reduce the feedback overhead in OFDM based systems.

Chapter VI compares the proposed feedback methods, i.e., the Givens rotation based channel quantization method (Chapter III) and the subspace based channel quantization method with the adaptive codebook design (Chapter II, IV). The feedback methods are compared in terms of their performance, their required feedback overhead, their computational complexity, etc. in an effort to provide guidelines on the selection of the proposed feedback schemes. The adaptive codebook method outperforms the differential Givens method in key performance indicators, i.e., the received SINR and the required uplink feedback overhead, at the cost of a higher complexity at the MS.

Chapter VII extends the previous work on the subspace based channel quantization method by introducing multiple antennas at the transmitter side (BS) and/or at the receiver side (MS) and by investigating the performance of the RBD precoding(/decoding) scheme as well as the ZF precoding scheme. The RBD precoding scheme requires the MS to compute the decoding matrix via a Singular Value Decomposition (SVD) in case of the multiple data stream transmission to each user. A cost-efficient decoding matrix quantization method is proposed which can avoid a complex computation at the MS while showing a slight degradation.

Chapter VIII describes a systematic scheduling mechanism at the network side which incorporates high resolution codebook operation.

Chapter IX concludes the thesis. The main achievements and conclusions are summarized. Moreover, the remaining open questions and directions for future research are pointed out.

Appendix A describes the symbols and the notation used in the thesis, and Appendix B provides the proof of Equation (2.14). Appendix C supplements Section 4.1 by explaining how to determine the coarse/fine feedback periods. Appendix D introduces a computational complexity comparison

of the differential Givens (D-Givens) and the adaptive codebook (aCB) based feedback schemes.

Throughout the thesis, the proposed methods have been verified by simulations for the single antenna BS and MS case, whereas the subspace based channel quantization method has been extended to the multiple antenna BS and MS case as well.

CHAPTER II

Subspace Based Channel Quantization Method

It is one of the biggest challenges of distributed cooperative antenna systems to provide BSs with downlink channel information for transmit filtering (precoding). Especially in cases of MU-MIMO systems, a well-known beamforming vector selection method cannot be directly applicable, as inter-user interferences should be taken into account. In this chapter we propose a novel feedback scheme via a subspace based Channel Quantization (CQ) method , which conveys channel state information explicitly, to address this issue. The proposed scheme adopts the chordal distance as a channel quantization criterion so as to capture channel characteristics represented by subspaces spanned by the channel matrix. The proposed method is tested for distributed cooperative antenna systems in terms of simulations. Simulation results show that the proposed subspace based channel quantization method outperforms the analog pilot retransmission method which was proposed in (*Zirwas et al.*, 2007, 2006) to address the challenges in providing downlink channel state information to the associated BSs.

This chapter is organized as follows. In Section 2.1, the challenges in distributed cooperative antenna systems are explained and the subspace based CQ method which addresses them is proposed in Sections 2.2, 2.3 and 2.4. In Section 2.5, a method which can enhance effectiveness of the proposed method is described. Section 2.6 presents simulation results and the conclusion can be found in Section 2.7.

The results of this chapter have been presented in (*Kim et al.*, 2007, 2008c).

2.1 Introduction

In cooperative antenna systems, several adjacent base stations (BSs) are cooperating so as to support multiple mobile stations (MSs) which are located in the corresponding cooperative area

(CA). Therefore, cooperative antenna systems can be regarded as a MU-MIMO system, in which multiple transmit antennas at the BS, which are conventionally considered to be located in one BS, are spread over several BSs. Cooperative antenna systems are based on the cooperation between multiple distributed BSs. This means that cooperative antenna systems need a fast and efficient backbone network as well as the central unit (CU) which manages the cooperation amongst associated BSs. The CU renders the overall network structure more complex by adding one more layer in the hierarchy, and eventually increases the costs.

One of the main challenges of the distributed cooperative antenna system is channel estimation for the downlink channel. All of the associated BSs in the CA need to know the full channel state information to calculate the corresponding precoding weight matrix. This information is needed to be transferred from MSs to BSs by using uplink resources. As several BSs and several MSs are involved in cooperative antenna systems and each BS and MS may be equipped with multiple antennas, the number of channel state parameters to be fed back is expected to be big. In an effort to reduce the amount of feedback, the analog pilot retransmission method has been suggested and tested in (*Zirwas et al.*, 2007, 2006), but the throughput of this method reaches only 40 % of that of the ideal case[1] due to noise enhancement effects, which require supplementary feedback schemes[2].

On the other hand, finite rate feedback strategies in MIMO systems have been extensively investigated recently. Beamforming codebook design methods are suggested based on Grassmannian packing (*Love et al.*, 2003) and systematic unitary design (*Hochwald et al.*, 2000), which guarantee substantial gains with just a small number of feedback bits. A precoding matrix codebook construction method, which is designed to maximize the mutual information, has been developed based on vector quantization (VQ) techniques (*Roh and Rao*, 2007, 2006b). These methods are designed to select a beamforming vector for the MISO case or a precoding matrix for the MIMO case from a set of codes. They are developed for point-to-point MIMO channels, in which the transmitter serves one receiver at a time. In the SU-MIMO case, it is known that even a small number of bits per antenna can be quite beneficial (*Love et al.*, 2004). In the multi-user MIMO (MU-MIMO) case, feedback rate scaling is required to achieve a throughput close to that with perfect feedback information in order to compensate for the interference between users (*Jindal*, 2006). The analysis in (*Jindal*, 2006) is based on the case when a user selects the precoding matrix by solely looking at its own channel without considering the interference to other users which is caused by adopting that

[1]Here the ideal case means the case in which the associated BSs have an access to the perfect downlink channel state information.

[2]Note that the simulation result at hand is valid for the specific scenario described in (*Zirwas et al.*, 2007, 2006), and there can be a different result depending on the exact scenario due to the factors like existence of the interference floor, etc.

precoding matrix. Hence, a better way to handle inter-user interference needs to be addressed.

In this chapter, we propose a subspace based CQ method which guarantees a much higher performance than the analog pilot retransmission method. It is different from other feedback methods mentioned above in the sense that it is designed to deal with multi-user interference. We also propose an iterative codebook design algorithm which converges to a locally optimum codebook. The proposed schemes can be used for cellular MU-MIMO systems as well, which involve one BS for downlink data transmission.

2.2 System Model and Motivation for Channel Quantization Method

We consider a precoded MU-MIMO system in which a group of BSs transmits data to multiple MSs simultaneously. Each of N_{BS} BSs and each of N_{MS} MSs have N_t and N_r antennas, respectively [3]. The data symbol block, $\mathbf{s} = [s_1, \ldots, s_{N_{tr}}]^{\mathcal{T}}$ with $N_{tr} = N_{\mathrm{MS}}N_r$, is precoded by a $N_{tt} \times N_{tr}$ matrix $\mathbf{W}$ with $N_{tt} = N_{\mathrm{BS}}N_t$, in case that the number of data streams for each user n_s ($n_s \leq N_r$) is N_r [4]. Here the first N_r data symbols are intended for the first user, the next N_r symbols for the second user, and so on. When denoting $i_{\mathrm{BS}}/i_{\mathrm{MS}}$ as the BS/MS index and i_t/i_r as the transmit/receive antenna index, respectively, we can denote $h_{i,j}$ where $i = N_r(i_{\mathrm{MS}} - 1) + i_r, j = N_t(i_{\mathrm{BS}} - 1) + i_t$ as the channel coefficient between the i_rth receive antenna of the i_{MS}th MS and the i_tth transmit antenna of the i_{BS}th BS. The $N_{tr}N_{tt}$ channel coefficients can be expressed as the $N_{tr} \times N_{tt}$ channel matrix $\mathbf{H}$ with $[\mathbf{H}]_{i,j} = h_{i,j}$. The received signals on N_{tr} receive antennas which are collected in the vector $\mathbf{y}$ can be formulated as

$$\mathbf{y} = \mathbf{HWs} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{n}$ is Additive White Gaussian Noise (AWGN). The signal model appears to be very similar to that of the single-user MIMO case at a first glance, but the difference lies in the fact that the channel matrix in our case contains elements belonging to multiple BSs and multiple MSs. The data estimate $\hat{\mathbf{s}}$ at the input of the quantization can be written as

$$\hat{\mathbf{s}} = g\mathbf{HWs} + g\mathbf{n}. \tag{2.2}$$

[3] In reality, MSs with different number of antennas can be present in a cooperative area. In this case, MSs with the same number of antennas can be grouped together. Then the proposed method can be applied to the MSs in the same group.

[4] In general, MSs experience different channel conditions, which leads to different number of serviceable data streams. This restriction can be overcome by grouping MSs with the same number of data streams together to get benefits from multi-user diversity.

The sum mean squared error (MSE) ϵ can be expressed as

$$\begin{aligned}\varepsilon &= \mathrm{tr}\left(\mathbb{E}[(\mathbf{s}-\hat{\mathbf{s}})(\mathbf{s}-\hat{\mathbf{s}})^{\mathcal{H}}]\right) \\ &= \mathrm{tr}(\mathbf{R_s}) - g\mathrm{tr}(\mathbf{HWR_s}) - g\mathrm{tr}(\mathbf{R_s}\mathbf{W}^{\mathcal{H}}\mathbf{H}^{\mathcal{H}}) + |g|^2\mathrm{tr}(\mathbf{HWR_s}\mathbf{W}^{\mathcal{H}}\mathbf{H}^{\mathcal{H}}) + |g|^2\mathrm{tr}(\mathbf{R_n}),\end{aligned} \tag{2.3}$$

where $\mathbf{R_n} = \mathbb{E}[\mathbf{n}\mathbf{n}^{\mathcal{H}}]$ and $\mathbf{R_s} = \mathbb{E}[\mathbf{s}\mathbf{s}^{\mathcal{H}}]$ are the covariance matrix of the noise and the covariance matrix of the data symbols, respectively.

Several techniques have been developed for downlink transmit filtering in MU-MIMO systems. Linear precoding techniques (e.g., transmit matched filter (TxMF), TxZF, and TxWF) have an advantage in terms of computational complexity (*Joham et al.*, 2005). Non-linear techniques (e.g., Tomlinson-Harashima Precoding (THP)) have a higher computational complexity but can usually provide a better performance than linear techniques (*Tomlinson*, 1971; *Harashima and Miyakawa*, 1972). Some linear techniques (e.g., BD and SMMSE precoding) are developed for the case in which there are multiple antennas at each receiver. The BD algorithm is designed to eliminate MUI (*Spencer et al.*, 2004b). BD outperforms the TxZF and asymptotically approaches the sum capacity of the channel at high SNRs. SMMSE performs better than some non-linear techniques (e.g., successive optimization (SO) THP and MMSE THP) with a relatively low computational complexity (*Stankovic and Haardt*, 2004). RBD shows the same throughput as SMMSE at low SNRs and as BD at high SNRs (*Stankovic and Haardt*, 2008). RBD provides a higher diversity than non-linear MMSE THP. For high data rates, RBD extracts the full diversity in the system. Moreover, RBD is less sensitive to channel estimation errors than non-linear techniques. In our case, we adopt the TxZF technique which completely suppresses the interference at the receiver (*Joham et al.*, 2005), and we focus on the single antenna receiver case. TxZF minimizes the sum MSE ε subject to a complete interference cancellation constraint, i.e., $g\mathbf{HW} = \mathbf{I}_{N_{tr}}$, as well as a transmit power constraint. In this case, the sum MSE ε in (2.3) simplifies to $\varepsilon = |g|^2\mathrm{tr}(\mathbf{R_n})$.

$$\begin{aligned}\{\mathbf{W}, g\} =& \arg\min_{\{\mathbf{W},g\}} |g|^2\mathrm{tr}(\mathbf{R_n}) \\ \text{s.t.: } & g\mathbf{HW} = \mathbf{I}_{N_{tr}} \text{ and } \mathrm{tr}\left(\mathbf{WR_s}\mathbf{W}^{\mathcal{H}}\right) = P_{tx}\end{aligned} \tag{2.4}$$

where P_{tx}, $\mathbf{R_n}$, and $\mathbf{R_s}$ are the maximum transmit power, the covariance matrix of the noise, and the covariance matrix of the data symbol, respectively. The TxZF strategy, while generally suboptimal, is known to achieve the same asymptotic sum capacity as that of dirty paper coding which is the optimal (channel capacity achieving) method, as the number of users goes to infinity (*Yoo and*

Goldsmith, 2006). The transmit precoding matrix $\mathbf{W}$ which satisfies the design criteria (2.4) takes the following form.

$$\begin{aligned} \mathbf{W} =& g^{-1}\mathbf{H}^{\mathcal{H}}\left(\mathbf{H}\mathbf{H}^{\mathcal{H}}\right)^{-1}, \\ \text{where } g =& \sqrt{\frac{\mathrm{tr}\left(\left(\mathbf{H}\mathbf{H}^{\mathcal{H}}\right)^{-1}\mathbf{R_s}\right)}{P_{tx}}} \end{aligned} \tag{2.5}$$

The challenge here is that BSs should know the downlink channel matrix $\mathbf{H}$ so as to construct the precoding matrix $\mathbf{W}$. In (*Zirwas et al.*, 2007, 2006; *Marzetta and Hochwald*, 2006; *Thomas et al.*, 2005), the analog pilot retransmission method has been proposed as a way of transferring channel state information to the BS. The analog pilot retransmission method is vulnerable to noise enhancement effects and this weakness of the analog method brings about a significant performance degradation, even though it is efficient in terms of the required resources[5]. As a way of combating noise, a digital method can be used instead of the analog method. A digital method implies that MSs measure the downlink channel and encode this information into a digital code and send it back to the BSs after performing appropriate digital signal processing (modulation, spreading, repetition, or channel coding, etc) to guarantee robust data transmission.

As it is explained in the previous section, most of the finite rate feedback strategies in MIMO systems are designed for the single user case, focusing on the selection and construction of the precoding matrix codebook. If this strategy is directly applied to the multi-user case, the performance will be degraded since the user is supposed to select the precoding matrix which is suitable in terms of its criterion (e.g., maximizing the mutual information or SNR), which may cause a severe interference to other users. Here we propose to quantize the channel from the MS side instead of quantizing the precoding matrix. Both methods are similar from the signal processing perspective in the sense that both schemes compress the information in a matrix, while the channel quantization method is better positioned to cope with inter-user interferences. The BSs, after receiving feedback messages from the MSs, can now build a precoding matrix with inter-user interferences taken into account, since the transmitters have the whole channel state information, albeit it is not perfect due to the limited feedback.

One way of quantizing the channel matrix is to view the channel matrix as a set of complex matrices, and to quantize every individual matrix by looking up a predefined codebook. As explained

[5]In (*Hahn et al.*, 2006), a more elaborated technique of this scheme has been proposed which utilizes multiple uplink subcarriers for retransmission of each downlink pilot in an effort to exploit frequency diversity. This technique has been shown to mitigate the adverse impact of the noise enhancements to some extent at the cost of the increased feedback resource.

above, the overall channel matrix is a $N_{\mathrm{MS}}N_r \times N_{\mathrm{BS}}N_t$ matrix, and is composed of the channel matrices for each user, which are of size $N_r \times N_{\mathrm{BS}}N_t$. Equation (2.6) depicts this relationship.

$$\mathbf{H} = [\mathbf{H}_1, \cdots, \mathbf{H}_j, \cdots, \mathbf{H}_{N_{\mathrm{MS}}}]^{\mathcal{T}}, \ j\text{: user index} \tag{2.6}$$

Here $\mathbf{H}_j$ is the transpose of the channel matrix for user j, which is a $N_{\mathrm{BS}}N_t \times N_r$ matrix. If we allocate n_{CB} bits for the codebook, we need $n_{\mathrm{CB}}N_{\mathrm{MS}}$ bits in total for every subcarrier. This method is suitable for the limited feedback in terms of required feedback bits, and the conventional VQ method can be applied with some modifications.

The system model is depicted in Fig. 2.1. The N_{BS} BSs need overall downlink channel state information $\mathbf{H}$ for the calculation of the precoding matrix $\mathbf{W}$ so as to form multiple spatial beams which enable independent and decoupled data streams for N_{MS} users. The individual user j estimates its portion of the channel $\mathbf{H}_j$ and quantizes it by finding the best candidate from the predefined set of codes $\mathbf{C}_i$. The index of the chosen code i_j is sent back to the BSs through the limited feedback channel. The BSs reconstruct the channel matrix $\hat{\mathbf{H}}$ by looking up the codebook, which is shared by transmitters and receivers. This reconstructed channel matrix is used for the calculation of the precoding matrix $\mathbf{W}$. We should note that in this case all of the N_{BS} associated BSs have the same channel matrix, as long as the feedback messages are received without errors. In case of the analog pilot retransmission method, the individual BS has its own version of the channel matrix, which is in general different from each other due to the nature of the analog transmission scheme, and this entails a significant performance degradation[6].

2.3 Subspace Based Channel Quantization Method

As proposed in the previous section, MS j is supposed to quantize its channel matrix $\mathbf{H}_j$. We view $\mathbf{H}_j$ not just as a complex matrix but as a subspace which is spanned by its columns. We perform a SVD to extract the unitary matrix $\mathbf{U}_j$ which includes the basis vectors $\mathbf{U}_j^{(\mathrm{S})}$ spanning the column space of $\mathbf{H}_j$ ($\mathbf{H}_j$: $N_{tt} \times N_r$, $\mathbf{U}_j$: $N_{tt} \times N_{tt}$, $\mathbf{U}_j^{(\mathrm{S})}$: $N_{tt} \times N_r$). Here, the superscripts (S) and (0) are used to denote a basis for the signal subspace and the left nullspace, respectively.

$$\mathbf{H}_j = \mathbf{U}_j \mathbf{\Sigma}_j \mathbf{V}_j^{\mathcal{H}}, \text{ where } \mathbf{U}_j = \begin{bmatrix} \mathbf{U}_j^{(\mathrm{S})} & \mathbf{U}_j^{(0)} \end{bmatrix} \tag{2.7}$$

[6]This could be overcome by introducing a fast and efficient backbone network exchanging the received channel information, but this is excluded as this thesis focuses on the distributed cooperative antenna system case.

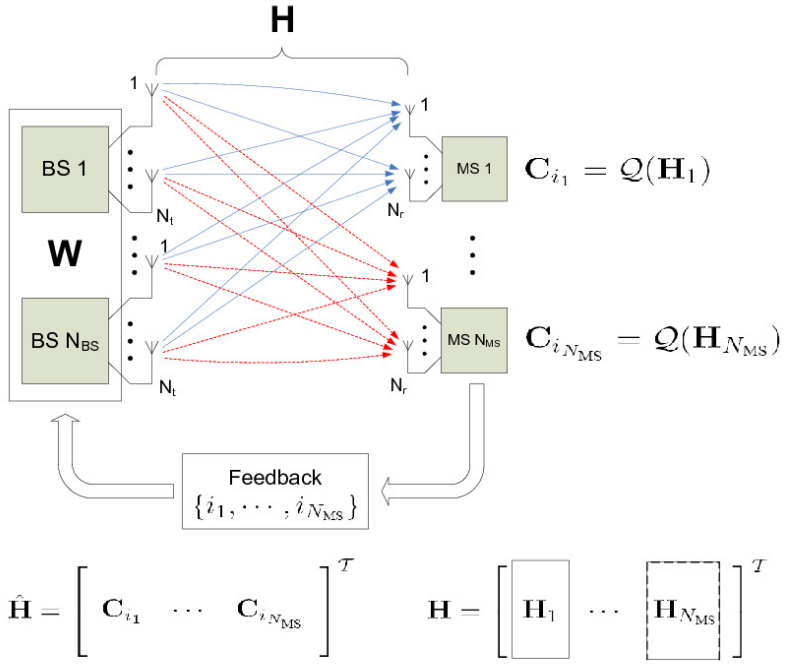


Figure 2.1: Cooperative antenna system downlink with N_{BS} cooperating base stations and N_{MS} mobile stations

The channel quantization uses the chordal distance as a distance metric, since we should measure the distance between subspaces. The subspace distance metric issue can be dealt with in the frame of quantization on the Grassmannian manifold (*Mondal et al.*, 2007). There are other subspace distance metrics (*Love and Heath*, 2005), but the chordal distance is the one which leads us to an analytic solution when designing the codebook (*Zhou and Li*, 2006).

The chordal distance is defined, based on the principal angles between the subspaces, as

$$d_c(\mathbf{T}_i, \mathbf{T}_j) = \frac{1}{\sqrt{2}} \|\mathbf{T}_i \mathbf{T}_i^{\mathcal{H}} - \mathbf{T}_j \mathbf{T}_j^{\mathcal{H}}\|_F = \|\sin \boldsymbol{\Theta}\|_F \tag{2.8}$$

for matrices $\mathbf{T}_i$, $\mathbf{T}_j \in \mathbb{C}^{N_{tt} \times N_r}$ which have orthonormal columns, where $\boldsymbol{\Theta} = \mathrm{diag}(\theta_1, \cdots, \theta_{N_r})$. $\{\theta_i\}_{i=1}^{N_r}$ are the principal angles between the subspaces spanned by the columns of $\mathbf{T}_i$ and $\mathbf{T}_j$: $\mathbf{T}_i^{\mathcal{H}} \mathbf{T}_j$ can be singular value decomposed as $\mathbf{T}_i^{\mathcal{H}} \mathbf{T}_j = \mathbf{U} \cdot (\cos \boldsymbol{\Theta}) \cdot \mathbf{V}^{\mathcal{H}}$, where $\cos \boldsymbol{\Theta} = \mathrm{diag}(\cos \theta_1, \cdots, \cos \theta_{N_r})$.

The quantized version of the column space basis vectors $\mathbf{U}_j^{(\mathrm{S})}$ is chosen to be the code which has the minimum chordal distance from it. Thus, the subspace quantization process can be written as

$$\hat{\mathbf{U}}_j^{(\mathrm{S})} = \mathcal{Q}(\mathbf{U}_j^{(\mathrm{S})}) = \arg \min_{\mathbf{C}_i \in \mathcal{C}} d_c(\mathbf{U}_j^{(\mathrm{S})}, \mathbf{C}_i) \tag{2.9}$$

where $\mathcal{C}$ is the codebook of size N_{cb} ($N_{cb} = 2^{n_{\text{CB}}}$) which has the code $\mathbf{C}_i \in \mathbb{C}^{N_{tt} \times N_r}$ as its elements. Here, $\mathbf{C}_i$ has unitary columns ($\mathbf{C}_i^{\mathcal{H}}\mathbf{C}_i = \mathbf{I}_{N_r}$), hence the code represents only the column space of $\mathbf{H}_j$. No channel magnitude information is fed back to the transmitter, since extra magnitude information does not improve the system performance for the single antenna MS case, compared with the case in which only the code index is provided to the transmitters when the link strengths (large-scale fading due to path loss and shadowing) are assumed to be known at the BSs[7]. The link strength information can be used for user grouping, but the user grouping scheme is not considered here. Note also that we focus on the TxZF method and the introduced methodologies are restricted to the TxZF strategy. In the case that channel magnitude information is to be fed back, the quantized version of the channel at the transmitter which takes this into account can be formulated as

$$\hat{\mathbf{H}}_j = \hat{\mathbf{U}}_j^{(\text{S})} \mathbf{\Sigma}_j^{(\text{S})}, \tag{2.10}$$

where $\mathbf{\Sigma}_j^{(\text{S})} \in \mathbb{R}_+^{N_r \times N_r}$ is a diagonal matrix which is composed of $N_r \times N_r$ elements in the upper left-hand corner of $\mathbf{\Sigma}_j$. The diagonal elements of $\mathbf{\Sigma}_j^{(\text{S})}$ constitute the channel magnitude information, which can be regarded as a refinement of the link strengths which are already available at the BSs by means of a separate feedback message, i.e., Channel Quality Information (CQI). This channel quantization model (2.10) can provide a better view of the channel, as it considers not only the channel directional information, but also the channel magnitude information (link strength refinement information[8], i.e., small-scale fading (fast fading)). In case multiple data streams are transmitted to a MS, the channel magnitude information could play an important role (*Spencer et al.*, 2004b). Here in this thesis we assume the single receiver antenna case, therefore it is excluded from consideration. The simulation results of the single antenna MS case show that this extra information does not enhance the system performance in terms of SINR, compared with the case in which only the channel directional information is provided while the link strength information is available at the BSs. Some of the simulation results can be found in Section 2.6. It is known that the channel magnitude information can provide a significant benefit to the user selection process, when a multi-antenna downlink system carrying more users than transmit antennas is considered (*Yoo et al.*, 2007). In our case, the multi-user diversity gain is not considered at the moment, so we focus on the directional information of the channel. The bottom line is that the MS needs to quantize the column space of $\mathbf{H}_j$ only. The channel magnitude information contained in $\mathbf{\Sigma}_j$ is, therefore, not required for the

[7]Link strength information is needed at the BSs for link adaptation, i.e., adaptive selection of the Modulation and Coding Scheme (MCS) in accordance with the channel quality information.

[8]This term can be a little bit misleading, as the eigenvalues of the channel matrix represent not only the link strength but also the channel correlation.

single antenna MS case.

In this case, the MS is supposed to send only n_{CB} bits of feedback. The channel quantization formula can be simplified as

$$\hat{\mathbf{H}}_j^d = \hat{\mathbf{U}}_j^{(\mathrm{S})} = \arg\min_{\mathbf{C}_i \in \mathcal{C}} d_c(\mathbf{U}_j^{(\mathrm{S})}, \mathbf{C}_i). \tag{2.11}$$

The superscript d implies that the channel is quantized in terms of the direction, with its magnitude information ignored.

The subspace based channel quantization method works as follows. MS j finds the code $\mathbf{C}_i$ which provides the minimum chordal distance with $\mathbf{U}_j^{(\mathrm{S})}$. Then it sends back a n_{CB} bit code index to all associated BSs. The reconstructed downlink channel matrix at the BSs is as follows.

$$\hat{\mathbf{H}} = \left[\hat{\mathbf{H}}_1^d, \cdots, \hat{\mathbf{H}}_j^d, \cdots, \hat{\mathbf{H}}_{N_{\mathrm{MS}}}^d\right]^{\mathcal{T}}, \ j\text{: user index} \tag{2.12}$$

Finally, the BSs calculate the TxZF precoding matrix[9] $\mathbf{W}$ by using the reconstructed channel matrix $\hat{\mathbf{H}}$.

$$\mathbf{W} = g^{-1}\hat{\mathbf{H}}^{\mathcal{H}}\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathcal{H}}\right)^{-1}, \tag{2.13}$$

where g is the normalization factor imposed by the transmit power constraint (2.5).

It is worth noticing that the channel quantization criterion (2.11) can be expressed as

$$\hat{\mathbf{h}}_j = \arg\max_{\mathbf{c}_i \in \mathcal{C}} |\langle \mathbf{v}_j, \mathbf{c}_i \rangle|, \text{ where } \mathbf{v}_j = \frac{\mathbf{h}_j}{\|\mathbf{h}_j\|_2}, \tag{2.14}$$

for the Multiple-Input Single-Output (MISO) case where the MS is equipped with one antenna. In this case, the task of quantizing the channel boils down to that of quantizing the channel vector, instead of the channel matrix. It basically selects the code of which the direction is closely aligned with the direction of the channel. Here the channel to be quantized, the directional information of the channel, and the corresponding code are all vectors of the same size ($\mathbf{h}_j, \mathbf{v}_j, \mathbf{c}_j \in \mathbb{C}^{N_{tt} \times 1}$). For the proof of this formula, please refer to the appendix.

[9] Actually, it is not a true zero-forcing precoding matrix in the strict sense, since the channel magnitude information is not considered. In this section, we use the term TxZF interchangeably for this particular case, assuming that readers are not to be confused.

2.4 Codebook Construction Based on Modified LBG VQ Algorithm

The Grassmannian subspace packing is optimal in terms of quantization for the uncorrelated Rayleigh fading channel (*Love et al.*, 2003). For the correlated Rayleigh fading channel, it is shown in (*Love and Heath*, 2004) that the design criteria based on the Grassmannian subspace packing in (*Love et al.*, 2003) are still applicable with a simple modification (rotation and normalization) in the presence of only receive correlation. In short, according to (*Love et al.*, 2003; *Love and Heath*, 2004) Grassmannian codebook design techniques are applicable to a wide range of realistic channel models, i.e., receive correlated Rayleigh fading channel as well as uncorrelated Rayleigh fading channel. The Grassmannian space $\mathcal{G}(m, n)$ is the set of all n-dimensional subspaces of the space $\mathbb{C}^m$, and the Grassmannian subspace packing problem is the problem of finding the best packing of N n-dimensional subspaces in $\mathbb{C}^m$. The best packing means that N points in $\mathcal{G}(m, n)$ are maximally spaced such that the minimal distance between any two of the subspaces is as large as possible.

In our case, the LBG vector quantization (VQ) algorithm (*Linde et al.*, 1980) is used to construct the codebook $\mathcal{C}$. The LBG VQ algorithm is an iterative algorithm based on the Lloyd algorithm which is known to provide an alternative systematic approach for the Grassmannian subspace packing problem (*Zhou and Li*, 2006). We in this section acquire the codebook through the iterative algorithm described in (*Zhou and Li*, 2006). The main difference of the proposed method is attributed to the fact that the codebooks in (*Zhou and Li*, 2006) are precoder codebooks, while the codebooks to be constructed here are channel quantization codebooks. The proposed algorithm aims at finding a tradeoff between good quantization properties and the Grassmannian subspace packing requirements by adopting the minimum chordal distance of the codebook as a decision criterion for iterations.

2.4.1 Design Issue

The LBG VQ based codebook $\mathcal{C}$ design problem can be stated as follows. For a given source vector, a given distortion measure, a given codebook evaluation measure, and given the size of the codebook, find a codebook and a partition[10] which result in maximizing the minimum chordal distance of the codebook. In other words, we want to find maximally spaced N_{cb} points in $\mathcal{G}(N_{tt}, N_r)$ with given channel realization samples.

Suppose that we have a training sequence[11] $\mathcal{T}$ to capture the statistical properties of the column

[10] The *partition* of the space is defined as the set of all encoding regions.

[11] For the numerical experiments, a training sequence has been generated by using the extended Spatial Channel

space basis vectors $\mathbf{U}_j^{(\mathrm{S})}$ of size $N_{tt} \times N_r$:

$$\mathcal{T} = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_M\} \tag{2.15}$$

where $\mathbf{X}_m \in \mathbb{C}^{N_{tt} \times N_r}$ is a sample of $\mathbf{U}_j^{(\mathrm{S})}$ which can be obtained by taking a SVD of the channel matrix $\mathbf{H}_j$. The codebook can be represented as follows.

$$\mathcal{C} = \{\mathbf{C}_1, \mathbf{C}_2, \cdots, \mathbf{C}_{N_{cb}}\} \tag{2.16}$$

The individual code is of the same size as a training matrix ($\mathbf{C}_n \in \mathbb{C}^{N_{tt} \times N_r}$). Let $\mathcal{R}_n$ be the encoding region[12] associated with the code $\mathbf{C}_n$ and let

$$\mathcal{P} = \{\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_{N_{cb}}\} \tag{2.17}$$

denote the partition of the space. If the source matrix $\mathbf{X}_m$ belongs to the encoding region $\mathcal{R}_n$, then it is quantized to $\mathbf{C}_n$:

$$\mathcal{Q}(\mathbf{X}_m) = \mathbf{C}_n, \text{ if } \mathbf{X}_m \in \mathcal{R}_n. \tag{2.18}$$

Our aim is to find a codebook of which the minimum chordal distance is maximized. There are several subspace distance metrics, e.g., the Fubini-Study distance, the projection two-norm distance, and the chordal distance metrics. It has been shown that the chordal distance is the only distance measure which makes the iterative algorithm feasible (*Zhou and Li*, 2006)[13]. The minimum chordal distance of the codebook is given by:

$$d_{c,\min}(\mathcal{C}) := \min d_c(\mathbf{C}_i, \mathbf{C}_j), \text{ for } \mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, \forall i \neq j. \tag{2.19}$$

Model (SCMe) described in (*Baum et al.*, 2005) in an effort to acquire more realistic channel samples. Note that the Grassmannian method is still applicable, as the transmit uncorrelated channel can be guaranteed for cooperative antenna systems with BSs having remotely placed antennas.
The SCMe, which is an extended version of Spatial Channel Model (SCM) to facilitate support for bandwidth up to 20 MHz, is used throughout this thesis in order to get as close to a true time varying channel as possible. We focus on the urban macro cell scenario in which the height of the BS is assumed to be well above the height of scatterers. Urban macro channel model specifies parameters for an urban environment with up to 3 km distance to a base station and Non Line Of Sight (NLOS). This implies that the received signal at the mobile antenna arrives from all directions after bouncing from the surrounding scatterers, but there is no line-of-sight (LOS) to the BS. In this sense, the urban macro channel of the SCMe can be considered as the closest channel model to the isotropically scattered channel (uncorrelated Rayleigh fading channel) compared with other SCMe channel models, i.e., suburban macro channel and urban micro channel.

[12]The encoding region is called a *Voronoi cell* in some publications.

[13]Zhou and Li have shown that the chordal distance can be formulated as $d_c^2(\mathbf{T}_i, \mathbf{T}_j) = \mathrm{tr}\left(\mathbf{I}_{N_r} - \mathbf{T}_j^{\mathcal{H}}\mathbf{T}_i\mathbf{T}_i^{\mathcal{H}}\mathbf{T}_j\right)$ which renders a simple analytical solution inside the iterations of the Lloyd algorithm as will be shown in (2.24).

The design problem can be stated as follows. Given $\mathcal{T}$ and N_{cb}, find $\mathcal{C}$ and $\mathcal{P}$ such that $d_{c,\min}(\mathcal{C})$ is maximized:

$$\mathcal{C}_{opt} = \arg\max_{\mathcal{C}} d_{c,\min}(\mathcal{C}). \tag{2.20}$$

2.4.2 Optimality Criteria

$\mathcal{C}$ and $\mathcal{P}$ must satisfy the following two criteria so as to be a solution to the above mentioned design problem (*Linde et al.*, 1980). We should note that the chordal distance is used as a distance metric.

- *Nearest Neighbor Condition*:

$$\mathcal{R}_n = \{\mathbf{X} : d_c(\mathbf{X}, \mathbf{C}_n) < d_c(\mathbf{X}, \mathbf{C}_{n'}), \forall n' \neq n\} \tag{2.21}$$

  This condition says that any channel sample $\mathbf{X}$, which is closer to the code $\mathbf{C}_n$ than any other codes in the chordal distance sense, should be assigned to the encoding region $\mathcal{R}_n$, and be represented by $\mathbf{C}_n$.

- *Centroid Condition*:

$$\mathbf{C}_n = \mathbf{U}_R \mathbf{I}_{N_{tt} \times N_r} \tag{2.22}$$

  where $\mathbf{I}_{N_{tt} \times N_r}$ is a selection matrix and $\mathbf{U}_R$ is an eigenvector matrix of the sample covariance matrix $\mathbf{R}$ which is defined as

$$\mathbf{R} := \frac{1}{N_{\mathcal{R}_n}} \sum_{\mathbf{X}_m \in \mathcal{R}_n} \mathbf{X}_m \mathbf{X}_m^{\mathcal{H}} \text{ where } N_{\mathcal{R}_n} = |\mathcal{R}_n|, \tag{2.23}$$

  provided that the eigenvalues in the eigenvalue matrix $\mathbf{\Sigma}_R$ of $\mathbf{R} = \mathbf{U}_R \mathbf{\Sigma}_R \mathbf{U}_R^{\mathcal{H}}$ are sorted in descending order.

  This condition means that the code $\mathbf{C}_n$ of the encoding region $\mathcal{R}_n$ should be the principal eigenvectors of the sample covariance matrix $\mathbf{R}$, meaning the N_r eigenvectors of $\mathbf{R}$ corresponding to the N_r largest eigenvalues. The centroid condition is designed to minimize the average distortion in the encoding region $\mathcal{R}_n$, when $\mathbf{C}_n^{opt}$ represents $\mathcal{R}_n$ (*Zhou and Li*, 2006).

This process is reproduced here for convenience.

$$\begin{aligned}\mathbf{C}_n^{opt} =& \arg\min_{\mathbf{C}} \frac{1}{N_{\mathcal{R}_n}} \sum_{\mathbf{X}_m \in \mathcal{R}_n} d_c^2(\mathbf{X}_m, \mathbf{C}) \\ =& \arg\min_{\mathbf{C}} \frac{1}{N_{\mathcal{R}_n}} \sum_{\mathbf{X}_m \in \mathcal{R}_n} \mathrm{tr}\left(\mathbf{I}_{N_r} - \mathbf{C}^{\mathcal{H}}\mathbf{X}_m\mathbf{X}_m^{\mathcal{H}}\mathbf{C}\right) \\ =& \arg\max_{\mathbf{C}} \mathrm{tr}\left(\mathbf{C}^{\mathcal{H}}\mathbf{R}\mathbf{C}\right)\end{aligned} \tag{2.24}$$

(2.22) is the optimum solution which minimizes the average distortion.

2.4.3 Modified LBG VQ Algorithm

The modified LBG (mLBG) VQ design algorithm is an iterative algorithm which finds the solution satisfying the two optimality criteria in Section 2.4.2. The algorithm requires an initial codebook $\mathcal{C}^{(0)}$. $\mathcal{C}^{(0)}$ is obtained by the splitting of an initial code, which is the centroid of the entire training sequence, into two codes. The iterative algorithm runs with these two codes as the initial codebook. The final two codes are split into four and the same process is repeated until the desired number of codes, which leads to the minimum chordal distance, is obtained. The codebook design steps are as follows for a given $\mathcal{T}$ and $\epsilon > 0$ ('small' number). $centroid(\mathcal{S})$ denotes the optimum code calculated for a given encoding region $\mathcal{S}$.

1. (Preparation) Let $N = 1$ and calculate $\mathbf{C}_1^* = centroid(\mathcal{T})$.

2. (Splitting) For $i = 1, 2, \ldots, N$, set $\mathbf{C}_i^{(0)} = (1+\epsilon)\mathbf{C}_i^*$, $\mathbf{C}_{N+i}^{(0)} = (1-\epsilon)\mathbf{C}_i^*$ and $N \to 2N$.

3. (Iteration) Set the iteration index $k = 0$ and calculate $d_{c,\min}^{(0)}(\mathcal{C})$.

   (a) Find $n^* = \arg\min_{n\in\{1,\ldots,N\}} d_c(\mathbf{X}_m, \mathbf{C}_n^{(k)})$ for $m = 1, \ldots, M$ and set $\mathcal{Q}(\mathbf{X}_m) = \mathbf{C}_{n^*}^{(k)}$.

   (b) Update the codes by finding the centroid $\mathbf{C}_n^{(k+1)} = centroid(\{\mathbf{X}_m : \mathcal{Q}(\mathbf{X}_m) = \mathbf{C}_n^{(k)}\})$ for $n = 1, \ldots, N$.

   (c) Set $k \to k + 1$.

   (d) Calculate $d_{c,\min}^{(k)}(\mathcal{C})$ and if $d_{c,\min}^{(k)}(\mathcal{C}) > d_{c,\min}^{(k-1)}(\mathcal{C})$, then save $d_{c,\min}^{(k)}(\mathcal{C})$ and $\mathbf{C}_n^{(k)}$ for $n = 1, \ldots, N$, and go back to step (a). Otherwise, go to step (e).

   (e) Set $\mathbf{C}_n^* = \mathbf{C}_n^{(k-1)}$ for $n = 1, \ldots, N$ as the final codes.

4. Repeat steps 2 and 3 until the desired number of codes are obtained, i.e., N is equal to N_{cb}.

In step 1, an initial code is acquired by calculating the centroid of the entire training sequence as preparation. In step 2, an initial codebook is obtained by splitting the initial code acquired in step 1. Splitting can be done by adding perturbation ϵ to the initial code. In step 3 (a), we map the individual training samples to the partition defined by the initial codebook. In step 3 (b), the codes are updated by finding the centroid of the training samples which belong to a certain partition. In step 3 (c), the iteration index is incremented. In step 3 (d), the minimum chordal distance of the acquired codebook is calculated and compared with that of the previous codebook. In case the minimum chordal distance is increased, we take this codebook and continue the iteration procedure. In case it is not increased, we take the previous codebook and proceed to the step 2 to split the codes further. We should repeat the steps 2 and 3 until the desired number of codes are obtained.

The minimum distances of the codebooks are collected in Table 2.1 as an example. A training sequence of the length $M = 50,000$ is used for obtaining the codebook. The SCMe is used to generate the Urban Macro channel samples. It shows that the codebooks acquired by the modified LBG VQ algorithm have better distance properties than the Grassmannian codebooks listed on the webpage about Grassmannian subspace packing.[14]

Table 2.1: The minimum codebook distances $d_{c,\min}(\mathcal{C})$

(N_{tt}, N_r)	n_{CB}	mLBG VQ	Grassmann
$(2,1)$	3	0.3895	0.3820
$(3,1)$	3	0.5706	0.5429
	4	0.4882	0.4167

2.5 Long term Channel Quantization Feedback based on Channel Covariance Matrix

In the previous section, we have investigated the channel quantization method under the assumption that this process is performed on a real time basis, meaning that the channel quantization (CQ) is performed on the instantaneous channel information. This is neither realizable nor required, due to the fact that the channel remains almost constant during the coherence time T_c. Here, we propose a long term CQ feedback method based on the channel covariance matrix $\mathbf{R}_{\mathbf{H}_j}$ as a practical solution for the proposed CQ method.

[14] http://cobweb.ecn.purdue.edu/~djlove/grass.html Note that the Grassmannian codes on this webpage are found by a brute force search and there is no guarantee that those codes are optimum. Hence it is possible to find the codes which are better than these codes in terms of the minimum codebook distance.

The feedback period is set as the coherence time T_c, and the CQ operation is performed not on the instantaneous value of the channel information, but on the average value by making use of the channel covariance matrix. $\mathbf{H}_j(i)\mathbf{H}_j^{\mathcal{H}}(i) = \left(\mathbf{U}_j(i)\mathbf{\Sigma}_j(i)\mathbf{V}_j^{\mathcal{H}}(i)\right)\left(\mathbf{U}_j(i)\mathbf{\Sigma}_j(i)\mathbf{V}_j^{\mathcal{H}}(i)\right)^{\mathcal{H}} = \mathbf{U}_j(i)\mathbf{\Sigma}_j^2(i)\mathbf{U}_j^{\mathcal{H}}(i)$ holds for the instantaneous channel (the time stamp is depicted by the index i, whereas j is the user index). The channel covariance matrix is defined as $\mathbf{R}_{\mathbf{H}_j} = \mathbb{E}\left[\mathbf{H}_j\mathbf{H}_j^{\mathcal{H}}\right] = \mathbf{U}_j\mathbf{\Sigma}_j^2\mathbf{U}_j^{\mathcal{H}}$. We assume that the channel sequence is an ergodic process, and replace $\mathbf{R}_{\mathbf{H}_j}$ with $\widehat{\mathbf{R}}_{\mathbf{H}_j} = \frac{1}{N_w}\sum_{i=1}^{N_w}\left(\mathbf{H}_j(i)\mathbf{H}_j^{\mathcal{H}}(i)\right) = \widehat{\mathbf{U}}_j\widehat{\mathbf{\Sigma}}_j^2\widehat{\mathbf{U}}_j^{\mathcal{H}}$, which calculates the estimate of the covariance matrix by taking channel information acquired at different time samples i. Here, the number of time samples N_w is set so that it coincides with the feedback period, which is taken as the coherence time T_c. The coherence time is set as $T_c = \sqrt{\frac{9}{16\pi f_D^2}} \approx \frac{0.423}{f_D}$, where $f_D = \frac{v_{\mathrm{MS}}}{c}f_c$. Here, f_D is the maximum Doppler frequency, v_{MS} is the mobile speed, and c is the radio propagation speed. Note that the MS j can simply perform an eigenvalue decomposition (EVD) of $\widehat{\mathbf{R}}_{\mathbf{H}_j}$ to extract the unitary matrix $\widehat{\mathbf{U}}_j$ of which the signal subspace spanning part $\widehat{\mathbf{U}}_j^{(\mathrm{S})}$ is used for the channel quantization process (2.9). The clear benefit of adopting a long term CQ feedback is that we can reduce the uplink resources required for CQ feedback without compromising the system performance. Long term CQ feedback also alleviates noise effects on the downlink channel estimation and quantization.

So far, we have assumed that the MS has the perfect downlink channel knowledge. This is not always true in reality. The downlink channel information can be acquired by various channel estimation schemes, and using pilot tones is one of them. In this case, the estimated channel of MS j at time stamp i, $\tilde{\mathbf{H}}_j(i)$, can be expressed as the actual channel information corrupted by AWGN matrix $\mathbf{N}_j(i)$ which is distributed by $\mathcal{CN}(0, \sigma_N^2\mathbf{I})$. In simulations, the noise variance σ_N^2 is acquired by the thermal noise equation $P_N = -174 + 10\log_{10}(W)$ [dBm] where W is the bandwidth for one subcarrier (15 kHz in accordance with the LTE parameters). $\mathbf{R}_{\mathbf{H}_j}$ should be re-formulated as follows.

$$\mathbf{R}_{\tilde{\mathbf{H}}_j} = \mathbb{E}\left[\tilde{\mathbf{H}}_j\tilde{\mathbf{H}}_j^{\mathcal{H}}\right] = \mathbf{U}_j\mathbf{\Sigma}_j^2\mathbf{U}_j^{\mathcal{H}} + \sigma_N^2\mathbf{I} = \mathbf{U}_j\left(\mathbf{\Sigma}_j^2 + \sigma_N^2\mathbf{I}\right)\mathbf{U}_j^{\mathcal{H}}, \text{ where } \tilde{\mathbf{H}}_j(i) = \mathbf{H}_j(i) + \mathbf{N}_j(i) \quad (2.25)$$

Thus EVD of $\mathbf{R}_{\tilde{\mathbf{H}}_j}$ provides us with the same $\mathbf{U}_j$ as the noise-free case provided that a sufficient number of realizations are collected . In our case, we calculate a time average instead. The number of time samples N_w should be large enough to have sufficient noise averaging effects, and be small enough to well represent the current channel status. Certainly there is a trade-off in deciding the feedback period. From (2.25) we can see that the long term CQ feedback scheme based on the channel covariance matrix is capable of reducing the impact of the noise as well as reducing the

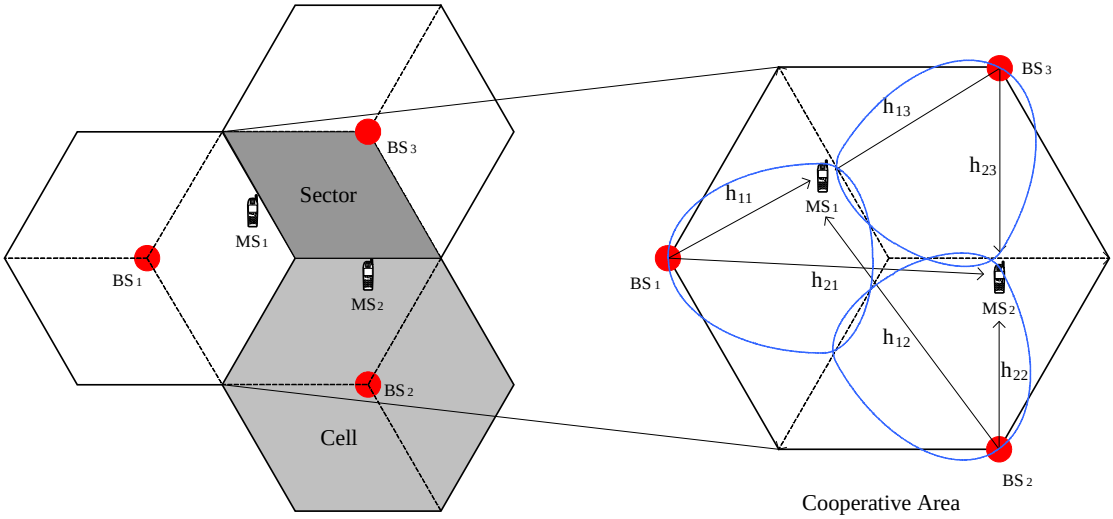


Figure 2.2: CA with 3 BSs and 2 MSs

uplink feedback resources.

2.6 Numerical Results

In this section we present numerical results. First, simulations have been performed for the 2 BSs - 2 MSs and 3 BSs - 2 MSs cases to evaluate the performance of the proposed channel quantization and the codebook construction method. Two (three) BSs are cooperating to transmit data signals to two MSs through the same resources at the same time. Both BSs and MSs have a single antenna, so it yields 2×2 and 2×3 overall channel matrices, respectively. We employ the transmit zero-forcing filter as an example beamforming scheme to prove the quality of the proposed quantization method. The SCME[15] in (*Baum et al.*, 2005) is used for the simulations and the proposed methods are tested for an Urban Macro channel with a mobile speed of 10 m/s. The system performance is evaluated in terms of the received SINR at the MS. Simulations are performed for 30,000 channel realizations and the cumulative distribution function (cdf) at one MS is obtained. OFDMA is assumed as the data transmission scheme and we focus on one subcarrier. The transmit power at the BS is set to be 10 W and it is equally allocated to 1201 subcarriers. The carrier frequency is 2.6 GHz.

The CA topology is same as depicted in Fig. 2.2. As in the conventional cellular topology, one

[15]The MATLAB code provided in the webpage (MATLAB implementation of the interim channel model for Beyond-3G systems (SCME); http://radio.aalto.fi/en/research/rf_applications_in_mobile_communications/propagation_research/matlab_scm_implementation/) supports a channel matrix generation function for links between multiple BSs and multiple MSs.

cell is composed of three sectors and the hexagonal area, which is composed of three sectors which are served by three BSs, forms a CA. Two MSs in the CA are served by three BSs simultaneously. In case of the 2 BSs - 2 MSs case, two BSs which maintain the strongest two links with MSs are chosen for downlink transmission. The cell radius[16] is 600 m and MSs are equally distributed in the CA for every drop[17].

The transmit zero-forcing filter formula follows (2.5), based on downlink channel information which is either perfect channel (pCh), or is provided by a downlink channel estimation method[18] which is shared by the BSs through a prompt and error free backbone network (centralized CA, cCA), or is acquired by the analog pilot retransmission method (distributed CA, dCA), or is captured and reconstructed by looking up an n bit codebook (n bit channel quantization, nbCQ). The cCA case assumes that the system employs a TDD scheme and the backbone network connecting associated BSs is delay free and error free. The downlink channel state information can be acquired by estimating the uplink channel by exploiting the uplink-downlink channel reciprocity, when there exists a direct link between a BS and a MS. In simulations, the uplink channel is assumed to be estimated by using the uplink pilot signal. The non-direct link channel information can be provided by BSs with direct links through prompt data communication over the backbone network. In the analog pilot retransmission method, the MS sends the received pilot which pertains to the downlink channel state information to all associated BSs over the uplink channel. In this case, two pilots are required on the uplink. One is for conveying the received pilot directly to the BSs (analog pilot retransmission), and the other is for estimating the uplink channel itself, which is necessary to compensate the retransmitted pilot for the uplink channel influence so as to acquire the downlink channel information. Therefore, these two pilots should be adjacent in time and frequency. Since the estimated version of the channel state information is used to peel off distortions caused by the uplink channel from the retransmitted pilot, this method is vulnerable to noise enhancement effects.

The BSs are assumed to be aware of the large-scale fading of the channel, and the channel quantization process (2.11) is based on true channel information[19], unless otherwise stated. The codebooks are acquired by the modified LBG VQ algorithm. The feedback link is error free and delay free.

[16]The inter-cell distance can be easily acquired by the CA topology as well, which is a cell radius multiplied by $\sqrt{3}$.

[17]As there is no restriction on distribution of MSs, near-far effects can be observed. However, it should be averaged out over a large number of simulation runs.

[18]In this case, the estimated downlink channel is formulated as an actual channel at the MS (with path loss and fading being considered) plus thermal noise.

[19]Some readers may find a direct comparison between cCA and nbCQ inadequate in the sense that cCA is based on the realistic channel estimation method while nbCQ is based on the ideal channel knowledge. However, the performance of the cCA case is provided here as a mere reference for the mapping of the performance of the proposed method in relation to an alternative method.

First, two proposals concerning the channel quantization model are evaluated. One model ($\hat{\mathbf{H}}_j^d$, Equation (2.11)) adopts the channel directional information only, and the other ($\hat{\mathbf{H}}_j$, Equation (2.10)) takes the channel magnitude information into account, as well as the channel directional information. Fig. 2.3 shows the cdf of the SINR for the 3 BSs - 2 MSs case[20]. The channel directional information based model (4bCDI, 5bCDI) performs closely to, or in the low SINR region even slightly better than the model which combines the directional and magnitude information (4bCDMI, 5bCDMI). We assume that the BSs have access to the link strengths (large-scale fading due to path loss and shadowing) for both cases, and the BSs have perfect knowledge of $\mathbf{\Sigma}_j^{(\mathrm{S})}$ for the latter case. Simulation results indicate that the extra channel magnitude information does not improve the SINR performance in these cases[21]. Please note that in this section the channel quantization (CQ) or the subspace based CQ refers to the channel directional information based model, unless otherwise mentioned.

Fig. 2.4 shows the cdf of the SINR for the 2 BSs - 2 MSs case. At 50 % outage SINR, the 3 bit channel quantization (3bCQ) shows 7.8 dB gain over the analog pilot retransmission case (dCA) and it is only 0.1 dB away from the centralized CA (cCA). The channel matrix at MS j $\mathbf{H}_j (j = 1, 2)$ is in this case a 2×1 complex vector and this is represented by a codebook of size $2^3 = 8$. Compared with the channel quantization method, the resource efficient dCA case requires 3 pilot tones per MS in case of FDD. Therefore, the proposed scheme performs much better than the pilot retransmission method without requiring extra resources. Fig. 2.5 deals with simulation results of the 3 BSs - 2 MSs case. The 3bCQ, 4bCQ, and 5bCQ cases have 3.2 dB, 5.0 dB, and 6.5 dB gains over the dCA case, respectively. In this case, the proposed method still has a lot of room for improvement even though the expected gain over the conventional method is not insignificant. The gap between the proposed method and the ideal case can be reduced by adopting a smart scheduling strategy like user grouping which selects users with orthogonal channel signatures so as to reduce interferences between different users (*Yoo and Goldsmith*, 2006).

The proposed method is to quantize the channel matrix based on the chordal distance, and the LBG VQ algorithm is modified as such. Conventional VQ methods use the Euclidean distance instead. Is the subspace based method better than the conventional method? A performance comparison result is shown in Fig. 2.6. The Euclidean distance based CQ ($n$beCQ) adopts the

[20] Note that MSs have a single antenna.

[21] We should be careful in interpreting the simulation results. We have simulated relatively low bit (4 and 5 bits) quantization cases. In this case, the precision of the channel directional information is more relevant to the system performance than the channel magnitude information. The channel magnitude information can play an important role for higher bit quantization cases, where the accuracy of the channel directional information is sufficiently high that only the magnitude information can help improve performance. In this section, we focus on the limited feedback case in which the channel directional information matters most.

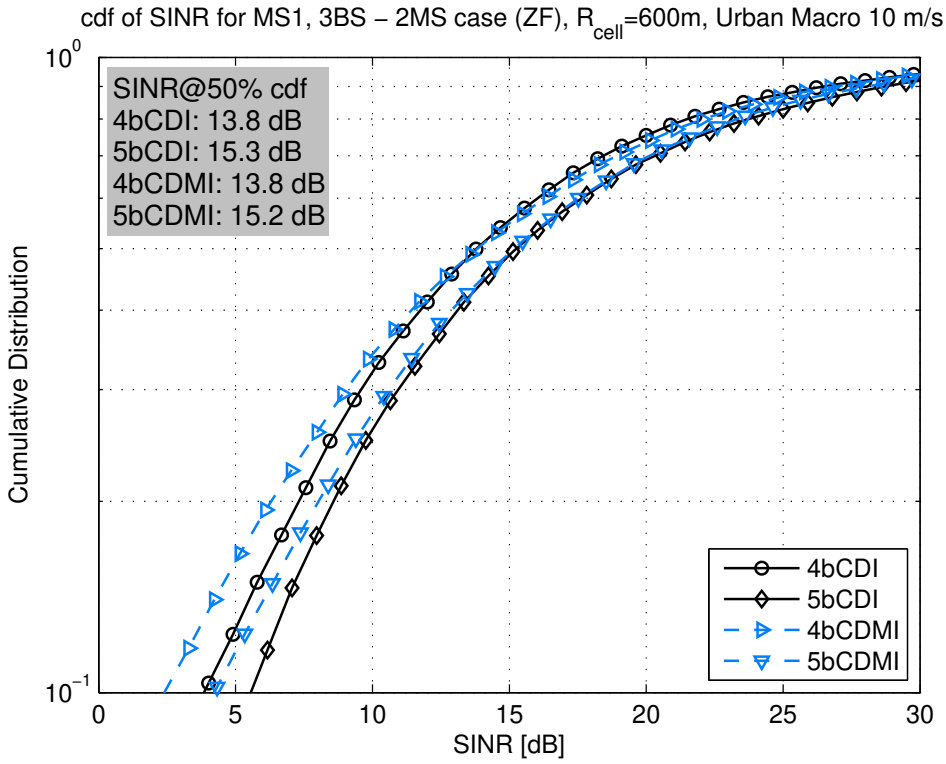


Figure 2.3: Performance comparison of channel quantization models (4bCDI: 4 bit channel directional information based quantization, 5bCDI: 5 bit CDI, 4bCDMI: 4 bit channel directional/magnitude information based quantization, 5bCDMI: 5 bit CDMI)

Euclidean distance as a distance metric for channel quantization. Its optimality criteria for codebook construction are as follows, accordingly.

- *Nearest Neighbor Condition*:

$$\mathcal{R}_n = \{\mathbf{X} : \|\mathbf{X} - \mathbf{C}_n\|_2^2 \leq \|\mathbf{X} - \mathbf{C}_{n'}\|_2^2, \forall n' \neq n\} \tag{2.26}$$

- *Centroid Condition*:

$$\mathbf{C}_n = \frac{\sum_{\mathbf{X}_m \in \mathcal{R}_n} \mathbf{X}_m}{N_{\mathcal{R}_n}}, \text{ where } N_{\mathcal{R}_n} = |\mathcal{R}_n| \tag{2.27}$$

The simulation results show that the subspace based CQ has a substantial gain over the Euclidean distance based CQ. At 50 % outage SINR, the 4bCQ and 5bCQ outperform the 4beCQ and 5beCQ by 2.9 dB and 3.0 dB, respectively.

There exists another CQ method which exploits Givens rotations (*Roh and Rao*, 2004). This method allows us to represent the column space basis vectors $\mathbf{U}^{(\mathrm{S})} \in \mathbb{C}^{t \times n}$ of the channel by

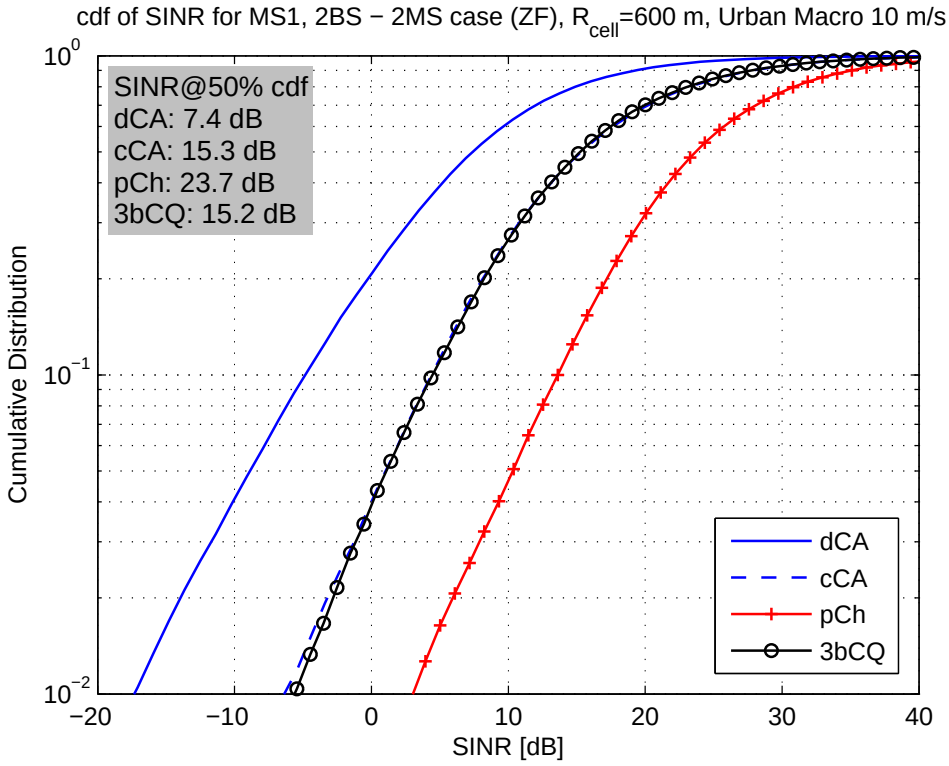


Figure 2.4: 2 BSs - 2 MSs case simulation results (dCA: distributed CA (analog pilot retransmission method), cCA: centralized CA, pCh: perfect channel, 3bCQ: 3 bit CQ)

$(2t-1)n - n^2$ real numbers. $t = 3, n = 1$ holds for the 3 BSs - 2 MSs case, and it requires 4 real number parameters ($\phi_{1,2}$, $\phi_{1,3}$, $\theta_{1,1}$, $\theta_{1,2}$) for channel matrix construction. The performance comparison result between the proposed method and the Givens rotation based channel matrix decomposition method is shown in Fig. 2.7. The Givens rotation based method with n bit feedback is denoted by nbGR. 4bGR allocates 1 bit for each parameter, and 5bGR assigns 2,2,1 and 0 bit(s) for $\phi_{1,2}$, $\phi_{1,3}$, $\theta_{1,1}$ and $\theta_{1,2}$, respectively[22]. At 50 % outage SINR, the 4bCQ case outperforms the 4bGR case by 2.5 dB, while the 5bCQ case shows comparable performance to the 5bGR case of which the computational complexity at MS is higher than that of the 5bCQ case. The Givens rotation based method is investigated in more detail in Chapter III.

Simulations have been performed to evaluate the long term CQ feedback (4bCQ) scheme based on the channel covariance matrix which is introduced in Section 2.5. Simulations have been done for various feedback periods by adjusting N_w and its result is shown in Fig. 2.8. N_w is the number of time samples which are used for channel covariance matrix calculation. We can conclude that the performance improves as N_w decreases due to the fact that as N_w increases the feedback cannot

[22]In this case, the value of $\theta_{1,2}$ is predefined and fixed.

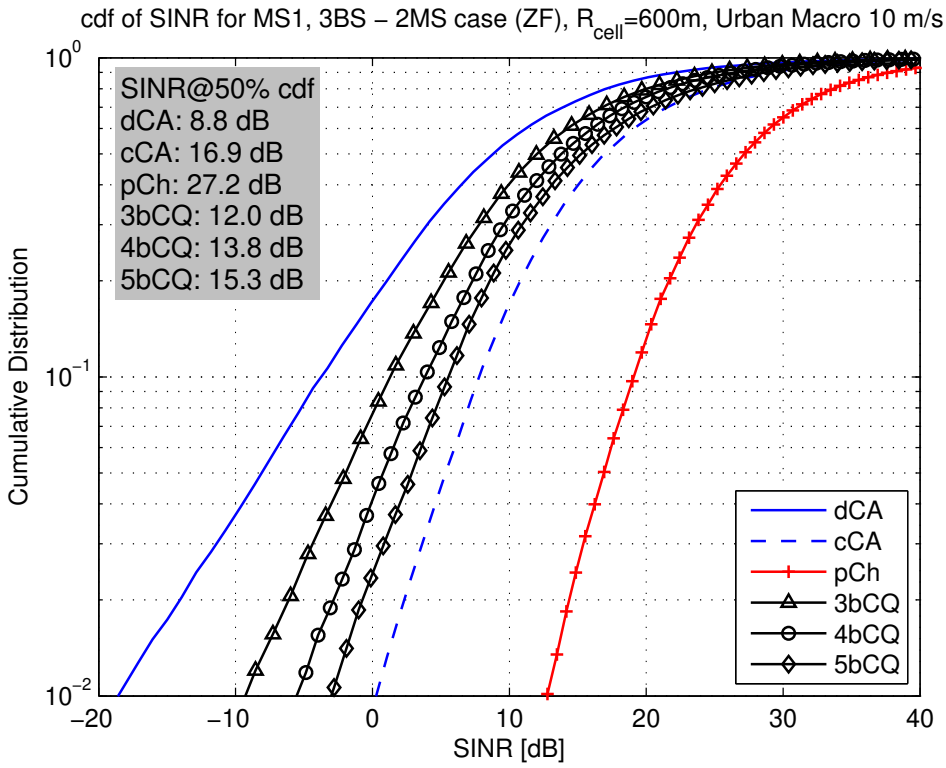


Figure 2.5: 3 BSs - 2 MSs case simulation results (dCA: distributed CA (analog pilot retransmission method), cCA: centralized CA, pCh: perfect channel, 3bCQ: 3 bit CQ, 4bCQ: 4 bit CQ, 5bCQ: 5 bit CQ)

capture time evolution effects of the channel in a timely manner. The $N_w = 1$ case can be considered as a reference case, since it sends back feedback messages for every symbol (instantaneous CQ feedback). The $N_w = 40$ case is 1.2 dB away from the reference case ($N_w = 1$) in terms of the 50 % outage SINR, whereas the $N_w = 20$ case is only 0.7 dB from the reference case. There certainly exists a trade-off between gains due to the reduced uplink resource by adopting long term feedback and the system performance. The necessary resource for CQ feedback can be reduced to $\frac{1}{20}$ for the $N_w = 20$ case compared with the $N_w = 1$ case, with only a 0.7 dB degradation.

As practical considerations, the channel estimation noise is taken into account in simulations for the 3 BSs - 2 MSs case and the result is depicted in Fig. 2.9. The channel covariance matrix based long term 4 bit CQ feedback scheme is employed with $N_w = 20$. The channel estimation noise is generated and hinders the channel estimation/quantization operation, but its effect is mitigated by noise averaging. The case without channel estimation noise is also shown to serve as a reference. We should note that there is only a 0.5 dB performance degradation at 50 % outage SINR, and the two cases are close together for the whole SINR region.

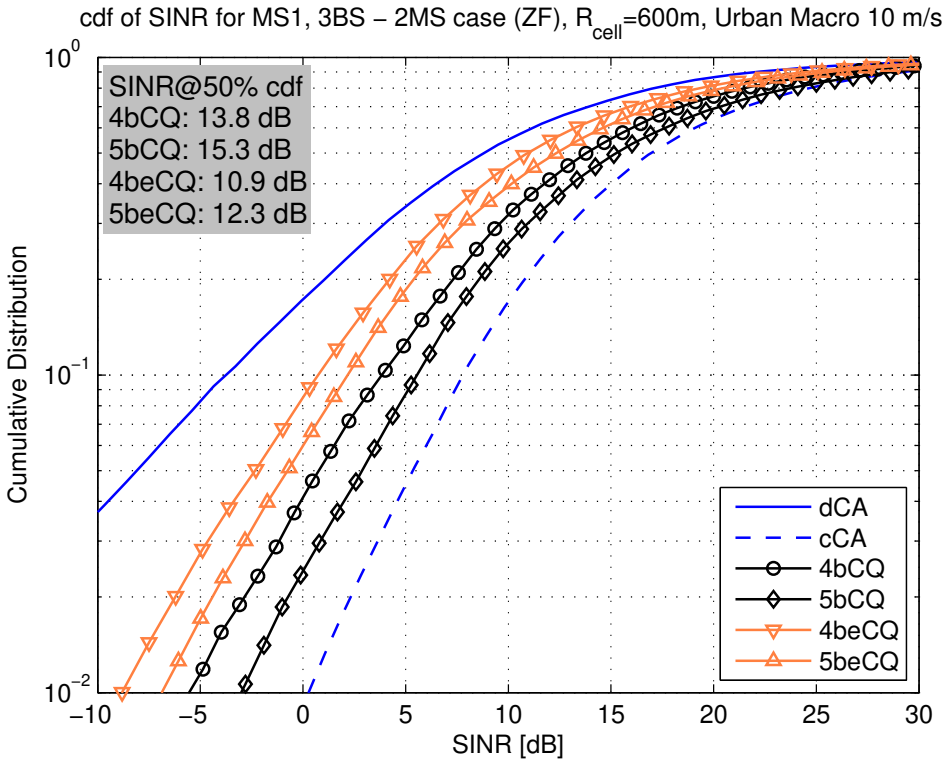


Figure 2.6: Performance comparison of the Euclidean distance based CQ and the Subspace based CQ (dCA: distributed CA (analog pilot retransmission method), cCA: centralized CA, 4bCQ: 4 bit subspace based CQ, 5bCQ: 5 bit subspace based CQ, 4beCQ: 4 bit Euclidean distance based CQ, 5beCQ: 5 bit Euclidean distance based CQ)

Fig. 2.10 shows the simulation results for the 2 BSs - 2 MSs case. 3 bit CQ feedback is tested for the following four cases: instantaneous 3bCQ feedback without channel estimation noise, instantaneous 3bCQ feedback with channel estimation noise, channel covariance matrix based long term 3bCQ feedback without channel estimation noise, and channel covariance matrix based long term 3bCQ feedback with channel estimation noise. The performance of the tested cases is summarized in Table 2.2, which can be interpreted as follows.

- Difference between instCh & instChwN: 2.6 dB

  A downlink channel estimation error caused by thermal noise is not ignorable when noise averaging is not adopted.

- Difference between instCh & statCh: 0.7 dB

  Both cases are without channel estimation noise. The degradation comes from channel time evolution, which is tolerable.

- Difference between statCh & statChwN: 0.7 dB

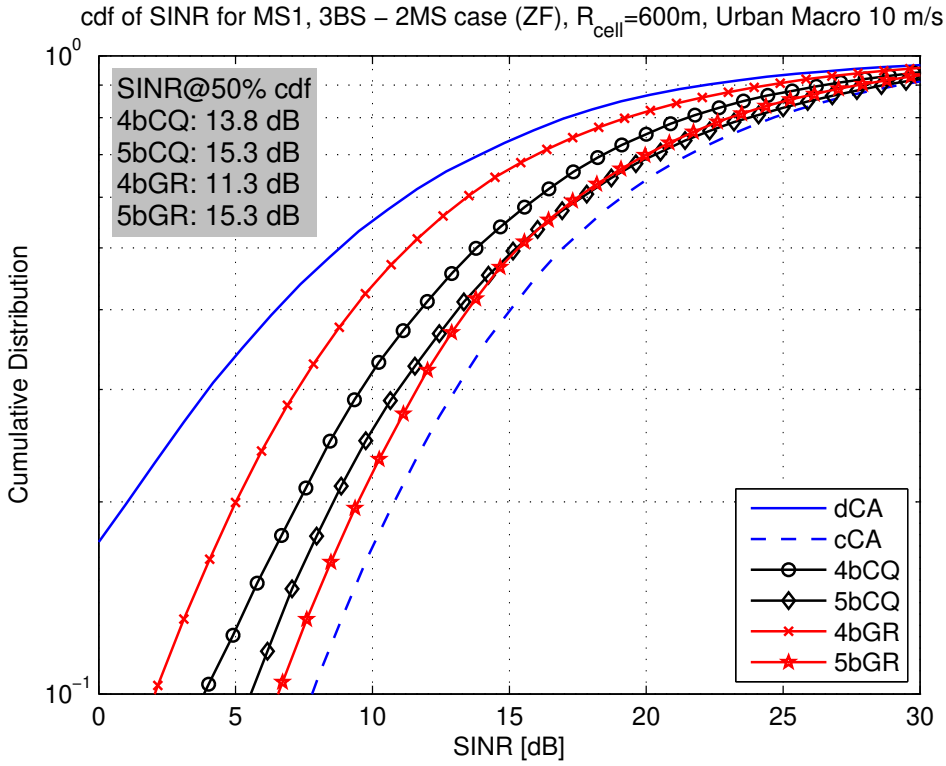


Figure 2.7: Performance comparison of the Givens rotation based CQ and the Subspace based CQ (dCA: distributed CA (analog pilot retransmission method), cCA: centralized CA, 4bCQ: 4 bit subspace based CA, 5bCQ: 5 bit subspace based CQ, 4bGR: 4 bit Givens rotation based CQ, 5bGR: 5 bit Givens rotation based CQ)

The difference is not as big as the difference between instCh & instChwN. The channel estimation error effects are ameliorated by introducing the channel covariance matrix.

- Difference between instCh & statChwN: 1.4 dB

  The ideal case vs. the reality - Considering that statChwN case reflects noisy channel estimation (accountable for 0.7 dB degradation) and channel evolution effects (accountable for another 0.7 dB degradation). Performance degradation is acceptable. Moreover, the feedback overhead of the statChwN case is 5 % of that of the instCh case. Even with a 1.4 dB degradation compared with instCh, statChwN shows a 6.5 dB gain over dCA case.

2.7 Conclusion

In this chapter, we have investigated precoded MU-MIMO systems with limited feedback. The subspace based channel quantization method is proposed as a way of providing BSs with downlink

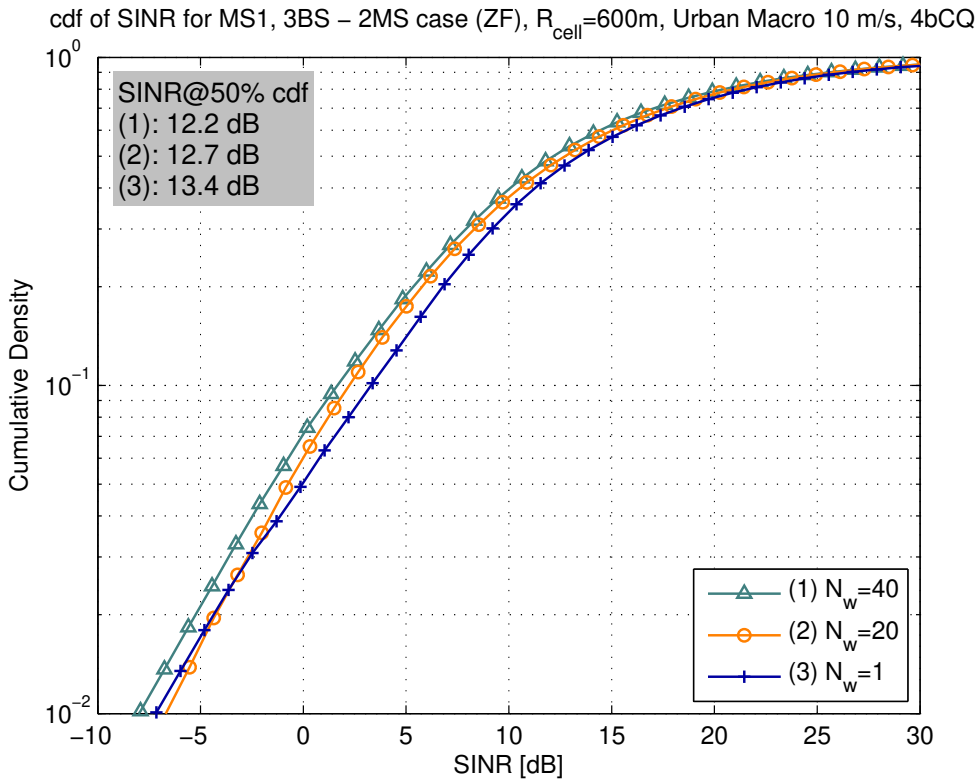


Figure 2.8: Performance comparison of the long term CQ feedback based on channel covariance matrix

channel state information in the presence of inter-user interference, which is applicable to the distributed cooperative antenna systems as well as MU-MIMO systems. The subspace based channel quantization improves the system performance significantly, compared to the analog pilot retransmission method with relatively small feedback overhead. We have also developed an efficient codebook construction algorithm which quantizes the channel directional information based on the well known LBG VQ by adopting the chordal distance and modifying the optimality criteria accordingly. The codebooks generated by the proposed algorithm have better distance properties than Grassmannian codebooks that are currently available.

We have also proposed a channel covariance matrix based long term channel quantization feedback method which takes the practical implementation issue into account. It is shown that the proposed method is robust against channel estimation noise which is inevitable in the field application, whereas it requires only a small amount of uplink feedback resources. Finding a way to further reduce the feedback overhead might be an interesting area of investigation.

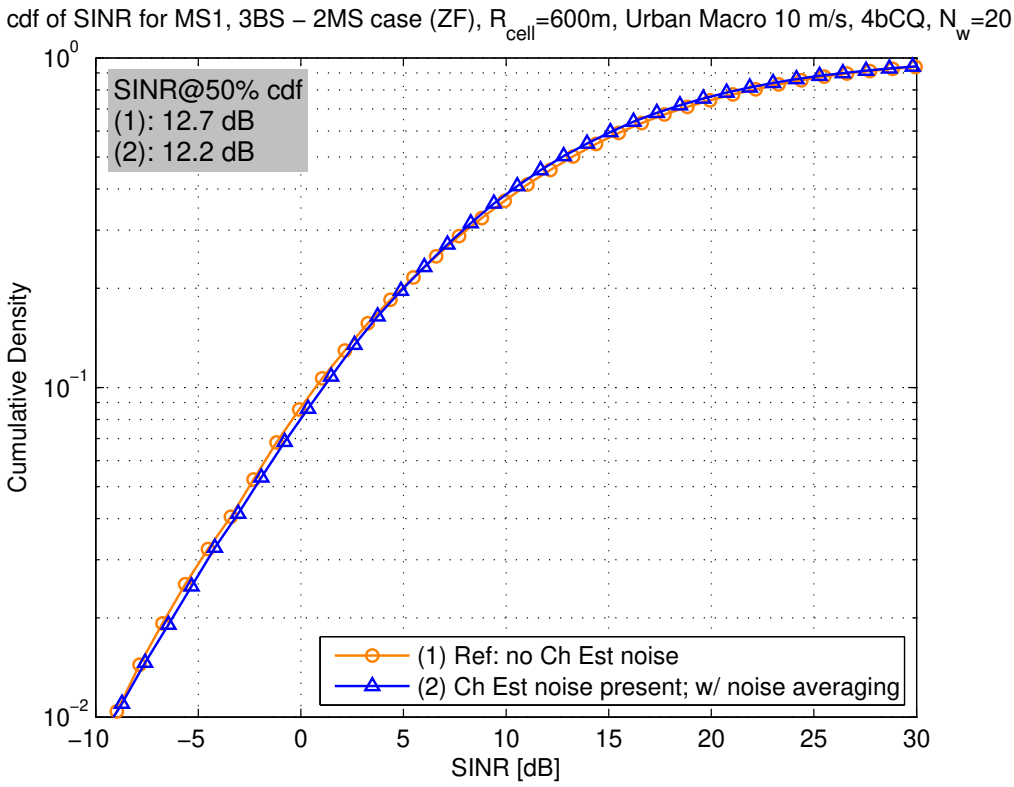


Figure 2.9: Performance comparison of the long term CQ feedback based on channel covariance matrix with/without channel estimation noise

Table 2.2: Instantaneous/long term CQ feedback (3bCQ) performance and feedback overhead, 2 BSs - 2 MSs case with v_{MS}=10 m/s

Case	50 % outage SINR [dB]	feedback overhead [bits/sym/subcarrier]
instCh	15.4	3.00
instChwN	12.8	3.00
statCh	14.7	0.15
statChwN	14.0	0.15

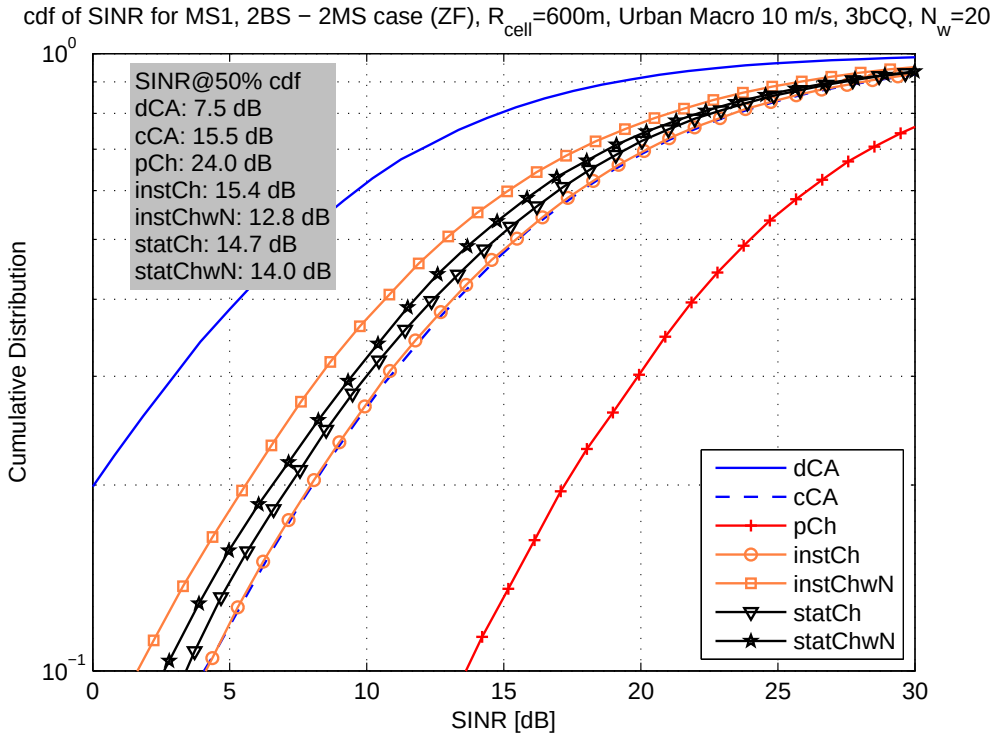


Figure 2.10: Performance comparison of the long term CQ feedback and the instantaneous CQ feedback with/without channel estimation noise (instCh: instantaneous 3bCQ feedback w/o channel estimation noise, instChwN: instantaneous 3bCQ feedback w/ channel estimation noise, statCh: long term 3bCQ feedback w/o channel estimation noise, statChwN: long term 3bCQ feedback w/ channel estimation noise)

CHAPTER III

Givens Rotation based Channel Quantization Method for Slowly Time-varying Channels

This chapter introduces an alternative method of reducing the amount of feedback required for channel quantization. This method, which is proposed in (*Roh and Rao*, 2004), is based on Givens parameterization and quantization of the channel parameters. In the thesis we propose a Givens parameter tracking scheme which addresses a practical deployment issue of the method. The channel matrix $\mathbf{H}$ can be decomposed to reveal its eigenvectors by using a SVD. As a result, we can acquire a unitary matrix which conveys channel spatial information. The geometrical structure of this unitary matrix is exploited to extract a set of parameters that has a one-to-one mapping to the original matrix. In this parameterization process, we use a Givens rotation based method, which has advantages over a Householder reflection based method in terms of the computational complexity, the number of the required parameters, and the quantization noise. The Givens parameter can be further quantized by using a simple 1-bit Adaptive Delta Modulation (ADM) to allow further reduction of redundancy in time/frequency, which is intended to exploit a time/frequency correlation of the channel. Since ADM provides a parameter tuning and tracking capability and a Givens parameter is a real number which has a limited value, Givens parameters can be quantized with only a small number of bits.

This chapter is organized as follows. The differential Givens (D-Givens) method, which consists of a differential Givens decomposition of the unitary matrix followed by an adaptive delta modulation, is briefly described in Section 3.1. Section 3.2 deals with the practical implementation issue, and the simulation results of the performance of the D-Givens method can be found in Section 3.3. The Givens parameter feedback frequency adaptation method, which is one of the major contributions of the author, is proposed and evaluated by simulations in Section 3.4.

The simulation results of this chapter have been in part shown in (*Kim et al.*, 2008c).

3.1 Channel Information Parameterization and Quantization

3.1.1 Givens Rotation

The channel matrix $\mathbf{H} \in \mathbb{C}^{r\times t}$ can be decomposed by using a SVD, $\mathbf{H} = \mathbf{U}_H\mathbf{\Sigma}_H\mathbf{V}_H^{\mathcal{H}}$, where $\mathbf{U}_H \in \mathbb{C}^{r\times r}$ and $\mathbf{V}_H \in \mathbb{C}^{t\times t}$ are unitary matrices and $\mathbf{\Sigma}_H \in \mathbb{R}^{r\times t}$ contains n singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$ of $\mathbf{H}$, when the rank of $\mathbf{H}$ is n. $\mathbf{H}$ can be rewritten by using a compact form, which is $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathcal{H}}$, where $\mathbf{U} \in \mathbb{C}^{r\times n}$, $\mathbf{V} \in \mathbb{C}^{t\times n}$ and $\mathbf{\Sigma} \in \mathbb{R}^{n\times n}$ are the first n columns of $\mathbf{U}_H$, the first n columns of $\mathbf{V}_H$, and the n by n upper left corner matrix of $\mathbf{\Sigma}_H$, respectively.

As the channel directional information at the receiver which is contained in $\mathbf{V}$ matters most, we now focus on how to extract essential parameters from the unitary matrix $\mathbf{V}$. Since the columns in $\mathbf{V}$ are geometrically structured, meaning that $\mathbf{V}$ can be expressed as a product of diagonal matrices and Givens matrices, the degrees of freedom in the matrix are much smaller than the number of real entries in the matrix. In (*Roh and Rao*, 2004), it is shown that the degrees of freedom in $\mathbf{V} \in \mathbb{C}^{t\times n}$ are $N = (2t-1)n - n^2$ (real numbers), which is smaller than $2tn$[1]. The next step is to extract a set of parameters that has a one-to-one mapping to the matrix $\mathbf{V}$. A Givens rotation based method provides a way in which the number of parameters is equal to the degree of freedom in the matrix (*Roh and Rao*, 2004).

A matrix $\mathbf{V} \in \mathbb{C}^{t\times n}(t \geq n)$ with orthonormal columns can be decomposed as

$$\mathbf{V} = \left[\prod_{k=1}^{n} \mathbf{D}_k(\phi_{k,k}, \cdots, \phi_{k,t}) \prod_{l=1}^{t-k} \mathbf{G}_{t-l,t-l+1}(\theta_{k,l})\right] \tilde{\mathbf{I}}, \tag{3.1}$$

where a diagonal matrix of size t by t

$$\mathbf{D}_k(\phi_{k,k}, \cdots, \phi_{k,t}) = \mathrm{diag}(\mathbf{1}_{k-1}, e^{j\phi_{k,k}}, \cdots, e^{j\phi_{k,t}}) \tag{3.2}$$

$\mathbf{1}_{k-1}$ is $(k-1)$ 1's and $\mathbf{G}_{p-1,p}(\theta)$ is the Givens matrix which operates in the $(p-1,p)$ coordinate

[1]The degrees of freedom is calculated as $N = 2tn - n - 2\binom{n}{2} - n$, in which $2tn$ is the number of real entries of a $t \times n$ complex matrix, from which n and $2\binom{n}{2}$ are subtracted due to the orthonormal column condition ($\mathbf{v}_i^{\mathcal{H}}\mathbf{v}_i = 1$ and $\mathbf{v}_i^{\mathcal{H}}\mathbf{v}_j = 0$ $(i \neq j)$ where $\mathbf{v}_i, \mathbf{v}_j$ are column vectors of $\mathbf{V}$.) and n is subtracted as one phase in each column can be made fixed such that the first row can be chosen to be a nonnegative real number for instance.

plane of the form

$$\mathbf{G}_{p-1,p}(\theta) = \begin{bmatrix} \mathbf{I}_{p-2} & & & \\ & c & -s & \\ & s & c & \\ & & & \mathbf{I}_{t-p} \end{bmatrix}, \tag{3.3}$$

$c = \cos\theta$, $s = \sin\theta$, and $t \times n$ matrix $\tilde{\mathbf{I}} = [\mathbf{I}_n, \mathbf{0}_{n,t-n}]^{\mathcal{T}}$. The number of the phases $\{\phi_{k,l}\}$ and the rotational angles $\{\theta_{k,l}\}$ are $n(2t-n+1)/2$ and $n(2t-n-1)/2$, respectively. Hence the total number of Givens parameters is $2tn - n^2$. Assume that $\mathbf{V}$ has nonnegative real-valued elements in the first row. In this case, the first parameter of $\mathbf{D}_k(\phi_{k,k}, \cdots, \phi_{k,t})$ is zero, i.e., $\phi_{k,k} = 0$ for all k. Therefore, the total number of parameters of $\mathbf{V}$ is $(2t-1)n - n^2$ and it is equal to the degrees of freedom.

The above decomposition procedure is explained with an example of a 4×3 orthonormal column matrix $\mathbf{V}$ (*Roh and Rao*, 2007).

$$\begin{aligned}
\mathbf{V} = &\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{D}_1{}^{\mathcal{H}}} \begin{bmatrix} |\times| & \times & \times \\ |\times| & \times & \times \\ |\times| & \times & \times \\ |\times| & \times & \times \end{bmatrix} \xrightarrow{\mathbf{G}_{3,4}{}^{\mathcal{H}}} \begin{bmatrix} |\times| & \times & \times \\ |\times| & \times & \times \\ |\times| & \times & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{\mathbf{G}_{2,3}{}^{\mathcal{H}}} \\
&\begin{bmatrix} |\times| & \times & \times \\ |\times| & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{\mathbf{G}_{1,2}{}^{\mathcal{H}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{\mathbf{D}_2{}^{\mathcal{H}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & |\times| & \times \\ 0 & |\times| & \times \\ 0 & |\times| & \times \end{bmatrix} \xrightarrow{\mathbf{G}_{3,4}{}^{\mathcal{H}}} \\
&\begin{bmatrix} 1 & 0 & 0 \\ 0 & |\times| & \times \\ 0 & |\times| & \times \\ 0 & 0 & \times \end{bmatrix} \xrightarrow{\mathbf{G}_{2,3}{}^{\mathcal{H}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \times \\ 0 & 0 & \times \end{bmatrix} \xrightarrow{\mathbf{D}_3{}^{\mathcal{H}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & |\times| \\ 0 & 0 & |\times| \end{bmatrix} \xrightarrow{\mathbf{G}_{3,4}{}^{\mathcal{H}}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \tilde{\mathbf{I}}
\end{aligned} \tag{3.4}$$

$|\times|$ denotes the magnitude of a particular element. Basically we want to convert $\mathbf{V}$ into $\tilde{\mathbf{I}}$ by applying several sequences of $\mathbf{D}_k$ and $\mathbf{G}_{p-1,p}$ matrices. As the first step, we should make all the entries in the first column under the first component zeros. This can be done first by having a real valued column by multiplying $\mathbf{V}$ by $\mathbf{D}_1^{\mathcal{H}}$, and then by applying a series of Givens matrices ($\mathbf{G}_{3,4}{}^{\mathcal{H}}$, $\mathbf{G}_{2,3}{}^{\mathcal{H}}$ and $\mathbf{G}_{1,2}{}^{\mathcal{H}}$) to make all entries under the (1,1) element zeros. As $\mathbf{V}$ is an orthonormal column

matrix and the Givens rotation preserves the length of vector, the (1,1) element becomes 1 and all the entries of the first row except the (1,1) element are made zeros. We continue similar procedures on the remaining columns one by one until we get $\tilde{\mathbf{I}}$. In the end, the matrix $\mathbf{V}$ can be factored as follows.

$$\mathbf{V} = \mathbf{D}_1(0, \phi_{1,2}, \phi_{1,3}, \phi_{1,4})\mathbf{G}_{3,4}(\theta_{1,1})\mathbf{G}_{2,3}(\theta_{1,2})\mathbf{G}_{1,2}(\theta_{1,3})\cdot\mathbf{D}_2(0, \phi_{2,3}, \phi_{2,4})\mathbf{G}_{3,4}(\theta_{2,1})\mathbf{G}_{2,3}(\theta_{2,2})\cdot\mathbf{D}_3(0, \phi_{3,4})\mathbf{G}_{3,4}(\theta_{3,1})\tilde{\mathbf{I}} \tag{3.5}$$

The bottom line is that once we have a set of parameters, i.e., the phases $\{\phi_{k,m}\}$ and the rotational angles $\{\theta_{k,l}\}$, the original orthonormal column matrix can be exactly reconstructed.

For another example, when $t = r = 2, n = 1$, $\mathbf{V} \in \mathbb{C}^{2\times 1}$ can be factored as

$$\mathbf{V} = \mathbf{D}_1(0, \phi_{1,2})\mathbf{G}_{1,2}(\theta_{1,1})\tilde{\mathbf{I}},$$
$$\text{where } \mathbf{D}_1(0, \phi_{1,2}) = \begin{bmatrix} 1 & 0 \\ 0 & e^{j\phi_{1,2}} \end{bmatrix}, \mathbf{G}_{1,2}(\theta_{1,1}) = \begin{bmatrix} \cos\theta_{1,1} & -\sin\theta_{1,1} \\ \sin\theta_{1,1} & \cos\theta_{1,1} \end{bmatrix}, \text{ and } \tilde{\mathbf{I}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{3.6}$$

In case $t = r = 3, n = 1$, $\mathbf{V} \in \mathbb{C}^{3\times 1}$ can be formulated as follows.

$$\mathbf{V} = \mathbf{D}_1(0, \phi_{1,2}, \phi_{1,3})\mathbf{G}_{2,3}(\theta_{1,1})\mathbf{G}_{1,2}(\theta_{1,2})\tilde{\mathbf{I}},$$
$$\text{where } \mathbf{D}_1(0, \phi_{1,2}, \phi_{1,3}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{j\phi_{1,2}} & 0 \\ 0 & 0 & e^{j\phi_{1,3}} \end{bmatrix}, \mathbf{G}_{2,3}(\theta_{1,1}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{1,1} & -\sin\theta_{1,1} \\ 0 & \sin\theta_{1,1} & \cos\theta_{1,1} \end{bmatrix}, \tag{3.7}$$
$$\mathbf{G}_{1,2}(\theta_{1,2}) = \begin{bmatrix} \cos\theta_{1,2} & -\sin\theta_{1,2} & 0 \\ \sin\theta_{1,2} & \cos\theta_{1,2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } \tilde{\mathbf{I}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Regarding the statistics of the parameters, the following is known (*Roh and Rao*, 2004). When the channel matrix $\mathbf{H}$ has *independent and identically distributed (i.i.d.)* $\mathcal{CN}(0, 1)$ entries, then all the Givens parameters are statistically independent. Moreover, the phase $\phi_{k,l}$ is uniformly distributed over $(-\pi, \pi]$ for all k and l, and the rotational angle $\theta_{k,l}$ has a probability density

$$\begin{aligned} p(\theta_{k,l}) =& 2l \sin^{2l-1}\theta_{k,l} \cos\theta_{k,l},\ 0 \le \theta_{k,l} < \frac{\pi}{2} \\ & \text{for } 1 \le k \le n, 1 \le l \le t - k. \end{aligned} \tag{3.8}$$

The graphical representation of this function is shown in Fig. 3.1. We can see that as the index l

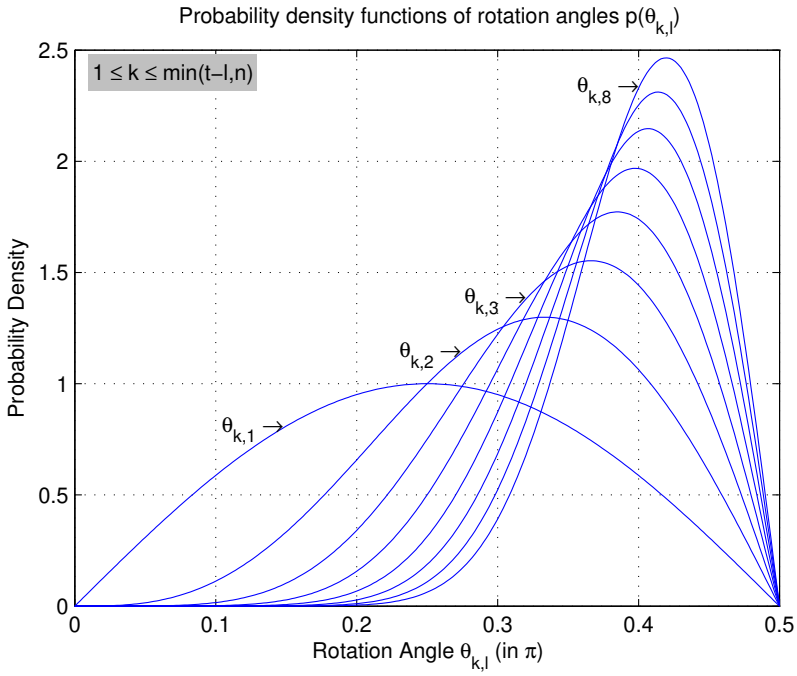


Figure 3.1: Probability density functions of the rotation angles

increases the distribution of the rotation angle θ is more concentrated, which can be exploited when designing a quantization.

3.1.2 Bit Allocation and Quantization

In the system set-up, we consider a distributed antenna system with an isotropically scattered channel. Hence all the Givens parameters are statistically independent, and the phases $\{\phi_{k,l}\}$ and the rotational angles $\{\theta_{k,l}\}$ are scalars which have bounded values; $\phi \in (-\pi, \pi]$, $\theta \in [0, \pi/2)$. Since the uplink resources are limited, we need to quantize the Givens parameters so as to reduce the required amount of feedback bits to represent the parameters. The statistical independence of all parameters being quantized significantly simplifies the bit allocation problem, eliminating any need for a complicated joint vector quantization scheme. In other words, we employ a scalar quantization of the individual parameters. The proposed scalar quantization has the following advantages. First, the number of parameters to quantize is minimal as it is equal to the degrees of freedom in the spatial information. Moreover, the parameters are all bounded quantities.

Regarding the bit allocation to the parameter, the optimum bit allocation function can be found in (*Dharamdial and Adve*, 2005). However, there are some restrictions in exploiting this optimum bit

allocation function. First of all, it requires the knowledge of the instantaneous channel realization, which is not available at the bit allocation phase. Secondly, it is dependent on the weighting factor which recognizes the fact that not all Givens parameters equally impact the reconstructed unitary matrix. It is very critical to get precise weighting factors, as will be shown in the following sections, but we cannot find a detailed explanation about how to get the weighting factors in (*Dharamdial and Adve*, 2005). Moreover, as will be shown in the numerical results, the adaptive delta modulation plays a more important role than the initial bit allocation to the performance. Therefore, we employ a uniform or random bit allocation rather than the optimum bit allocation for the initial parameter quantization.

Concerning the parameter quantization for the given bit, we employ a uniform and statistical quantization for the phase and the rotational angle, respectively. As mentioned in the previous section, the phase $\phi_{k,j}$ is uniformly distributed over $(-\pi, \pi]$ for all k and j, and the distribution of the rotational angle $\theta_{k,l}$ follows the probability density function $p(\theta_{k,l})$ (3.8), when the fully scattered channel is considered. Our strategy is, hence, to use the uniformly distributed quantization points for the phase, and to use the non-uniformly distributed quantization points for the rotational angle. We can realize a non-uniform distribution of the quantization points by taking the probability density function into account or by applying the training based codebook generation algorithm like LBG algorithm (*Linde et al.*, 1980).

3.1.3 Adaptive Delta Modulation

The ADM is widely used to quantize a slowly varying scalar value. In slowly time-varying channels, the corresponding Givens parameters are also slowly and continuously changing most of the time, which makes the ADM eligible to quantize the Givens parameters. The encoder of the ADM consists of a simple accumulator and a one-bit quantization. The working principle of ADM is as follows (*Aldajani and Sayed*, 2001).

Assume that we want to construct a signal $\hat{\phi}[k]$ that tracks a signal $\phi[k]$. This can be achieved according to the following construction. At each instant of time, we start with the value $\hat{\phi}[k-1]$ and update it to $\hat{\phi}[k]$ so that this new value is closer to $\phi[k]$ than its old value. The update is based on the difference between $\phi[k]$ and $\hat{\phi}[k-1]$, defined by

$$e_a[k] = \phi[k] - \hat{\phi}[k-1]. \tag{3.9}$$

The signal $\hat{\phi}[k-1]$ is increased or decreased by a positive amount $\Delta[k]$ depending on the encoder

output history and the sign of the error (3.9). The step-size $\Delta[k]$ of the one-bit quantization is adaptively changing to better track the dynamics of the signal. The step-size is increased if the consequent two encoded bits are same, and decreased otherwise, that is,

$$\Delta[k] = \begin{cases} \alpha\Delta[k-1] & \text{if } c[k] = c[k-1] \\ \frac{1}{\alpha}\Delta[k-1] & \text{if } c[k] \neq c[k-1] \end{cases} \tag{3.10}$$

where $\Delta[k]$ and $c[k] = \text{sign}[e_a[k]] \in \{-1, +1\}$ are the step-size and the encoded bit for the k-th sample, respectively. α is a system parameter, which satisfies $\alpha > 1$. The sign of the error, $e_a[k]$, decides whether $\hat{\phi}[k-1]$ increases or decreases at each time instant. Thus, the signal $\hat{\phi}[k]$ is varied according to the adaptation rule:

$$\hat{\phi}[k] = \hat{\phi}[k-1] + \text{sign}[e_a[k]]\Delta[k]. \tag{3.11}$$

An observation of the step-size $\Delta[k]$ reveals the following equivalent form.

$$\Delta[k] = \alpha^{w[k]}\Delta[0], \tag{3.12}$$

where

$$w[k] = w[k-1] + q[k] \tag{3.13}$$

and

$$q[k] = c[k]c[k-1]. \tag{3.14}$$

Here $w[k]$ ($w[0] = 0$) can be considered as a weight factor which reflects the history of the delta modulation trend (3.13) and is used to decide the step-size (3.12). This alternative representation allows us to describe the scheme for updating $\hat{\phi}[k]$ in a block diagram form, as shown in Fig. 3.2. The upper and lower parts of the figure implement equations (3.12) and (3.11), respectively. Notice the fact that system parameters α and $\Delta[k]$ are known to the ADM encoder and decoder, and only the encoded bit $c[k]$ is required for the receiver to decode the quantized value $\hat{\phi}[k]$. The ADM is a low-rate scalar quantization scheme (as low as one bit per parameter), and it has inherently a channel tracking feature for the slowly varying channels over time. Note that tracking is vulnerable to feedback errors which can lead to error propagation. As these effects affect the step size as well as the decoded value itself, its impact can be significant.

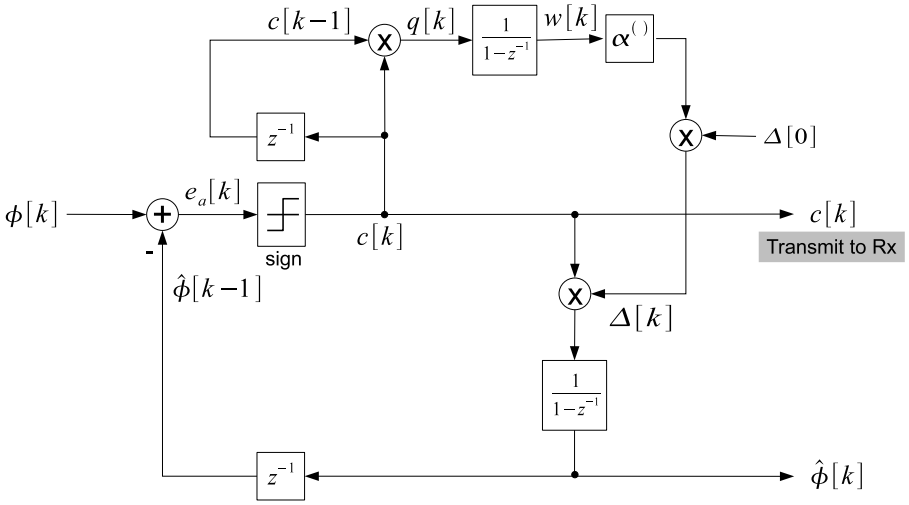


Figure 3.2: Conventional ADM Encoder Schematic

3.2 Implementation Issue

The ADM is an efficient scheme to track the correlated signal with a limited resource. In case of tracking the Givens parameters, which are angular values, it seems to be an appropriate method since they have bounded values, e.g., phase $\phi \in (-\pi, \pi]$, and rotational angle $\theta \in [0, \pi/2)$, and there exists correlation over time or frequency[2] (in case of multi-carrier transmission scheme like OFDM).

However, observing the time variance of the phase value ϕ reveals the fact that it shows discontinuities when it approaches the border (π or $-\pi$). It disappears at a certain point and promptly appears on the other side of the border. This behavior comes from the fact that π and $-\pi$ are equivalent in terms of the angular value. The problem here is that even when the signal itself is slowly varying, the bounded value representing it into a scalar can show abrupt changes, which may fail the effort of the ADM encoder to encode the signal differences. This *cyclic overflow*[3], which is caused by the modular feature of the phase value ($\theta = \theta + 2\pi n$, where $\theta \in (-\pi, +\pi]$ and n is integer), should be taken into account when designing the ADM encoder/decoder in case of tracking the phase value ϕ of the Givens parameters. We can simply eliminate the adverse effect of cyclic overflow by making the ϕ graph continuous prior to the quantization procedure by adding or subtracting 2π at cyclic overflow points. In this case, the value of ϕ is no longer bounded in $(-\pi, \pi]$. Please refer to Section 3.4.2.1 for more information.

[2] In this section, we focus on exploiting temporal correlation only.

[3] Cyclic periodicity can be a better term in that it indicates the cause of its behavior. However, we stick to the term "cyclic overflow" throughout this document, as it better captures its behavior observed in the polar coordinate.

3.3 Numerical Results

Numerical results are presented in this section. Simulations have been performed for the 3 BSs - 2 MSs and 2 BSs - 2 MSs cases to evaluate the performance of the proposed differential Givens (D-Givens) method. Three or two BSs are cooperating to transmit a data signal for two MSs through the same resources at the same time. Both BSs and MSs have a single antenna, so it yields 3×2 and 2×2 overall channel matrices, which correspond with 3×1 and 2×1 channel unitary matrices[4] at the MS, respectively. The Givens decomposition of the channel unitary matrix $\mathbf{V}$ provides 4 Givens parameters, i.e., $\phi_{1,2}$, $\phi_{1,3}$, $\theta_{1,1}$ and $\theta_{1,2}$, in case of $\mathbf{V} \in \mathbb{C}^{3\times1}$, and 2 Givens parameters, i.e., $\phi_{1,2}$ and $\theta_{1,1}$, in case of $\mathbf{V} \in \mathbb{C}^{2\times1}$.

We employ the transmit zero-forcing filter as a beamforming scheme. The 3GPP SCMe is used for the simulations and the proposed methods are tested for an Urban Macro channel with a mobile speed of 3 m/s and 10 m/s. The system performance is evaluated in terms of the received SINR at the MS. Simulations are performed for 100,000 channel realizations and the cdf at one MS is obtained. OFDMA is assumed as the data transmission scheme and we focus on one subcarrier. The transmit power at the BS is set to be 10 W and it is equally allocated to 1201 subcarriers. The carrier frequency is 2.6 GHz. A feedback link is assumed to be error-free and delay-free.

Fig. 3.3 shows the simulation results conducted to compare performances between the Givens rotation based method and the channel vector quantization method for the $\mathbf{V} \in \mathbb{C}^{2\times1}$ case. The mobile speed is set to be 10 m/s which is equivalent to the maximum Doppler frequency $f_D = 86.7$ Hz. The Givens rotation based parameter extraction and quantization method is denoted by GR with the number of allocated bits to the individual Givens parameters indicated by $(n_{\phi_{1,2}}, n_{\theta_{1,1}})$ where n_X denotes the number of bits allocated to quantize a Givens parameter X. The subspace based channel vector quantization method is denoted by CQ. In both methods, quantization is performed per every symbol and the ADM is not employed in case of the GR. The feedback overhead is hence same for both methods, i.e., GR and CQ[5].

The simulation results shown in Fig. 3.3 lead us to the following conclusions.

- System performance improves as the total number of allocated bits increases.

  In Fig. 3.3, we can check that the 5bGR cases outperform the 4bGR cases which in turn show the better result than the 3bGR case. It is well expected as the higher number of total allocated bits allows the higher precision for Givens parameter quantization. The impact of

[4] A channel unitary matrix can be acquired by performing SVD on the channel matrix.

[5] Note that the total number of allocated bits for quantization is indicated by a preceding number at the graph label, e.g., N_{tot} in N_{tot}bGR or N_{tot}bCQ

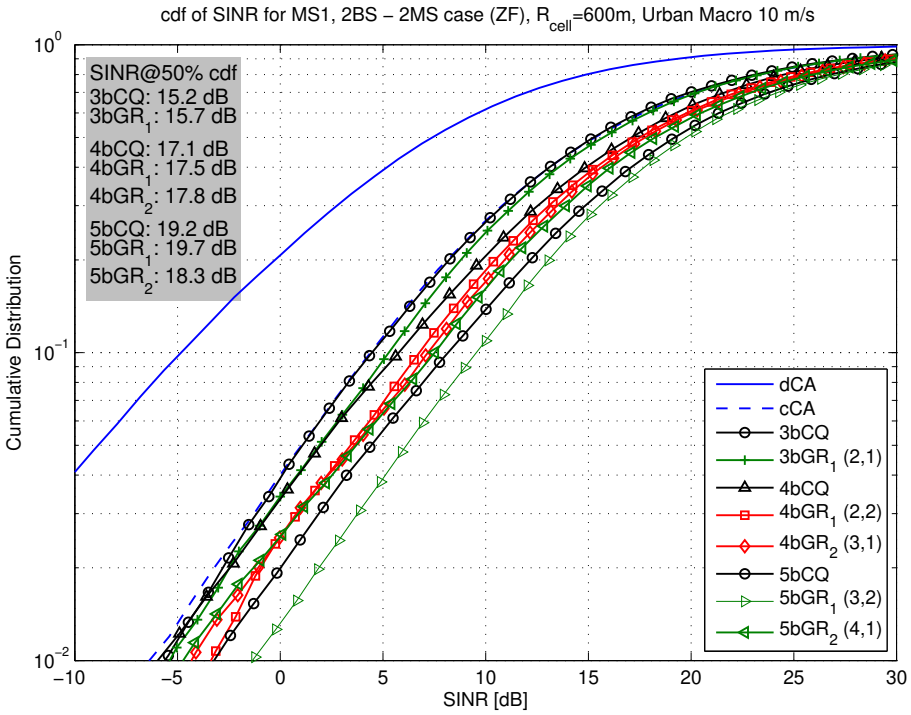


Figure 3.3: Performance comparison: Givens rotation based method and channel vector quantization method, $\mathbf{V} \in \mathbb{C}^{2\times 1}$ case (dCA: distributed CA, cCA: centralized CA, nbCQ: n bit subspace based Channel Quantization method, nbGR ($n_{\phi_{1,2}}, n_{\theta_{1,1}}$): n bit Givens Rotation based method with ($n_{\phi_{1,2}}, n_{\theta_{1,1}}$) bit allocation for individual Givens parameters)

the individual Givens parameter bit allocation for the given total budget will be discussed shortly.

- The Givens rotation based method outperforms, or at least is comparable with the subspace based channel quantization method in terms of the overall performance for the correctly selected Givens parameter bit allocation case.

  In most cases the GR case outperforms the CQ case over the entire SINR region, but the performance gap is not more than 0.7 dB in terms of 50 % outage SINR. The performance of the GR case varies over different combinations of the Givens parameter bit allocation for the given total bit budget, e.g., 5bGR$_1$ shows 1.4 dB gain over 5bGR$_2$ at 50 % outage SINR in Fig. 3.3.

- The Givens parameter bit allocation poses a noticeable impact to the performance in case that the total bit budget is limited in relation with the dimension of the channel matrix to be quantized.

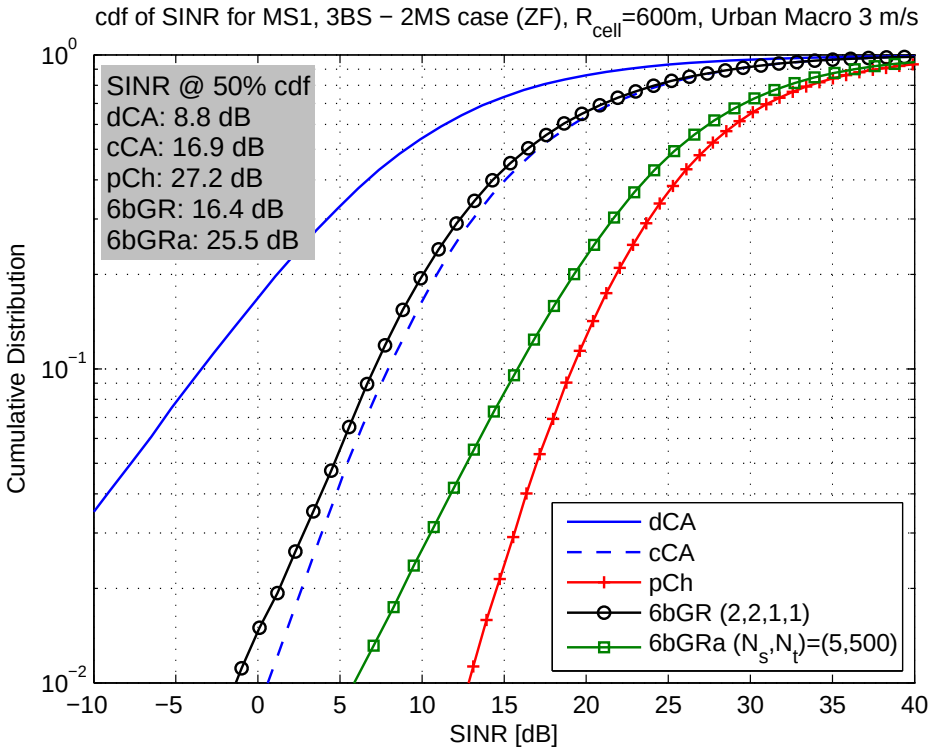


Figure 3.4: Performance comparison: D-Givens method and Givens method, $\mathbf{V} \in \mathbb{C}^{3\times 1}$ case (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, 6bGR $(n_{\phi_{1,2}}, n_{\phi_{1,3}}, n_{\theta_{1,1}}, n_{\theta_{1,2}})$: 6 bit Givens rotation based method with $(n_{\phi_{1,2}}, n_{\phi_{1,3}}, n_{\theta_{1,1}}, n_{\theta_{1,2}})$ bit allocation for individual Givens parameters, 6bGRa (N_s, N_t): 6 bit Givens rotation based method followed by ADM with parameter tracking interval N_s and parameter quantization interval N_t)

For example, the 5bGR_1 (3,2) case shows a 1.4 dB gain over the 5bGR_2 (4,1) case at 50 % outage SINR.

As is discussed in the previous section, it is not a trivial task to find the optimum bit allocation function, which requires the knowledge of the instantaneous channel realization and the Givens parameter weighting factor according to (*Dharamdial and Adve*, 2005). Observation on the simulation results reveals the following. Givens parameter bit allocation plays an important role when available bit resource is limited relative to the channel dimension.

Fig. 3.4 and Fig. 3.5 show the simulation results conducted to compare performances between the D-Givens method and the Givens method for the $\mathbf{V} \in \mathbb{C}^{3\times 1}$ case and the $\mathbf{V} \in \mathbb{C}^{2\times 1}$ case, respectively. The mobile speed is set to be 3 m/s ($f_D = 26.0$ Hz) and 10 m/s ($f_D = 86.7$ Hz), respectively. The Givens rotation based parameter extraction and quantization method is denoted by GR with the number of allocated bits to individual Givens parameters indicated by $(n_{\phi_{1,2}}, n_{\phi_{1,3}}, n_{\theta_{1,1}}, n_{\theta_{1,2}})$ in

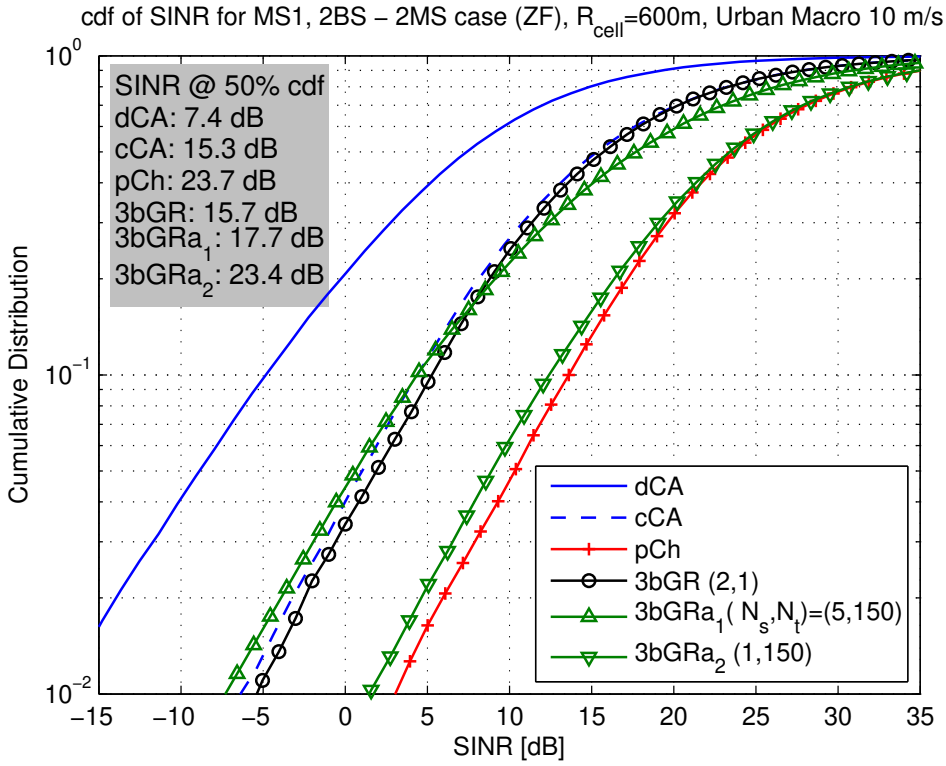


Figure 3.5: Performance comparison: D-Givens method and Givens method, $\mathbf{V} \in \mathbb{C}^{2\times 1}$ case (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, 3bGR $(n_{\phi_{1,2}}, n_{\theta_{1,1}})$: 3 bit Givens rotation based method with $(n_{\phi_{1,2}}, n_{\theta_{1,1}})$ bit allocation for individual Givens parameters, 3bGRa (N_s, N_t): 3 bit Givens rotation based method followed by ADM with parameter tracking interval N_s and parameter quantization interval N_t)

case of $\mathbf{V} \in \mathbb{C}^{3\times 1}$ and $(n_{\phi_{1,2}}, n_{\theta_{1,1}})$ in case of $\mathbf{V} \in \mathbb{C}^{2\times 1}$. In this case, quantization is performed once per symbol and the feedback overhead is N_{tot} bits/sym[6] in case of N_{tot}bGR. The differential Givens method, in which n_X bit Givens parameter quantization is followed by a 1 bit ADM, is denoted by GRa with the parameter quantization interval indicated by N_t and the parameter tracking interval indicated by N_s (unit: the number of OFDM symbols, i.e., multiple of the OFDM symbol duration $(T_s = 71.37\mu s)$). The operation practice is as follows. At first, Givens parameter extraction is performed and an individual Givens parameter X is quantized by using n_X bits. From the next feedback instance on, Givens parameter extraction is performed and the Givens parameter X is modulated by using 1 bit adaptive delta modulation scheme taking the previous Givens parameter value as one of the inputs. Just in case the ADM scheme loses track of the parameter values, the Givens parameter quantization is performed every N_t symbols. For example, 6bGRa (5, 500) case

[6]The precise unit should be bits/sym/subcarrier, as we focus on one subcarrier at the moment. In this section we use bits/sym for brevity. The feedback overhead reduction method for multiple subcarrier case will be introduced later.

in Fig. 3.4 requires 6 bits per every 500 OFDM symbols and 4 bits per every 5 OFDM symbols (1 bit for each Givens parameter).

The simulation results shown in Fig. 3.4 and Fig. 3.5 lead us to the following conclusions.

- The D-Givens method performs much better than the Givens method, and requires less feedback overhead.

  Fig. 3.4 shows that the 6bGRa (5,500) case outperforms the 6bGR case with a 9.1 dB gain at 50 % outage SINR. The required feedback overhead for the 6bGRa (5,500) case is 0.804 bits/sym, while 6.0 bits/sym is required for the 6bGR case. In Fig. 3.5, we can check that the 3bGRa_2 (1,150) case performs better than the 3bGRa_1 (5,150) case by 5.7 dB at 50 % outage SINR and the 3bGRa_1 (5,150) case in turn outperforms the 3bGR case by 2.0 dB at 50 % outage SINR. The required feedback overhead is 2.0067 bits/sym, 0.4067 bits/sym and 3.0 bits/sym for the 3bGRa_2 (1,150), the 3bGRa_1 (5,150) and the 3bGR case, respectively. In short, the GRa case (D-Givens method) can achieve a much better performance than the GR case (Givens method) with less feedback overhead. The D-Givens method yields a better performance, as the ADM scheme is better positioned in tracking the parameter value in the presence of the strong correlation amongst parameter samples. As a result the ADM scheme is capable of providing a value which is closer to the actual parameter value than quantizing the parameter directly without exploiting correlation with the previous value.

  We can also find out that the parameter tracking interval N_s plays a crucial role in the performance, e.g., the 3bGRa_2 (1,150) case outperforms the 3bGRa_1 (5,150) case by 5.7 dB at 50 % outage SINR, which leads us to the following conclusion.

- The parameter tracking interval N_s plays a crucial role in the performance in case of the D-Givens method.

  The smaller the parameter tracking interval becomes, the higher the correlation between parameter samples becomes, and the more likely it is that the D-Givens method provides a closer value to the actual Givens parameter value.

  The ADM scheme adopted in the D-Givens method shows a parameter tuning capability as well as the tracking capability, and the tuning capability is dependent on the parameter tracking interval N_s.

Additional simulations have been performed to assess the impact of the initial bit allocation and the Givens parameter tracking interval for the D-Givens case. The objective is to find out

which factor has a bigger impact to the performance in case the D-Givens method is employed. The simulation set-up is as follows: $\mathbf{V} \in \mathbb{C}^{2\times 1}$ with v_{MS}=10 m/s. The D-Givens methods with various N_s values for the fixed N_t value have been simulated for the two different initial bit allocation cases, i.e., (2,1) and (3,2). Table 3.1 depicts the performance result in terms of 50 % and 5 % outage SINR, and Table 3.2 summarizes the required feedback overhead for the corresponding cases. The column-wise difference Δ_{col} depicts the gap between different initial bit allocation cases, and the row-wise difference Δ_{row} shows the largest difference among the various parameter tracking interval N_t cases. The graphical presentation of the simulation results can be found in Fig. 3.6. The simulation results of the reference Givens method are summarized in Table 3.3.

The results of the Givens method in Table 3.3 provide us with the impression of the bit allocation impact, as the ADM scheme is not adopted in this case. In short, the performance gap at 50 % outage SINR between the 3(2,1) case and the 5(3,2) case is 4.0 dB and the feedback overhead difference is 66.67 %. The corresponding values for the D-Givens case are 1.0 dB and 3.27 %, respectively. Please note that the selected values are the biggest values among the simulation results and they can be as low as 0.3 dB (when N_s=2) and 0.66 % (when N_s=1). We can conclude that the difference in the key performance indicators (50 % outage SINR and feedback overhead) becomes smaller when adopting the ADM scheme to the Givens method.

On the other hand, Table 3.1 and Table 3.2 yield the following conclusion. Comparing Δ_{row} and Δ_{col} shows that the difference over the various parameter tracking intervals (Δ_{row}) is much bigger than the difference over the different initial bit allocation cases (Δ_{col}). For example, according to Table 3.1 Δ_{row} for the 3(2,1) case is 5.7/8.3 dB, whereas Δ_{col} for the N_s=2 case is just 0.3/0.1 dB. Table 3.2 also reveals that the Δ_{row} for 3(2,1) case is 393.41 %, while Δ_{col} for N_s=1 case is only 0.66 %.

Based on the simulation result analysis stated above, we can draw the following conclusion.

- For systems adopting the D-Givens method, the Givens parameter tracking interval has a profound impact on the performance and the feedback overhead amount, whereas the initial bit allocation to Givens parameters plays a minimal role.

  By observing Fig. 3.6, we can confirm that the Givens parameter tracking interval rather than the initial bit allocation is a decisive factor in the performance. There are in large two groups of graphs, which are a better performing group (3bGRa$_2$ (1,150) and 5bGRa$_2$ (1,150)) and a worse performing group (3bGRa$_1$ (5,150) and 5bGRa$_1$ (5,150)). The performance deciding factor is the parameter tracking interval N_s, and the initial bit allocation brings about only a

minimal difference in performance.

Two Givens parameter tracking interval decision methods will be discussed in the following section.

Table 3.1: D-Givens (GRa) performance (50 %/5 % outage SINR in dB) over various N_s values, $\mathbf{V} \in \mathbb{C}^{2\times 1}$ with v_{MS}=10 m/s, N_t=150

	N_s				
$N_{tot}(n_{\phi_{1,2}}, n_{\theta_{1,1}})$	1	2	3	5	Δ_{row} [dB]
3(2,1)	23.4/8.9	22.6/7.5	20.6/4.3	17.7/0.6	5.7/8.3
5(3,2)	23.7/9.7	22.9/7.6	21.6/5.5	18.7/1.9	5.0/7.8
Δ_{col} [dB]	0.3/0.8	0.3/0.1	1.0/1.2	1.0/1.3	

Table 3.2: D-Givens (GRa) feedback overhead (bits/sym) over various N_s values, $\mathbf{V} \in \mathbb{C}^{2\times 1}$ with v_{MS}=10 m/s, N_t=150

	N_s				
$N_{tot}(n_{\phi_{1,2}}, n_{\theta_{1,1}})$	1	2	3	5	Δ_{row} [%]
3(2,1)	2.0067	1.0067	0.6733	0.4067	393.41
5(3,2)	2.02	1.02	0.6867	0.4200	380.95
Δ_{col} [%]	0.66	1.32	1.99	3.27	

Table 3.3: Givens (GR) performance and feedback overhead, $\mathbf{V} \in \mathbb{C}^{2\times 1}$ with v_{MS}=10 m/s

$N_{tot}(n_{\phi_{1,2}}, n_{\theta_{1,1}})$	50 % outage SINR	feedback overhead
3(2,1)	15.7 dB	3 bits/sym
5(3,2)	19.7 dB	5 bits/sym
Δ_{col}	4.0 dB	66.67 %

3.4 Givens Parameter Tracking

This section presents the Givens parameter feedback frequency adaptation method that exploits the empirical normalized autocovariance function (nACvF) of the Givens parameters over time, which is one of the major contributions of the author. This method helps us to find the basic requirements for the feedback frequencies of the Givens parameters which guarantee the same system performance as in the reference case, even if the comparison case has a different channel profile (different mobile speed, for example), provided that the empirical nACvFs of the Givens parameters are known. The proposed method is evaluated by means of simulations.

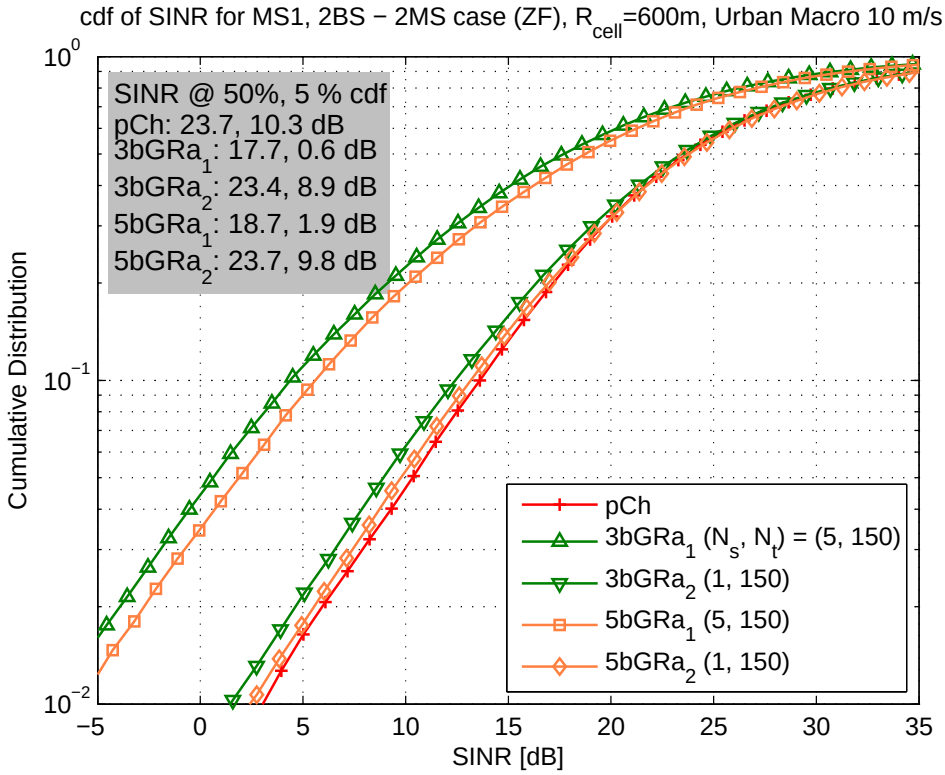


Figure 3.6: D-Givens performance comparison: impact of initial bit allocation and Givens parameter tracking interval, $\mathbf{V} \in \mathbb{C}^{2\times 1}$ case (pCh: perfect channel, nbGRa (N_s, N_t): n bit Givens rotation based method followed by ADM with parameter tracking interval N_s and parameter quantization interval N_t)

3.4.1 Motivation

The Givens parameter extraction method followed by the adaptive delta modulation (ADM) based 1 bit quantization is a resource efficient feedback strategy for transmit filtering in MU-MIMO systems (*Roh and Rao*, 2004). This method has proven to perform very close to the perfect feedback case due to the channel tuning and tracking capability of the ADM, in case that the 1 bit quantization feedback frequency is high enough to track the temporal change of the channel (slowly-varying channel case).

The question here is how precisely we should quantize and track these parameter values. The accuracy of the Givens parameter quantization is related with the number of bits allocated to the parameters, and the Givens parameter tracking performance is dependent on the feedback frequency. The phase ϕ is uniformly distributed over $(-\pi, \pi]$ and the rotational angle θ is distributed with a certain probability density function over $[0, \pi/2)$. Furthermore, all parameters are statistically

independent, which simplifies the bit allocation problem (*Dharamdial and Adve*, 2005). Regarding the feedback frequency problem, we have to take into account the parameter variation over time. This temporal variation may show a different behavior for different parameters, since individual Givens parameters, which compose the channel unitary matrix after reconstruction, may vary differently over time even though the channel matrix experiences temporal change as a whole.

Fig. 3.7 and Fig. 3.8 show snapshots of the temporal behavior of the Givens parameters. In both cases, we can observe that the range of the phase ϕ value is $(-\pi, \pi]$ whereas that of the rotational angle θ value is $[0, \pi/2)$. We can also see that parameters show different temporal variation. At first glance, it appears to be that the phase changes more rapidly over time compared with the rotational angle. On the other hand, looks can be deceiving and we have to take into account the range of the parameters in which they are varying, when comparing the amount of the temporal changes of each parameter. In other words, we need to find an appropriate measure of the temporal variation in order to make full use of this property for finding an optimum feedback period of each parameter. As a result, we can organize the feedback frequency of the individual Givens parameter in such a way that the system performance is maximized in terms of the received SINR with minimal expense of the feedback resources. In this section, we focus on the investigation on decision of the feedback frequency of the Givens parameters, since the parameter bit allocation affects the system performance less than the feedback frequency. For a slowly-varying channel, the channel parameter tuning/tracking capability by 1 bit ADM feedback plays a more important role to the performance than the initial parameter quantization resolution.

3.4.2 Decision of Givens Parameter Tracking Feedback Frequency

In this section, we explain a method to quantify the temporal correlation of the Givens parameters and propose a feedback frequency decision strategy in use of this method.

3.4.2.1 Quantification of Temporal Correlation of Givens Parameters

Our goal is to develop a method which provides a strategy to decide a feedback frequency for each Givens parameter, and it is required to have a tool for numerical comparison of the temporal correlation between the parameters. Hence, we need to quantify the temporal correlation of the Givens parameters. For this purpose, we employ a normalized autocovariance function (nACvF) which is also known as the autocorrelation coefficient function (*Dunn*, 2010). The autocovariance can be thought of as a measure of how similar a signal is to a time-shifted version of itself. When

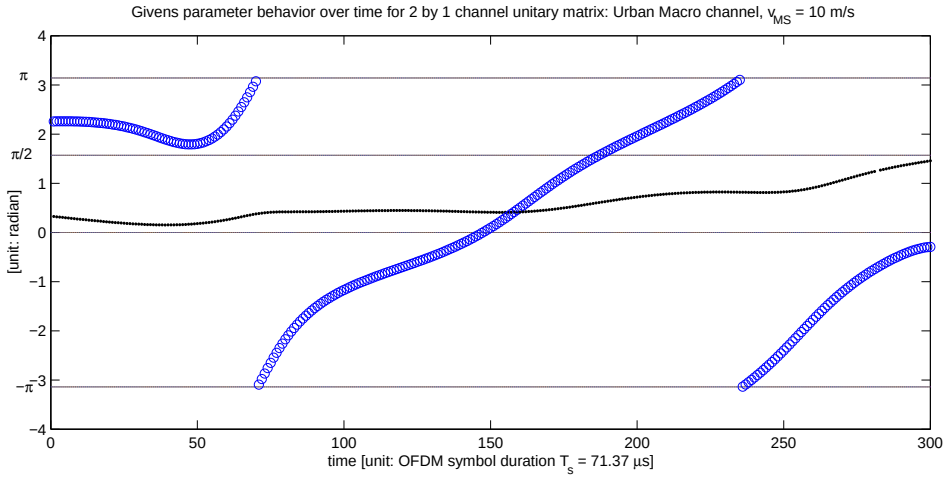


Figure 3.7: Givens parameter behavior over time for the channel unitary matrix $\mathbf{V} \in \mathbb{C}^{2\times 1}$ case, ◯: $\phi_{1,2}$, •: $\theta_{1,1}$

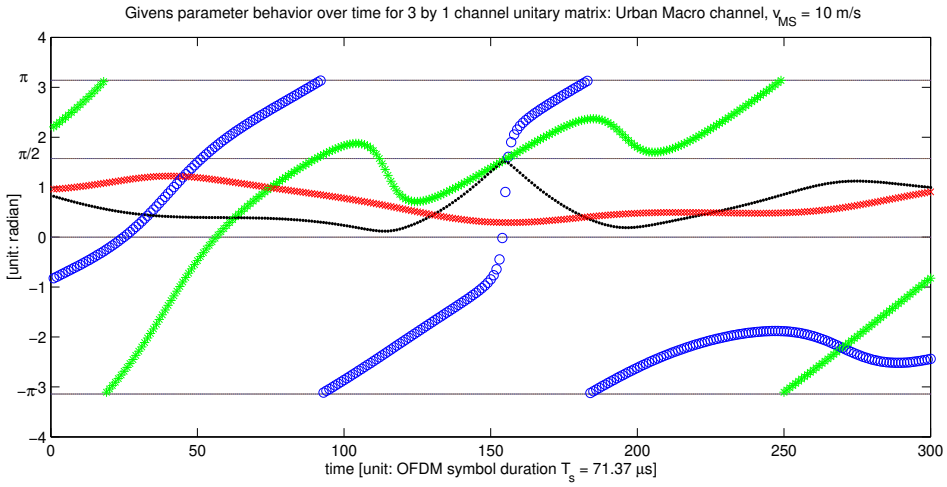


Figure 3.8: Givens parameter behavior over time for the channel unitary matrix $\mathbf{V} \in \mathbb{C}^{3\times 1}$ case, ◯: $\phi_{1,2}$, ∗: $\phi_{1,3}$, •: $\theta_{1,1}$, ×: $\theta_{1,2}$

the value of the autocovariance is equal to that of its variance, it indicates perfect correlation at that lag. The normalization with the variance will put this into the range $[-1, 1]$, with 1 indicating perfect correlation and -1 indicating perfect anti-correlation.

Let $X(t)$ be the value of the process at time t (where t may be an integer for a discrete-time process or a real number for a continuous-time process). Here we assume that $X(t)$ is wide-sense stationary, which means that the nACvF depends only on the time difference k and its mean value can be expressed as $\mathbb{E}[X(t)] = \mathbb{E}[X(t+k)] = \mu_X$. If $X(t)$ has a mean μ_X and a variance σ_X^2 then the definition of the nACvF of the real value process $X(t)$ is

$$c_{XX}(k) = \frac{\mathbb{E}[(X(t) - \mu_X)(X(t+k) - \mu_X)]}{\sigma_X^2}, \tag{3.15}$$

where $\mathbb{E}[\cdot]$ is the expected value operator. For a discrete time series of length N with known mean and variance, an estimate of the normalized autocovariance[7] may be obtained as follows.

$$\hat{c}_{XX}(k) = \frac{1}{N\sigma_X^2} \sum_{t=1}^{N} (X(t) - \mu_X)(X(t+k) - \mu_X) \tag{3.16}$$

The nACvF of the Givens parameters can be obtained by simply replacing X with the Givens parameter of our interest (ϕ or θ) in (3.16). We should note that the phase ϕ appears to have a discontinuous value as can be seen in Fig. 3.7 and 3.8, but it should be considered that its value has cyclic periodicity. In other words, two phase sample points, which are located nearby $-\pi$ and π borders respectively, have similar values even though they look like being far apart from each other in the graph. In our case, we solve this problem by reconstructing the phase graph, denoted by $\phi_{\mathrm{rc}}(t)$, in a way that it shows continuous temporal advancement without distorting its value. Fig. 3.9 and Fig. 3.10 show snapshots of the temporal behavior of the phase values: before and after the ϕ-reconstruction process. As a result, we can obtain continuously varying ϕ data of which the value has been unchanged due to the 2π periodicity property ($\phi = \phi + 2\pi n$, where n is integer) of the angular distance.

The nACvF of the Givens parameters (θ and ϕ) can be formulated by replacing X with θ or ϕ_{rc} in (3.16), and they are introduced below for convenience. Here, k is a positive integer representing the temporal separation or delay, which is multiple of the OFDM symbol duration ($T_s = 71.37\mu s$). Its unit is the number of the OFDM symbols: $k = [\tau] = \lfloor \tau / T_s \rfloor$, where τ is the temporal separation

[7] In this section, the terms 'normalized autocovariance' and 'estimate of the normalized autocovariance' are used interchangeably. The normalized autocovariance function refers to the equation (3.16) from this point on.

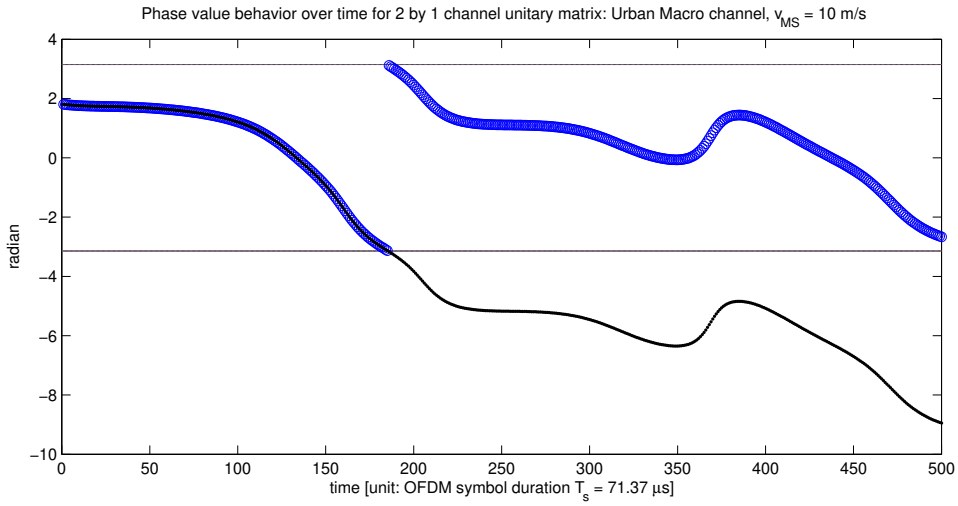


Figure 3.9: ϕ-reconstruction result for $\phi(t)$ with 1 discontinuity case, ◯: ϕ, •: ϕ_{rc}

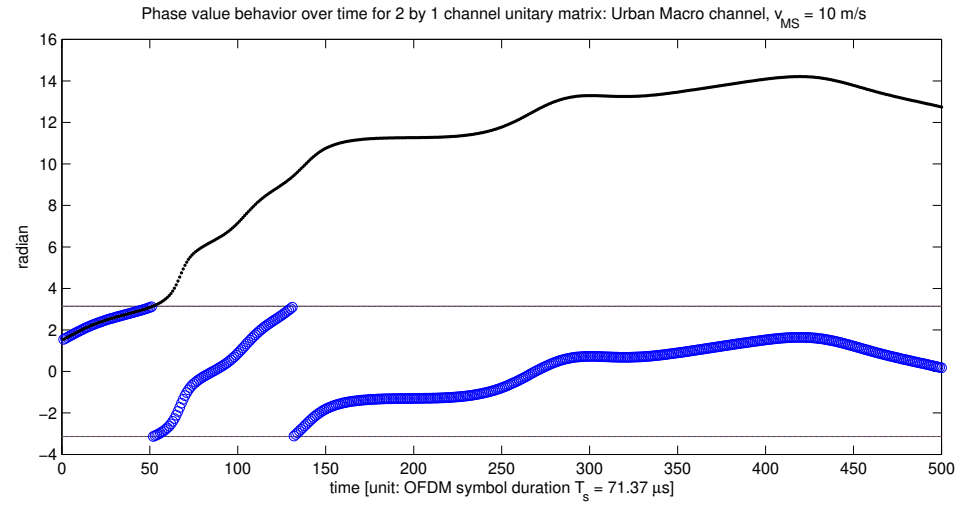


Figure 3.10: ϕ-reconstruction result for $\phi(t)$ with 2 discontinuities case, ◯: ϕ, •: ϕ_{rc}

in seconds.

$$\hat{c}_{\theta\theta}(k) = \frac{1}{N\sigma_\theta^2}\sum_{t=1}^{N}(\theta(t)-\mu_\theta)(\theta(t+k)-\mu_\theta) \tag{3.17}$$

$$\hat{c}_{\phi\phi}(k) = \frac{1}{N\sigma_{\phi_{\mathrm{rc}}}^2}\sum_{t=1}^{N}(\phi_{\mathrm{rc}}(t)-\mu_{\phi_{\mathrm{rc}}})(\phi_{\mathrm{rc}}(t+k)-\mu_{\phi_{\mathrm{rc}}}) \tag{3.18}$$

The normalized autocovariance functions of the Givens parameters, which are obtained by (3.17) and (3.18), can be found in Fig. 3.11 and 3.12. Monte Carlo techniques incorporating the SCMe channel model are used to collect 100,000 samples for two different mobile speed cases (3 m/s and 10 m/s). We can state the following properties, which are derived from observing the empirical normalized autocovariance function.

- The normalized autocovariance function of the Givens parameters has its maximum value at the origin (at zero temporal separation point), and exhibits monotonically decreasing behavior, as expected. This property can be formulated as follows.

$$|\hat{c}_{XX}(k)| \le \hat{c}_{XX}(0), \text{ where } X = \theta \text{ or } \phi \tag{3.19}$$

$$\hat{c}_{XX}(k) \ge \hat{c}_{XX}(k'), \text{ for } k \le k' \tag{3.20}$$

- The normalized autocovariance function $\hat{c}_{XX}(k)$ does not have a periodic component, so we can conclude that the Givens parameter function $\theta(k)$ or $\phi(k)$ does not have a periodic component by the properties of the normalized autocovariance function of the stationary processes.

- The rotational angle θ varies faster over time than the phase ϕ, contrary to our expectations.

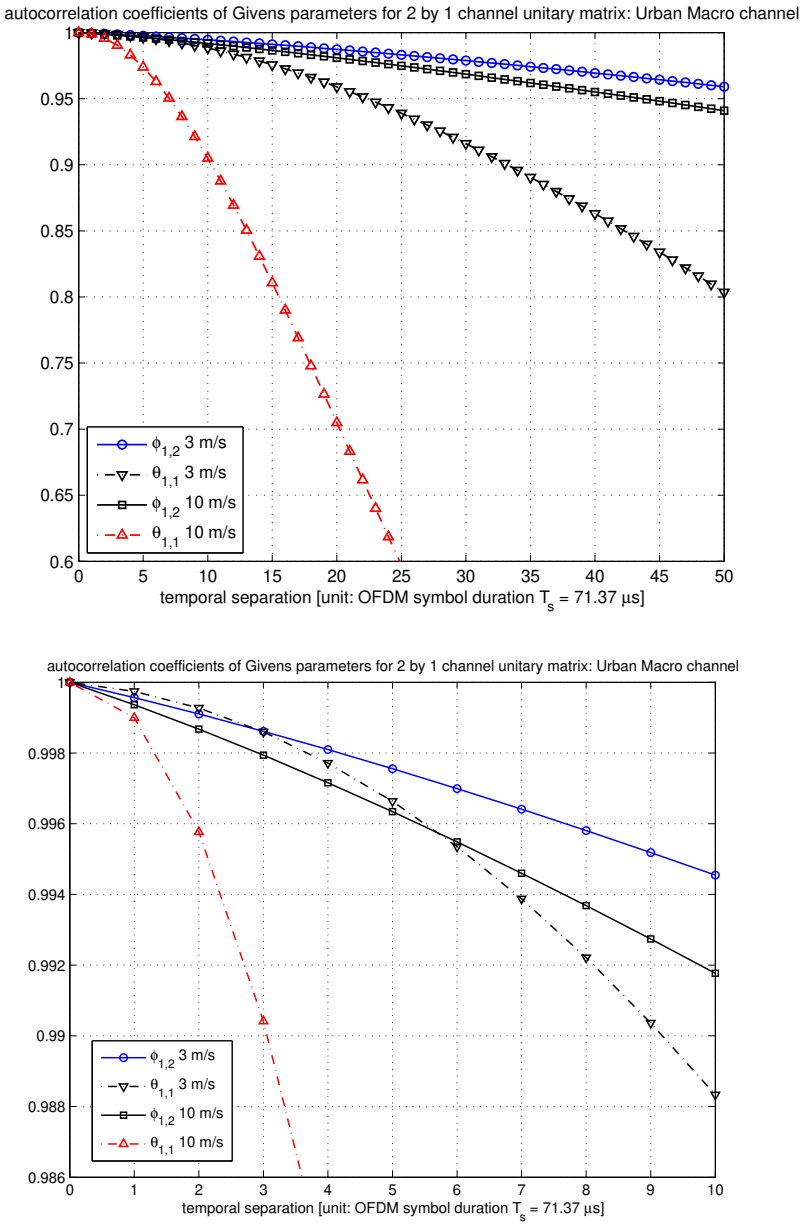


Figure 3.11: Normalized autocovariance of Givens parameters for the channel unitary matrix $\mathbf{V} \in \mathbb{C}^{2\times 1}$ case

3.4.2.2 Feedback Frequency Decision Strategy

In the previous section, we have explained how to quantify the temporal correlation of the Givens parameters by using the normalized autocovariance function. The question is how to exploit the empirical nACvF in deciding the feedback frequency of the Givens parameters. We can come up with the following feedback frequency decision strategies. They are supposed work under the condition

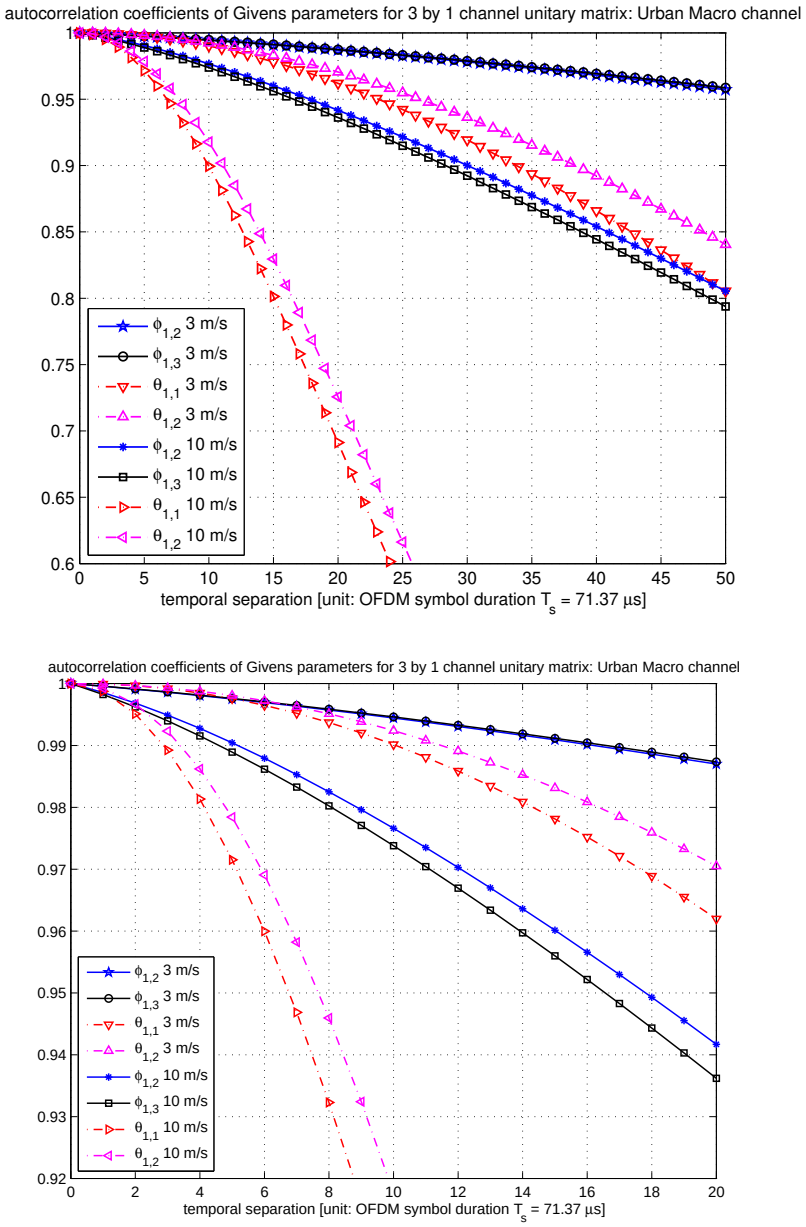


Figure 3.12: Normalized autocovariance of Givens parameters for the channel unitary matrix $\mathbf{V} \in \mathbb{C}^{3\times 1}$ case

that the Givens parameter tracking by ADM works properly that the deviation of the reconstructed one from the original Givens parameter is negligible. It means, in turn, that they are applicable to relatively high frequency feedback cases, which guarantee a high performance in tracking the Givens parameters. To utilize the feedback resources in a smart way, we focus on the smart utilization of the 1 bit ADM feedback frequency of the Givens parameter, and we do not deal with the initial bit

allocation problem.

1. *Feedback frequency utilization strategy*

   The purpose of this strategy is to best utilize feedback resources for a given channel profile, e.g., the dimension of the channel matrix, the channel type, and the mobile speed. Once the channel profile is provided, we can get the empirical nACvF for the corresponding Givens parameter set. Our feedback frequency strategy is to decide the feedback frequencies of the Givens parameters such that all parameters have the approximately same normalized autocovariance level. This strategy is, therefore, based on the assumption that all Givens parameters equally impact the reconstructed unitary matrix. In reality, as it will be shown, the impact of the Givens parameters is not equal, and the requirement of the normalized autocovariance level of certain Givens parameters seems to be higher than others. Hence, we need to introduce a weight factor which recognizes the different importance of the individual Givens parameter.

EVALUATION EXAMPLE 3.1

For the given channel profile (channel unitary matrix $\mathbf{V} \in \mathbb{C}^{t \times n}$, Urban Macro channel, and mobile speed v_{MS}), we need to best utilize the feedback period of the Givens parameters when employing a 1 bit parameter tracking scheme by the ADM. If all the Givens parameters are equally important for the reconstruction of the unitary matrix $\mathbf{V}$, it will be a best strategy to organize feedback periods such that the normalized autocovariance of all the Givens parameters should be same. To evaluate this idea, we need to compare the nACvF based feedback period decision case with the reference case which adopts the straightforward equal feedback period assignment method with the feedback period $[\tau_{\mathrm{ref}}]$[8]. To make a fair comparison, we need to find the feedback period N_{GP}-tuple $([\tau_1], [\tau_2], \cdots, [\tau_{N_{\mathrm{GP}}}])$, where $N_{\mathrm{GP}} = (2t-1)n - n^2$ is the number of Givens parameters, $[\tau_n]$ is the feedback period of X_n, and X_n is the nth Givens parameter (ϕ or θ), which satisfies

$$\hat{c}_{X_1X_1}([\tau_1]) \approx \hat{c}_{X_2X_2}([\tau_2]) \approx \cdots \approx \hat{c}_{X_{N_{\mathrm{GP}}}X_{N_{\mathrm{GP}}}}([\tau_{N_{\mathrm{GP}}}]), \tag{3.21}$$

such that

$$\frac{1}{[\tau_1]} + \frac{1}{[\tau_2]} + \cdots + \frac{1}{[\tau_{N_{\mathrm{GP}}}]} \approx \frac{N_{\mathrm{GP}}}{[\tau_{\mathrm{ref}}]}, \text{ where } [\tau] \text{ is integer.} \tag{3.22}$$

[8]The unit of feedback period is the number of OFDM symbols: $[\tau] = m$, where $\tau = m \cdot T_s$ ($T_s = 71.37\mu$s).

Here, $\hat{c}_{X_n X_n}([\tau_n])$ is the normalized autocovariance function of nth Givens parameter at temporal separation $[\tau_n]$. The basic idea is that the nACvF based feedback period decision method will show a better performance than the equal period assignment method, provided that both cases have the same (or very close) amount of average feedback overhead. The evaluation result can be found in Section 3.4.3.

---

2. *Feedback frequency adaptation strategy*

   This strategy is proposed to provide a method to adapt the feedback frequencies for different channel profiles, for example, different mobile speeds. We can acquire the feedback frequency requirement of the Givens parameters which guarantees the same system performance for different channel profiles, provided that the empirical nACvFs of the Givens parameter set in comparison are known. Our feedback frequency adaptation strategy is to adapt feedback frequencies of the Givens parameters, in comparison to a given reference case, e.g., for different mobile speed, such that the corresponding parameters of the cases have the same normalized autocovariance level in hope that this leads to the same performances. This strategy is, therefore, based on the assumption that the same normalized autocovariance level of the corresponding Givens parameters in different cases promises the same or comparable system performance. Provided that a reference performance is given, we can set a feedback frequency which guarantees the same performance for different conditions (mobile speed, for instance) of which the cases share the same channel matrix dimension. This strategy is proposed to bypass the challenges of finding optimum weights of the Givens parameters by referring to a known reference case.

Evaluation Example 3.2

For the given reference feedback case characterized by the channel profile (channel unitary matrix $\mathbf{V} \in \mathbb{C}^{t \times n}$, Urban Macro channel, and mobile speed v_{MS}) and feedback periods for tracking the Givens parameters, we need to calculate a feedback period N_{GP}-tuple of the Givens parameters $([\tau_1], [\tau_2], \cdots, [\tau_{N_{\mathrm{GP}}}])$, which provides the same performance as the reference feedback case, for the case characterized by a different channel profile element, i.e., mobile speed v'_{MS}, when both cases have the same dimension of the channel unitary matrix. We assume that the

system performance is governed by the precision of the reconstructed Givens parameters and the precision of them, in turn, is dependent on their feedback period. Then the system performance of the comparison case can be kept at the same level as the reference case by adapting the feedback period of each Givens parameter such that its normalized autocovariance value is same as that of its counterpart in the reference case. To evaluate this idea, we compare the nACvF based feedback period adaptation case of mobile speed v'_{MS} with the reference case of mobile speed v_{MS}, as well as with the straightforward case which adopts the linear scaling method in deciding the feedback periods. The evaluation process can be formulated as follows.

The straightforward case with mobile speed v'_{MS} takes the following as feedback periods of the Givens parameters, which could lead to the same system performance as the reference case with mobile speed v_{MS}.

$$\left(\left\langle \frac{\tau_{\mathrm{ref}}}{r} \right\rangle, \cdots, \left\langle \frac{\tau_{\mathrm{ref}}}{r} \right\rangle\right), \text{ where } r = \frac{v'_{\mathrm{MS}}}{v_{\mathrm{MS}}} \tag{3.23}$$

Here, $\langle \cdot \rangle$ is a rounding operator and τ_{ref} is the Givens parameter feedback period of the reference case with mobile speed v_{MS}. The reference case is assumed to share a feedback period for all Givens parameters. It is based on the assumption that the nACvF of the Givens parameters is linearly scalable with respect to the mobile speed, which means

$$\hat{c}^{(v_{\mathrm{MS}})}_{X_n X_n}([\tau_n]) = \hat{c}^{(v'_{\mathrm{MS}})}_{X_n X_n}\left(\left[\frac{\tau_n}{r}\right]\right) = \hat{c}^{(r v_{\mathrm{MS}})}_{X_n X_n}\left(\left[\frac{\tau_n}{r}\right]\right), \text{ where } r = \frac{v'_{\mathrm{MS}}}{v_{\mathrm{MS}}}. \tag{3.24}$$

$\hat{c}^{(v_{\mathrm{MS}})}_{X_n X_n}([\tau_n])$ is the normalized autocovariance function of the nth Givens parameter with mobile speed v_{MS} at temporal separation τ_n.

The comparison case with mobile speed v'_{MS} does not take the linear scalability of the nACvF (3.24) for granted. Fig. 3.11 shows that the linear scalability does not hold for the given case, whereas it seems to hold for the case shown in Fig. 3.12[9]. In this case, it is required to acquire individual nACvFs for the comparison case with mobile speed v'_{MS} as well as the reference case with mobile speed v_{MS}. Then we can find the feedback periods, which guarantee the same performance as the reference case, to be

$$([\tau_1], [\tau_2], \cdots, [\tau_{N_{\mathrm{GP}}}]), \tag{3.25}$$

[9]This can be easily checked by comparing normalized autocovariance values of the same Givens parameter at temporal separation 3 and 10 for $v_{\mathrm{MS}} = 10$ m/s and $v_{\mathrm{MS}} = 3$ m/s cases, respectively.

such that

$$\hat{c}^{(v'_{\mathrm{MS}})}_{X_n X_n}([\tau_n]) = \hat{c}^{(v_{\mathrm{MS}})}_{X_n X_n}([\tau_{\mathrm{ref}}]), \text{ for } n \in \{1, 2, \cdots, N_{\mathrm{GP}}\} \tag{3.26}$$

The evaluation result can be found in Section 3.4.3.

3.4.3 Numerical Results

In this section we present numerical results. Simulations have been performed for the 2 BSs - 2 MSs case to evaluate the feedback frequency decision strategies proposed in the previous section. Two BSs are cooperating to transmit a data signal for two MSs through the same resources at the same time. Both BSs and MSs have a single antenna, so it yields a 2×2 overall channel matrix, which corresponds with 2×1 channel unitary matrix at the MS. The Givens decomposition of the channel unitary matrix provides 2 Givens parameters, which are $\phi_{1,2}$ and $\theta_{1,1}$, and ϕ and θ are used to denote them for simplicity, respectively. We allocate a total of 3 bits for initial quantization of the Givens parameters, of which 2 bits are used to quantize ϕ and 1 bit for θ. This initial parameter quantization phase is directly followed by the parameter tracking phase which is performed by modulating the difference of each parameter by using 1 bit feedback for each parameter. The proposed methods attempt to reduce the amount of feedback in the tracking phase by choosing the feedback periods in a smart way.

We employ the transmit zero-forcing filter as an example beamforming scheme. The 3GPP SCMe is used for the simulations and the proposed methods are tested for an Urban Macro channel with a mobile speed of 3 m/s and 10 m/s. The system performance is evaluated in terms of the received SINR at the MS. Simulations are performed for 100,000 channel realizations and the cdf at one MS is obtained. For each drop, 500 and 150 consecutive channel realizations are generated for a mobile speed of 3 m/s and 10 m/s, respectively, to evaluate the channel tracking capability of the proposed scheme. OFDMA is assumed as the data transmission scheme and we focus on one subcarrier. The transmit power at the BS is set to be 10 W and it is equally allocated to 1201 subcarriers. The carrier frequency is 2.6 GHz.

3.4.3.1 Feedback Frequency Utilization Strategy

For a given Urban Macro MIMO channel of the channel unitary matrix $\mathbf{V} \in \mathbb{C}^{2\times 1}$ with mobile speed $v_{\mathrm{MS}} = 3$ m/s, we need to best utilize feedback periods of the Givens parameters for the

parameter tracking phase. We have two Givens parameters (ϕ and θ) to track by 1 bit ADM. We organize their feedback periods such that their normalized autocovariance values should be same, assuming that both ϕ and θ are equally important for the reconstruction of the unitary matrix $\mathbf{V}$. Fig. 3.11 reveals the fact that normalized autocovariance of ϕ at temporal separation of 9 OFDM symbols has the same level as the normalized autocovariance of θ at a temporal separation of 6 OFDM symbols in case of a mobile speed of $v_{\mathrm{MS}} = 3$ m/s ($\hat{c}^{(3)}_{\phi\phi}(9) \approx \hat{c}^{(3)}_{\theta\theta}(6)$). Therefore, we can set the feedback period pair $([\tau_\phi], [\tau_\theta])$ as $(9, 6)$, which means that we send an 1 bit ADM signal every 9 OFDM symbols to track ϕ and every 6 OFDM symbols to track θ. In this case, the feedback overhead is $1/6 + 1/9 = 0.2778$ [bits/symbol]. On the other hand, the reference case using a straightforward equal feedback period assignment method employs $([\tau_\phi], [\tau_\theta]) = (8, 8)$, of which the feedback overhead is $1/8 + 1/8 = 0.25$. The $(9, 6)$ case is expected to outperform the reference case with a $(8, 8)$ feedback period pair, since the former is designed to keep a normalized autocovariance of the Givens parameters same level, which guarantees at the same level of precision of the reconstructed Givens parameters, if all the parameters are equally important. Fig. 3.13 shows the simulation result. We can see that the $(8, 8)$ case, which uses less feedback resources and adopts the simple equal feedback period assignment scheme, outperforms the $(9, 6)$ case, contrary to our expectations. It implies that not all the Givens parameters are equally important. In our case, ϕ can be more important than θ, but it is not trivial to find exact weighting of the Givens parameters. It is, hence, difficult to come up with the optimum feedback frequency decision strategy.

Fig. 3.14 shows another simulation result which confirms the different impact of the individual Givens parameter. The reference point of the $(3, 3)$ feedback period pair is provided for an Urban Macro MIMO channel of the channel unitary matrix $\mathbf{V} \in \mathbb{C}^{2\times 1}$ with mobile speed $v_{\mathrm{MS}} = 10$ m/s. Let us say that we have extra feedback resource and we are allowed to increase the feedback frequency of one parameter from 1/3 (every 3 symbols) to 1/2 (every other symbol). Our question is which parameter we should choose to increase performance. Fig. 3.11 indicates that we can gain more in terms of the normalized autocovariance level when choosing θ over ϕ. However, the simulation result, depicted in Fig. 3.14, shows that there is no meaningful difference in performance between the $(3, 2)$ case ($[\tau_\theta] = 2$) and the $(2, 3)$ case ($[\tau_\phi] = 2$). This can be another counterexample of all-Givens-parameters-are-equally-important assumption.

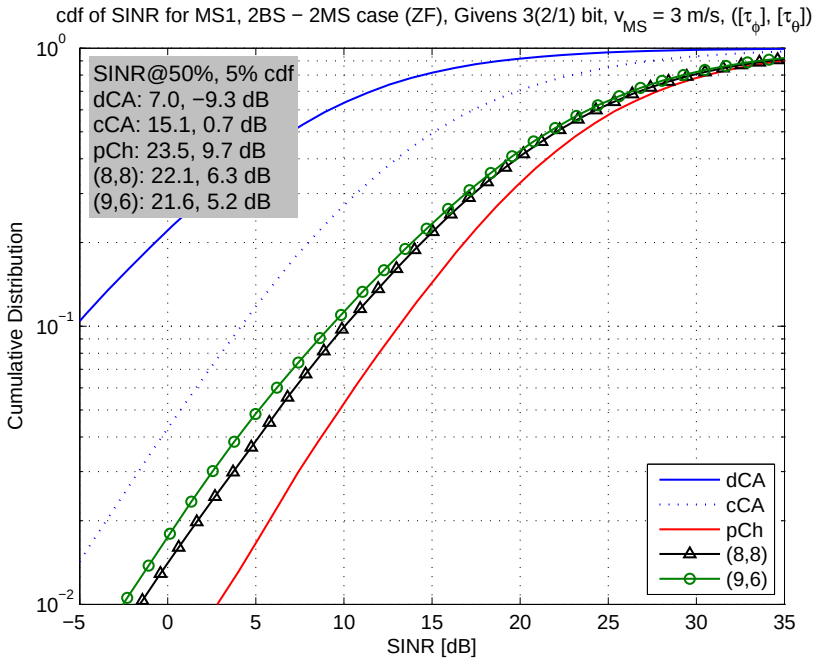


Figure 3.13: Performance of feedback frequency utilization scheme for $v_{\mathrm{MS}} = 3$ [m/s] case (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, $([\tau_\phi], [\tau_\theta])$: D-Givens method with ϕ tracking interval $[\tau_\phi]$ OFDM symbols and θ tracking interval $[\tau_\theta]$ OFDM symbols)

3.4.3.2 Feedback Frequency Adaptation Strategy

For two reference feedback cases characterized by the Urban Macro MIMO channel of the channel unitary matrix $\mathbf{V} \in \mathbb{C}^{2\times 1}$ with mobile speed $v_{\mathrm{MS}} = 10$ m/s and feedback period pairs $(2, 2)$ and $(3, 3)$ for tracking the Givens parameters, we need to calculate the feedback period pairs of the Givens parameters $([\tau_\phi], [\tau_\theta])$ for the comparison case defined by mobile speed $v'_{\mathrm{MS}} = 3$ m/s which provides the same performance as the reference feedback cases, respectively. In short, we need to get the Givens parameter feedback pairs at a mobile speed of 3 m/s, which guarantee the same system performance as the feedback pairs (2,2) and (3,3) at a mobile speed of 10 m/s. We adapt the feedback period pairs of the Givens parameters such that their normalized autocovariance values should be same as that of their counterparts in the reference case. Fig. 3.11 shows that ϕ's normalized autocovariance value of $v_{\mathrm{MS}} = 10$ m/s at a temporal separation of 2 OFDM symbols is close to that of $v_{\mathrm{MS}} = 3$ m/s at a temporal separation of 3 OFDM symbols ($\hat{c}^{(10)}_{\phi\phi}(2) \approx \hat{c}^{(3)}_{\phi\phi}(3)$), and the same relationship holds for θ's nACvF of $v_{\mathrm{MS}} = 10$ m/s at 2 OFDM symbols delay and that of $v_{\mathrm{MS}} = 3$ m/s at 6 OFDM symbols delay ($\hat{c}^{(10)}_{\theta\theta}(2) \approx \hat{c}^{(3)}_{\theta\theta}(6)$). On the contrary, the straightforward case adopting

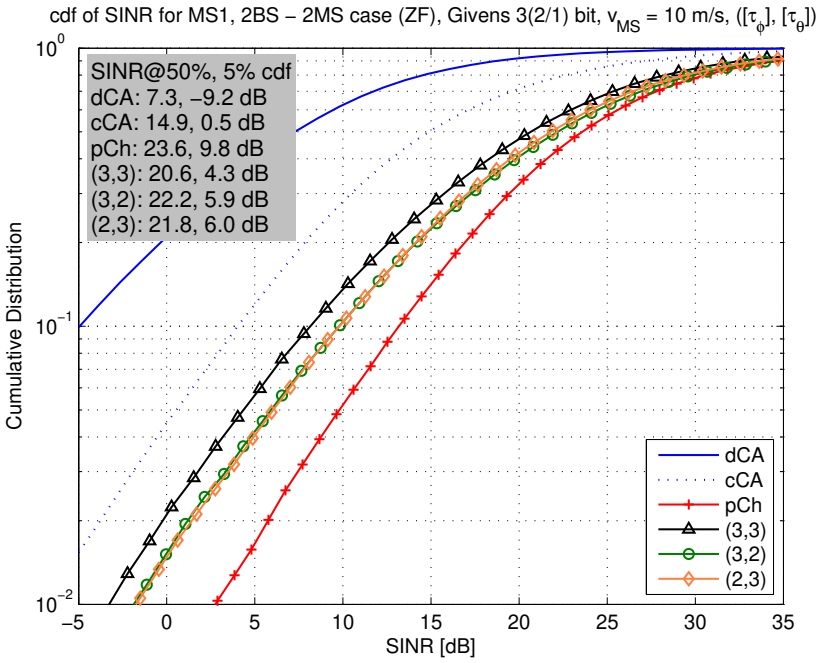


Figure 3.14: Performance of feedback frequency utilization scheme for $v_{\mathrm{MS}} = 10$ [m/s] case (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, $([\tau_\phi], [\tau_\theta])$: D-Givens method with ϕ tracking interval $[\tau_\phi]$ OFDM symbols and θ tracking interval $[\tau_\theta]$ OFDM symbols)

the linear scaling method does not require a normalized autocovariance function. According to (3.23), we can get $(\langle\frac{20}{3}\rangle, \langle\frac{20}{3}\rangle) = (7, 7)$ as a feedback pair of the straightforward case. Regarding the reference case of $v_{\mathrm{MS}} = 10$ m/s with a (3,3) feedback pair, we can get (4,9) as an adaptive feedback pair and (10,10) as a straightforward case feedback pair in case of $v_{\mathrm{MS}} = 3$ m/s. Fig. 3.15 and Fig. 3.16 show the simulation results. In both cases, the normalized autocovariance based adaptive feedback case guarantees the minimum performance degradation in comparison to the reference case ($0.2 \sim 0.5$ dB performance gain at 50% outage SINR), whereas the straightforward case suffers from the performance degradation ($0.9 \sim 1.6$ dB degradation at 50% outage SINR). The simulation results confirm the fact that the feedback frequency adaptation strategy holds.

3.4.4 Conclusion

In this section, the Givens parameter tracking feedback decision method is proposed and evaluated. First, the empirical normalized autocovariance function (nACvF) is obtained based on channel realizations in an effort to quantify the temporal correlation of the Givens parameters. Two feed-

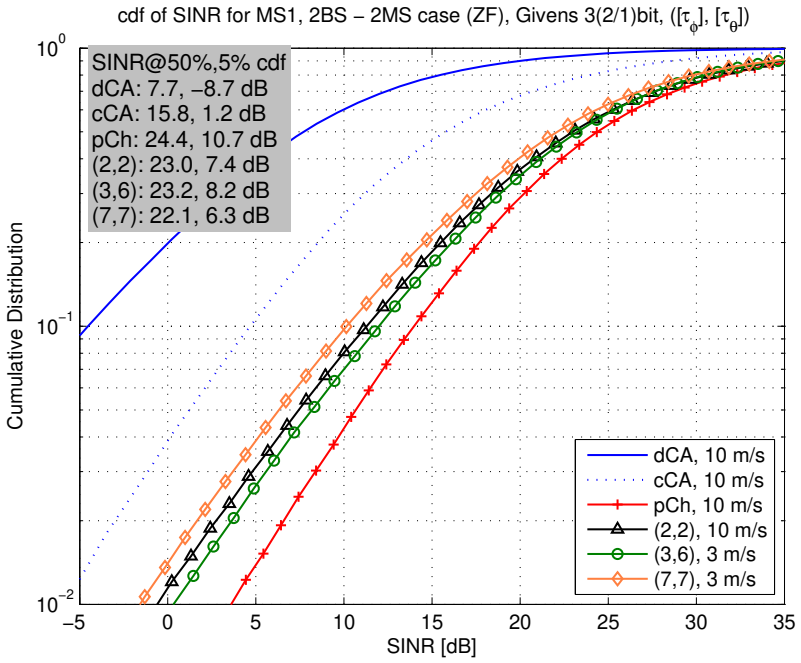


Figure 3.15: Performance of feedback frequency adaptation scheme for the reference case $([\tau_\phi], [\tau_\theta]) = (2, 2)$ at $v_{\mathrm{MS}} = 10$ [m/s] (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, $([\tau_\phi], [\tau_\theta])$: D-Givens method with ϕ tracking interval $[\tau_\phi]$ OFDM symbols and θ tracking interval $[\tau_\theta]$ OFDM symbols)

back frequency decision strategies exploiting the nACvF, which are the feedback frequency utilization strategy and the feedback frequency adaptation strategy, are proposed. The former is to decide the feedback frequencies of the Givens parameters such that all parameters have the approximately same normalized autocovariance level. The latter is to adapt feedback frequencies of the Givens parameters, in comparison to a given reference case, e.g., for different mobile speed, such that the corresponding parameters of the cases have the same normalized autocovariance level, hence the same performances. The nACvF based feedback frequency adaptation strategy is proven to be of practical value by the simulation results.

3.5 Summary

In this chapter, the differential Givens method by Roh and Rao (*Roh and Rao*, 2004) has been examined as an alternative method of reducing the amount of feedback required for channel quantization, and its performance has been evaluated in terms of numerical simulations. To address a practical deployment issue of the method, the author has proposed the Givens parameter feedback

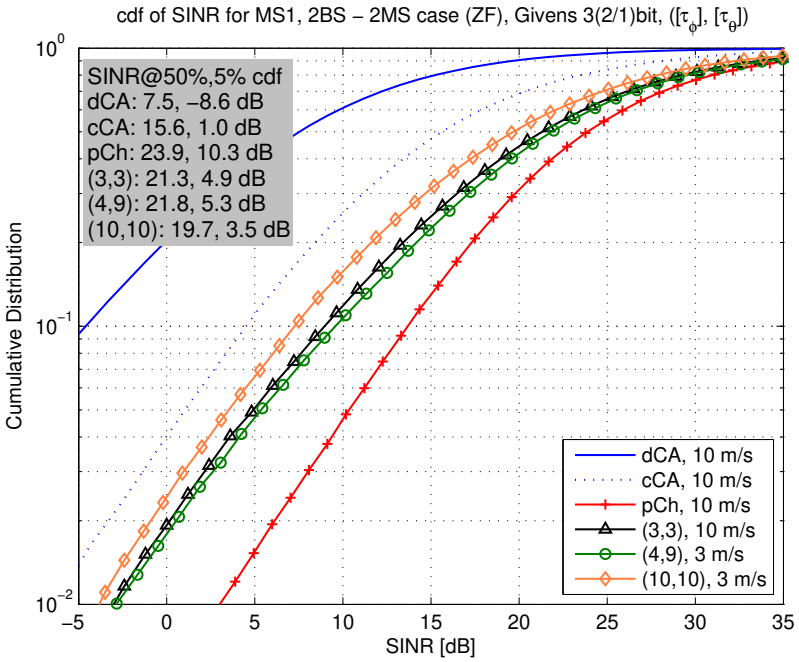


Figure 3.16: Performance of feedback frequency adaptation scheme for the reference case $([\tau_\phi], [\tau_\theta]) = (3, 3)$ at $v_{\mathrm{MS}} = 10$ [m/s] (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, $([\tau_\phi], [\tau_\theta])$: D-Givens method with ϕ tracking interval $[\tau_\phi]$ OFDM symbols and θ tracking interval $[\tau_\theta]$ OFDM symbols)

frequency adaptation method. The main idea is to formulate and to use the empirical autocorrelation normalized autocovariance function in an effort to quantify the temporal correlation of the Givens parameters and to exploit it for practical deployment. The performance of the proposed normalized autocovariance function (nACvF)-based scheme has also been evaluated in terms of numerical simulations. As a further investigation topic, it would be interesting to investigate the sensitivity of the individual Givens parameters and their impact on the system performance.

CHAPTER IV

Feedback Overhead Reduction Methods for Temporally-correlated Channels

This chapter introduces three feedback overhead reduction methods which exploit the temporal correlation of the channel, i.e., a combined codebook design method, a recursive codebook design method and an adaptive codebook design method. As presented in Chapter II, MU-MIMO systems require higher resolution codebooks, which leads to an increase of the feedback overhead. The author addresses this issue by proposing novel feedback overhead reduction methods as follows.

The combined codebook (cCB) design method introduces a hierarchical structure in existing subspace-based channel quantization codebooks. The combined codebook has a two-layered structure, which entails a two-layered feedback scheme, e.g., coarse feedback and fine feedback. We also propose a method for best utilizing feedback periods.

The recursive codebook (rCB) design introduces the subspace tracking feedback in an effort to reduce the feedback overhead. The subspace tracking feedback encoding, which is also termed a recursive feedback, selects the best precoding/channel matrix codes from a small size adaptive codebook, depending on the previous decisions. With respect to the operation scenario, the proposed feedback scheme operates in two modes, which are the subspace adjusting mode and the subspace tracking mode. The subspace tracking mode is operated by the recursive codebook, which is a subset of the complete codebook.

The adaptive codebook (aCB) design introduces a transition probability based codeword sorting scheme in an effort to reduce the feedback overhead. The adaptive codebook sorts the codewords in the increasing order of the chordal distance with respect to the initial codeword, which is the codeword chosen at the previous time step. The adaptive codebook method in conjunction with the data compression scheme requires much less feedback overhead compared with other techniques at

the cost of slightly increased memory requirements.

This chapter is organized as follows. The mechanism of the proposed feedback overhead reduction methods, i.e., cCB, rCB, and aCB, is explained in detail together with corresponding simulation results in Sections 4.1, 4.2, and 4.3, respectively. Section 4.4 summarizes the chapter.

This research has been partially published in (*Kim et al.*, 2008c,d,e). The adaptive codebook method, the most efficient of the proposed algorithms, has been granted as an European patent (*Kim et al.*, 2010).

4.1 Combined Codebook: Hierarchical Codebook Design Method

4.1.1 Background

The channel quantization method and corresponding codebook generation method used in this section are based on the subspace based channel quantization method which is explained in Chapter II. It should be noted that the subspace based channel quantization method is applicable to general multi-user MIMO cases as well as the distributed cooperative antenna systems. The *combined codebook* design method in this section is also applicable to the general MU-MIMO case.

4.1.1.1 Codebook Construction Based on Modified LBG VQ Algorithm

The Linde-Buzo-Gray vector quantization (LBG VQ) based codebook $\mathcal{C}$ design problem can be stated as follows. Given a source vector, given a distortion measure, given a codebook evaluation measure, and given the size of the codebook, find a codebook and a partition[1] which result in maximizing the minimum chordal distance of the codebook.

Suppose that we have a training sequence $\mathcal{T}$ to capture the distribution of the column space basis vectors $\mathbf{U}_j^{(\mathrm{S})}$ of size $N_{tt} \times N_r$ to estimate the mean and covariance matrix:

$$\mathcal{T} = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_M\}, \tag{4.1}$$

where $\mathbf{X}_m \in \mathbb{C}^{N_{tt} \times N_r}$ is a sample of $\mathbf{U}_j^{(\mathrm{S})}$ which can be obtained by taking a SVD of the channel matrix sample $\mathbf{H}_j$. The channel matrix samples are generated by Monte Carlo simulations using the SCMe. The number of channel samples M is assumed to be large (e.g., $M \geq 1000 N_{cb}$, where $N_{cb} = 2^{n_{\mathrm{CB}}}$ is the codebook size), so that the mean and covariance matrix of the source are well

[1] The *partition* of the space is defined as the set of all encoding regions.

estimated by the training sequence. The codebook can be represented as follows.

$$\mathcal{C} = \{\mathbf{C}_1, \mathbf{C}_2, \cdots, \mathbf{C}_{N_{cb}}\} \tag{4.2}$$

Each code has the same size as a training matrix ($\mathbf{C}_n \in \mathbb{C}^{N_{tt} \times N_r}$). Let $\mathcal{R}_n$ be the encoding region associated with the code $\mathbf{C}_n$ and let

$$\mathcal{P} = \{\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_{N_{cb}}\} \tag{4.3}$$

denote the partition of the space. If the source matrix $\mathbf{X}_m$ belongs to the encoding region $\mathcal{R}_n$, then it is quantized to $\mathbf{C}_n$:

$$\mathcal{Q}(\mathbf{X}_m) = \mathbf{C}_n, \text{ if } \mathbf{X}_m \in \mathcal{R}_n. \tag{4.4}$$

Our aim is to find a codebook such that the minimum chordal distance is maximized. The minimum chordal distance of the codebook is given by:

$$d_{c,\min}(\mathcal{C}) := \min d_c(\mathbf{C}_i, \mathbf{C}_j), \text{ for } \mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, \forall i \neq j. \tag{4.5}$$

The design problem can be stated as follows: Given $\mathcal{T}$ and N_{cb}, find $\mathcal{C}$ and $\mathcal{P}$ such that $d_{c,\min}(\mathcal{C})$ is maximized.

4.1.2 Motivation

In this section, we propose a hierarchical codebook design method, which exploits the temporal correlation of the channel as a way of reducing the feedback overhead. It is known that the wireless channel does not change radically within the coherence time T_c. Accordingly, the code index would not change so often during this time period, since the code index can be considered as a channel state indicator. On the other hand, we can easily draw the conclusion that the codebook index transition rate over time is dependent on the size of the codebook, which decides the resolution of the channel quantization. It means that for a given channel, a bigger size codebook (fine codebook) has a higher capability of differentiating encoding regions than a smaller size codebook (coarse codebook). The period during which the coarse codebook provides the same code index (let us call this a non-transition period) can be composed of several shorter non-transition periods when the fine codebook is used to quantize the channel.

Thus, if we are able to design a codebook hierarchically so that a codebook has several layers,

say two layers, one of which represents coarse encoding regions and another provides fine encoding regions, we can achieve the performance of a fine channel quantization by using much smaller feedback resources. This can be achieved by organizing the coarse/fine codebook feedback periods in a smart way to take advantage of non-transition periods of the coarse/fine codebook.

The underlying assumptions are as follows. First, every encoding region of the coarse codebook is the set of smaller encoding regions of the corresponding fine codebook. This is depicted in Fig. 4.1. This is only a conceptual figure in which the encoding region spans a two-dimensional space. The coarse encoding region in this example is divided into 4 fine encoding regions. Second, the non-transition period of the coarse codebook is composed of several non-transition periods of the fine codebook which is associated with the encoding region of the corresponding coarse code.

The operation scenario of the combined codebook is as follows. As a preparation, we need to design the combined codebook which has a two-layer structure, namely a n_c bit coarse codebook and a n_f bit fine codebook. Corresponding feedback periods should be decided, based on the statistical properties of the non-transition time. The $n_t = n_c + n_f$ bit combined codebook, as a whole, is designed to be composed of 2^{n_c} groups of the fine codes, and each fine codebook group consists of 2^{n_f} fine codes. The feedback operation works as follows. For every coarse feedback period τ_c of the n_c bit coarse codebook, the MS sends the coarse codebook index i_c (n_c bit) back to the BSs to indicate the fine codebook group index to which the subsequent fine code indices belong. At the same time, the MS sends the fine code index i_f (n_f bit) to indicate the fine code index of the chosen fine codebook group. This subsequent feedback is done for every fine feedback period τ_f of the n_t bit fine codebook. Since that n_c bit feedback is sent back for every τ_c and only n_f bit extra feedback is needed for every τ_f, we can save uplink resource when compared to the case of sending back n_t bit feedback for every time. Interested readers can consult Fig. 4.1 for better understanding.

4.1.3 Hierarchical Codebook Construction

The main design problem of the hierarchical codebook construction is to divide a coarse encoding region into equally probable fine encoding regions. Here the term 'equally probable' means that the probability that a channel sample falls into a certain encoding region is same for all candidate encoding regions. If not, the fine codebook feedback period should be varied adaptively to reflect this unequal probability. Consequently the equally probable encoding regions allow us to fix the feedback period for given channel-dependent constraints.

The modified LBG VQ (mLBG VQ) algorithm, which is used for codebook construction, gener-

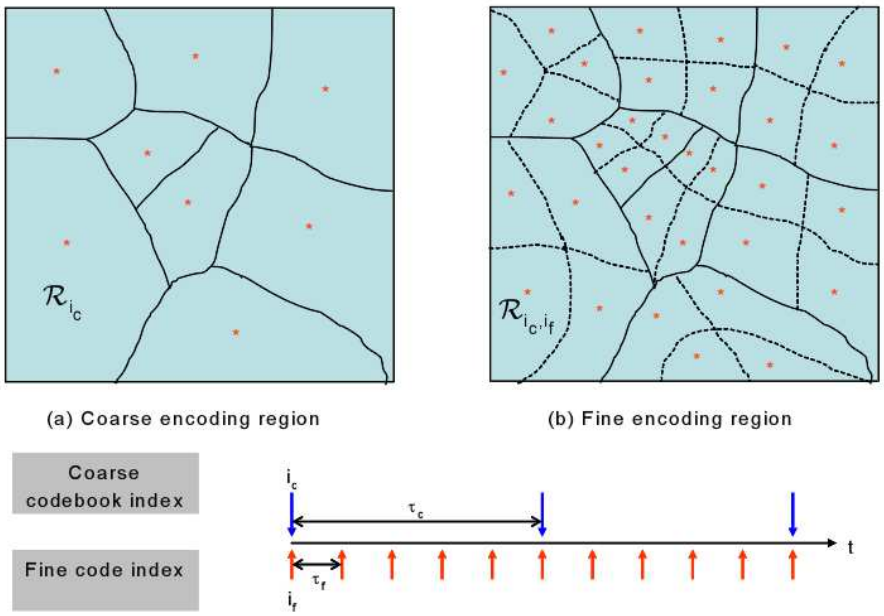


Figure 4.1: Coarse/fine encoding regions and feedback timeframe (n_c=3, n_f=2 with τ_c=5τ_f)

ates the codebook which has this property. The resulting codes are maximally spaced codes of which an individual code is designed to provide the minimum mean squared chordal distance between the code and the channel samples in that encoding region. This criterion places finer encoding regions in densely populated areas, and the resulting encoding regions are asymptotically equally probable. This is shown to be true in (*Huang et al.*, 2006b) as well, for codes generated by the Lloyd algorithm based codebook construction method.

The design problem of the hierarchical codebook construction with a n_c bit coarse and a n_t bit fine codebook is to divide the given channel space which is a subspace of $\mathbb{C}^{N_{tt} \times N_r}$ into 2^{n_c} equally probable coarse encoding regions and to divide each coarse encoding region into 2^{n_f} equally probable fine encoding regions. In the end, we want to have 2^{n_c} groups of codebooks each of which is composed of 2^{n_f} codes. This can be solved as follows. We first perform the mLBG VQ algorithm to get $N_c = 2^{n_c}$ coarse codes and the corresponding coarse encoding regions. These encoding regions should be equally probable ($P_n = 1/N_c, \forall n \in \{1, \cdots, N_c\}$). Then, we perform the mLBG VQ algorithm for channel samples which belong to each coarse encoding region, individually. As a result of N_c parallel codebook generation processes for each coarse encoding region, we can acquire $N_{cb} = 2^{n_t} = 2^{n_c+n_f}$ fine codes with corresponding equally probable fine encoding regions

$(P_n = 1/N_f, \forall n \in \{1, \cdots, N_f\})$. Each coarse encoding region consists of $N_f = 2^{n_f}$ fine encoding regions. The codebook construction process can be formulated as follows.

Assume that the following parameters are given: training sequence $\mathcal{T}$, the number of feedback bits for the coarse and the fine codebooks n_c and n_f, respectively. We perform the mLBG VQ algorithm for a given $\mathcal{T}$ to get $N_c = 2^{n_c}$ coarse encoding regions. The mLBG VQ algorithm takes $\mathcal{T}$ and N_c as input parameters and provides a coarse codebook[2] $\mathcal{C}^c$ and a partition $\mathcal{P}$ as an output. Thus, this can be formulated as follows.

$$(\mathcal{C}^c, \mathcal{P}) = \text{mLBG}(\mathcal{T}, N_c) \tag{4.6}$$

The partition of the space is defined as the set of all encoding regions:

$$\mathcal{P} = \{\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_{i_c}, \cdots, \mathcal{R}_{N_c}\}, \tag{4.7}$$

where i_c is the coarse encoding region index and N_c is the total number of coarse encoding regions. We now perform the mLBG VQ for each coarse encoding region $\mathcal{R}_{i_c}$ so as to get 2^{n_f} fine encoding regions per each $\mathcal{R}_{i_c}$. This process is formulated as $(\mathcal{C}^c_{i_c}, \mathcal{P}_{i_c}) = \text{mLBG}(\mathcal{R}_{i_c}, N_f)$, for $i_c \in \{1, 2, \cdots, N_c\}$, where $\mathcal{C}^c_{i_c}$ is the codebook associated with the partition $\mathcal{P}_{i_c}$. The partition $\mathcal{P}_{i_c} = \{\mathcal{R}_{i_c,1}, \mathcal{R}_{i_c,2}, \cdots, \mathcal{R}_{i_c,i_f}, \cdots, \mathcal{R}_{i_c,N_f}\}$ is the set of $N_f = 2^{n_f}$ fine encoding regions. The partition $\mathcal{P}_{i_c}$ is equal to $\mathcal{R}_{i_c}$ in terms of elements: $\mathcal{P}_{i_c} = \mathcal{R}_{i_c}$.

The second step of the codebook generation process provides us with N_{cb} fine codes in total. The corresponding codebook $\mathcal{C}$ can be regarded as a set of codebooks $\mathcal{C}^c_{i_c}$, where i_c is the codebook index[3]. The overall codebook $\mathcal{C}$ and the i_cth fine codebook $\mathcal{C}^c_{i_c}$ can be expressed as:

$$\mathcal{C} = \{\mathcal{C}^c_1, \mathcal{C}^c_2, \cdots, \mathcal{C}^c_{N_c}\} = \{\mathbf{C}_1, \mathbf{C}_2, \cdots, \mathbf{C}_{N_{cb}}\}, \tag{4.8}$$

$$\mathcal{C}^c_{i_c} = \{\mathbf{C}^{i_c}_1, \mathbf{C}^{i_c}_2, \cdots, \mathbf{C}^{i_c}_{i_f}, \cdots, \mathbf{C}^{i_c}_{N_f}\}, \text{ for } i_c \in \{1, 2, \cdots, N_c\}, \tag{4.9}$$

where the fine codes are arranged in such an order that the condition $\mathbf{C}^{i_c}_{i_f} = \mathbf{C}_{(i_c-1)N_f+i_f} \in \mathbb{C}^{N_{tt} \times N_r}$ is satisfied and i_f is the fine code index within the coarse encoding region $\mathcal{R}_{i_c}$. It becomes clear at this point that the resulting codebook has a hierarchical structure. This is the reason why it is

[2]The coarse codebook is not used in the end, as n_c bit is used only to indicate the fine codebook, which is the group of the fine codes. The combined codebook is a set of codebooks, each of which is in turn a set of fine codes.

[3]Here, we differentiate the term code and codebook, in such a way that a code indicates an individual code, whereas a codebook indicates a set of codes. Thus i_c indicates not an individual code index, but a codebook index to which a fine code belongs. It means that the codebook $\mathcal{C}^c_{i_c}$ is constructed based on the i_cth coarse encoding region $\mathcal{R}_{i_c}$. The elements of $\mathcal{C}^c_{i_c}$ are fine codes.

called a hierarchical codebook.

For example, Fig. 4.1 shows the case with $n_c = 3$ and $n_f = 2$. The partition of the channel sample space consists of $N_c = 8$ coarse encoding regions, the i_cth of which is denoted as $\mathcal{R}_{i_c}$, as a result of the mLBG VQ procedure. Each individual coarse encoding region is again decomposed into $N_f = 4$ fine encoding regions, the i_fth of which is $\mathcal{R}_{i_c,i_f}$ in case that it is based on $\mathcal{R}_{i_c}$. In the end, we get $N_t = 32$ fine codes associated with corresponding fine encoding regions.

4.1.4 Operation Scenario

The operation scenario of the hierarchical codebook deployment, which is also termed a *combined codebook* in this section, is as follows.

- *Coarse Feedback*: Feedback of the codebook index i_c

  For every coarse feedback period τ_c, the MS sends the n_c bit codebook index i_c back to the associated BSs to indicate the chosen codebook. Based on the channel information observed for the time period τ_c, the MS quantizes the channel matrix and finds the best code in terms of the chordal distance. The index of the chosen code $\mathbf{C}_{(i_c-1)N_f+i_f} \in \mathbb{C}^{N_{tt}\times N_r}$ can be decomposed into two parts, e.g., the coarse index i_c and the fine index i_f. The coarse feedback involves sending back i_c.

- *Fine Feedback*: Feedback of the code index i_f

  For every fine feedback period τ_f within τ_c, the MS sends the n_f bit code index i_f back to the associated BSs so as to indicate the chosen code. It means that the MS performs the channel quantization for every τ_f, but the scope of candidate codes is restricted within the codebook $\mathcal{C}^c_{i_c}$. This can save a lot of computational burden for the MS, since the number of candidate codes is N_f instead of N_cN_f. The fine feedback involves sending back i_f.

The BSs collect the coarse and fine feedback messages, and combine these information to find the chosen code, which is one of $N_{cb} = 2^{n_c+n_f}$ fine codes which are predefined and shared by both BSs and MSs. There are several points to be worthy of our attention.

1. The feedback periods τ_c and τ_f have a significant effect on the system performance.

   It is a challenging task to find an analytical solution for calculating an optimum feedback period. The optimization problem is supposed to minimize the performance degradation compared with the ideal case with a given feedback resource, and it is a function not only of the dimension of the quantized channel matrix, the speed of the MS, and the carrier frequency,

but also of the number of feedback bits n_c, n_f which decide the resolution of the channel quantization. For example, even if the other parameters remain the same except for n_c and n_f, a bigger n_c and n_f (a bigger codebook size) should be accompanied by a smaller τ_c and τ_f (more frequent feedback) to make full use of its potential, since they increase the quantization resolution. A semi-analytical way of deciding feedback periods is introduced in Appendix C.

2. Once found, τ_c and τ_f can have a fixed value for given channel-related parameters (the channel matrix dimension, the carrier frequency and the speed of the MS) and a given number of feedback bits. They are still able to guarantee the target performance.

   In other words, the feedback period does not have to be adaptive to the specific codebook or code, since the encoding regions are equally probable. If all parameters other than the mobile speed remain the same, we only need to scale the feedback period with respect to the mobile speed.

3. Within the coarse feedback period τ_c, the resulting codebook has a limited scope, since it selects the best code from the chosen codebook only.

   If the actual channel realization falls into a different coarse encoding region, the quantization error would be significant.

4. The coarse feedback period should be an integer multiple of the fine feedback period.

   This condition is not required for the optimum performance, but it is rather imposed to facilitate an easy deployment of the proposed method.

The BS may benefit from keeping the values of the coarse codebook $\mathcal{C}^c = \{\mathbf{C}_1^c, \mathbf{C}_2^c, \cdots, \mathbf{C}_{N_c}^c\}$, just in case that MSs are temporarily disabled to send fine feedback messages to BSs. In this case, the BSs are accessible to coarse feedback indices only. However, the BSs can still reconstruct the channel if the coarse codebook is available at the BSs. The BS is supposed to have two codebooks, one of which is the coarse codebook of size N_c, and another is the combined codebook of size N_{cb}. Only the combined codebook needs to be saved on the MS side.

4.1.5 Numerical Results

In this section we present numerical results. Simulations have been performed for the 2 BSs - 2 MSs case in which the two BSs cooperate to transmit data for two MSs using the same resources. Both BSs and MSs have a single antenna, so we have a 2×2 overall channel matrix. We employ

the transmit zero-forcing filter as an example beamforming scheme. It is based on downlink channel information which is either perfect (pCh), or is provided by a downlink channel estimation method which is shared by the BSs through a prompt and error free backbone network (centralized CA, cCA), or is captured and reconstructed by looking up an n_t bit full size codebook (n_t bit channel quantization, n_tbCQ). The 3GPP SCMe is used for the simulations and the proposed methods are tested for an Urban Macro channel with a 10 m/s mobile speed and a 2.6 GHz carrier frequency. The system performance is evaluated in terms of the received SINR at the MS. Simulations are performed for 30,000 channel realizations and the cdf at one MS is obtained.

The performance of the combined codebook is shown in Fig. 4.2. The combined codebooks are provided by the modified LBG VQ algorithm and the hierarchical codebook construction method. The combined codebook of the n_c bit coarse and n_f bit fine feedback with corresponding feedback periods τ_c and τ_f is denoted by n_c+n_fbCQ ($[\tau_c]$,$[\tau_f]$), where the unit of the feedback period is the number of OFDM symbols: $[\tau_c] = \frac{\tau_c}{T_s}$, $[\tau_f] = \frac{\tau_f}{T_s}$. The feedback periods are determined by the feedback period decision method in Appendix C. At 50 % outage SINR, the 5 bits codebook case (5bCQ), which is generated by the hierarchical codebook construction method, shows a 3.9 dB gain over the centralized CA case (cCA). In this case, the 5 bits feedback is sent back for every symbol. The 3+2 bit combined codebook with the feedback period pair ($[\tau_c]$,$[\tau_f]$)=(10,5), (3+2bCQ$_1$), is less than 0.1 dB away from the performance of the 5bCQ case. In terms of the required resources, the 5bCQ case requires 5 bits/symbol for feedback, while the 3+2bCQ$_1$ case needs only 0.7 bits/symbol. Thus, the 3+2 bit combined codebook can achieve the performance of the 5 bits codebook with a negligible degradation by using just 14 % of the feedback resources.

Both the (10,10) and the (20,5) cases show degradations in performance. The degradation of the (20,5) case with a longer coarse feedback period is 0.1 dB bigger than the (10,10) case with a longer fine feedback period, even though the former requires 0.55 bits/symbol feedback overhead while the latter needs 0.5 bits/symbol overhead. It is due to the fact that even though both τ_c and τ_f are important in deciding the performance, τ_c has a more profound impact than τ_f, since the coarse codebook is associated with the bigger encoding region which leads to a bigger error if the feedback is outdated.

4.1.6 Conclusion

In this section, we proposed the hierarchical codebook design method (combined codebook) and its operation scheme in practice. The combined codebook contains two layers of feedback, i.e.,

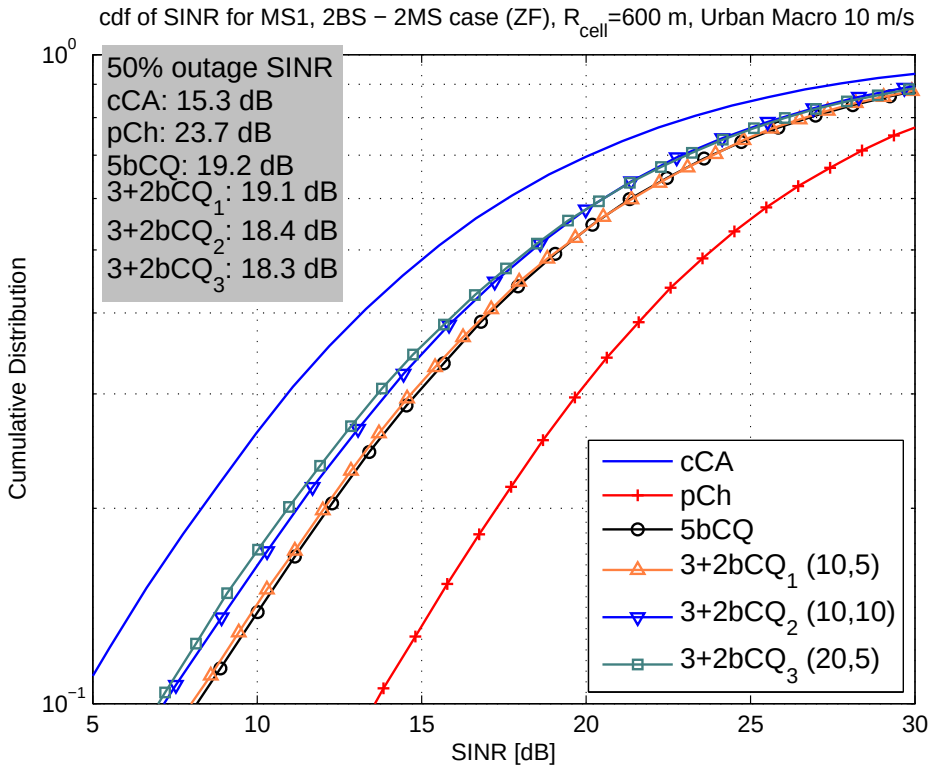


Figure 4.2: Combined codebook simulation results for $(n_c{+}n_f\text{b})=(3{+}2\text{b})$ case (cCA: centralized CA, pCh: perfect channel, 5bCQ: 5 bit CQ, 3+2bCQ$_i$ ($[\tau_c]$,$[\tau_f]$): 3 bits coarse and 2 bits fine feedback with a feedback period pair $([\tau_c],[\tau_f]) \in \{(10,5),(10,10),(20,5)\}$)

the coarse feedback and the fine feedback. We have found the decision of the feedback period is critical in the system performance and proposed in Appendix C a semi-analytic method of deciding a feedback period for a given reference point which is known to achieve the target performance. The combined codebook is proven to reach the reference performance and to be very efficient in terms of the required resources. The MS can also benefit from the proposed methods in terms of the smaller computational burden in the sense that the combined codebook with a n_f bit fine feedback has only 2^{n_f} candidate codes to search for quantization[4], instead of 2^{n_t} candidate codes for the $n_t = n_c + n_f$ bit full codebook.

[4]This is valid during the fine feedback period only, and the n_t bit full codebook needs to be searched for the coarse feedback. This needs to be done less frequently, though.

4.2 Recursive Codebook: Subspace Tracking Codebook Design Method

4.2.1 Motivation

The feedback reduction method has been introduced in Section 4.1 which is termed *combined codebook*. Once the combined codebook is constructed, the hierarchical structure of the codebook is solidly determined and leaves no room for change. This rigidness makes the combined codebook vulnerable to the border effect. The border effect here refers to a phenomenon that the possibility for the next channel quantization (CQ) result to fall into a different coarse codebook is higher in case that the current CQ result points to the code located nearer to the border. In the design of the combined codebook, the decision of the coarse feedback period addresses the border effect. However, the resulting strict constraint on the coarse feedback period can be alleviated by allowing an additional degree of freedom for the codebook structure, which is explained in the following subsection.

4.2.2 Recursive Codebook Construction

The *recursive codebook* design method is proposed in this section. A recursive feedback encoding scheme is introduced which provides the codebook with the subspace tracking capability. This feature leads to a further reduction of the feedback overhead compared with the combined codebook. The recursive method induces flexibility into the formation of the codebook to address the border effect problem of the combined codebook.

- Analytical Framework

  The basic idea of the recursive codebook is to design a small size adaptive codebook which is able to track the channel subspace, so as to address the border effect of the combined codebook. The recursive codebook corresponds to recursive vector quantization. Recursive Vector Quantization (RVQ) is a vector quantization with memory, in which the quantization output depends on prior inputs as well as the current input (*Gersho and Gray*, 1992).

  The downlink temporally correlated channel state sequence can be modeled using a first-order finite-state Markov chain (*Huang et al.*, 2009)[5]. RVQ can be modeled as a first-order finite-

[5]Huang et al. have modeled the time-varying quantized CSI as a first-order finite-state Markov chain and validated it by simulations. They have shown that the first-order Markov chain provides a good approximation of Clarke's fading model for the spatially correlated channel as well as the *i.i.d.* Rayleigh fading (spatially uncorrelated) channel, and the first-order Markov curve converges to Clarke's fading model curve as the time separation increases. It has been also suggested that higher-order feedback-state Markov chains provide no practical improvement on modeling accuracy, unless it is a 10 or higher order Markov chain. According to their simulation results, for the given system parameters of ours, i.e., $f_c = 2.6$ GHz and $T_s = 71.37$ μs, the first-order Markov chain model provides very good approximation

state Markov chain in the same way. We use state variables to depict the influence of the past on the current operation of the quantization. As a result, the RVQ is effectively formulated as the state transition and state dependent encoding, which can be considered as a Markov chain. The important notationwhich will appear in the process of analysis is as follows:

$\mathbf{H}_p$	downlink channel state information (CSI) at time p, $\mathbf{H}_p \in \mathbb{C}^{N_{tt} \times N_r}$
J_p	channel state (quantized channel index) at time p, $J_p \in \mathbb{J} = \{1, \cdots, N_{cb}\}$ where N_{cb} is the codebook size
$\mathcal{C}$	codebook containing N_{cb} channel profiles[6]
$\mathbf{C}_i$	$\mathbb{C}^{N_{tt} \times N_r}$ ith member of the codebook $\mathcal{C}$, $i \in \{1, \cdots, N_{cb}\}$
$\mathcal{C}^r$	recursive codebook, $\mathcal{C}^r \subset \mathcal{C}$,

where N_{cb} is the size of the codebook which corresponds with the number of channel states. The CSI quantization process is summarized as follows. The CSI quantization maps the downlink CSI $\mathbf{H}_p$ at time p onto a member of the codebook $\mathcal{C}$ and provides its index J_p as an output. The codebook $\mathcal{C}$ is acquired by partitioning the channel space into N_{cb} regions, which is called Voronoi cells/encoding regions $\{\mathcal{R}_i\}$. The codebook members (codewords) and the encoding regions have a one-to-one relationship. Therefore, the CSI quantization operation (4.10) implies (4.11) in practice.

$$\mathbf{C}_i = \mathcal{Q}(\mathbf{H}_p), i \in \{1, \cdots, N_{cb}\} \tag{4.10}$$

$$J_p = i \text{ if } \mathbf{H}_p \in \mathcal{R}_i, i \in \{1, \cdots, N_{cb}\} \tag{4.11}$$

The stationary and transition probabilities of the first-order Markov chain are defined as

$$\begin{aligned} &\text{(stationary probability)} \quad P_i = \Pr\{J_p = i\} = \Pr\{\mathbf{H}_p \in \mathcal{R}_i\}, \\ &\text{(transition probability)} \quad P_{i,j} = \Pr\{J_p = i | J_{p-1} = j\} = \Pr\{\mathbf{H}_p \in \mathcal{R}_i | \mathbf{H}_{p-1} \in \mathcal{R}_j\}, \end{aligned} \tag{4.12}$$

where $i, j \in \{1, \cdots, N_{cb}\}$. Now RVQ is modeled by a first-order Markov chain. Therefore, the downlink temporally correlated channel can be treated as a Markov source and makes it

to Clarke's model with a time separation up to 100 symbols and 30 symbols for the $v_{\text{MS}} = 3$ m/s case and the $v_{\text{MS}} = 10$ m/s case, respectively. It is within the boundary of our operation scenarios and hence the simulation space as well.

[6] Channel profile is defined as a channel related quantization result, which can be a precoding matrix or a channel matrix.

eligible for data compression techniques. Here a sub-optimal but simple and effective feedback compression method is proposed, which introduces a small size subspace tracking codebook.

First of all, we define the ϵ-neighborhood of the Markov state i as

$$\mathcal{N}_i(\epsilon) = \{j \in \mathbb{J} | P_{i,j} \geq \epsilon\}, \quad i \in \{1, \cdots, N_{cb}\}, \tag{4.13}$$

where $P_{i,j}$ is the transition probability and ϵ a small positive number. It is illustrated in Fig. 4.3[7]. Given a N_{cb}-region partition of the channel space, each feedback requires

$$n_t = \log_2 N_{cb} \text{ bits} \tag{4.14}$$

for identifying the current channel state. In case that we narrow our scope into the small portion of the entire channel space which has a significant probability ($> \epsilon$) only, the required number of feedback bits for identifying the current channel state $J_p = i$ is

$$n_r(i) = \lceil \log_2 |\mathcal{N}_i(\epsilon)| \rceil \text{ bits}, \quad i \in \{1, \cdots, N_{cb}\}, \tag{4.15}$$

where $\lceil a \rceil$ is the smallest integer larger than a. From (4.13), $|\mathcal{N}_i(\epsilon)| \leq N_{cb}$ holds and hence $n_r(i) \leq n_t$. We should notice the fact that $\mathcal{N}_i(\epsilon)$ can be acquired only when the current Markov state i is given. In other words, $\mathcal{N}_i(\epsilon)$ is ϵ-neighborhood at the current time, say p, under the condition that $J_p = i$. However, this cannot be directly applied to the recursive codebook design problem, since the recursive codebook should be constructed at time $p-1$. The recursive feedback encoding is performed in a causal fashion, and the information at the next time step cannot influence the operation in the previous time step. Therefore, we have to define the *next step* neighborhood instead, for the design of the recursive codebook. The next step ϵ-neighborhood of the Markov state j at time $p-1$ is defined as

$$\mathcal{N}'_j(\epsilon) = \{i \in \mathbb{J} | P_{i,j} \geq \epsilon\}, \quad j \in \{1, \cdots, N_{cb}\}. \tag{4.16}$$

$\mathcal{N}'_j(\epsilon)$ is the set of Markov states at time p for a given Markov state j at time $p-1$, while $\mathcal{N}_i(\epsilon)$ is the set of Markov states at time $p-1$ which leads to the current (at time p) Markov state i, see Fig. 4.3 for a graphical representation of this concept. The recursive codebook should be

[7] As it is shown in (*Huang et al.*, 2006b), the graph of the Markov chain is depicted as a mesh network on Grassmann manifold, due to the congruency of the Voronoi cells and the fact that the channel shape (directional information) is isotropic.

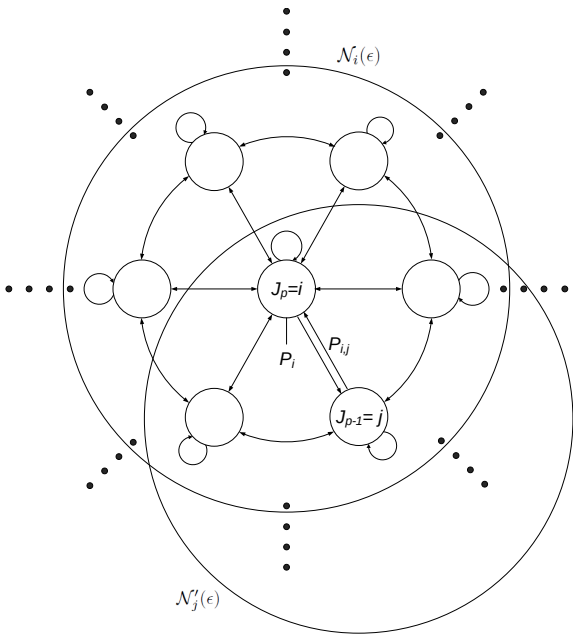


Figure 4.3: Markov chain diagram (Circles: Markov states, Arrows: positive transition probabilities, $\mathcal{N}_i(\epsilon)$: ϵ-neighborhood of the Markov state i at time p, $\mathcal{N}'_j(\epsilon)$: next step ϵ-neighborhood of the Markov state j at time $p-1$)

based on the next step ϵ-neighborhood of the Markov state j, and this requires the information about the channel transition probability. Our proposed method is based on heuristics, but it is simple and effective.

- Recursive Codebook Design

  The proposed recursive codebook design method does not make use of any channel statistical information. It takes an existing codebook $\mathcal{C}$ of size N_{cb} and reorganize it such that there exists a small size codebook which is decided by the current state for each codeword. This small size adaptive codebook $\mathcal{C}^r$, termed *recursive codebook*, is a subset of the whole codebook $\mathcal{C}$. The construction/operation procedure of the recursive codebook is as follows.

  1. **(Codebook Construction Phase)** We perform the codebook construction algorithm for a given training sequence $\mathcal{T}$ to get $N_{cb} = 2^{n_t}$ codewords. In our case, we use the mLBG algorithm for the codebook generation. The mLBG algorithm takes $\mathcal{T}$ and N_{cb} as input parameters, and provides a codebook $\mathcal{C}$ as an output. Thus, this can be formulated as follows.

  $$(\mathcal{C}) = \text{mLBG}(\mathcal{T}, N_{cb}) \tag{4.17}$$

2. **(Subspace Adjusting Mode)** Set the time index $p = 0$. The initial quantization is performed with the full size codebook $\mathcal{C}$, which requires n_t feedback bits. The quantization procedure with $\mathcal{C}$ can be rewritten as

$$\mathbf{C}_{J_p}^{\text{opt}} = \mathcal{Q}(\mathbf{H}_p, \mathcal{C}) = \arg\min_{\mathbf{C}_i \in \mathcal{C}} d_c(\mathbf{U}_p^{(\text{S})}, \mathbf{C}_i). \tag{4.18}$$

Here $d_c()$ is the chordal distance and its definition can be found in (2.8). This step is repeated for every τ_a (subspace adjusting feedback period) in an effort to prevent the codebook from losing track of the channel state.

3. **(Subspace Tracking Mode)**

   (a) We construct a recursive codebook $\mathcal{C}^r[J_p]$ for the given Markov state J_p at time p as a subset of $\mathcal{C}$ as follows:

$$\mathcal{C}^r[J_p] = \{\mathbf{C}_i | \forall i \in \mathcal{N}'_j(\epsilon) \text{ for } j = J_p\}. \tag{4.19}$$

   Here, the next step ϵ-neighborhood of the Markov state j, $\mathcal{N}'_j(\epsilon)$, is redefined as

$$\mathcal{N}'_j(\epsilon) = \{i \in \mathbb{J} | d_c(\mathbf{C}_i, \mathbf{C}_j) \leq \epsilon\}, \quad j \in \{1, \cdots, N_{cb}\}, \tag{4.20}$$

   where ϵ is taken such that the condition $|\mathcal{N}'_j(\epsilon)| = N_{rc}$ is satisfied. $\mathcal{N}'_j(\epsilon)$ is the set of Markov states at time p for the given Markov state j at time $p-1$. $N_{rc} = 2^{n_r}$ is the size of the recursive codebook, which has a fixed value for convenience of implementation. Therefore, $\mathcal{C}^r[J_p]$ is a collection of N_{rc} codewords chosen from $\mathcal{C}$ which are closest to $\mathbf{C}_{J_p}$. In this case, we have chosen the chordal distance $d_c(\mathbf{T}_i, \mathbf{T}_j)$ as the distance measure. We should note that the recursive codebook $\mathcal{C}^r[J_p]$ is centered around $\mathbf{C}_{J_p}$ and includes itself. The recursive codebook is dependent on the current state J_p and is of the same size irrespective of J_p, and it has a tracking capability (state-dependent codebook), which is illustrated in Fig. 4.4.

   (b) This procedure is repeated for every τ_t (subspace tracking feedback period). If τ_t elapses, then set $p \to p + 1$.

   (c) The quantization procedure is performed with the recursive codebook $\mathcal{C}^r[J_{p-1}]$ which is acquired by step (a). We should note that it requires only n_r feedback bits instead

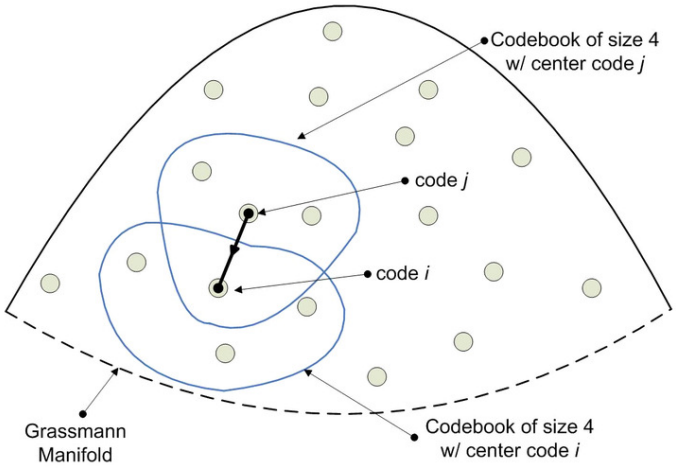


Figure 4.4: Recursive Codebook which is composed of $N_{rc} = 4$ neighboring codewords

of n_t bits.

$$\mathbf{C}_{J_p}^{\text{opt}} = \mathcal{Q}(\mathbf{H}_p, \mathcal{C}^r[J_{p-1}]) = \arg \min_{\mathbf{C}_i \in \mathcal{C}^r[J_{p-1}]} d_c(\mathbf{U}_p^{(\text{S})}, \mathbf{C}_i) \tag{4.21}$$

(d) If the actual time $t[p]$ at the time index p reaches a subspace adjusting feedback period τ_a, go back to step 2 (subspace adjusting mode). Fig. 4.5 depicts the corresponding time frame. Otherwise, set the state transition as (4.22) and go back to step (a) (subspace tracking mode), which allows the recursive codebook $\mathcal{C}^r[J_p]$ at the next time index to be centered around the most recent codeword $\mathbf{C}_{J_p}^{\text{opt}}$.

$$J_p = j, \text{ if } \mathbf{C}_{J_p}^{\text{opt}} = \mathbf{C}_j \tag{4.22}$$

There are several points worth our attention.

– *The codebooks can be constructed offline.*

The full size codebook and the recursive codebooks, which are corresponding to the step 1. and 3.(a), respectively, can be generated offline by simulation or based on the measured data. The BS and the MS do not have to construct the codebook at run time, but only need to store the predefined codebook tables.

– *The recursive codebook is a small size state-dependent codebook with channel tracking capability.*

The recursive codebook has a smaller size than the full size codebook $\mathcal{C}$, but it effectively

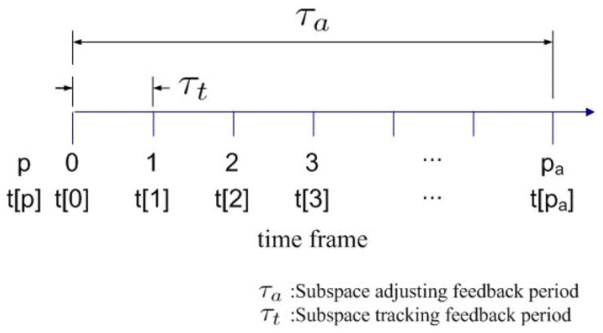


Figure 4.5: Feedback Time Frame: time index p, subspace tracking/adjusting feedback period τ_t/τ_a

exploits the temporal correlation of the channel by updating the codebook formation based on the previous channel quantization result. Its channel tracking capability allows it to become robust against the border effect.

– *The subspace adjusting feedback is required to reposition the codebook in case it loses track.* If the channel response changes slowly over time, the recursive codebook would lead to the same performance as the full size codebook. However, if the channel changes abruptly, the codebook could lose track of the channel and this error would propagate from that point on. Hence it is necessary to refer to the full size codebook so as to adjust to the current channel as frequently as needed.

4.2.3 Operation Scenario

The full size codebook $\mathcal{C}$ of size N_{cb} and the N_{cb} recursive codebooks $\mathcal{C}^r$ of size N_{rc} are constructed prior to the system deployment. The BS and the MS are supposed to store the full size codebook and tables defining the recursive codebooks.

The feedback operation consists of two modes, which are the *subspace adjusting mode* and the *subspace tracking mode*. The operation procedure is carried out as explained in the previous section. We assume that the time instants for sending back the codeword index J_p are determined by the periodic feedback scheme, and the subspace adjusting feedback period τ_a and the subspace tracking feedback period τ_t are defined accordingly.

4.2.4 On Decision of the Feedback Period

The decision of the feedback period which achieves the target performance is not a trivial issue. Semi-analytic methods are developed for the combined codebook case in Appendix C. In principle the methodology of calculating the fine feedback period τ_f should be directly applicable when finding the subspace tracking feedback period τ_t, since both of them are dependent on the size N_{cb} of the full size codebook.

In case of the subspace adjusting feedback period τ_a, it is expected to be longer than the coarse feedback period τ_c. The recursive codebook has a subspace tracking capability, which makes it robust against the border effect. Therefore the feedback overhead of the recursive scheme can be reduced compared with the combined feedback case by sending the subspace adjusting feedback less frequently than the coarse feedback. We should note that the recursive codebook should have a sizable number of codewords in order to have enough capacity to track the channel for a sufficiently long time period. The simulation result in Section 4.2.5 suggests that the recursive codebook size N_{rc} be at least 8, which may be dependent on the system configuration.

4.2.5 Numerical Results

The performance of the proposed recursive codebook method has been evaluated by Monte Carlo simulations. Simulations have been performed for the 2 BSs - 2 MSs case. The same simulation environment and the same parameters are used as Section 4.1.5. The performance of the recursive codebook is shown in Fig. 4.6. The recursive codebook of the n_t bit subspace adjusting and n_r bit subspace tracking feedback with corresponding feedback periods τ_a and τ_t is denoted by n_t-n_rbCQ ($[\tau_a]$,$[\tau_t]$), in which the unit of $[\tau_a]$ and $[\tau_t]$ is the number of OFDM symbols.

For $n_t = 5$ bits case, the feedback profiles include the following cases: 5bCQ - full size codebook as a reference case, 5-2bCQ - recursive codebook of size $N_{rc} = 4$, and 5-3bCQ - recursive codebook of size $N_{rc} = 8$. In case of $n_t = 6$ bits: 6bCQ - full size codebook as a reference and 6-3bCQ - recursive codebook of size $N_{rc} = 8$. The recursive codebook performances in terms of 50 % outage SINR are summarized in Table 4.1 and Table 4.2 for the case of $n_t = 5$ bits and $n_t = 6$ bits, respectively. The results are sorted by performance. The unit of the feedback overhead is bits per OFDM symbol per subcarrier, which is denoted by [b/sym/sc][8].

Simulation results show that there is no big gain when the size of the recursive codebook is 4 (5-2bCQ). In this case, the tracking capability of the recursive codebook is limited so that the

[8]In Section 4.1.5, [bits/sym/subcarrier] is used. This section uses [b/sym/sc] due to the space limitation in tables.

case with a feedback period pair $([\tau_a], [\tau_t]) = (20, 5)$ already suffers from a performance degradation compared with the (10,5) case. The channel tracking capability takes its effect when adopting the recursive codebook of size $N_{rc} = 8$ (5-3bCQ), which achieves the same performance as that of the (10,5) case with smaller feedback overhead by extending the subspace adjusting feedback period τ_a fifteen-fold from 10 to 150. The feedback overhead reduction ratio in [b/sym/sc] in comparison with the combined codebook case is in this case 12.4 % (Case 1 in Table 4.3). In case that a more relaxed target is allowed (Case 2 in Table 4.3) or a higher quantization level is required (Case 3 in Table 4.3), the feedback overhead-saving ratio can be over 20 %, up to 36 % according to the simulation results. The relevant simulation results selected from Fig. 4.6 can be found in Fig. 4.7.

Table 4.1: Recursive codebook performance and feedback overhead for $n_t = 5$ bits case (FB: feedback)

Rank	Profile	50 % outage SINR [dB]	FB overhead [b/sym/sc]
1	5bCQ	19.4	5.00
2	5-2bCQ$_1$ (10,5)	19.1	0.70
	5-3bCQ$_1$ (15,5)	19.1	0.73
	5-3bCQ$_2$ (100,5)	19.1	0.62
	5-3bCQ$_3$ (150,5)	19.1	0.61
3	5-3bCQ$_5$ (200,5)	18.7	0.61
4	5-2bCQ$_2$ (20,5)	18.6	0.55
5	5-3bCQ$_4$ (100,10)	18.4	0.32
6	5-2bCQ$_3$ (40,5)	18.0	0.48

Table 4.2: Recursive codebook performance and feedback overhead for $n_t = 6$ bits case (FB: feedback)

Rank	Profile	50 % outage SINR [dB]	FB overhead [b/sym/sc]
1	6bCQ	20.8	6.00
2	6-3bCQ$_1$ (9,3)	20.7	1.33
3	6-3bCQ$_2$ (150,3)	20.6	1.02

Table 4.3: Feedback overhead reduction (rCB: recursive codebook, cCB: combined codebook, Perf.: 50 % outage SINR, FB: feedback, Overh.: overhead [b/sym/sc], c ∈ [a,b]: $c \in \{c | a \leq c \leq b\}$)

Case	Perf. [dB]	rCB		cCB		FB reduction [%]
		Profile	Overh.	Profile	Overh.	
1	19.1	5-3bCQ$_3$ (150,5)	0.61	3+2bCQ$_1$ (10,5)	0.70	**12.4**
2	18.4	5-3bCQ$_4$ (100,10)	0.32	3+2bCQ$_2$ (10,10)	0.50	**36.0**
3	[20.4, 20.6]	6-3bCQ$_2$ (150,3)	1.02	3+3bCQ$_1$ (9,3)	1.33	**23.5**

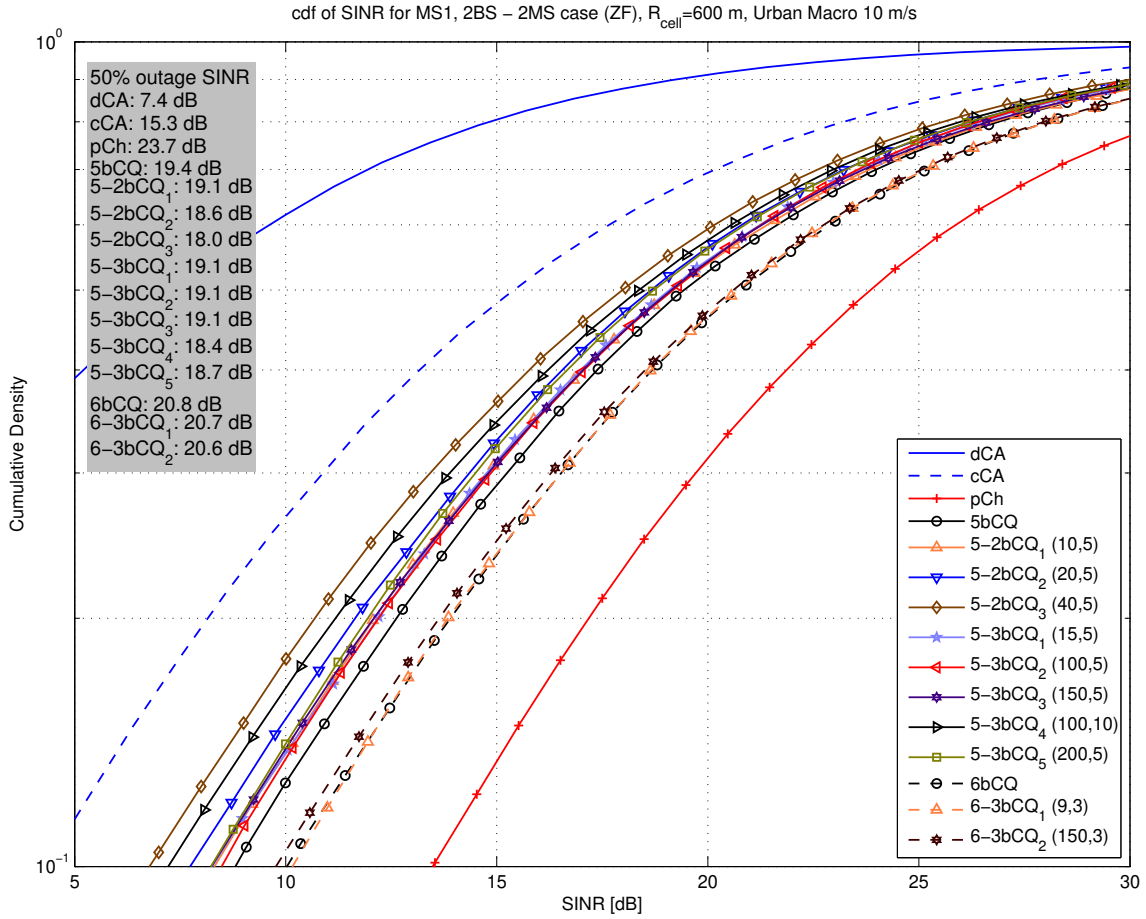


Figure 4.6: Recursive codebook simulation results for $(n_t$-n_rb)$\in\{$(5-2b),(5-3b),(6-3b)$\}$ case (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, 5bCQ: 5 bits CQ, n_t-n_rbCQ$_i$ ($[\tau_a]$,$[\tau_t]$): n_t bit subspace adjusting and n_r bit subspace tracking feedback with a feedback period pair $([\tau_a],[\tau_t])$)

4.2.6 Conclusion

In this section, the recursive codebook design method is proposed together with its application which operates in two modes, i.e., the subspace adjusting feedback mode and the subspace tracking feedback mode. The recursive codebook, which is capable of tracking the channel subspace, requires much less feedback overhead compared with the combined codebook scheme in return for a slightly increased cost of storing the table to be used for mapping the recursive code index[9].

[9]It has been also shown in an independent later work that the recursive codebook design concept can be directly applied to LTE downlink systems to reduce the feedback overhead.(*Hsu et al.*, 2013)

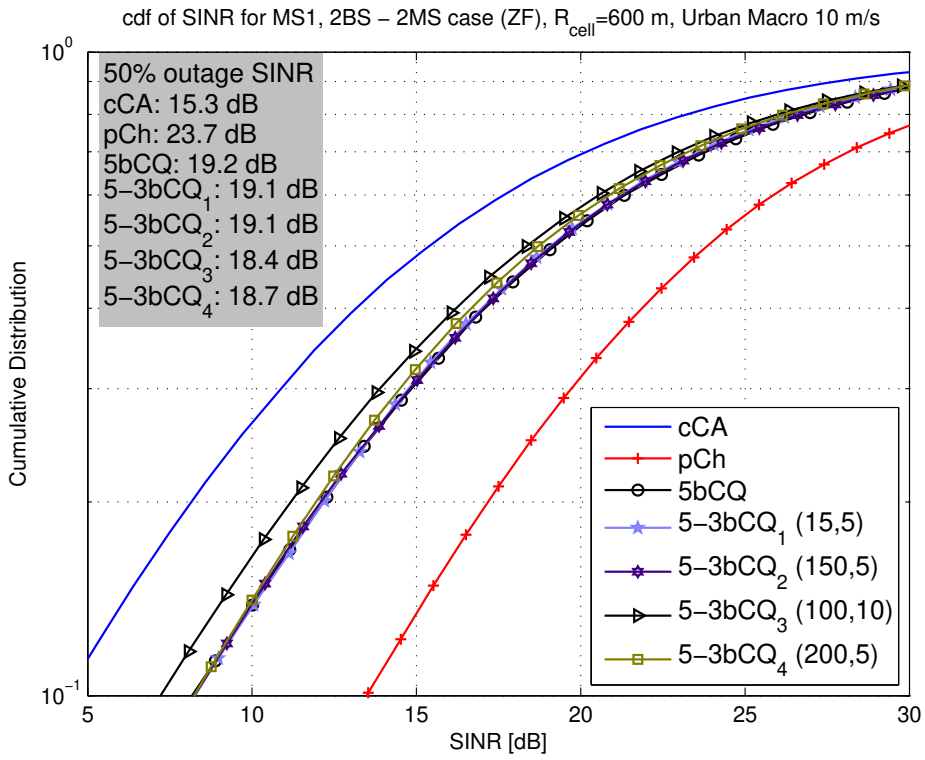


Figure 4.7: Recursive codebook simulation results for $(n_t$-n_rb)=(5-3b) case (cCA: centralized CA, pCh: perfect channel, 5bCQ: 5 bits CQ, 5-3bCQ$_i$ ($[\tau_a]$,$[\tau_t]$): 5 bits subspace adjusting and 3 bits subspace tracking feedback with a feedback period pair $([\tau_a],[\tau_t]) \in \{(15,5),(150,5),(100,10),(200,5)\}$)

4.3 Adaptive Codebook: Channel Adaptive Codebook Organization Method

4.3.1 Motivation

The feedback reduction methods which are introduced in the previous sections are *combined codebook* and *recursive codebook*. Both methods are inefficient due to the following reasons.

- Both methods incorporate a periodic feedback scheme.

  It is more efficient to send feedback messages only when it is required. Therefore, an aperiodic feedback scheme is better.

- Both methods do equal bit allocation for code index encoding.

  We can use a variable bit allocation by adopting a lossless data compression method in an effort to encode the feedback message efficiently (*Wade*, 1994).

4.3.2 Adaptive Codebook Construction

The *adaptive codebook* design method and the corresponding feedback encoding scheme are proposed in this section. We introduce an adaptive feedback encoding scheme which sorts the codebook entries according to their transition probabilities with respect to the current codebook entry[10]. As a result, the codebook entries are adaptively re-ordered in decreasing order of the transition probability. The most probable code in the next time step is placed in the earliest part of the codebook, and so on. Code index 1 is reported most frequently, and it is followed by code indices 2, 3, $\cdots$, in order. This allows us to consult source-coding schemes for data compression. It allocates a small number of bits to the frequently observed code index and more bits to the less frequent code indices. The required feedback rate can be greatly reduced by adopting the proposed method.

- Analytical Framework

  The theoretical framework used in this section is identical to that of the recursive codebook in many ways. Adaptive Vector Quantization (AVQ) is a vector quantization with memory, in which the quantization output depends on prior inputs as well as the current input. The AVQ is effectively formulated as the state transition and state dependent encoding, which can be considered as a first-order finite-state Markov chain. The important definitions which will appear in the process of analysis are as follows:

$\mathbf{H}_p$	downlink CSI at time p, $\mathbf{H}_p \in \mathbb{C}^{N_{tt} \times N_r}$
J_p	channel state (quantized channel index) at time p, $J_p \in \mathbb{J} = \{1, \cdots, N_{cb}\}$
	where N_{cb} is the codebook size
$\mathcal{C}$	codebook containing N_{cb} channel profiles
$\mathbf{C}_i$	$\mathbb{C}^{N_{tt} \times N_r}$ ith member of the codebook $\mathcal{C}$, $i \in \{1, \cdots, N_{cb}\}$
$\mathcal{C}^a$	adaptive codebook, $\mathcal{C}^a = \mathcal{C}$
$\mathbf{C}_i^a$	ith member of the adaptive codebook $\mathcal{C}^a$, $i \in \{1, \cdots, N_{cb}\}$,

  where N_{cb} is the size of the codebook which corresponds to the number of channel states. We should note that the adaptive codebook $\mathcal{C}^a$ is equal to the complete codebook $\mathcal{C}$ in that they share the same codes as their elements, but the individual code index of $\mathcal{C}^a$ is different from

[10] In (*Simon and Leus*, 2008a), a similar feedback compression method (entropy coding) has been proposed. The difference lies in the fact that the adaptive codebook construction method uses the chordal distance metric as an alternative to the transition probability metric, whereas the authors in (*Simon and Leus*, 2008a) have proposed and adopted an adaptive transition probability estimation method which is supposed to perform at runtime. In comparison, the proposed adaptive codebook can be constructed offline prior to runtime and is much more effective when the size of the codebook is large.

that of $\mathcal{C}$. In other words, the adaptive codebook $\mathcal{C}^a$ can be acquired by sorting codebook entries of $\mathcal{C}$ with respect to a certain criterion.

Regarding the CSI quantization operation, equations (4.10) and (4.11) in Section 4.2.2 hold. The stationary and transition probabilities of the first-order Markov chain are defined as (4.12) in Section 4.2.2.

Now AVQ is modeled by a first-order Markov chain[11]. Therefore, the downlink temporally correlated channel can be treated as a Markov source and makes it eligible for data compression techniques. The basic idea of the adaptive codebook is to re-arrange the order of the codewords based on the current chosen codeword such that the most probable next state codeword is located at the beginning of the codebook. The codewords are sorted in the decreasing order of their transition probabilities. Hence, we can use the source encoding scheme to perform the efficient representation of the feedback messages as in Section 4.3.3. Simply speaking, if some source symbols are known to be more probable than others, then we may exploit this feature in the generation of a source code by assigning short code words to frequent source symbols, and long code words to rare source symbols (variable-length code).

An adaptive codebook $\mathcal{C}^a[J_p]$ for the given Markov state J_p at time p can be constructed as follows.

$$\mathcal{C}^a[J_p] = \{\mathbf{C}_i | \forall i \in \mathbb{J}, \mathbf{C}_1^a = \mathbf{C}_{J_p}, P_{i,J_p} > P_{j,J_p} \text{ for } i,j \in \{i,j \in \mathbb{J} | i < j\}\} \tag{4.23}$$

The first codeword of the adaptive codebook is the code $\mathbf{C}_{J_p}$ chosen at time p, which has the largest transition probability from the current state, and the other codewords are sorted in the decreasing order of the transition probability. The order of the transition probability can be acquired without knowledge of the exact transition probabilities which are difficult to estimate. The codewords can be sorted in the order of the transition probabilities by using a certain criterion. For example, the subspace distance measure, e.g., the chordal distance, can be used for ordering.

The resulting adaptive codebook construction procedure is as follows.

1. **(Codebook Construction Phase)** We perform the codebook construction algorithm for a given training sequence $\mathcal{T}$ to get $N_{cb} = 2^{n_t}$ codewords. For example, we can use the mLBG algorithm for the codebook generation. The mLBG algorithm takes $\mathcal{T}$ and N_{cb} as

[11] Refer to Fig. 4.3 for graphical presentation of this concept.

input parameters, and provides a codebook $\mathcal{C}$ as an output. Thus this can be formulated as follows.

$$(\mathcal{C}) = \text{mLBG}(\mathcal{T}, N_{cb}) \tag{4.24}$$

2. **(Adaptive Codebook Construction Phase)** We construct an adaptive codebook $\mathcal{C}^a[J_p]$ for the given Markov state J_p at time p as a sorted version of $\mathcal{C}$ as follows:

$$\mathcal{C}^a[J_p] = \{\mathbf{C}_i | \forall i \in \mathbb{J}, \mathbf{C}_1^a = \mathbf{C}_{J_p}, \Delta d_c(\mathbf{C}_i^a) < \Delta d_c(\mathbf{C}_j^a) \text{ for } i, j \in \{i, j \in \mathbb{J} | i < j\}\},$$

$$\text{where } \Delta d_c(\mathbf{C}_k) := d_c(\mathbf{C}_1, \mathbf{C}_k). \tag{4.25}$$

In this case, we choose the chordal distance as the distance measure as defined in (2.8), but this can be replaced by any distance measure as long as it corresponds with the metric used for the codebook construction and the quantization. The comparison with (4.23) reveals the fact that the transition probability is replaced by the chordal distance, which simplifies the process.

The codewords of $\mathcal{C}^a[J_p]$ are sorted such that their chordal distances from the initial codeword $\mathbf{C}_{J_p}$ are in an increasing order. This order is equal to the decreasing order of the transition probabilities. Simulation results in Section 4.3.5.1 confirm its validity. We should note that the adaptive codebook can be constructed offline and only the table of the code indices is required in addition to the codebook for the aCB operation, once the elements of the codebook $\mathcal{C}$ are given.

The adaptive codebook operation procedure is as follows.

1. **(Initial Code Assignment Mode)** Set the time index $p = 0$. The initial codeword assignment is performed with the codebook $\mathcal{C}$, which requires n_{CB} feedback bits ($|\mathcal{C}| = 2^{n_{\text{CB}}}$). The quantization procedure with $\mathcal{C}$ can be rewritten as

$$\mathbf{C}_{J_p}^{\text{opt}} = \mathcal{Q}(\mathbf{H}_p, \mathcal{C}) = \arg \min_{\mathbf{C}_i \in \mathcal{C}} d_c(\mathbf{U}_p^{(\text{S})}, \mathbf{C}_i). \tag{4.26}$$

This step is required only when the adaptive codebook operation starts or the code transition tracking mode has failed by interruption or disruption.

2. **(Code Transition Tracking Mode)**

   (a) The adaptive codebook $\mathcal{C}^a[J_p]$ for the given Markov state J_p at time p can be acquired

by looking up the predefined table which is provided by (4.25).

(b) This procedure is repeated for every code transition tracking time. Set $p \rightarrow p+1$.

(c) The quantization procedure is performed with the adaptive codebook $\mathcal{C}^a[J_{p-1}]$ which is acquired by the step (a). For the slowly varying channel, we expect the adaptive code index to be 1 for most cases, which indicates that the same code is used as at the previous time. Otherwise it is restricted to the small number. The code index 1 can be encoded to a zero message (no feedback is transferred), and we can encode it in a smart way by using a source-coding scheme to reduce the feedback overhead. This will be discussed in detail in Section 4.3.3.

$$\mathbf{C}_{J_p}^{\text{opt}} = \mathbf{C}_{J_p^a}^{a\ \text{opt}} = \mathcal{Q}(\mathbf{H}_p, \mathcal{C}^a[J_{p-1}]) = \arg \min_{\mathbf{C}_i^a \in \mathcal{C}^a[J_{p-1}]} d_c(\mathbf{U}_p^{(\text{S})}, \mathbf{C}_i^a) \tag{4.27}$$

Here J_p^a is the optimum code index of the adaptive codebook $\mathcal{C}^a[J_{p-1}]$, which corresponds to the code index J_p of the original codebook $\mathcal{C}$. J_p^a needs to be fed back, and our design goal of the adaptive codebook is to ensure $J_p^a \ll J_p$.

(d) If any reason to initialize the adaptive codebook operation is detected, go to step 1 (initial code assignment mode). Otherwise, set the state transition as in (4.28) and go back to step (a) (code transition tracking mode), which allows the adaptive codebook $\mathcal{C}^a[J_p]$ at the next time index to be ordered with the most recent codeword $\mathbf{C}_{J_p}^{\text{opt}}$ as the initial codeword.

$$J_p = j, \text{ if } \mathbf{C}_{J_p}^{\text{opt}} = \mathbf{C}_j \tag{4.28}$$

4.3.3 Feedback Encoding

The feedback message (code index) is encoded before it is sent over the feedback link. For the feedback encoding process, we can exploit a source-coding technique to reduce the average feedback rate by using variable-length codes (*Simon and Leus*, 2008a; *Wade*, 1994). The schemes introduced in (*Simon and Leus*, 2008a; *Wade*, 1994) are lossless data compression techniques, whereas prior arts (combined codebook and recursive codebook) adopt a fixed-length codeword scheme in which the feedback message is encoded in a lossy fashion.

We assume two different underlying system models. The first model adopts a non-dedicated feedback channel. In this case, there is no dedicated uplink resource for the feedback, and the feedback link can be also used for data transmission. As a result, the receiver needs to know when

the codeword ends and when the data part starts. The second model assumes that there exists a dedicated feedback channel which is used to carry the feedback messages only. In this case, we can reduce the required energy for the feedback by reducing the average feedback overhead.

Two models are briefly explained below. Each model imposes different requirements for conducting a source-coding scheme. The second model is assumed for the proposed method, since it requires only the order of the transition probabilities for the encoding process and it shows better feedback overhead reduction performances.

- Non-dedicated Feedback Channel

  A non-dedicated channel requires the feedback codes to satisfy the prefix condition (*Haykin*, 2009). A codeword is not allowed to contain any other codeword as a prefix. However, the complete information about the transition probabilities is necessary to design a prefix-free code. One important class of prefix-free codes is the Huffman code.

- Dedicated Feedback Channel

  In case of the dedicated feedback channel model, we do not have to use prefix-free codes, as the beginning and the end of the codeword can be easily determined[12]. This property allows us to use non-prefix-free (NPF) codes which do not satisfy the prefix condition. NPF codes do not require the knowledge of the precise transition probabilities, but only need the order of the transition probabilities. The decreasing order of the transition probabilities is known to be same as the increasing order of the chordal distances with respect to the initial codeword. Therefore, the adaptive codebook indices can be encoded by using the NPF coding scheme, in the given increasing order. The basic principle of the NPF encoding is as follows. We assign no codeword (null) to the most probable codebook entry[13], and gradually longer codewords to the other codebook entries in the decreasing order of probability. The most probable codebook entry of the adaptive codebook is the code which is chosen at the previous step (code index 1), and other codes are sorted in the decreasing order of probability (increasing order of the chordal distance), which makes the application of the NPF coding straightforward.

  The NPF encoding example of the adaptive codebook indices is depicted in Table 4.4. The feedback overhead in bits, which is denoted by n, can be found as follows.

$$n = \lfloor \log_2(i+1) - \varepsilon \rfloor, \tag{4.29}$$

[12]The drawback of this method is that resources dedicated to the feedback channel cannot be used for data transmission.

[13]Note that no bit transmission is a certain state which has to be detected as well at the receiver.

Table 4.4: Adaptive codebook index encoding example

Code index (i)	NPF code (b)	Required resource (n) [bit]
1	/	0
2	0	1
3	1	1
4	00	2
5	01	2
6	10	2
7	11	2
8	000	3
9	001	3
10	010	3
11	011	3
12	100	3
⋮	⋮	⋮

where i is the code index of the adaptive codebook and ε is a small number which satisfies $\varepsilon > 0$. The NPF code corresponding to the code index i is formulated as follows.

$$b = bin(i - 2^n, n) \tag{4.30}$$

$bin(a, n)$ converts decimal number a to binary number, using at least n bits.

4.3.4 Operation Scenario

One exemplary implementation is described below, which follows the principles of the proposed method explained in Section 4.3.2. The base stations and the mobile stations have knowledge of the predefined codebook $\mathcal{C}$ and the corresponding adaptive codebook $\mathcal{C}^a$. The adaptive codebook can be acquired offline, according to (4.23) or (4.25). The BSs and MSs need to save the codebook $\mathcal{C}$ and the table of the adaptive codebook indices only. The dedicated feedback channel in the uplink is assumed.

At the initial phase (initial code assignment mode), the MS needs to perform the channel quantization process by using the whole codebook $\mathcal{C}$. In this case, n_t bits are required for the feedback. At the next time for the channel quantization (code transition tracking mode), the MS uses the adaptive codebook, the first codeword of which is the codeword chosen at the previous time. When the chosen codeword is same as the previous one (code index = 1), no feedback message is sent back. The feedback message is encoded according to the NPF coding scheme.

The required feedback resource is reduced, thanks to the properties of the adaptive codebook

and the data compression scheme. On the other hand, the BS is able to decode the feedback message without any extra information, since it can trace the adaptive behavior of the codebook. The code transition tracking mode can be continued as long as there is no interruption of feedback messages from the MS. When there is no reception of the feedback messages for a long time which is enough to lose track of the channel, we should switch to the initial code assignment mode.

4.3.5 Performance Evaluation

4.3.5.1 Numerical Results

The performance of the proposed adaptive codebook method has been evaluated by Monte Carlo simulations. Simulations have been performed for the 2 BSs - 2 MSs case. For the simulation environment and the used parameters please refer to Section 4.1.5. The proposed methods are tested for the Urban Macro channel with a mobile speed of 3 m/s and 10 m/s. The system performance is evaluated in terms of the received SINR at the MS. The adaptive codebook is denoted by aCB, and the simulation result for the recursive codebook case is also presented for the comparison.

Fig. 4.8 and Fig. 4.9 show the simulation results for the mobile speed 3 m/s case and the 10 m/s case, respectively. In case of the 3 m/s case (Doppler spread $f_D = 26$ Hz), the adaptive codebook (aCB) performs very closely to the perfect channel case (pCh) with only 0.7 dB degradation in terms of 50 % outage SINR, which costs just 0.23 bits/symbol. This means that the adaptive codebook scheme achieves the same performance as the 10 bits codebook case (10 bits/symbol) with less than 2.3 % of the feedback overhead when $v_{\mathrm{MS}} = 3$ m/s. In case of the 10 m/s case ($f_D = 86.7$ Hz), the aCB case performs 0.8 dB away from the pCh case with 0.66 bits/symbol feedback overhead. In comparison with the recursive codebook case (5-3bCQ (150,5)), the adaptive codebook has a 4.2 dB gain compared with the recursive codebook with just 8.2 % more feedback expenses. For the 3 m/s case, it achieves 3.7 dB gain with the same amount of feedback. Thus, the simulation results prove an efficiency of the proposed scheme.

Fig. 4.10 shows the histogram of the adaptive codeword indices, which have been acquired during the process of simulations depicted in Fig. 4.8 and Fig. 4.9. We have used the Monte Carlo simulation to collect 90,000 code indices which represent the 10 bits channel ($\mathbf{H} \in \mathbb{C}^{2\times 1}$) quantization results for the Urban Macro channel. According to the working mechanism of the adaptive codebook, the code index 1 means no codeword transition with respect to the previously selected codeword. Simulation results show that the code index transition events do not happen for 87.8 % and 62.5 % of time, in case of $v_{\mathrm{MS}} = 3$ m/s and $v_{\mathrm{MS}} = 10$ m/s, respectively. We should note that we do not have to

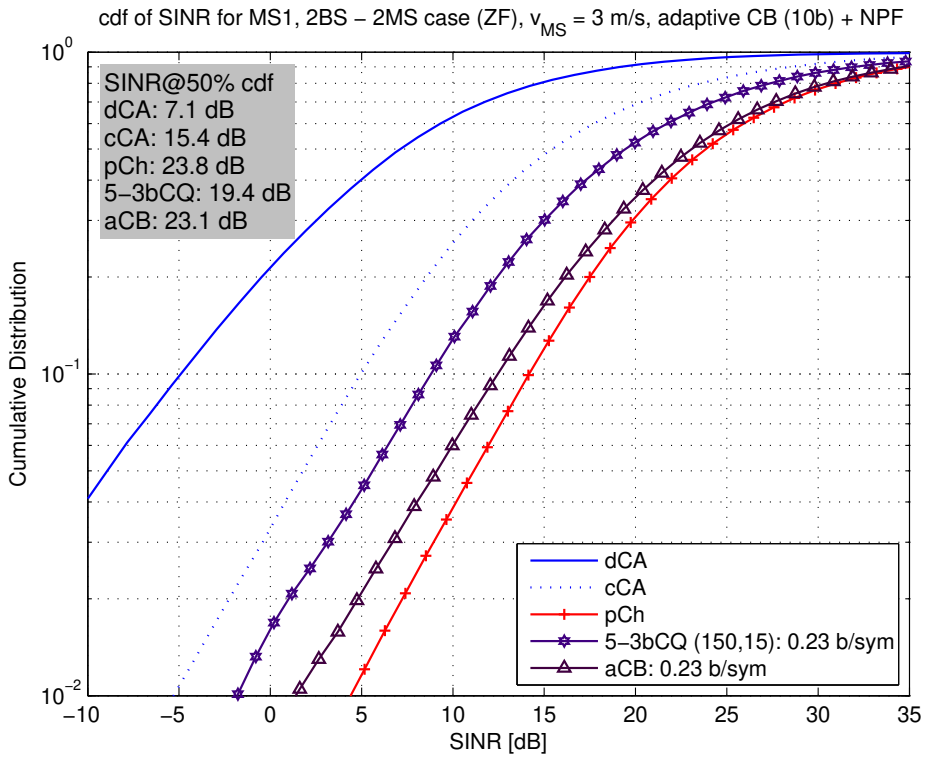


Figure 4.8: Adaptive codebook simulation results for $v_{\mathrm{MS}} = 3$ m/s case (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, n_t-n_rbCQ ($[\tau_a]$,$[\tau_t]$): n_t bit subspace adjusting and n_r bit subspace tracking feedback with a feedback period pair ($[\tau_a]$, $[\tau_t]$), aCB: adaptive codebook)

allocate any feedback resource[14] to the code index 1, when adopting a dedicated feedback channel.

In Fig. 4.10, the code index of the adaptive codebook is sorted in the decreasing order of its frequency (transition probability), up to the code index 11 ($v_{\mathrm{MS}} = 3$ m/s case) and 21 ($v_{\mathrm{MS}} = 10$ m/s case). The exceptional cases beyond this point can be neglected due to their rare incidence. Therefore, the histograms confirm the validity of using the chordal distance to order the codewords of the adaptive codebook, instead of the actual transition probability.

In case of $v_{\mathrm{MS}} = 3$ m/s, Fig. 4.11 and 4.12 illustrate typical code index behaviors over time for before and after aCB encoding cases, respectively. It is shown that the code index, which usually covers the whole code index range, can be mapped into a much smaller region by the transformation performed by the adaptive codebook encoding.

[14]This no feedback state needs to be detected at the receiver.

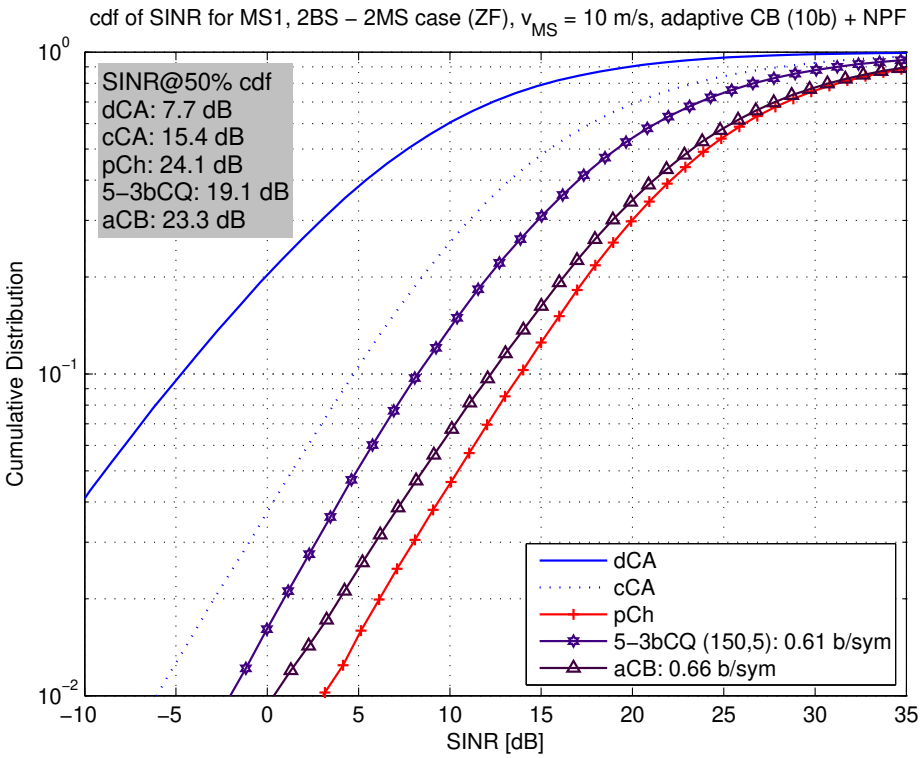


Figure 4.9: Adaptive codebook simulation results for $v_{\mathrm{MS}} = 10$ m/s case (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, n_t-n_rbCQ ($[\tau_a]$,$[\tau_t]$): n_t bit subspace adjusting and n_r bit subspace tracking feedback with a feedback period pair ($[\tau_a]$, $[\tau_t]$), aCB: adaptive codebook)

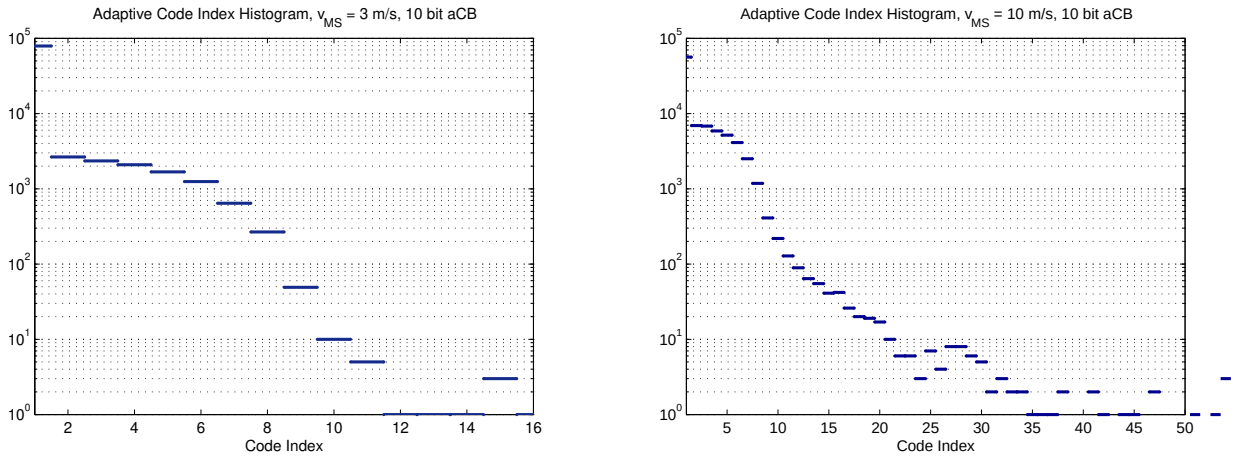
4.3.5.2 Advantages of the Adaptive Codebook

This section describes how the proposed method improves earlier solutions.

As explained in the previous sections, the proposed adaptive codebook can be built offline prior to the system deployment. The cost for the system is extra data storage at the MS and the BS for the adaptive code index table of size $N_{cb} \times N_{cb}$ (N_{cb} is the codebook size). The advantages which can be acquired in return for this low cost are as follows.

- The proposed adaptive codebook scheme is *very efficient*. For example, our simulation results show that we can achieve the 10 bits codebook performance with only 2.3 % of the required feedback bits (0.23 bits/symbol) by adopting the proposed scheme.

- The proposed adaptive codebook scheme is applicable to any system as long as there exists *correlation in the channel*. We can exploit the proposed scheme to reduce the feedback overhead



(a) Case I: $v_{\mathrm{MS}} = 3$ m/s ($f_D = 26$ Hz)

(b) Case II: $v_{\mathrm{MS}} = 10$ m/s ($f_D = 86.7$ Hz)

Figure 4.10: Adaptive code index histogram for 90,000 collected samples: $\mathbf{H} \in \mathbb{C}^{2\times 1}$, Urban Macro channel, 10 bits adaptive codebook (aCB) $|\mathcal{C}^a| = 2^{10}$

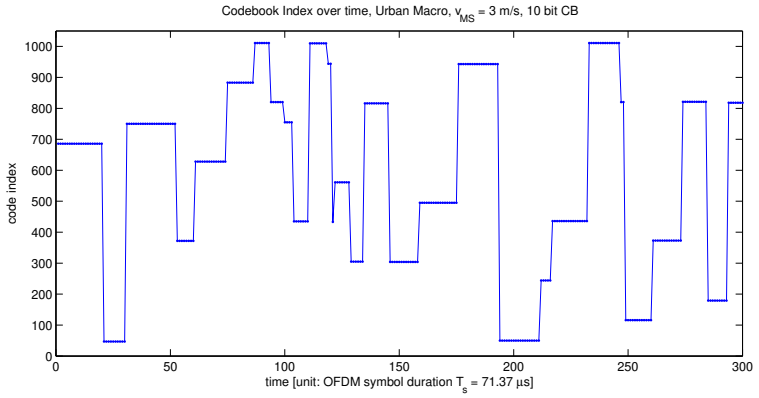


Figure 4.11: Typical normal (fixed arrangement) code index behavior over time

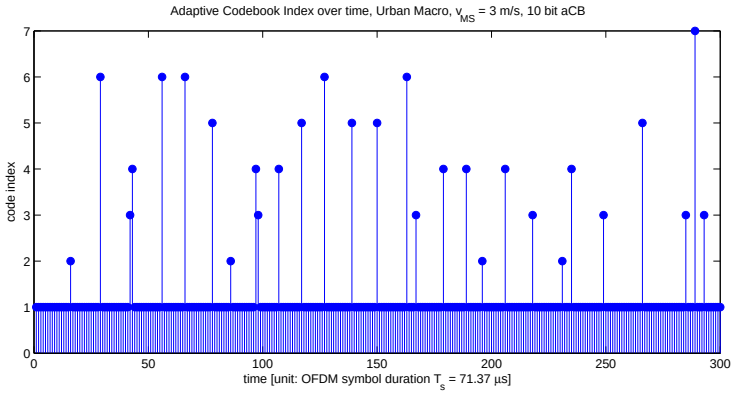


Figure 4.12: Typical adaptive code index behavior over time

for systems with a temporally correlated channel or for multi-carrier transmission systems like OFDM. In the latter case, we can apply the adaptive codebook scheme over subcarriers to make use of the correlation over adjacent carrier frequencies, which is illustrated in Chapter V.

- The proposed adaptive codebook scheme requires *no subspace adjusting procedure* which is necessary for the recursive codebook scheme as explained in Section 4.2. In case of the recursive codebook, we need to send the subspace adjusting feedback in order to avoid losing track of the subspace. On the contrary, the adaptive codebook scheme makes the adjusting procedure obsolete.

4.3.6 Conclusion

In this section, an adaptive codebook design method is proposed together with the feedback encoding method, as a means to reduce the feedback overhead by exploiting the temporal correlation of the channel. The adaptive codebook sorts the codewords in the increasing order of the chordal distance with respect to the initial codeword, which is the codeword chosen at the previous time step. The adaptive codebook design method converts the code index behavior from being widely spread over the whole code index spectrum to being concentrated in the initial part of the spectrum. This allows us to use the lossless data compression method to encode the feedback messages (*Simon and Leus*, 2008a; *Wade*, 1994). As a result, the adaptive codebook scheme in cooperation with the NPF encoding requires much less feedback overhead compared with other techniques, e.g., the recursive codebook scheme, in return for a slightly increased cost of storing the tables defining the adaptive codebook.

4.4 Summary

In this chapter, three feedback overhead reduction methods have been proposed which exploit the temporal correlation of the channel, i.e., the combined codebook design method, the recursive codebook design method and the adaptive codebook design method.

The combined codebook design method (cCB) is based on the hierarchical codebook structure. Its two-layered codebook structure definitely helps in reducing the feedback amount in comparison with the conventional full size feedback scheme, but its rigidness entails a border effect which limits its performance. The recursive codebook design method (rCB) addresses the border effect problem of the cCB by introducing a subspace tracking capability. Simulation results show that rCB is up

to 36 % more efficient in saving feedback overhead than cCB. However, rCB still bears the risk of losing track of the channel especially when there exist abrupt channel changes.

The methods mentioned above, i.e., cCB and rCB, are constrained by the following system requirements: periodic feedback and equal bit allocation for feedback. These system requirements make cCB and rCB eligible to be directly applicable in the existing communication system, as most of the systemcurrently in use adopt periodic fixed-size feedback, e.g., 3GPP LTE. However, they also make cCB and rCB inefficient. The adaptive codebook design method (aCB) is based on the assumption of employing an aperiodic variable-size feedback, which enables a better operation of the given full size codebook by incorporating a transition probability based codeword sorting scheme. It has been shown that aCB achieves the performance of a 10 bits full size codebook by using only 2.3 % of the uplink resource and aCB has a 4.2 dB gain compared with rCB with just 8.2 % more feedback expenses.

In short, aCB turns out to be the most efficient method of all three methods proposed in this chapter, but it requires a minor modification of the legacy communication system of incorporating an aperiodic variable-size feedback.

It should be noted that in this chapter an error-free and delay-free feedback link has been assumed for evaluation. It can be an interesting topic for further study to design and incorporate an UL error model in an effort to investigate a realistic feedback link and its impact to the overall system performance.

CHAPTER V

Adaptive Codebook for MIMO-OFDM: Two-Dimensional Adaptive Feedback Encoding with CSI Codeword Interpolation Method

This chapter suggests how to extend the adaptive codebook (aCB) encoding method, which is built on a subspace based quantization method, for MIMO-ODFM systems by proposing the two-dimensional adaptive feedback encoding method, the CSI codeword interpolation method, and how to combine them.

As is shown in Chapter IV, the aCB encoding method can reduce the feedback amount by exploiting the correlation of the channel. The aCB encoding method, which is proposed for narrowband channels, can be easily extended to frequency selective channels in case that orthogonal frequency division multiplexing (OFDM) is employed (*Chang*, 1966; *Weinstein and Ebert*, 1971). OFDM converts a broadband (frequency-selective) channel into a parallel collection of narrowband (frequency flat) subchannels. In OFDM system the channel correlation exists not only in the time domain but also in the frequency domain. In this case, the adaptive encoding scheme can be configured in two dimensions, i.e., the time domain and the frequency domain, and the same aCB encoding method can be applied in the frequency domain as in the time domain. In principle, the aCB encoding method can be performed for every active subcarrier in order to provide BSs with the complete channel information in the frequency domain. However, the feedback requirement grows in proportion to the number of subcarriers, and the total number of required feedback bits can be prohibitively high for the system even with a relatively small number of subcarriers. As the channel state information at the neighboring subcarriers is substantially correlated, an extended aCB encoding method is proposed in an effort to reduce the feedback requirements, which sends back only a fraction of the aCB code indices to the transmitter and reconstructs the CSI for all subcarriers through in-

terpolation at the transmitter. A related state of the art technique, i.e., a transmit beamformer interpolation method for MIMO-OFDM systems with limited feedback, is presented in (*Choi and Heath*, 2005). In this paper a spherical interpolator has been developed that introduces parameters for phase rotation to satisfy the phase invariance and unit norm properties of the transmit beamforming vectors. The phase rotation value needs to be fed back to the BS by the MS, as it is required for the beamforming vector interpolation at the BS. Unfortunately the cost function which is used to determine the phase rotation value is not directly applicable to our CSI codeword interpolation task, as the overall channel information, which is required for the phase value calculation, is not available at an individual MS[1]. In this chapter we introduce the CSI codeword interpolation method together with a corresponding cost function to determine a phase rotation value which is modified to make it directly applicable to the codeword interpolation task. Note that the proposed phase value calculation method can be performed at the BS as well as at the MS, as long as the channel information is available at the BS.

This chapter is organized as follows. In Section 5.1, the basic principle of the two-dimensional aCB encoding method is outlined. Section 5.2 briefly introduces the state of the art technique. Section 5.2 also provides the author's contributions, i.e., a novel CSI codeword interpolation method and a cost function for phase parameter calculation. Section 5.3 presents numerical evaluation results, and Section 5.4 finally concludes the chapter.

5.1 Adaptive Feedback Encoding over Frequency and Time

3GPP LTE, which is our target system, does not employ a physical layer (PHY) preamble to facilitate carrier offset estimation, channel estimation, timing synchronization, etc. Instead, special reference signals are embedded in the Physical Resource Block (PRB) pair as shown in Fig. 5.1[2]. A PRB is the smallest addressable unit, and it is defined as 12 consecutive subcarriers in frequency and 7 consecutive symbols in time. Every square within the grid represents a single subcarrier for one symbol period and is referred to as a Resource Element (RE). Reference signals, which are denoted as R in Fig. 5.1, are transmitted during the first and fifth OFDM symbols of each slot and every sixth subcarrier. Note that reference signals are staggered in both time and frequency. The channel response on subcarriers bearing the reference signals can be computed directly, and the aCB encoding can be done based on the CSI acquired at these reference signals. Let us introduce the term

[1] Note that in the MU-MIMO case an individual MS does not have access to the overall channel information but the channel information for the corresponding MS only.

[2] Refer to (*3GPP*, 2012) for more information about the LTE frame structure.

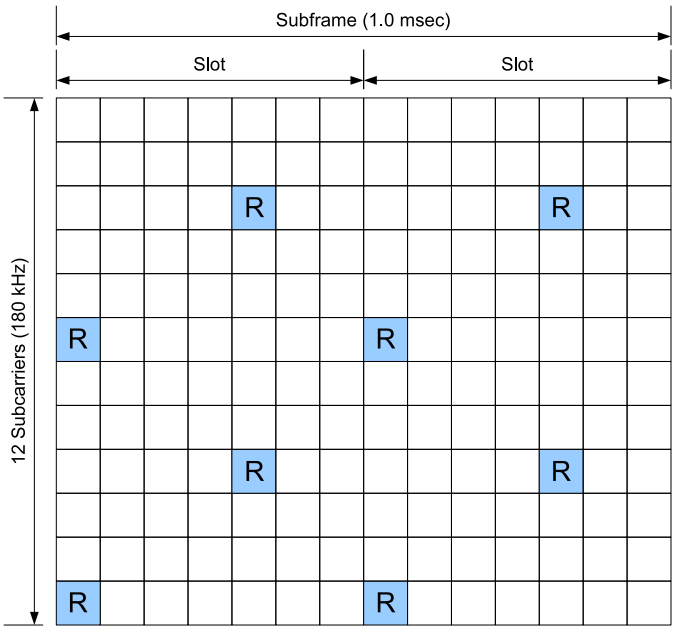


Figure 5.1: LTE reference signal arrangement (denoted as R) in the case of normal CP length for one antenna port: interspersed among resource elements

node here to indicate a reference signal allocated resource element at which the channel measurement and the adaptive feedback encoding takes place[3]. As LTE reference signals are interspersed in two dimensions, i.e., frequency and time, and the channel is correlated over frequency and time, the aCB encoding can proceed in two dimensions. The question is: in which sequence of nodes shall the aCB encoding be done?

The optimum aCB encoding sequence in terms of the required feedback resource can be found by an exhaustive search method. In short it generates all possible aCB encoding sequences and tests them to select the one which requires the minimum feedback resource. This method guarantees the optimum aCB encoding sequence, but is not practical as it requires too much computation at the MS. It should be also taken into account that it requires an additional feedback for the optimum encoding sequence indication, as the aCB encoding sequence should be signaled to the transmitter as well. In an attempt to reduce the additional feedback for aCB encoding sequence indication, in this chapter we propose a two-dimensional aCB encoding method with predefined encoding sequences, which allows two encoding sequences only. As there are two possible encoding sequences, the required feedback amount to indicate the aCB encoding sequence is only 1 bit. The graphical presentation of this concept can be found in Fig. 5.2. For example, as depicted in Case I of Fig. 5.2, the aCB

[3]Here it is assumed that not all reference signals are used for adaptive feedback encoding to reduce the required feedback amount. See Fig. 5.3 in which nodes are marked in red. Only the red colored reference signals are used for feedback encoding in this case.

encoding proceeds in parallel in the frequency domain and these multiple aCB encoding sequences over frequency are branched out from a single aCB encoding sequence in the time domain of which its encoding direction is perpendicular to the frequency domain sequences, and vice versa for Case II. You can draw analogy to a domino sequence starting from the middle left point (origin node in Fig. 5.2).

The two-dimensional (frequency and time) adaptive feedback encoding can be implemented as follows, depending on in which domain a higher correlation exists (frequency or time). Refer to Fig. 5.2 for a better understanding[4], and note that time/frequency correlation here is correlation of the channel estimates between nodes spread over time/frequency[5]. In short, the adaptive feedback encoding should be applied more frequently in a domain with a higher correlation in an effort to minimize the required aCB encoding feedback resources. If the frequency correlation is higher than the time correlation (Case I in Fig. 5.2), we can proceed aCB encodings in parallel in the frequency domain at different time instances (different symbols), and the adaptive encoding over time needs to be done only to provide a link between parallel aCB encoding sequences[6]. On the other hand, if the time correlation is higher than the frequency correlation (Case II in Fig. 5.2), the adaptive encoding over time should be performed in parallel at different frequency instances (different subcarriers), and the adaptive encoding over frequency needs to be done only to provide a link between the parallel aCB encoding sequences.

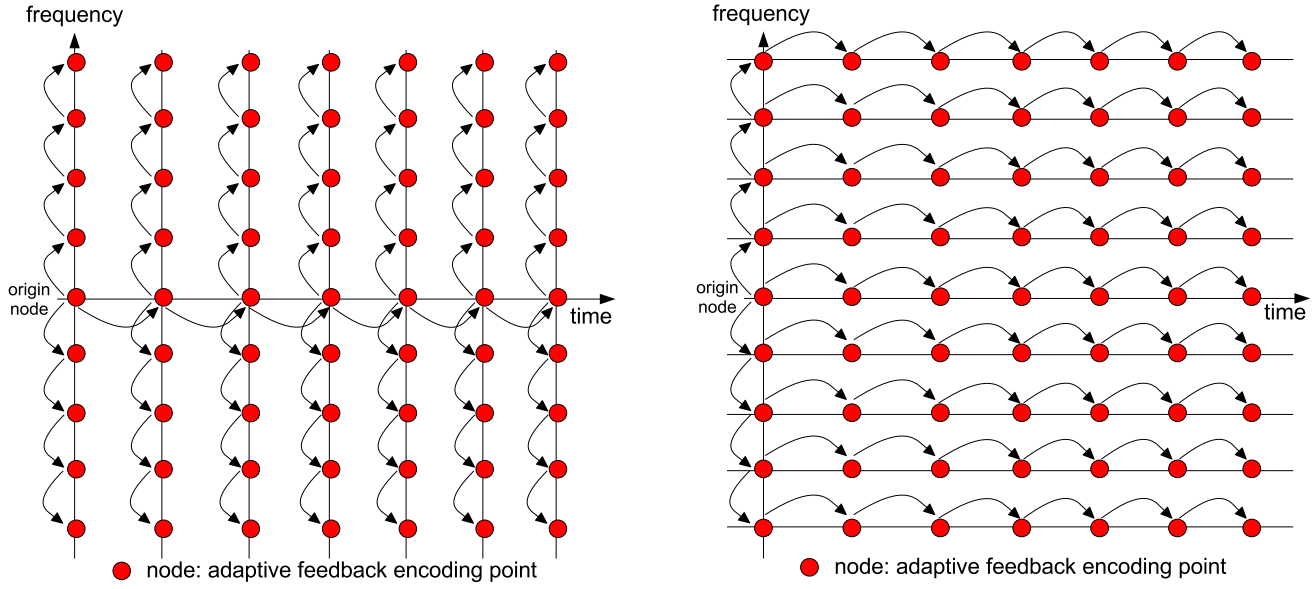
5.2 Multi-carrier Strategy

As explained in Section 5.1, the LTE frame structure allows us to acquire the CSI at every sixth subcarrier, and we need to calculate the appropriate CSI in-between subcarriers. In OFDM, the neighboring subcarriers are usually significantly correlated, and this results in high correlation between neighboring channel codewords. One way of using the correlation of the channel codewords is to combine the neighboring subcarriers into a cluster and use the same channel codeword correspond-

[4]The topology of the selected nodes follows that of Fig. 5.3. Note that in this case the reference signals which are aligned in time and frequency axis are selected for the nodes.

[5]How large the frequency/time encoding space can depend on the resource allocation status. The minimum possible size is one PRB pair, i.e., 12 subcarriers in frequency and 14 OFDM symbols in time. The maximum possible size can be 100 PRBs (1200 subcarriers) in frequency and up to several minutes in time.

[6]In Fig. 5.2 (a), the time domain adaptive feedback encoding starts from the origin node and proceeds in both directions, in an effort to reduce the possible impact of error propagation. The same logic applies to the case of Fig. 5.2 (b). In principle, its starting point can be located at any node, as the aCB encoding is a lossless encoding scheme and encoding error might happen only when a lossy encoding scheme, e.g., cCB or rCB, is employed. When a lossless encoding scheme is employed, the position of the starting point does not have any impact to the performance in terms of the required feedback resource, as the required feedback resource depends not on the length of the encoding sequence but on the channel correlation between the directly adjacent nodes. However, the aCB encoding scheme, when modified to restrict the search space for practical use, can be converted to a lossy scheme. In this case, it would be better to reduce the length of the encoding sequence.



(a) Case I: frequency correlation > time correlation

(b) Case II: time correlation > frequency correlation

Figure 5.2: Adaptive feedback encoding in two dimensions - frequency and time (arrow indicates a direction of the aCB encoding sequence, and red circle indicates a node)

ing to the center subcarrier in the cluster (*Grünheid et al.*, 2001). This method will be referred to as *clustering* in this chapter. If we combine K subcarriers[7] into one cluster, the amount of feedback information is reduced to $1/K$. As the cluster size K increases, however, the system performance is significantly degraded due to the distortion experienced by subcarriers near the cluster boundary. Hence, there is a trade-off between the performance and the feedback amount reduction ratio. In (*Choi and Heath*, 2005), an interpolation based method has been proposed as an alternative to reduce feedback requirements, which can be applied to our case with some modifications.

5.2.1 Transmit Beamforming Interpolation Method for SU-MIMO-OFDM with Limited Feedback

In (*Choi and Heath*, 2005), a limited feedback architecture is proposed which combines beamforming vector quantization and smart vector interpolation for single user MIMO OFDM systems. In the proposed system, the receiver sends a fraction of information about the optimal beamforming vectors to the transmitter and the transmitter computes the beamforming vectors for all subcarriers through interpolation. The spherical interpolation algorithm is used to preserve the unit norm con-

[7] The number of subcarriers in one cluster K can be 6, for example, as the reference signals are located in every 6 subcarrier.

straint of the beamforming vector, and a phase parameter is introduced to take the lack of uniqueness of the optimal beamforming vector into account. The beamforming vectors and phase parameters are quantized at the receiver and the quantized information is provided to the transmitter.

The interpolator in (*Choi and Heath*, 2005), which performs a weighted average of the quantized beamforming vectors and re-normalizes the result to place it on the unit sphere, is expressed as follows for given $\{\widehat{\mathbf{w}}(lK'+1), 0 \le l \le N/K'-1\}$. Here, N is the total number of subcarriers, K' is the size of the interpolation area in terms of the number of subcarriers, and $\widehat{\mathbf{w}}(k)$ is the beamformer vector at the k-th subcarrier[8] which is determined by selecting the codeword maximizing the effective channel gain[9].

$$\widetilde{\mathbf{w}}(lK'+k;\theta_l) = \frac{(1-c_k)\widehat{\mathbf{w}}(lK'+1) + c_k\{e^{j\theta_l}\widehat{\mathbf{w}}((l+1)K'+1)\}}{\|(1-c_k)\widehat{\mathbf{w}}(lK'+1) + c_k\{e^{j\theta_l}\widehat{\mathbf{w}}((l+1)K'+1)\}\|} \tag{5.1}$$

where $c_k = (k-1)/K'$ is the linear weight value, θ_l is a parameter for phase rotation with $0 \le l \le N/K'-1$. While the spherical interpolator only utilizes $\widehat{\mathbf{w}}(lK'+1)$ and $\widehat{\mathbf{w}}((l+1)K'+1)$, the modified interpolator evaluates the beamforming vector from $\widehat{\mathbf{w}}(lK'+1)$ and $e^{j\theta_l}\widehat{\mathbf{w}}((l+1)K'+1)$. The role of θ_l is to remove the distortion caused by the arbitrary phase rotation of the optimal beamforming vector[10].

To maximize the performance of the interpolator, the receiver calculates the optimal phase $\{\theta_l, 0 \le l \le N/K'-1\}$ based on a performance metric which is intended to maximize an effective channel gain. The optimal θ_l maximizing the minimum effective channel gain can be found as

$$\theta_l = \arg\max_{\theta} \min_{lK+1 \le k \le (l+1)K} \|\mathbf{H}(k)\widetilde{\mathbf{w}}(k;\theta)\|^2. \tag{5.2}$$

The normalization factor in (5.1), however, makes it difficult to solve (5.2) analytically. Hence a modified cost function is provided which uses a numerical grid search as follows.

$$\theta_l = \arg\max_{\theta\in\Theta} \|\mathbf{H}(lK'+K'/2+1)\widetilde{\mathbf{w}}(lK'+K'/2+1;\theta)\|^2 \tag{5.3}$$

[8] We focus on frequency interpolation here, but the same interpolation method can be applied in the time domain.

[9] The closed form formulae of the transmit beamforming vectors of various types, i.e., Zero Forcing filter, Wiener filter and Matched filter, can be found in (*Joham et al.*, 2005).

[10] A typical pth-order spherical interpolator is formulated as $\widetilde{\mathbf{v}} = \frac{\sum_{i=0}^{p} b_i \mathbf{v}_i}{\|\sum_{i=0}^{p} b_i \mathbf{v}_i\|}$ where $\|\mathbf{v}_i\| = 1$ and $\{b_i\}$ are weight values such that each $b_i \ge 0$ and $\sum_{i=0}^{p} b_i = 1$. When $p = 1$, the interpolated vector $\widetilde{\mathbf{v}}$ is a point on the spherical line between points $\mathbf{v}_0$ and $\mathbf{v}_1$. This algorithm cannot be applied to the beamformer interpolation directly due to the lack of uniqueness of the optimal beamforming vector. That is to say: when $\mathbf{w}(k)$ is the optimal beamforming vector maximizing the effective channel gain $\Gamma(k) = \|\mathbf{H}(k)\mathbf{w}(k)\|^2$, $e^{j\theta}\mathbf{w}(k)$ also maximizes the effective channel gain. In other words, the optimal beamforming vector is not a unique point but represents a complex line. The phase θ has a dramatic impact on the resulting interpolated vectors, and this makes it difficult to directly apply spherical interpolation to perform beamformer interpolation.

where $\Theta = \{0, \frac{2\pi}{P}, \frac{4\pi}{P}, \cdots, \frac{2(P-1)\pi}{P}\}$, and P is the number of quantized levels. The additional phase parameter θ requires $\log_2 P$ extra feedback bits per interpolation area. We should note that (5.2) is approximated by (5.3) based on the assumption that the subcarrier $(lK' + K'/2 + 1)$ has the worst effective channel gain, as the subcarrier $(lK' + K'/2 + 1)$ is the farthest from the subcarriers $(lK' + 1)$ and $((l + 1)K' + 1)$ of which the beamforming vectors are available via feedback. It is also worthwhile to note that the cost function to find $\{\theta_l\}$ requires the overall channel matrix at subcarrier[11] $(lK' + K'/2 + 1)$, i.e., $\mathbf{H}(lK' + K'/2 + 1)$, which is available for the single-user case as in (*Choi and Heath*, 2005), unlike for the multi-user case.

The receiver transmits $\{\theta_l\}$ along with the indices of the selected beamforming vectors $\{\widehat{\mathbf{w}}(lK' + 1); 0 \leq l \leq N/K' - 1\}$ to the transmitter. At the transmitter, the beamforming vectors for all the subcarriers are computed by (5.1).

5.2.2 CSI Codeword Interpolation Method for MU-MIMO-OFDM with Limited Feedback

We should note that the beamformer interpolation method explained in Section 5.2.1 is based on the single-user case. In the multi-user case, the overall channel matrix can be formulated as

$$\mathbf{H} = [\mathbf{H}_1, \cdots, \mathbf{H}_j, \cdots, \mathbf{H}_{N_{\mathrm{MS}}}]^{\mathcal{T}}, \ j\text{: user index}, \tag{5.4}$$

where $\mathbf{H}$ is the overall channel matrix, $\mathbf{H}_j$ is the transpose of the channel matrix for user j, and N_{MS} is the number of MSs. Equation (5.4) reveals the fact that the overall channel matrix $\mathbf{H}$ is composed of the channel matrices of the individual MSs $\mathbf{H}_j$. Therefore the precoding matrix at the transmitter should be calculated by taking into account the overall channel matrix $\mathbf{H}$, not just the relevant MS's channel matrix $\mathbf{H}_j$ in order to suppress the inter-user interferences. In Chapter II, the channel quantization method is proposed as a way of combating the inter-user interference. In short, each individual MS quantizes its channel by using a predefined codebook and sends its index back to the BS. Then the BS collects the channel informations which are provided by MSs in the form of codewords and recomposes the overall channel matrix, which is used to calculate the transmit precoding matrix.

[11] Even though it is not explicitly mentioned in the reference, the basic assumption is that the estimated channel is available at this subcarrier. Note that the reference is based on the general OFDM systems, not necessarily LTE.

The channel codeword interpolator, which is inspired by (5.1), can be formulated as follows.

$$\widetilde{\mathbf{h}}_j(lK'+k;\theta_l) = \frac{(1-c_k)\widehat{\mathbf{h}}_{j,l} + c_k e^{j\theta_l}\widehat{\mathbf{h}}_{j,l+1}}{\|(1-c_k)\widehat{\mathbf{h}}_{j,l} + c_k e^{j\theta_l}\widehat{\mathbf{h}}_{j,l+1}\|}, \tag{5.5}$$

where $\widetilde{\mathbf{h}}$ is the interpolated channel codeword at MS j, $\widehat{\mathbf{h}}_{j,l} = \mathcal{Q}(\mathbf{H}_j(lK'+1))$ is the quantized CSI at MS j and $c_k = (k-1)/K'$ is the linear weight value in which K' is the size of the interpolation area. Here, $\mathcal{Q}(\cdot)$ is a channel quantization function which is expressed as $\mathcal{Q}(\mathbf{H}_j) = \arg\min_{\mathbf{C}_i \in \mathcal{C}} d_c(\mathbf{U}_j^{(\mathrm{S})}, \mathbf{C}_i)$, where $\mathbf{U}_j^{(\mathrm{S})}$ is a column space basis vector of $\mathbf{H}_j$, and $d_c(\mathbf{A},\mathbf{B})$ is the chordal distance between $\mathbf{A}$ and $\mathbf{B}$. It means that $\widehat{\mathbf{h}}_{j,l}$ is selected from the given codebook $\mathcal{C}$. Note that the proposed codeword interpolation method is designed for the case that $\widehat{\mathbf{h}}_{j,l}$ is a vector[12], i.e., $\widehat{\mathbf{h}}_{j,l} \in \mathbb{C}^{N_{tt}\times 1}$. Hence its application is at the moment restricted to the case that the MS has a single antenna, i.e., $N_r = 1$.

Regarding how to find the optimal θ_l, we can think of the novel cost function (5.6) which is motivated by the fact that the CSI for MS j at subcarrier $(lK'+K'/2+1)$, i.e., $\mathbf{H}_j(lK'+K'/2+1)$, can be presented in two different ways, i.e., through quantization ($\widehat{\mathbf{h}}_{j,l+0.5}$) or interpolation ($\widetilde{\mathbf{h}}((l+0.5)K'+1;\theta_l)$), which makes the BS as well as the MS capable of finding θ_l. The idea is to find θ_l which leads to the minimum distortion between the interpolated channel information and the quantized channel information at the middle of the interpolation area. We propose to choose the chordal distance as a distortion metric to keep a unified approach. Hence θ_l can be acquired by

$$\begin{aligned}\theta_l &= \arg\min_{\theta\in[0,2\pi)} d_c\left(\widehat{\mathbf{h}}_{j,l+0.5}, \widetilde{\mathbf{h}}_j((l+0.5)K'+1;\theta)\right) \\ &= \arg\min_{\theta\in[0,2\pi)} d_c\left(\widehat{\mathbf{h}}_j\left(lK'+\frac{K'}{2}+1\right), \widetilde{\mathbf{h}}_j\left(lK'+\frac{K'}{2}+1;\theta\right)\right).\end{aligned} \tag{5.6}$$

The interpolated channel information at the middle of the interpolation area can be formulated by substituting $k = 0.5K'+1$ in (5.5) as

$$\widetilde{\mathbf{h}}_j((l+0.5)K'+1;\theta) = \frac{0.5\widehat{\mathbf{h}}_{j,l} + 0.5e^{j\theta}\widehat{\mathbf{h}}_{j,l+1}}{\|0.5\widehat{\mathbf{h}}_{j,l} + 0.5e^{j\theta}\widehat{\mathbf{h}}_{j,l+1}\|}. \tag{5.7}$$

Note that θ_l can be found by a numerical method at the BS as well, as the MS j conveys the quantized CSI at the mid-point of the interpolation area ($\widehat{\mathbf{h}}_{j,l+0.5}$) to the BS. As discussed in Section 4.3, we can reduce the feedback overhead to a reasonable amount by adopting the adaptive encoding scheme,

[12] In (*Choi and Heath*, 2005), it has been shown that the beamformer interpolation method can be extended to the system with two receive antennas at the MS, when channel impulse responses between different transmit and receive antenna pairs are independent. However this cannot be applied to our case, as it is feasible only for the dominant eigenmode transmission case.

when there exists channel correlation in the frequency domain. Therefore we propose that the MS should transmit the indices of the selected channel codewords $\{\widehat{\mathbf{h}}_{j,l}, \widehat{\mathbf{h}}_{j,l+0.5}; 0 \leq l \leq N/K'-1\}$ to the transmitter, instead of transmitting the phase values $\{\theta_l; 0 \leq l \leq N/K'-1\}$ separately. Since the MS j should transmit the indices of the selected channel codewords to the transmitter anyway, it is not required to transmit the phase values separately. With two-dimensional aCB encoding in operation, we can get the full benefit of the CSI codeword interpolation method at the expense of the slightly increased complexity at the BS without introducing any change of operation at the MS. As the CSI should be available at the middle point of the interpolation area, the minimum of the half size of the interpolation area is the distance between adjacent reference signals, as the estimated channel state information can be obtained only at resource elements in which the reference signals are located. However, the channel information at the in-between subcarriers can also be acquired by adopting Wiener filtering which has been widely used for two-dimensional pilot aided channel interpolation in OFDM systems (*Hoeher et al.*, 1997). This would allow a finer codeword interpolation at the expense of an increased computational complexity.

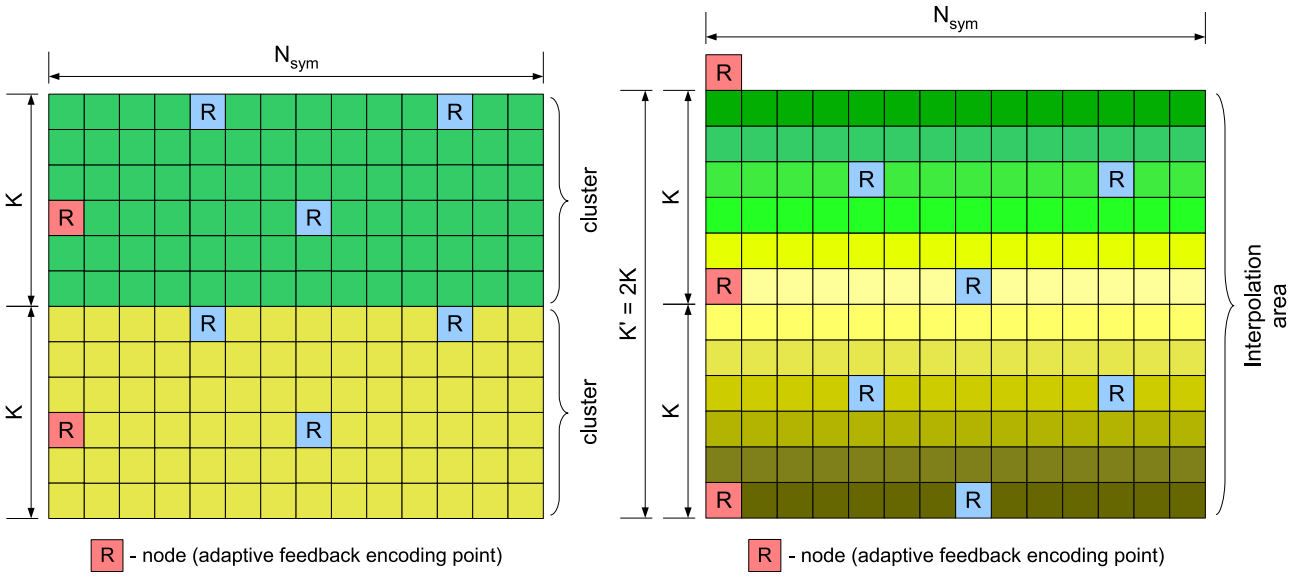
5.3 Numerical Results

5.3.1 Simulation Set-up

The proposed schemes are evaluated by Monte Carlo simulation methods. Our target system is 3GPP LTE, and the frame structure of its Physical Resource Block (PRB) is shown in Fig. 5.1. As explained in the previous sections, our multi-carrier strategy is composed of two methods, i.e., a clustering method and a codeword interpolation method, which are evaluated by simulations incorporating the LTE PRB structure.

The frame structure of the simulation, which is based on the LTE PRB, is depicted in Fig. 5.3. A square represents a resource element, and a square with a letter R, i.e., R, depicts a reference signal at which the receiver can estimate the channel state information. A reference signal notation in red color means a node at which the receiver performs the adaptive feedback encoding[13]. In Fig. 5.3 (a), K is the size of one cluster over the frequency axis, and its unit is the number of subcarriers. As a cluster should have at least one reference signal for channel estimation, the lower bound of K is D, i.e., the distance between two adjacent reference signals interspersed in frequency. In our simulation, we assume the case in which K is a multiple of D ($K = nD$ for $n \in \{1, 2, 3, \cdots\}$).

[13]Not all reference signals are used as nodes in Fig. 5.3. Yet the CSI at the non-node reference signals which are marked in blue can be used to improve the channel estimation at the MS and aCB encoding at the nodes.



(a) Clustering Method

(b) Codeword Interpolation Method

Figure 5.3: Feedback reduction methods for multi-carrier transmission systems (boxed R with blue background: reference signal, boxed R with red background: reference signal which is selected as node, box with greenish background: resource element to be served by the background color indicated precoding matrix)

Fig. 5.3 (a) shows the case of $K = D = 6$. N_{sym} is the feedback period, distance between nodes in the time domain, (unit: the number of OFDM symbols). In other words, it is a time duration during which the same precoding matrix is applied for a certain cluster (clustering method) or a subcarrier (codeword interpolation method). We presume a feedback channel to be error-free and delay-free, meaning that the transmit filtering calculated by the channel estimation and quantization result at the node can be directly applied to the corresponding cluster or the interpolation area. The resource elements which are served by the same precoding matrix are indicated by the same color in Fig. 5.3. We should note that the same precoding matrix is applied to the whole cluster (Fig. 5.3 (a)), whereas the interpolated precoding matrices, which are depicted by gradient colors, are applied to the individual subcarriers in the interpolation area (Fig. 5.3 (b)). K' in Fig. 5.3 (b) is the size of one interpolation area. It should be noted that the condition $K' = 2K$ should be fulfilled to make a fair performance comparison between the clustering method and the codeword interpolation method. In short, the comparison should be performed with respect to the same amount of feedback overhead, in other words, the same number of nodes. The feedback messages are encoded over nodes in time and frequency as depicted in Fig. 5.2, using the adaptive scheme.

5.3.2 Simulation Results

First, simulations have been performed to compare the required feedback overheads when encoding over time or over frequency. A cooperative antenna system with 2 BSs and 2 MSs case is assumed. Both BSs and MSs have a single antenna, and the ZF method is used for transmit filtering (precoding). Two BSs cooperate to serve two MSs at the same time, and the cdf of the SINR at MS1 is shown in Fig. 5.4. An urban macro channel is used and the MS speed is 3 m/s. The carrier frequency is 2.6 GHz. A clustering method is used for multi-carrier processing. The total number of subcarriers N is 72, and the size of a cluster K is 6 or 12. A 10 bit adaptive codebook is employed and the feedback period is chosen to be 14 OFDM symbols (1 ms)[14]. Each drop simulates 30 subframes, i.e., 420 consecutive OFDM symbols. Simulation results have been collected after 300 drops. The amount of the required feedback is calculated in bits/sym for both cases: adaptive codebook encoding over time (t) and frequency (f).

The simulation results in Fig. 5.4 show that the feedback encoding over time case leads to less feedback overhead in comparison with the encoding over frequency case, in both cases (aCB_1 (72,6) and aCB_2 (72,12)). It means that the time correlation is stronger than the frequency correlation in this case. The adaptive encoding over time with $(N, K) = (72, 6)$ case where N is the total number of subcarriers and K is the cluster size, aCB_1 (72,6), is only 1.4 dB away from the perfect channel information at the transmitter case, pCh, at the 50 % outage SINR. It requires less than 27 (1.90*14=26.6) bits over one Transmit Time Interval (TTI) (1 ms) for the whole 72 subcarriers. In case of aCB_2 (72,12), the 50 % outage SINR performance shows 2.9 dB degradation with less than 14 bits feedback overhead. We should note that the aCB_1 (72,6) case requires less than two times of the feedback overhead of the aCB_2 (72,12) case, when the feedback encoding is done over frequency. It is because the aCB_1 (72,6) case can exploit higher channel correlation than the aCB_2 (72,12) case. Therefore the former requires less than two times of the feedback overhead of the latter, even though the former has two times of the number of nodes[15] of the latter.

Fig. 5.5 shows the reference performance of the adaptive feedback encoding scheme. We set $K = 1$ and $N_{sym} = 1$, meaning that all resource elements are dedicated to the adaptive feedback encoding. This case is of course not a realistic scenario, but it gives us some insight on the maximum performance we can achieve by adopting the two-dimensional adaptive feedback scheme. The simulation result shows that the reference case is only 0.5 dB and 0.3 dB away from the prefect channel knowledge case for 50 % and 5 % outage SINR, respectively. The required feedback overhead,

[14]The shortest feedback period defined in 3GPP LTE is the duration of one subframe (1 ms).
[15]A node is defined as the adaptive feedback encoding point.

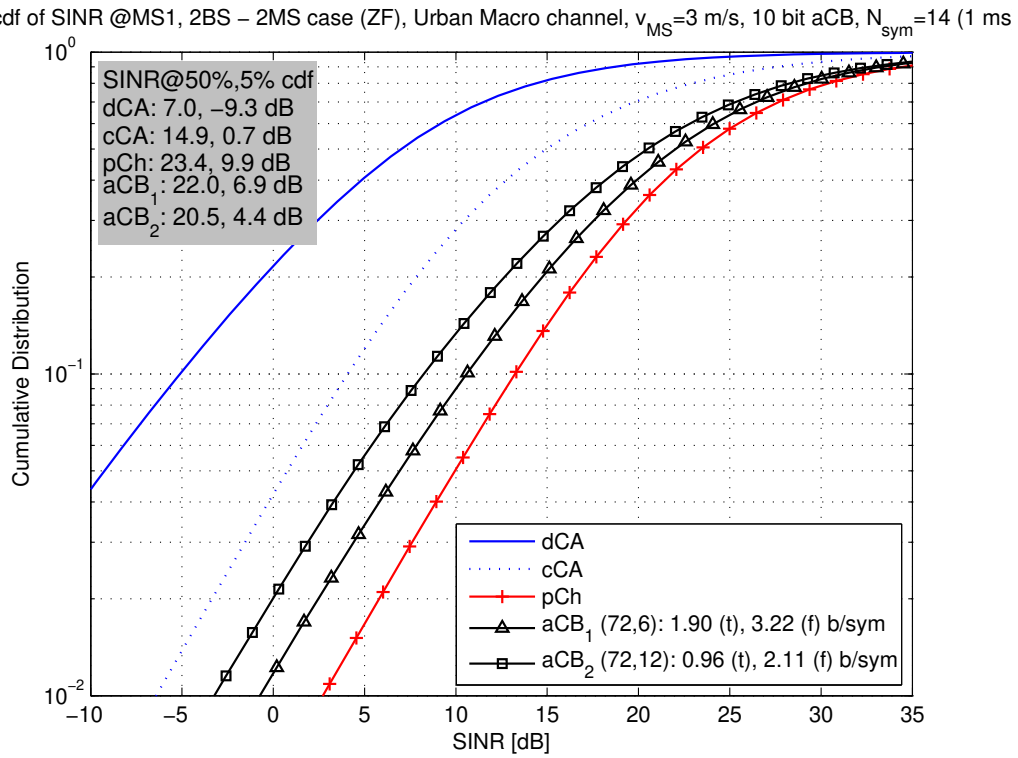


Figure 5.4: Adaptive feedback encoding in two directions - time (t) and frequency (f) (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, aCB (N, K): adaptive codebook where N is the total number of subcarriers and K is the cluster size)

however, is prohibitively high. It should be noticed that the aCB_1 (72,6) case in Fig. 5.4 performs with 1.1 dB degradation in comparison with the reference case, but it requires only 13.6 % of the feedback overhead of the reference case.

In Fig. 5.6 we compare the adaptive feedback performances under various feedback periods: aCB_1 $N_{sym} = 14$ and aCB_2 $N_{sym} = 5$ for the total number of subcarriers $N = 72$ and the cluster size $K = 6$. The mobile speed is set as 10 m/s, more than three times of the previous simulations. The aCB_1 $N_{sym} = 14$ case shows a 2.0 dB performance degradation compared with the mobile speed 3 m/s case, even with higher feedback overload. We observe that the reduced channel temporal correlation caused by a higher mobile speed leads to the increase of the feedback overhead encoded over time (from 1.90 bits/sym to 4.13 bits/sym). On the other hand, the feedback overhead encoded over frequency remains at a similar level (from 3.22 bits/sym to 3.40 bits/sym), which makes it more effective than the feedback encoding over time. The more frequent feedback ($N_{sym} = 5$) leads to a better performance in return for the increased feedback overhead.

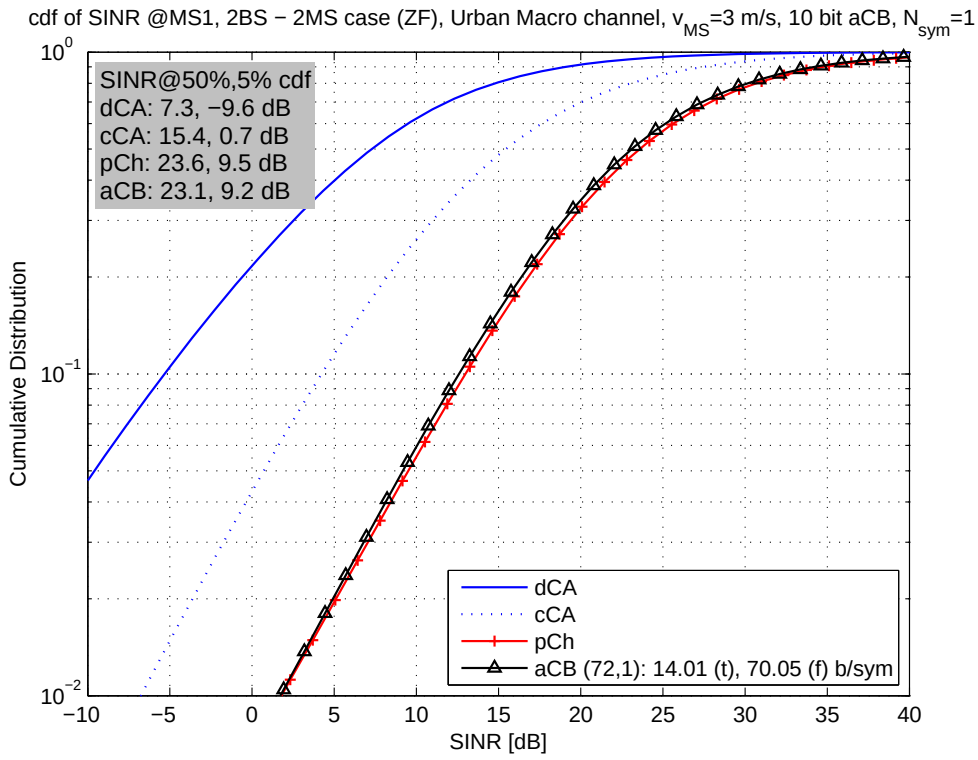


Figure 5.5: Reference performance of the adaptive feedback encoding scheme (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, aCB (N,K): adaptive codebook with N the total number of subcarriers & K the cluster size)

In Fig. 5.7 we compare the adaptive feedback performances between a codeword interpolation based method and a clustering method. An urban macro channel with a mobile speed of 3 m/s is used, and feedback messages are sent every 14 symbol, as in Fig. 5.4. For the interpolation method, we assume that total $N = 72$ subcarriers are grouped by equal size interpolation areas of size $K' = 12$ subcarriers, which is denoted by aCB_2 $[72, 12]$. We compare this interpolation based method with the clustering method, where a total of $N = 72$ subcarriers composed of 12 clusters of size $K = 6$ subcarriers are used, which is denoted by aCB_1 $(72, 6)$. These two cases are selected to make a fair comparison in terms of the required feedback overhead.

We have observed that the interpolation method (aCB_2 $[72, 12]$) outperforms the clustering method by 0.5 dB and 1.5 dB at 50 % and 5 % outage SINR, respectively. Both methods require approximately the same amount of feedback, as they use the exactly same feedback encoding scheme. As an alternative method, the MS can calculate θ_l and feedback this information by using an

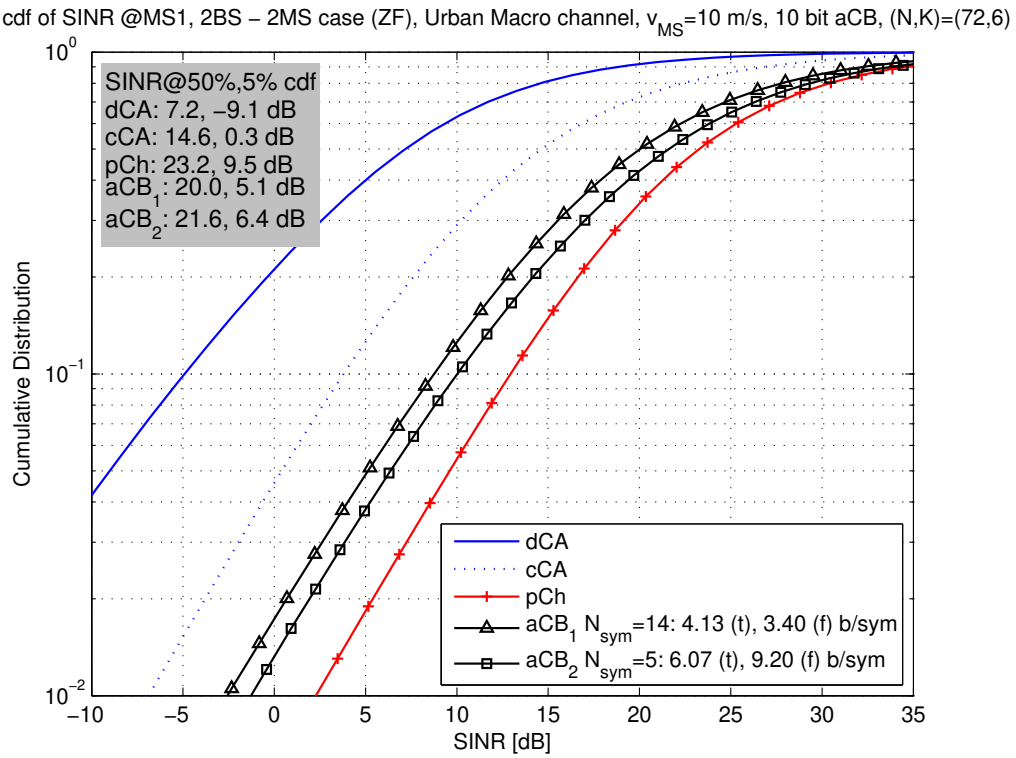


Figure 5.6: Adaptive feedback with various feedback periods (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, aCB $N_{sym} = n$: adaptive codebook with feedback period of n OFDM symbols)

efficient encoding scheme[16]. The required feedback overhead of the aCB_2 $[72, 12]$ case, which adopts the adaptive codebook, is (1.9 [bits/symbol] $*$ 14 [symbols/TTI])/12 [nodes/TTI] = 2.2 [bits/node], meaning that we need about 2.2 bits to encode one node on the average. If we come up with a method which requires less than 2.2 bits to encode θ_l, we can reduce the feedback overhead of the interpolation based method, compared with the clustering method[17]. As θ_l is a scalar, the adaptive delta modulation (ADM) can be used in case the value of θ_l has correlation over time or frequency (*Aldajani and Sayed*, 2001). By adopting ADM for tracing θ_l we can allocate 1 bit to encode θ_l as long as a high correlation is guaranteed between sequential samples.

[16]The MS, in comparison with the BS, is at a better position to find θ_l, in the sense that it can use the estimated channel instead of the quantized channel in (5.6) when finding θ_l.

[17]The required number of bits per node for aCB encoding can become slightly higher than 2.2 [bits/node] when θ_l is encoded separately. In this case the distance between the adjacent nodes in the frequency domain is doubled, and the channel correlation between the nodes becomes lower accordingly.

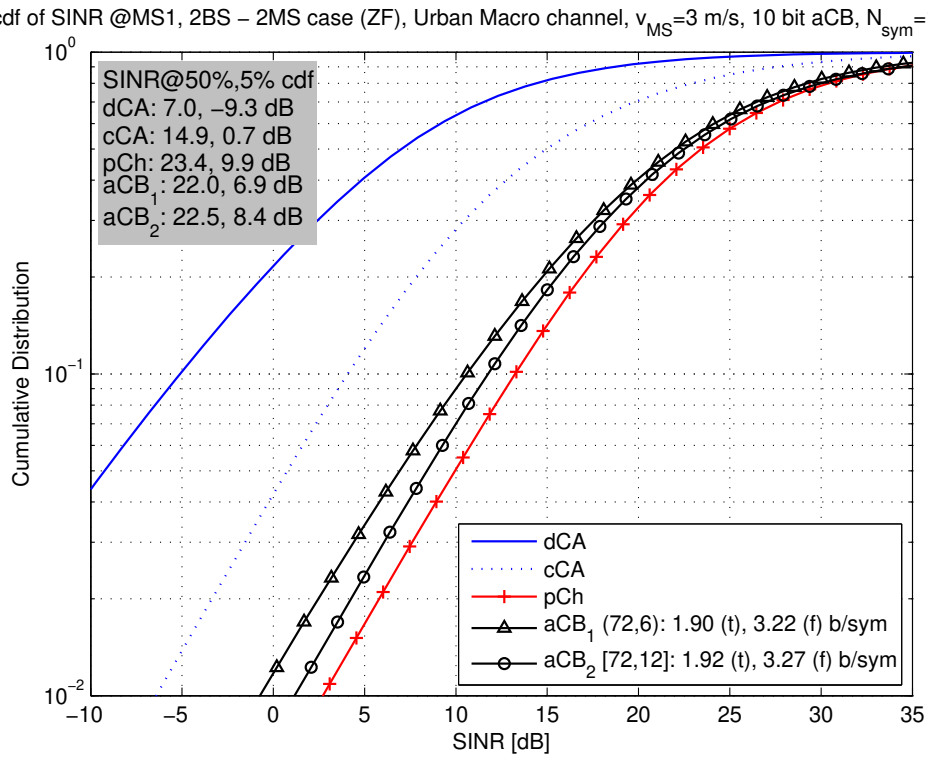


Figure 5.7: Performance comparison of interpolation method and clustering method (dCA: distributed CA, cCA: centralized CA, pCh: perfect channel, aCB$_1$ (N, K): clustering with N the total number of subcarriers & K the cluster size, aCB$_2$ $[N, K']$: interpolation with N the total number of subcarriers & K' the interpolation area size)

5.4 Conclusion

In this chapter, adaptive codebook encoding methods are extended to two dimensions: time and frequency. In reality, the channel information cannot be acquired for each subcarrier and for each OFDM symbol as the reference signals are interspersed among resource elements. The clustering method and the codeword interpolation method are evaluated by taking into account the LTE parameters taken into account. Simulation results show that the codeword interpolation method outperforms the clustering method while using the same amount of feedback. The codeword interpolation method achieves a performance of being only 0.9 dB away from the perfect channel case with 1.92 bits/sym feedback overhead. It would be an interesting future research topic to find a way to generalize the codeword interpolation method for mobile stations with multiple antennas.

CHAPTER VI

Comparison of the Proposed Methods

In this chapter, we compare two limited feedback schemes for MU-MIMO systems, especially targeting the distributed cooperative antenna systems. The first is introduced as the existing state of the art technique, and the second is proposed as the author's contribution in the dissertation. The feedback schemes to be compared are as follows.

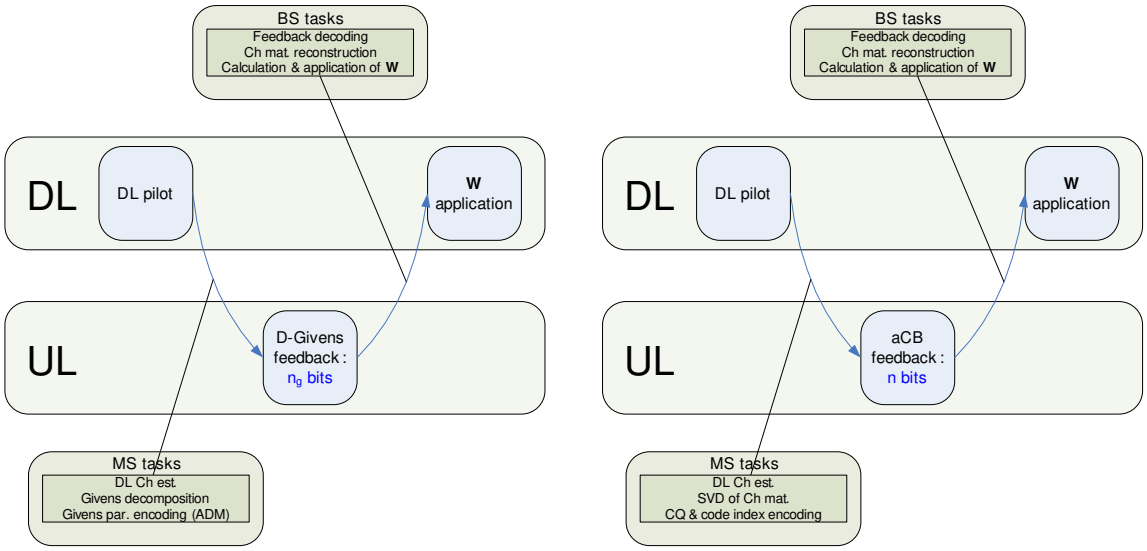
- Givens Rotation based Channel Decomposition Method with Adaptive Delta Modulation (ADM)

  This method is described in Chapter III. The method is composed of the Givens decomposition of the unitary matrix followed by the adaptive delta modulation which is introduced in (*Roh and Rao*, 2004) as a way to reduce the channel quantization feedback data rate in the presence of channel correlation. The method is called a differential Givens method, which is denoted in this chapter as 'D-Givens'.

- Subspace based Channel Quantization Method with an Adaptive Codebook

  This method is described in Chapters II, IV and V. According to the method, downlink channel information is captured and reconstructed by looking up a pre-defined codebook. Regarding the feedback scheme, the channel adaptive codebook organization and corresponding feedback encoding scheme is proposed as a way of compressing the feedback data for temporally correlated channels. This method can be extended over the frequency domain as long as there exists channel correlation in frequency, and it is evaluated in Chapter V. It is denoted in this chapter as 'aCB'.

This chapter is organized as follows. The required uplink resources (feedback overhead) together with a comparison of the main features of the feedback schemes can be found in Sections 6.1 and



(a) Givens rotation based channel quantization method with ADM

(b) Subspace based channel quantization method with aCB

Figure 6.1: Transmission sequence diagram of the proposed feedback schemes (both methods operate in FDD mode)

6.2. Section 6.3 presents a performance comparison of the proposed feedback methods, and Section 6.4 concludes the chapter by providing guidelines on the selection of the feedback scheme.

6.1 Givens Rotation based Channel Decomposition Method with ADM

The D-Givens method requires n_g feedback bits, where n_g is the number of Givens parameters to track by the ADM. The feedback resource required for the initial Givens parameter quantization is not considered, as its impact to the required feedback resource calculation is not significant in case that a feedback scheme is in operation for a sufficiently long time without disruptions.

Fig. 6.1 (a) shows how the D-Givens method works. In the first phase of the feedback operation, the associated BSs send a DL subframe, in which DL pilots are interspersed, to the MSs. The MSs estimate the DL channel matrix by exploiting DL pilots, perform a Givens decomposition of the channel matrix to extract n_g Givens parameters and encode the Givens parameters by using ADM. In the second phase of the feedback operation, n_g bit D-Givens feedback information is delivered to the associated BSs. At reception, the associated BSs decode the received feedback message, reconstruct the DL channel matrix and calculate the precoding matrix $\mathbf{W}$.

6.2 Subspace Based Channel Quantization Method with Adaptive Codebook

The aCB method adopts an aperiodic feedback scheme with a flexible number of feedback bit. In Fig. 6.1 (b), the average number of bits used per feedback is assumed to be n.

Fig. 6.1 (b) shows how the aCB method works. In the first phase of the feedback operation, the associated BSs send a DL subframe, in which DL pilots are interspersed, to the MSs. The MSs estimate the DL channel matrix by exploiting the DL pilots, perform a SVD of the channel matrix to extract channel directional information, perform channel quantization and encode the acquired code index by using non-prefix-free (NPF) coding. In the second phase of the feedback operation, aCB feedback information is delivered to the associated BSs. At reception, the associated BSs decode the received feedback message, reconstruct the DL channel matrix and calculate the precoding matrix $\mathbf{W}$.

6.3 Comparison of the Performance of Proposed Feedback Methods

The performance of the proposed feedback schemes which have been evaluated by Monte Carlo simulation methods in previous chapters are depicted in Fig. 6.2. Simulations have been performed for the 2 BSs - 2 MSs case in which the two BSs cooperate to transmit data for two MSs using the same resources. Both BSs and MSs have a single antenna, so we have a 2×2 overall channel matrix. We employ the transmit zero-forcing filter as an example beamforming scheme. It is based on downlink channel information which is either perfect (pCh), or is provided by a downlink channel estimation method which is shared by the BSs through a prompt and error free backbone network (centralized CA, cCA), or is acquired by analog pilot re-transmission method (distributed CA, dCA). The limited feedback methods in comparison are a Givens rotation based channel decomposition method with ADM (D-Givens) and a subspace based channel quantization method with adaptive codebook (aCB). The D-Givens method has been tested with a parameter set (2/1)-(10,10), i.e., $n_{\phi_{1,2}} = 2$, $n_{\theta_{1,1}} = 1$ bit allocation for Givens parameters $\phi_{1,2}$, $\theta_{1,1}$ followed by ADM with parameter tracking interval $[\tau_{\phi_{1,2}}] = [\tau_{\theta_{1,1}}] = 10$ (unit: number of OFDM symbols) for both parameters. A D-Givens performance has been evaluated on subcarrier basis and channel temporal evolution has been taken into account[1]. The aCB method has been evaluated with a multi-carrier processing in place by employing an interpolation method with $[72, 12]$, i.e., total $N = 72$ subcarriers are grouped by

[1]The frequency correlation has not been exploited in (*Roh and Rao*, 2004).

equal size interpolation areas of size $K' = 12$ subcarriers. The aCB method has been evaluated for an OFDM system with 72 subcarriers, and the frequency correlation has been taken into account for feedback encoding. The pilot tone topology follows the frame structure of 3GPP LTE, as explained in Chapter V. The channel temporal evolution has been considered. The 3GPP SCMe is used for the simulations and the proposed methods are tested for an Urban Macro channel with a 3 m/s mobile speed and a 2.6 GHz carrier frequency. The system performance is evaluated in terms of the received SINR at the MS and the required uplink feedback resource in [bits/subcarrier/feedback] in which the feedback is sent back per every transmit time interval (TTI) (one subframe, 1 ms, 14 OFDM symbols). Simulations are performed for 30,000 channel realizations and the cdf of the effective SINR at one MS is obtained.

It is worth noting the following simulation assumptions. Regarding the downlink channel estimation, both D-Givens and aCB methods assume perfect channel estimation. Concerning the temporal evolution of the channel, both D-Givens and aCB methods have been tested with channel temporal evolution taken into account.

The simulation results shown in Fig. 6.2 lead to the following conclusion.

- The aCB method outperforms the D-Givens method in terms of the received SINR and the required feedback resources.

  The aCB method performs very well for the overall SINR region, both low and high, with about 1.0 dB degradation from the reference case (pCh). This performance has been achieved by spending only 0.37 bits/subcarrier/feedback uplink overhead which is 13.2 % of the required resource in case of the D-Givens method. It can be thought to be not a fair comparison since the aCB method exploits state of the art techniques, i.e., the feedback reduction method over time and frequency and NPF encoding, whereas the D-Givens method in comparison does not resort to the schemes exploiting the frequency correlation of the channel. However, even for the hypothetical case that the D-Givens method could be extended to the multi-carrier case with a clustering size of 6 subcarriers without loss of performance[2], the aCB method would still outperform the D-Givens method in terms of both the received SINR and the required feedback resources. The Key Performance Indicators (KPIs) are summarized in Table 6.1.

- The D-Givens method suffers from a higher deviation from the reference case in a lower SINR region than in a higher SINR region.

[2]In this hypothetical case, the required feedback overhead would be reduced by 6 folds, i.e., $2.8/6 \simeq 0.47$ bits/subcarrier/feedback.

Table 6.1: KPIs: performance comparison of the feedback schemes

FB scheme	50 %/5 % outage SINR [dB]	FB overhead [bits/subcarrier/feedback]
aCB	22.5/8.4	0.37
D-Givens	20.8/3.6	2.80

The lower received SINR the greater the deviation from the perfect channel case in case of the D-Givens method. The aCB method maintains almost constant gap with respect to the perfect channel case, showing 1.1 dB and 0.8 dB degradation at 5 % and 50 % outage SINR, respectively. In case of the D-Givens method, the degradation is 5.9 dB and 2.5 dB, respectively.

6.4 Guidelines on the Selection of the Feedback Scheme

In this section, we intend to provide guidelines on the selection of the feedback scheme. Since the aCB method outperforms the D-Givens method in key performance indicators, i.e., the received SINR and the required uplink feedback overhead, the aCB method seems to be the most favored candidate for the feedback scheme. However, the high performance of the aCB method does not come without cost. The aCB method employs an aperiodic feedback scheme and variable size feedback bits which requires a major coordination effort when intended to be used in existing broadband mobile communication standards like 3GPP LTE. On the other hand, the D-Givens method can be relatively easily adopted in the current standards and it still provides a significant gain in comparison with the dCA case. The comparison of the proposed feedback schemes is summarized in Table 6.2.

Table 6.2: Comparison of the feedback schemes

FB scheme	Pros	Cons
D-Givens	Easy adaptation to standard Lower complexity at MS	Lower SINR performance per a FB overhead
aCB	Higher SINR performance per a FB overhead for overall SINR region	Higher complexity at MS Coordination w/standard required

We should also pay attention to the complexity aspect. Both D-Givens and aCB methods are involved with a computation intensive operation, i.e., a SVD of the channel matrix. Computational complexity is one of the critical issues especially for the MS, as MS is limited in its computation capability and power. A SVD calculating module is required for both D-Givens and aCB methods to acquire the right singular vector, and it is followed by an ADM module and a chordal distance

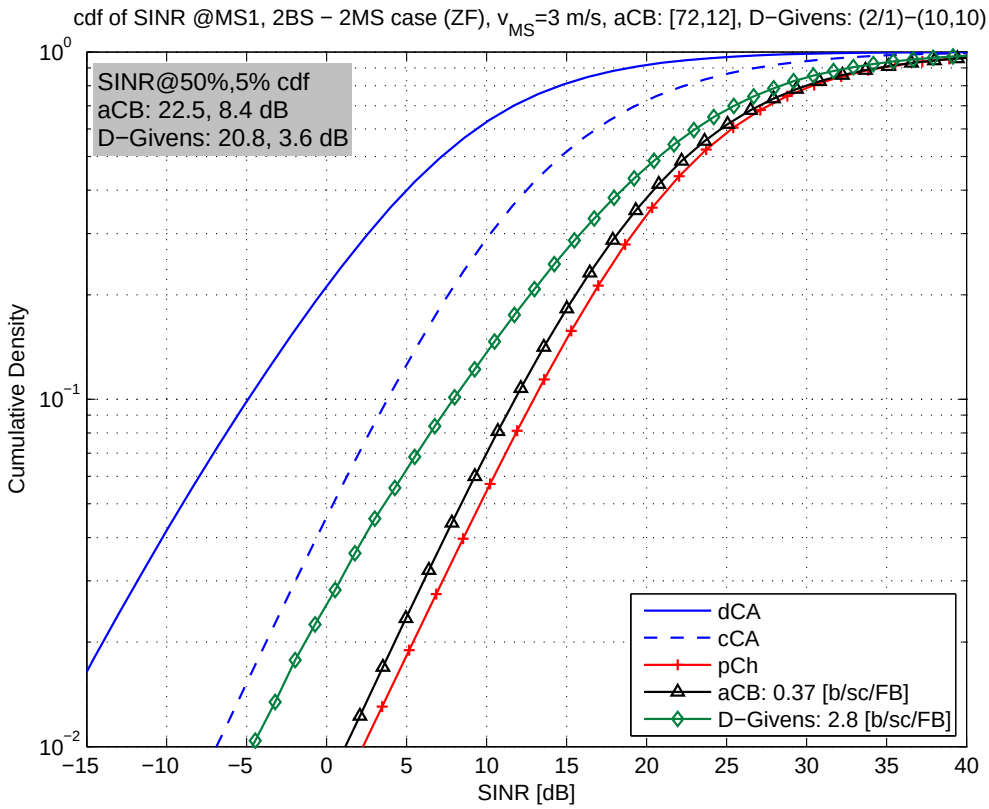


Figure 6.2: Performance comparison of the proposed feedback schemes for 2 BSs - 2 MSs case (dCA: distributed CA; analog pilot retransmission method, cCA: centralized CA; realistic channel estimation at the BSs which are connected with each other via a delay-free backbone network [TDD], pCh: perfect channel, aCB: adaptive codebook with [72,12], i.e., interpolation with $N = 72$ the total number of subcarriers & $K' = 12$ the interpolation area size and the feedback period of 14 symbols, D-Givens: differential Givens with (2/1)-(10,10), i.e., $n_{\phi_{1,2}} = 2$, $n_{\theta_{1,1}} = 1$ bit allocation for Givens parameters $\phi_{1,2}$, $\theta_{1,1}$ followed by ADM with parameter tracking interval $[\tau_{\phi_{1,2}}] = [\tau_{\theta_{1,1}}] = 10$ symbols for both parameters)

based codeword selection module in case of the D-Givens method and the aCB method, respectively. Since the codeword selection procedure (2.11) is in general computationally more demanding than the ADM procedure (3.9, 3.10, 3.11) and should be executed for every candidate codeword, the complexity of the aCB method is higher than that of the D-Givens method[3]. Clearly, there exists a trade-off between the system performance and the complexity at the MS.

[3]Detailed complexity analysis can be found in Appendix D.

6.5 Conclusion

In this chapter, two major limited feedback schemes for MU-MIMO systems, the differential Givens method (*Roh and Rao*, 2004) and the proposed adaptive codebook (aCB) method, have been summarized and compared in terms of the required feedback overhead, the computational complexity, and the numerical performance results. Based on the comparison analysis, guidelines on the selection of the feedback scheme have also been provided. The aCB method provides a better performance with a reduced feedback overhead at the cost of a higher computational complexity at the MS. It might be interesting to investigate both methods for more various antenna combinations at the BS(s) and at the MS(s) as well.

CHAPTER VII

Multiple Antenna Extension of the Subspace Based Channel Quantization Method

This chapter extends the previous work by introducing multiple antennas at the transmitter side (BS) and/or at the receiver side (MS). In this thesis the overall investigation has been focusing on the ZF precoding (TxZF) method which is designed to support single data stream transmission to each user in which each user is equipped with only one receive antenna. TxZF is beneficial for low noise or high power situations, but due to the stringent zero-interference requirements an ill conditioned channel matrix will require a large normalization factor, which can lead to the dramatic reduction of SNR at the receivers. To overcome this drawback and limitation of TxZF, several other linear transmit processing techniques for MU-MIMO systems have been proposed such as block diagonalization (BD) (*Spencer et al.*, 2004b), and regularized block diagonalization (RBD) (*Stankovic and Haardt*, 2008), which are briefly explained below.

A BD method is developed for the case in which there are multiple antennas at the receiver (*Spencer et al.*, 2004b). It is a generalization of channel inversion which optimizes the power transfer to a group of antennas of the same user rather than an individual single antenna in an effort to eliminate multi-user interference (MUI). Compared to TxZF, BD approaches the optimal solution at high SNRs while allowing multiple receive antennas at each user, but the zero MUI constraint can lead to a significant capacity loss when the users' subspaces largely overlap. A RBD method is designed in (*Stankovic and Haardt*, 2008) such that each user transmits on the eigenmodes of the combined channel of all other users but with the power which is inversely proportional to the regularized singular values of this matrix. The system performance at low SNRs is limited by noise whereas it is limited by MUI at high SNRs. Accordingly, at low SNRs the user transmits in the nullspace and a part of the signal subspace of all other users, balancing the MUI in order to better

use the available subspace. At high SNRs the user transmits only in the nullspace of all other users to minimize the MUI. As a result, RBD has the same capacity as successive MMSE (SMMSE) (*Stankovic and Haardt*, 2004) at low SNRs and as BD at high SNRs. Compared to BD, RBD improves the performance at low SNRs. For high data rates, RBD extracts the full diversity in the system (*Stankovic and Haardt*, 2008).

In an effort to bring the above mentioned linear precoding schemes further into practical use, the authors in (*Ravindran and Jindal*, 2007) and (*Song et al.*, 2008) have proposed limited feedback schemes for BD and RBD, respectively. In (*Ravindran and Jindal*, 2007), a limited feedback system employing the BD precoding at the BS has been considered in which each receiver knows its channel perfectly, but the transmitter is only provided with a finite number of channel feedback bits from each receiver. The channel quantization method based on the chordal distance metric has been proposed which represents the spatial direction of each user's channel, but due to the difficulty of designing optimal quantization codebooks, they have provided an analysis of the maximum rate loss of the finite rate feedback with BD by using a random vector quantization argument. The authors in (*Ravindran and Jindal*, 2007) have shown that it is required to scale the number of feedback bits linearly with the system SNR to maintain a bounded rate loss, which turns out to be also valid for RBD in (*Song et al.*, 2008). In (*Song et al.*, 2008), a limited feedback system with RBD precoding has been investigated. For RBD precoding, the transmitter requires the channel magnitude information as well as the channel direction information. The latter refers to the knowledge of subspaces spanned by the user's channel matrices. The authors in (*Song et al.*, 2008) have proposed a new channel feedback scheme for systems with RBD precoding to convey both channel direction and magnitude information to the transmitter, which can achieve a bounded rate loss by linearly increasing the number of feedback bits with the system SNR. They have also proposed a codebook generation scheme which is based on a modified LBG vector quantization method and showed that it can significantly reduce the number of feedback bits compared to random vector quantization.

As we want to extend our investigation to the case of multiple data steams per user with multiple receive antennas, we have conducted a performance evaluation for one of the state of the art linear precoding techniques which has been proposed to deal with this case, i.e., the Regularized Block Diagonalization (RBD) method (*Stankovic and Haardt*, 2008), as well as the ZF precoding method. Both precoding schemes, i.e., ZF and RBD, have been tested in terms of simulations for various system conditions, e.g., with or without multi-user interference (MUI) where multiple data streams per user are allocated whenever they can be supported, under the distributed cooperative antenna

systems framework. The subspace based channel quantization (CQ) method which is proposed in Chapter II is used for the ZF precoding performance evaluation while the channel quantization scheme proposed in (*Song et al.*, 2008), which further develops the CQ proposed in this thesis and a corresponding codebook generation method, is adopted for the RBD precoding scheme.

Simulation results show that the RBD precoding scheme outperforms the ZF precoding scheme when there exists MUI at the cost of the increased computational complexity both at the BS and the MS. The RBD precoding scheme requires the receiver to compute the decoding matrix which is associated with the precoding matrix in use at the transmitter. This computation involves a singular value decomposition (SVD) of the effective channel matrix, of which the computational complexity is high. In this regard a brief investigation has been performed on how to deliver the decoding matrix information from the BS to the MS in an effort to reduce the computational burden at the MS. A simple yet cost-effective method is proposed for the BS to quantize the decoding matrix for a certain MS and to provide the MS with the selected code index over the air, which would incur the decoding matrix quantization error while it shall greatly simplify the receiver processing. Simulation results show that this method results in a slight degradation, as expected, but the benefit seems to outweigh the cost.

This chapter is organized as follows. In Section 7.1, some background information about the RBD precoding method is provided with its associated channel quantization scheme. In Section 7.2, the proposed scheme of delivering the decoding matrix information is described for the RBD precoding method with the presence of multiple data stream transmission to each user. Section 7.3 presents simulation results and the conclusions can be found in Section 7.4.

7.1 RBD Precoding Method and the associated Channel Quantization Scheme

The ZF precoding aims at nulling the MUI at the cost of losing some signal gain, and it mainly targets systems transmitting a single data stream to each user with one receive antenna (*Vojčić and Jang*, 1998). The block diagonalization (BD) method has been proposed in an effort to generalize the channel inversion scheme when there are multiple antennas at each receiver (*Spencer et al.*, 2004b). However, the BD scheme as well as the ZF scheme assumes zero MUI which imposes a dimensionality constraint regarding to the total number of antennas at the BS and the users. This zero MUI constraint can lead to a significant capacity loss when the users' subspaces largely overlap (*Spencer et al.*, 2004b).

Regularized block diagonalization (RBD) (*Stankovic and Haardt*, 2008) was proposed to improve the performance in the low SNR region, while using as much as possible of the available spatial degrees of freedom and minimizing the interference between different users at the same time. The RBD precoding technique has gained substantial attention (*Stankovic and Haardt*, 2004, 2008; *Song et al.*, 2008), since it promises a significantly improved sum rate and diversity order compared to all previously proposed linear precoding techniques, i.e., ZF, MMSE, and BD. In order to correctly perform RBD, the transmitter requires not only the channel direction information, but also the channel magnitude information which is used to avoid the noise enhancement and improve the diversity (*Song et al.*, 2008). In (*Song et al.*, 2008), the authors proposed a novel channel quantization scheme to provide both channel direction and magnitude information, which is described in more detail below.

7.1.1 System Model

We consider a multi-user MIMO system with cooperating multiple N_{BS} BSs and N_{MS} MSs, where each BS has N_t transmit antennas and each MS has N_r receive antennas[1]. The associated BSs separate the data streams of multiple users by using RBD precoding. The received signal of the ith user is expressed as

$$\mathbf{y}_i = \mathbf{G}_i \left(\mathbf{H}_i \sum_{k=1}^{N_{\mathrm{MS}}} \mathbf{F}_k \mathbf{s}_k + \mathbf{n}_i \right) \tag{7.1}$$

where the vector $\mathbf{s}_k \in \mathbb{C}^{n_s \times 1}$ contains the data symbols and n_s ($\leq N_r$) represents the number of data streams per user. The matrix $\mathbf{F}_k \in \mathbb{C}^{N_{tt} \times n_s}$ denotes the RBD precoding matrix at the transmitter (BS), and the matrix $\mathbf{H}_i \in \mathbb{C}^{N_r \times N_{tt}}$ is the channel matrix from the associated BSs to the ith user where N_{tt} is the total number of transmit antennas in the CA, $N_{tt} = N_{\mathrm{BS}} N_t$. The vector $\mathbf{n}_i \in \mathbb{C}^{N_r \times 1}$ represents the complex Gaussian noise vector with the variance σ_n^2, which is independent of $\mathbf{s}_k$. The matrix $\mathbf{G}_i \in \mathbb{C}^{n_s \times N_r}$ denotes the decoding matrix at the receiver (MS) and $\mathbf{y}_i \in \mathbb{C}^{n_s \times 1}$ is the receive signal vector of user i.

It is assumed that the total transmit power of all users is constrained by P_{tx}, i.e., $\mathrm{E}\left\{ \left| \sum_{k=1}^{N_{\mathrm{MS}}} \mathbf{F}_k \mathbf{s}_k \right|^2 \right\} \leq P_{tx}$ ($P_{tx} \geq 0$). Each of the receivers is assumed to have perfect and instantaneous knowledge of its own channel matrix.

[1]The RBD scheme supports a flexible number of receive antennas per MS, but here it has been assumed that each MS is equipped with the same number of antennas, to make it consistent with the system requirements adopted in this thesis.

7.1.2 Regularized Block Diagonalization (RBD)

The RBD precoding matrix can be found as (*Stankovic and Haardt*, 2008)

$$\mathbf{F}_i = \beta \mathbf{F}_{a_i} \mathbf{F}_{b_i} \in \mathbb{C}^{N_{tt} \times n_s} \tag{7.2}$$

where

$$\mathbf{F}_{a_i} = \tilde{\mathbf{V}}_i \left(\tilde{\mathbf{\Sigma}}_i^2 + \alpha \mathbf{I}_{N_{tt}} \right)^{-1/2} \in \mathbb{C}^{N_{tt} \times N_{tt}} \tag{7.3}$$

in which $\tilde{\mathbf{V}}_i$ and $\tilde{\mathbf{\Sigma}}_i$ are the matrix of the right singular vectors and the diagonal matrix of the singular values of the combined channel matrix of all other users $\tilde{\mathbf{H}}_i = \left[\mathbf{H}_1^{\mathcal{T}} \cdots \mathbf{H}_{i-1}^{\mathcal{T}} \mathbf{H}_{i+1}^{\mathcal{T}} \cdots \mathbf{H}_{N_{\mathrm{MS}}}^{\mathcal{T}} \right]^{\mathcal{T}} \in \mathbb{C}^{(N_{tr} - N_r) \times N_{tt}}$ (i.e., $\tilde{\mathbf{H}}_i = \tilde{\mathbf{U}}_i \tilde{\mathbf{\Sigma}}_i \tilde{\mathbf{V}}_i^{\mathcal{H}}$), respectively, and $\alpha = \left(\frac{P_{tx}}{N_{tr} \sigma_n^2} \right)^{-1}$. Here N_{tr} is the total number of the receive antennas in the CA, i.e., $N_{tr} = N_{\mathrm{MS}} N_r$. The second part can be calculated as

$$\mathbf{F}_{b_i} = \mathbf{V}_i \mathbf{D}_i \in \mathbb{C}^{N_{tt} \times n_s} \tag{7.4}$$

where the matrix $\mathbf{V}_i \in \mathbb{C}^{N_{tt} \times N_{tt}}$ is the right singular vector matrix of $\mathbf{H}_i \mathbf{F}_{a_i}$ (i.e., $\mathbf{H}_i \mathbf{F}_{a_i} = \mathbf{U}_i \mathbf{\Sigma}_i \mathbf{V}_i^{\mathcal{H}}$) and the matrix $\mathbf{D}_i \in \mathbb{C}^{N_{tt} \times n_s}$ is the power loading matrix ($\mathbf{D}_i = \mathbf{I}_{N_{tt} \times n_s}$ in case of the equal power loading). Then the parameter β is used to set the total transmit power to P_{tx}. The decoding matrix of user i, i.e., $\mathbf{G}_i$, is chosen as $\mathbf{U}_i^{\mathcal{H}} \in \mathbb{C}^{n_s \times N_r}$, which is the left singular vector matrix of $\mathbf{H}_i \mathbf{F}_{a_i}$. Note here that the decoding matrix calculation at user i requires knowledge of $\mathbf{H}_i \mathbf{F}_{a_i}$ which can be acquired by estimation of the effective channel via known pilot signals embedded in the downlink subframe.

Equation (7.3) indicates that each user transmits on the eigenmodes of the combined channel matrix of all other users with the power that is inversely proportional to the singular values of the combined channel matrix of these users. At high SNRs each user transmits only in the nullspace of all other users. In Equation (7.4), we can see that the system performance is further optimized by the specific optimization criterion assuming a set of parallel single user MIMO channels.

7.1.3 Channel Quantization Scheme for RBD

The description in Section 7.1.2 shows that for the RBD precoding the transmitter additionally requires the channel magnitude information which defines the strength of the eigenmodes of the users' channel matrices, in contrast to the ZF (*Jindal*, 2006) or the BD precoding (*Ravindran and Jindal*, 2007) in which the transmitter only requires the channel direction information which refers

to the knowledge of subspaces spanned by the users' channel matrices. In (*Song et al.*, 2008), the authors proposed a channel quantization scheme which addresses this requirement by quantizing the stacked vector of the channel matrix to keep the relative magnitude information for the columns of the channel matrix and to avoid the loss of the channel magnitude information which can be caused when quantizing the channel matrix as a unitary codeword matrix directly. The proposed channel quantization scheme is formulated as the following equation

$$\mathrm{vec}\{\hat{\mathbf{H}}_i\} = \arg\min_{\mathbf{w}_j \in \mathcal{C}} d_c^2(\mathrm{vec}\{\mathbf{H}_i\}, \mathbf{w}_j) \tag{7.5}$$

where the vector $\mathbf{w}_j \in \mathbb{C}^{N_r \cdot N_{tt} \times 1}$ is a codeword of the quantization codebook $\mathcal{C}$ used at the ith user, which is predefined by offline processing and known to the transmitter and user i. Here, $\mathrm{vec}\{\mathbf{H}_i\}$ denotes the stacked vector of the channel matrix $\mathbf{H}_i$ and $d_c(\cdot,\cdot)$ is the chordal distance metric (refer to Equation (2.8)). The matrix $\hat{\mathbf{H}}_i \in \mathbb{C}^{N_r \times N_{tt}}$ is the quantized version of the channel matrix $\mathbf{H}_i$, which can be acquired by reshaping $\mathrm{vec}\{\hat{\mathbf{H}}_i\}$ to a matrix form.

The channel quantization codebook $\mathcal{C}$ consists of $2^{n_{\mathrm{CB}}}$ unit norm vectors ($\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_{2^{n_{\mathrm{CB}}}}\}$), in which n_{CB} is the size of feedback per user in bits. The modified Linde-Buzo-Gray (mLBG) based codebook construction method, which is basically the same as the method introduced in Section 2.4 except for using a vectorized form instead of a matrix form, has proven to outperform the random vector quantization based method in terms of the required feedback overhead which leads to the same throughput in (*Song et al.*, 2008).

Please note that the methods in (*Song et al.*, 2008) take a notion of channel quantization rather than the beamforming vector selection as the methods proposed in this thesis. The only difference is to use the stacked vector of the channel matrix instead of taking the channel matrix as it is.

7.2 On Delivery of the Decoding Matrix Information for RBD Precoding Method

The RBD precoding scheme requires the MS to calculate a decoding matrix $\mathbf{G}_i \in \mathbb{C}^{n_s \times N_r}$ which is associated with the precoding matrix in use $\mathbf{F}_i \in \mathbb{C}^{N_{tt} \times n_s}$ in case that multiple data streams are transmitted to each user. It means that the MS should perform a SVD operation[2] on $\mathbf{H}_i\mathbf{F}_{a_i} \in$

[2]The MS might be able to use simpler decoding schemes instead, e.g., MMSE, for the sake of lower computational complexity at the MS (*Chong et al.*, 2014). However, their performances are expected to be lower than a SVD-based method. Simulation results in Section 7.3 verify this.

$\mathbb{C}^{N_r \times N_{tt}}$ ($N_{tt} \geq N_r$) to get $\mathbf{U}_i$, which requires $8(4N_{tt}^2 N_r + 13N_r^3)$ flops[3] (*Golub and Van Loan*, 2012). It can be quite challenging for the MS to execute a real time processing of this decoding matrix computation task via SVD which should be performed after the channel estimation, especially when a broadband transmission is considered as in Chapter V. Therefore, it would be beneficial if this computation effort at the MS can be saved.

One way to achieve this is for the BS to quantize the decoding matrix for a certain MS by using a predefined codebook and to provide the MS with its index. As the decoding matrix is an unitary matrix which contains the (precoded) channel directional information, the following quantization formula can be used

$$\text{vec}\{\hat{\mathbf{G}}_i\} = \arg \min_{\mathbf{g}_j \in \mathcal{C}^g} d_c^2(\text{vec}\{\mathbf{G}_i\}, \mathbf{g}_j) \tag{7.6}$$

where the vector $\mathbf{g}_j \in \mathbb{C}^{n_s \cdot N_r \times 1}$ is a codeword of the quantization codebook $\mathcal{C}^g$ used for the ith user. Using a stacked vector form facilitates the reuse of the predefined vector form codebook for the decoding matrix quantization as well as for the channel quantization, which is very important for saving memory and power consumption at the MS. It is assumed that using a stacked vector would not incur a significant performance degradation, but this requires further study. By adopting another quantization procedure of Equation (7.6) which shall take place this time at the BS, it is expected to have a performance degradation to some extent by resorting to this method. However, this network assisted decoding method would greatly simplify the receiver processing compared to the default method of the MS calculating the decoding matrix via SVD. Simulation results comparing both methods can be found in Section 7.3 which follows below.

7.3 Numerical Results

For performance simulations, a cooperative antenna system with 2 BSs and 1 or 2 MS(s) case is assumed. BSs and MS(s) can have a single or two antenna(s), and two BSs cooperate to serve one or two MS(s) at the same time. In case of serving a MS with two antennas, two uncoded data streams are transmitted to the MS with an equal power loading being used[4]. An Urban Macro channel of the 3GPP SCMe is used and the MS speed is 3 m/s. The carrier frequency is 2.6 GHz. Simulation

[3] A flop is a real floating point number add, subtract, multiply, or divide. In case of $N_{tt} = 4, N_r = 2$, the number of required flops for this operation is 1856. Alternatively, a MMSE decoding on $\mathbf{H}_i\mathbf{F}_i \in \mathbb{C}^{N_r \times n_s}$ can be used to reduce the complexity, which requires $8(n_s-1)^3 + 4(8N_r-1)(n_s-1)^2 - 2(2N_r-1)(n_s-1)$ flops. In case of $N_r = 2, n_s = 2$, the number of required flops is 62.

[4] This requires a full rank channel condition, and in the simulations the channel rank has been checked to satisfy this requirement. When we draw channels randomly, it should be satisfied with probability approaching 1.

results have been collected after 200 drops in which each drop simulates 30 subframes[5] , i.e., 420 consecutive OFDM symbols. In case of the simulations of the channel quantization schemes, an n bit adaptive codebook is employed and the feedback period is chosen to be 14 OFDM symbols (1 ms). The cdf of the SINR at MS1 for the first data stream is shown as a performance indicator, unless otherwise stated. A error-free and delay-free feedback link is assumed, wherever applicable.

First, simulations have been performed to compare the ZF precoding and the RBD precoding schemes for various channel configurations, e.g., with or without MUI. The SINR distribution assuming the perfect DL channel knowledge at the BSs is shown in Fig. 7.1. The results can be interpreted as:

- There is no performance difference between the ZF and the RBD precoding scheme when they are deployed without MUI (only one user is served) and without multiple data stream transmission to a MS.

  The ZF and RBD precoding schemes show exactly same performance for {1}×{1,1} and {1}×{2,2} configurations in which only one MS which is equipped with a single antenna is served, i.e., there is no MUI and no inter data stream interference. When comparing the {1}×{1,1} configuration and the {1}×{2,2} configuration, we can see a performance improvement thanks to the transmit diversity gain by doubling the number of total transmit antennas while keeping the total transmit power of all users same. A performance difference is observed when the MUI is imposed ({1,1}×{1,1}, {1,1}×{2,2}).

- The RBD scheme outperforms the ZF scheme when multiple users are served.

  The RBD precoding scheme performs better than the ZF precoding scheme for {1,1}×{1,1} and {1,1}×{2,2} configurations in which two MSs with a single antenna are served at the same time, i.e., with the presence of MUI but without inter data stream interference, which becomes more evident in the higher SINR region. The addition of one transmit antenna at each BS results in the transmit diversity gain. This observation holds true for per user performance comparision which can be found in Fig. 7.2. It shows that there is practically no performance gap between MS1 and MS2 irrespective of the precoding schemes, when only one data stream is transmitted to a certain user.

- The RBD scheme provides a balance between a MUI suppression and a MIMO processing gain at each user with multiple antennas/data streams.

[5] In the time domain, LTE transmissions are organized into (radio) frames of length 10 ms, each of which is divided into ten equally sized subframes of length 1 ms.

Whereas the ZF precoding scheme is designed for a multi-user DL transmission with MSs using a single antenna, the RBD precoding/decoding scheme shall be able to sort out the multiple data streams at the MSs with multiple antennas as well as to mitigate interference between users. Performance results of the TxRBD cases for the {2}×{2,2} and {2,2}×{2,2} configurations verify this to some degree, in which two data streams with equal power allocated to each data stream are transmitted to each user with two antennas while possibly serving two MSs at the same time, i.e., with the presence of the inter data stream interference and/or with MUI.

First, when there is no MUI while the MS is equipped with two receive antennas ({2}×{2,2}), its performance depicted in Fig. 7.1 appears to show not only no degradation compared to the single receive antenna case ({1}×{2,2}), but also a slight improvement. The cdf of the per-user/per-data stream SINR can be found in Fig. 7.3. Observing the SINR behavior of the second data stream reveals the fact that there is a significant received SINR gap between the first and the second data stream, which indicates that it would be recommended to apply an adaptive selection of the modulation and coding scheme (MCS) per data stream in accordance with the available SINR. This SINR gap could be narrowed by using a smarter power allocation scheme rather than an equal power loading. When one additional user is served, both the single receive antenna user case ($\{1,1\}$×{2,2}) and the dual receive antenna case ($\{2,2\}$×{2,2}) show a degradation compared to their single user counterparts, i.e., {1}×{2,2} and {2}×{2,2}, respectively. In this case, the dual receive antenna (dual data streams) case is more severely affected by the MUI, compared with the single user case, in the sense that not only a SINR degradation is observed for the first data stream, but also the received SINR gap between two data streams for a certain user is increased, compared with the single user case. This might be due to the fact that the single receive antenna case could benefit from the transmit diversity which can alleviate the inter-user interference, while the dual receive antenna case cannot fully utilize this transmit diversity due to the diversity/multiplexing tradeoff. It means that for the dual receive antenna case, not all of the space-time dimensions can be used for diversity gain only, since some of them should be used for multiplexing gain, i.e., to decompose the MIMO channel as two data streams intended for a certain user shall be separated[6]. It is worth to note that the performance enhancement of the RBD scheme does not come without the cost of an increased computational complexity at the MS.

[6]The single antenna user sees a 1×4 channel matrix, and the dual antenna user sees a 2×4 channel matrix in this case.

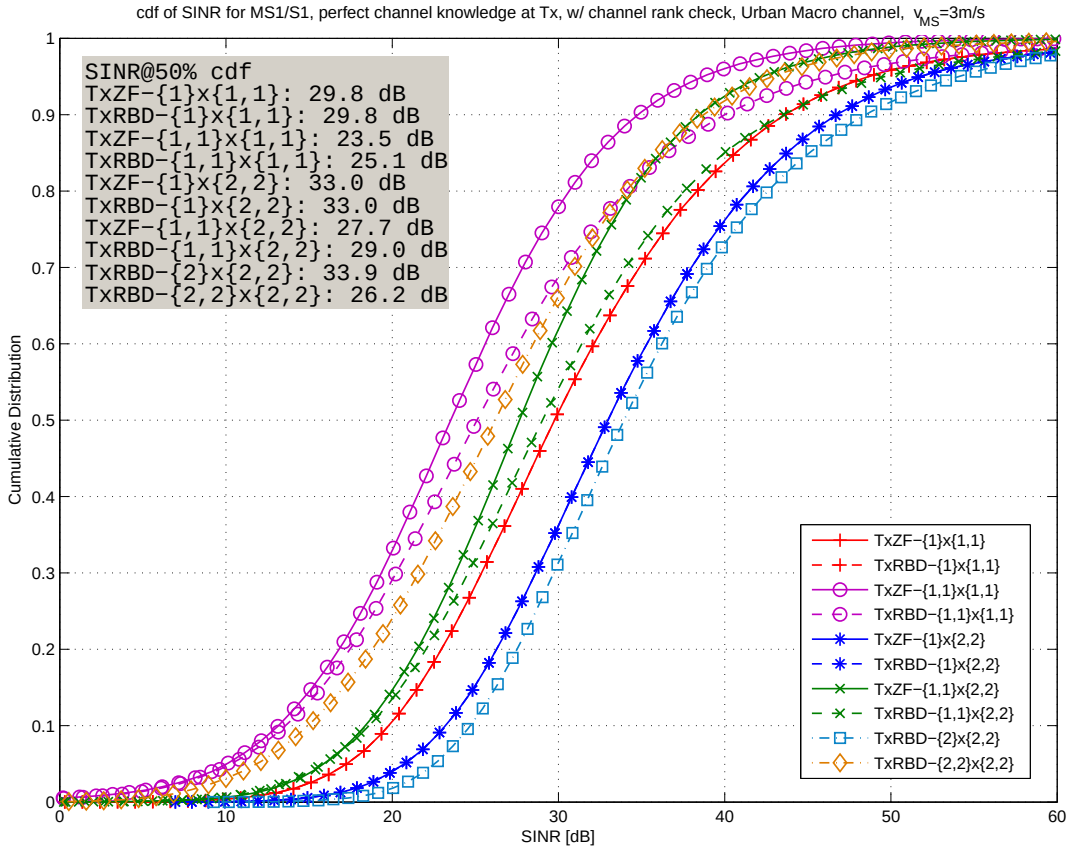


Figure 7.1: Performance comparison of the ZF and the RBD precoding methods with perfect DL channel information at BSs (TxZF-$\{N_r\}\times\{N_t,N_t\}/\{N_r,N_r\}\times\{N_t,N_t\}$: ZF precoding for 2 cooperating BSs with N_t transmit antennas per each BS transmitting to one MS with N_r receive antenna/two MSs with N_r receive antennas per each MS, TxRBD-$\{N_r\}\times\{N_t,N_t\}/\{N_r,N_r\}\times\{N_t,N_t\}$: RBD precoding)

The performance of the mLBG codebook based CQ for a single data stream transmission case is depicted in Fig. 7.4. First, as expected, we can see that the cases with the presence of MUI, i.e., the $\{1,1\}\times\{1,1\}$ and $\{1,1\}\times\{2,2\}$ configurations, require a higher resolution codebook compared with the non-MUI cases, i.e., the $\{1\}\times\{1,1\}$ and $\{1\}\times\{2,2\}$ configurations, respectively, to achieve the same level of relative performance with respect to the reference (perfect channel information at the transmitter) case. For example, the ZF or RBD precoding case for the $\{1\}\times\{1,1\}$ configuration can achieve a performance with a loss of 0.1 dB with respect to the reference at 50 % cdf of SINR already with a 4 bit codebook, while for the $\{1,1\}\times\{1,1\}$ configuration even with a 10 bit codebook we should suffer a loss of 0.7 dB and 1.5 dB with respect to the reference in case of the ZF and the RBD precoding, respectively.

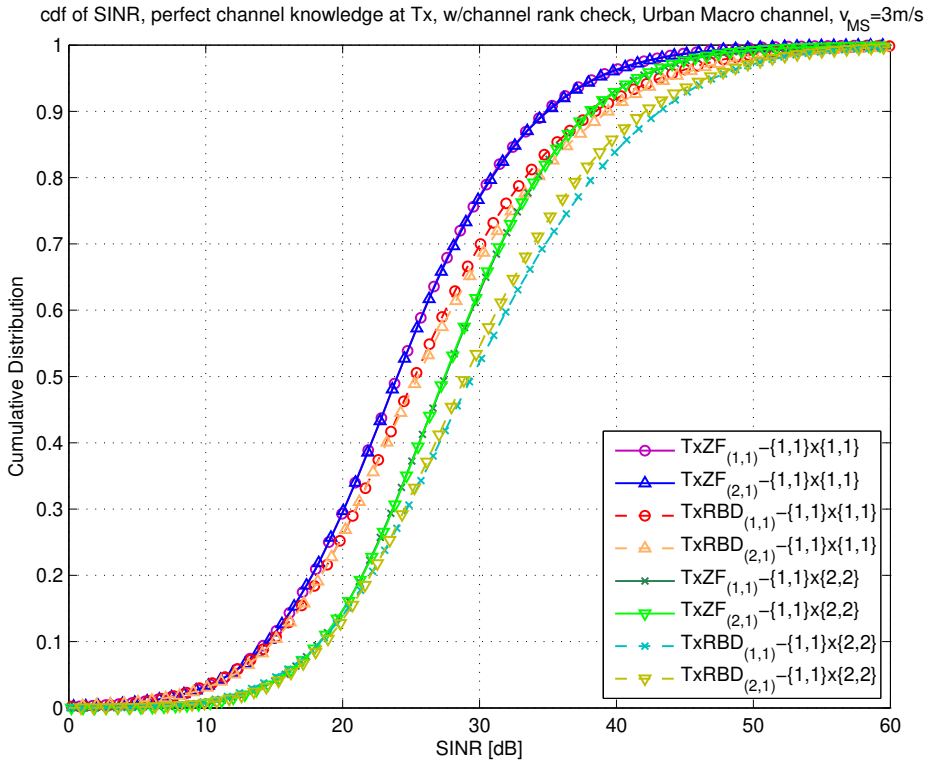


Figure 7.2: Per-user performance of the ZF and the RBD precoding methods when multiple users are served (TxZF$_{(i,1)}$-$\{1,1\}\times\{N_t,N_t\}$: MSi when using ZF precoding for 2 cooperating BSs with N_t transmit antennas per each BS transmitting to two MSs with 1 receive antenna per each MS, TxRBD$_{(i,1)}$-$\{1,1\}\times\{N_t,N_t\}$: RBD precoding)

Second, we can see a performance gap between the ZF and the RBD precoding schemes for the channel quantization case as well as for the perfect channel knowledge at the transmitter case. For example, in Fig. 7.4 (b), we can achieve a better performance with a 10 bit codebook by using the RBD precoding scheme than by using the ZF precoding scheme with a 12 bit codebook in the presence of MUI ($\{1,1\}\times\{2,2\}$ configuration).

Fig. 7.5 shows simulation results comparing two alternative strategies of acquiring the RBD decoding matrix information described in Section 7.2. The first alternative is for the MS to compute the decoding matrix via a SVD, which is denoted as nbCB in Fig. 7.5. The second alternative is for the BS to quantize the decoding matrix for a certain MS by using a prefined n bit codebook[7] and to provide it, which is denoted as nb/nbCB-DecMatQ in Fig. 7.5. In comparison with the former method, the latter method would entail a performance degradation stemming from the decoding

[7]Here the decoding codebook size does not have to be the same as the channel quantization codebook size, meaning that the higher resolution codebook can be used for the decoding matrix to reduce the possible loss which can be caused by the quantization procedure.

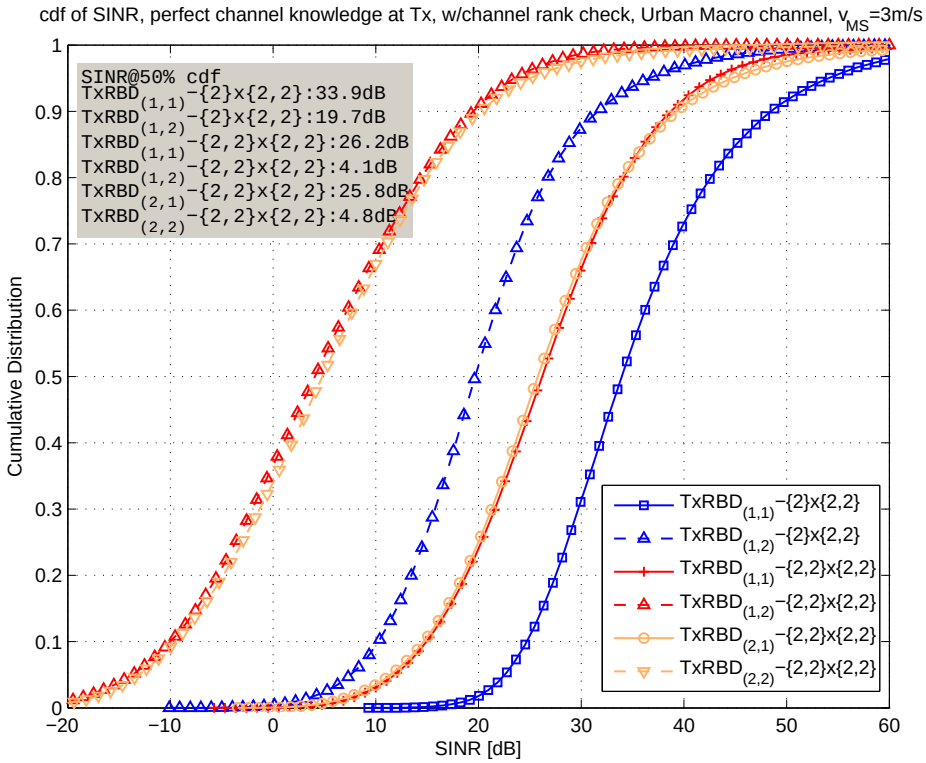


Figure 7.3: Per-user/-data stream performance of the RBD precoding methods for MS with multiple receive antennas (TxRBD$_{(1,j)}$-{2}×{2,2}: data stream j when using RBD precoding for 2 cooperating BSs with 2 transmit antennas per each BS transmitting 2 data streams to one MS with 2 receive antenna, TxRBD$_{(i,j)}$-{2,2}×{2,2}: data stream j at MSi when using RBD precoding for 2 cooperating BSs with 2 transmit antennas per each BS serving 2 MSs at a time by transmitting 2 data streams to each individual MS with 2 receive antennas)

matrix quantization procedure, which is less than 1.0 dB at 50 % cdf of SINR for the first data stream cases. Considering the high computational complexity of performing a SVD operation at the MS which is required by the first alternative[8] and the channel estimation error which is inevitable in practice, the second alternative can be a viable option. Therefore the second alternative provides a good trade-off between the performance and the required DL feedback overhead as well as the computational complexity at the MS. It is worth noting that a simpler decoding scheme, e.g., MMSE, might be used at the MS for the sake of the reduced complexity when the RBD precoding scheme is used at the BS. Its performance which is denoted as nbCB-MMSEdec in Fig. 7.5, however, shows a significant degradation for the first data stream compared with other alternative methods. The

[8]In case of the second alternative, the decoding matrix quantization operation is required instead, not at the MS but at the BS in which the system requirements are much more relaxed in various aspects, e.g., power consumption, die size, memory size, computational power, etc.

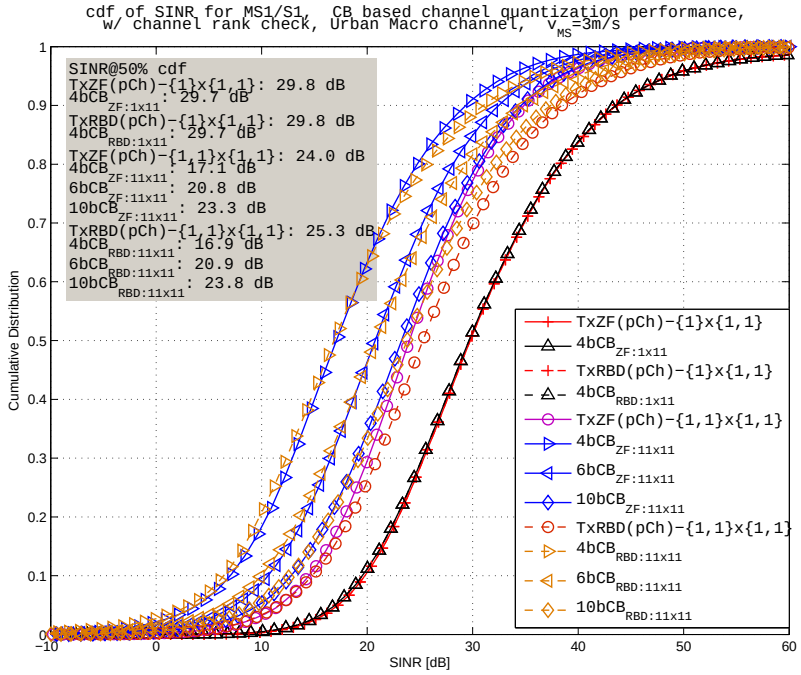
performance of the MMSE decoding scheme for the second data stream is on par with that of the RBD decoding scheme.

7.4 Conclusion

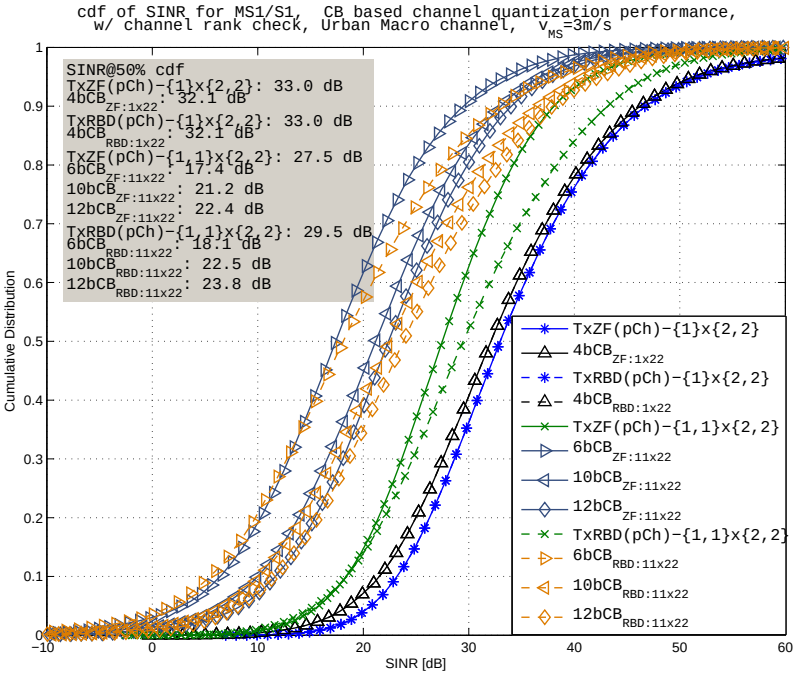
In this chapter, the previous work on the subspace based channel quantization method is extended by introducing multiple antennas at the transmitter side (BS) and/or at the receiver side (MS) and by investigating the performance of the RBD precoding(/decoding) scheme as well as the ZF precoding scheme. Both precoding schemes, i.e., ZF and RBD, have been tested in terms of simulations for various system configurations, e.g., with or without multi-user interference (MUI) where multiple data streams per user are allocated whenever they can be supported. The simulation results show that the RBD precoding scheme outperforms the ZF precoding scheme when there exists MUI, i.e., multiple users are served at the same time, and it can transmit multiple data streams to each user, at the cost of the increased computational complexity both at the BS and the MS. The RBD precoding scheme requires the MS to compute the decoding matrix via a SVD in case of the multiple data stream transmission to each user. A cost-efficient decoding matrix quantization method is proposed which can avoid a complex computation at the MS while showing only a slight degradation.

The beauty of the subspace based channel quantization method lies in the fact that it renders the MS precoding scheme-agnostic, meaning that the MS can rely on the same channel quantization procedure irrespective of the precoding scheme to be deployed. The BS(s) can determine the appropriate precoding method considering the channel configuration and other parameters like complexity and expected gain, etc., after collecting feedbacks from the users. If the RBD precoding scheme is used, it can be indicated to the MS on the control information, optionally together with the associated decoding matrix information to support an appropriate decoding operation.

As a future research topic, it might be worthwhile to extend the scope of the research by taking into account other decoding schemes as well, e.g., sphere decoder (*Hassibi and Vikalo*, 2005; *Vikalo and Hassibi*, 2005).

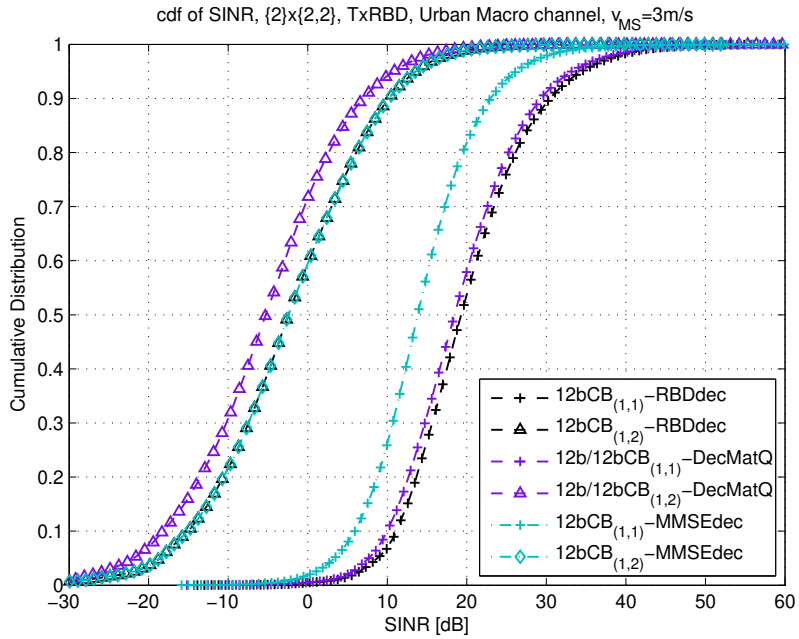


(a) Transmitters with single antenna ({1}×{1,1} and {1,1}×{1,1})

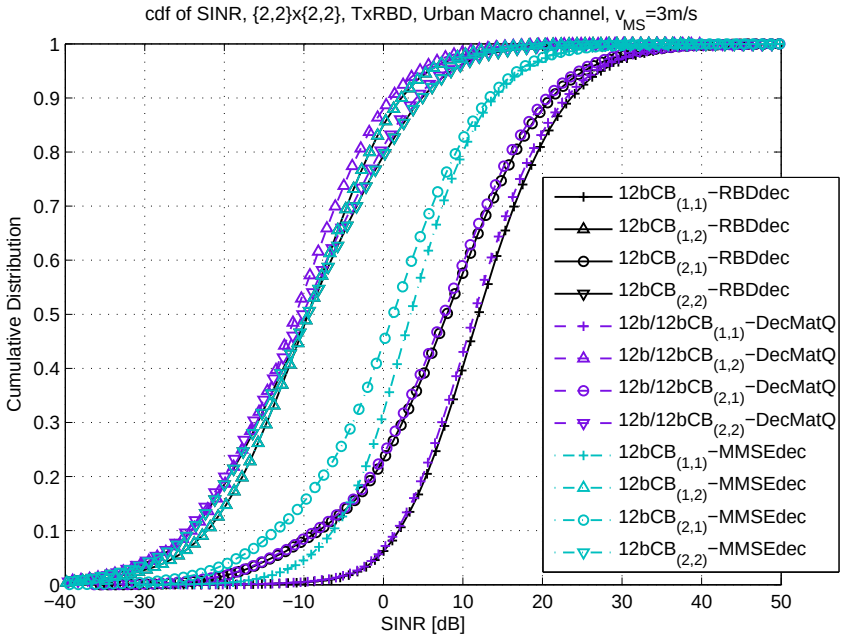


(b) Transmitters with two antennas ({1}×{2,2} and {1,1}×{2,2})

Figure 7.4: Performance comparison of the mLBG codebook for single data stream transmission case (TxZF/RBD(pCh): ZF/RBD precoding with perfect channel information at transmitter, nbCB$_{\mathrm{ZF/RBD}}$: n bit mLBG codebook for ZF/RBD precoding)



(a) Single user case ({2}×{2,2})



(b) Two user case ({2,2}×{2,2})

Figure 7.5: Performance comparison of the various decoding schemes for RBD precoding (nbCB$_{(i,j)}$-RBDdec: data stream j at MSi when using n bit FB for RBD precoding at BS while RBD decoding done at MS, nb/nbCB$_{(i,j)}$-DecMatQ: when using n bit FB for RBD precoding at BS while n bit CB used for RBD decoding matrix quantization at BS and delivered to MS, nbCB$_{(i,j)}$-MMSEdec: when using n bit FB for RBD precoding at BS while MMSE decoding done at MS)

CHAPTER VIII

High Granularity Codebook Operation Scheme for MU-MIMO Systems

As presented in Chapter II as well as other investigation results (*Motorola*, 2010c; *Kim et al.*, 2008c), MU-MIMO systems require channel feedback information with high accuracy[1]. The high accuracy of the feedback can be achieved by introducing a high granularity (resolution) codebook which will lead to a high feedback overhead as well as a high computational complexity at the MS. To address this issue, many proposals to reduce the feedback overhead as well as the complexity at the MS have been suggested in 3GPP standardization meetings (*Motorola*, 2010c) and in academia (*Kim et al.*, 2008d,e). Nevertheless, the current 3GPP LTE standard does not include usage of the high granularity codebook for MU-MIMO systems. However, it is envisioned that in future cellular systems, e.g., 5G, a high granularity codebook and its efficient deployment schemes, e.g., an adaptive codebook as described in Chapter IV, (*Kim et al.*, 2008e) or a differential codebook (*Samsung Electronics Co. Ltd.*, 2010a), should be indispensable for successful utilization of the MU-MIMO mode. In this case, the scheduling scheme at the network sides (BS) can be complicated, as different codebook sets can be used depending on the scenarios, e.g., low resolution codebook for MSs in the SU-MIMO mode while a high resolution codebook should be used for MSs in MU-MIMO mode. In this chapter we propose a unified and systematic scheduling mechanism which can be used for operations at the network sides. The proposed scheme provides a hierarchical user categorization method for SU-and MU-MIMO systems, of which the operation can be triggered by timer expiration events. By setting the corresponding timer values differently based on their needs, e.g., by setting the user grouping timer larger than the tracking mode entrance qualification timer which is larger than the exit qualification timer, the complexity of the scheduling scheme at the BS can be managed

[1] This holds for FDD systems. In case of TDD systems, downlink channel information can be estimated by looking into the uplink channel due to channel reciprocity.

at a reasonable level.

This chapter is organized as follows. The high level conceptual overview of the proposed method can be found in Section 8.1. Major benefits of the proposed scheme are presented in Section 8.2. The mechanism of the proposed scheme, i.e., the SU-/MU-MIMO user categorization and the corresponding timer value setting, is explained in detail in Section 8.3. Section 8.4 concludes the chapter.

This work has been filed as an European patent (*Kim*, 2016).

8.1 Overview of the Hierarchical SU-/MU-MIMO Scheduling Algorithm

The user grouping, i.e., SU- or MU-MIMO decision, and the tracking mode decision (to find MSs with low mobility, which are eligible for low overhead/complexity feedback scheme) for a given MU-MIMO group are performed in a hierarchical manner by introducing timers with different expiration time. Here, state of the art user grouping schemes can be used, e.g., a user selection algorithm to form a Space Division Multiple Access (SDMA) group (*Fuchs et al.*, 2007; *Cheng et al.*, 2014). First, a user grouping is performed at the expiration of the *user grouping timer* in order to select MSs which are eligible for the MU-MIMO processing. The channel directions of the MSs which belong to the same user group are as far spatially separated to each other as possible. In this case they are least likely to interfere with each other when MU-MIMO is performed (best companion to each other). The resulting set of MSs is here denoted as a MUS (MU-MIMO User Service) group. Please refer to Fig. 8.1. At this stage, MSs report with low resolution feedback. Second, for those selected MSs (by the first step above) which belong to the same group, MSs which would be eligible for the tracking mode are selected. In the tracking mode, a high resolution codebook can be used with a reduced feedback overhead/complexity scheme. Tracking mode is determined based on MS's mobility at the expiration of the *tracking mode entrance qualification timer* (its timer value is shorter than the user grouping timer). These MSs enter into a tracking mode by being commanded to report with high resolution feedback. Third, for the MSs which belong to the same group and are operating in the tracking mode, a kind of qualification test is executed to check if they are still eligible for the tracking mode and MU-MIMO operation, at the expiration of the *exit qualification timer* (its timer value is shorter than the tracking mode entrance qualification timer). An overview of the scheduling algorithm is depicted in Fig. 8.2.

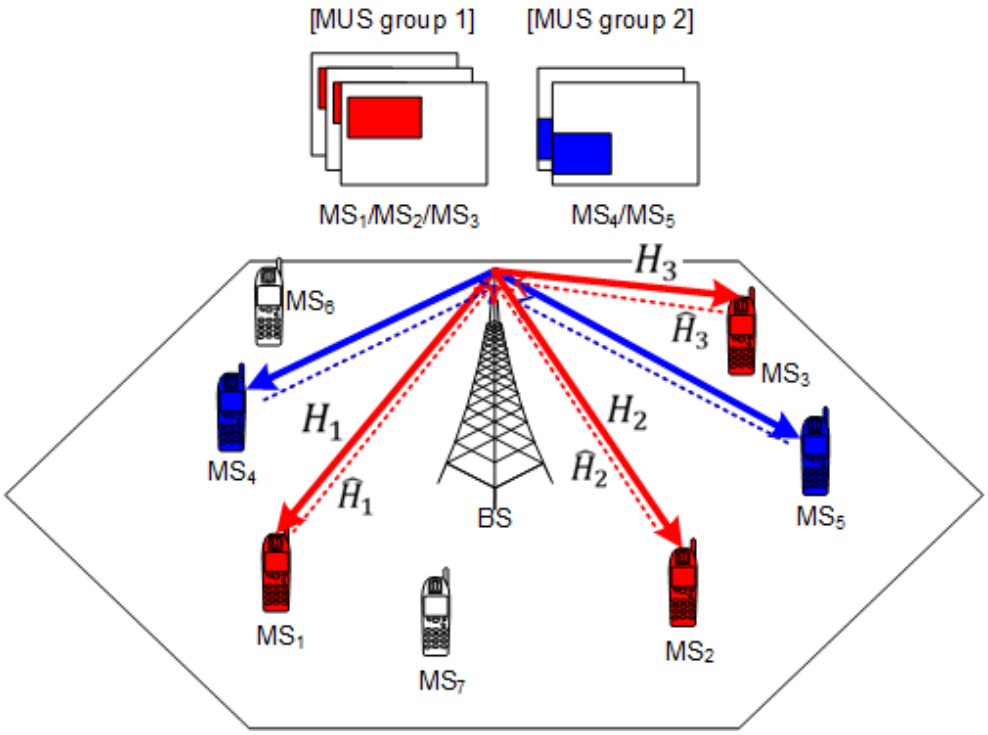


Figure 8.1: MUS (MU-MIMO User Service) group deployment example

8.2 Benefit of the Proposed Scheme

From the cellular network operation point of view, the BS(s) need to categorize MSs into several groups, e.g., those which are eligible for SU-MIMO or for MU-MIMO[2], and those which are eligible for reduced feedback among MU-MIMO qualified MSs. This categorization can be costly in terms of computational complexity at the BS(s), hence it should be avoided to perform a categorization operation too frequently. On the other hand, as the MSs move around, their propagation channels vary accordingly. Therefore the categorization operation should take place frequently enough to reflect this channel variation. In short, we need to come up with a scheme which can facilitate to achieve a trade-off by tuning the system parameters. The categorization can be designed to be performed in a step-by-step manner, starting from the user grouping, then to the tracking mode qualification, and finally to the user group/tracking mode exit qualification with an increasing frequency of operation. It is due to the fact that the user grouping, which sorts out MSs for its eligibility to SU- or MU-MIMO, can be done less frequently based on the long term channel information only (low reso-

[2]In principle, MSs of which the DL channel are orthogonal to each other can be grouped together for MU-MIMO. A user which requires high throughput can be served by SU-MIMO scheme.

lution feedback should be sufficient) while the tracking mode decision, which finds MSs (belonging to the same MU-MIMO group) which have a temporally correlated channel, should be done more frequently as its decision is based on the user mobility. Once in tracking mode, the MSs should report their channel information by using a high resolution codebook via a reduced feedback scheme (Chapter IV) for the sake of reduced feedback overhead/complexity at the MS. By setting the user grouping timer larger than the tracking mode entrance qualification timer which is larger than the exit qualification timer, the overall feedback overhead can be managed efficiently at the system level, and SU- and MU-MIMO grouping together with the deployment of high resolution feedback can be performed in a hierarchical and systematic manner, with a reasonable scheduling complexity at the BS. A more detailed description of the operation can be found in Section 8.3.

8.3 Detailed Description of the Proposed Scheme

Fig. 8.1 depicts the typical operation scenario which the proposed scheme can address. One or more BS(s) serve multiple MSs, and the DL channel varies for each MS in the cell due to the different channel propagation environment and MS mobility. Some MSs might be selected for SU-MIMO, whereas it may be beneficial from the system capacity perspective if some other MSs can be grouped together to be served using the same time-/frequency-resources at the same time (MU-MIMO). In order to exploit the benefit of MU-MIMO, the channel matrix of each MS should be examined and compared to other MSs' channel matrices to find a set of suitable MSs with channel signatures that are as different to each other as possible (best companion to each other). One example of the state of the art user grouping schemes which can be used is the user selection algorithm (*Fuchs et al.*, 2007) which can treat space, time, and frequency jointly, or its extended scheme to the multi-carrier case (*Cheng et al.*, 2014). The resulting set of MSs is termed MU-MIMO User Service (MUS) group. In short, the MSs which belong to the same MUS group are least likely to interfere with each other when a MU-MIMO scheme is applied[3]. In order to do this user grouping at the BS, the BS needs to know the DL channel matrix of each MS ($\mathbf{H}_i$ (i is the MS index) in Fig. 8.1), and this can be provided by the MS in a form of the quantized channel information based on the codebook of the limited size, which is denoted as $\hat{\mathbf{H}}_i$ in Fig. 8.1. In Fig. 8.1, two MUS groups have been assumed, in which three MSs (MS_1, MS_2, MS_3) are served with the same resource (marked as red box) via a MU-MIMO scheme while two MSs (MS_4, MS_5) are served with another

[3]Note that the size of the MUS group is a parameter which can determine a trade off between an average power per user and a multiplexing gain. After MUS grouping is performed, the rest of the MSs which have not been assigned to any MUS group can be served by a SU-MIMO scheme.

resource (marked as blue box) via another MU-MIMO scheme. As stated earlier, a high resolution codebook is required to get the full benefit of MU-MIMO. For slowly varying channels, the MSs can still be able to provide a high resolution channel information via a feedback encoding scheme which exploits channel temporal correlation without increasing the UL signaling overhead/computational complexity at the MS. Therefore it is beneficial when MU-MIMO is applied to the MSs with a high temporal correlation in the same MUS group. Note that in this particular example MS_6 and MS_7 do not belong to any MUS group, which means they might be suitable for SU-MIMO. Note also that it is assumed here that the BS generates and transmits precoded reference signals/pilot symbols in the DL frame. Hence the BS does not have to choose from a predefined precoding matrix set, but rather freely compute and apply the whatever precoding matrix which deems suitable for the current transmission environment, based on the user channel reports and the associated computation.

One potential issue here is that the BS needs to regularly examine the DL channel information from every MS in the cell to update the MUS grouping and so on, as the DL channel can change quickly due to the mobility of the MS. This might be challenging in the sense that there can be a high number of active MSs (> 100) in the cell at a certain time, and the MUS grouping can be very computationally complex. Hence, it is desirable to come up with an idea how to organize this scheduling algorithm so that it can be handled in a reasonable manner. The proposed scheme tries to address this challenge by performing this scheduling at the BS in a hierarchical manner, i.e., to perform the most channel-temporal-change-sensitive scheduling most frequently. Based on this analysis, the MSs in a cell can be grouped into different sets (see Fig. 8.2).

First, all the MSs in the cell automatically belong to the superset *Candidates* group at first. Its sensitivity to channel temporal change is not identified yet, so the BS commands a basic reporting of the channel information to Candidates, e.g., CQI, low resolution channel directional information (CDI), etc., in the least frequent manner. Among these superset Candidates, there can be serval sets of *MUS groups* of which the members are eligible for the MU-MIMO scheme. In each MUS group, there can be one subset *Tracking Mode (TM)* set of which the members are eligible for the encoded feedback scheme (high resolution feedback with reduced overhead/complexity). It depends on each network vendor's choice, but it is believed that it makes sense to apply MU-MIMO to the MSs belonging to the same Tracking Mode set. Other MSs, which do not belong to TM set, are basically considered to be eligible for SU-MIMO or other less spectrally efficient transmission schemes. Updating these sets can be performed with a different periodicity based on different timer value setting. Fig. 8.2 depicts a flow chart view of the proposed scheduling algorithm.

As a prerequisite condition (①)[4], the BS commands all the MSs in the cell (*Candidates*) to report the quantized channel information by using a low resolution codebook. The BS collects and monitors this low resolution channel information $\hat{\mathbf{H}}_i^l$. Optionally, the BS may request the MSs to report the Doppler spread estimate ($f_{DoS} = \frac{vf_c}{c} \cdot cos(\theta)$, where f_c: carrier frequency, v: speed of MS, c: speed of radio propagation, θ: angle of the velocity vector). This can be also used for future assessment of the MS mobility with respect to Tracking Mode grouping. When the *User Grouping Timer* is expired (a few [s]; system parameter[5]) (②), the BS performs the user grouping based on the collected low resolution channel information $\hat{\mathbf{H}}_i^l$ to find the MSs eligible for the MU-MIMO scheme (③). Regarding user grouping schemes, one might use the user grouping methods proposed in (*Yoo and Goldsmith*, 2006; *Fuchs et al.*, 2007) for example. As a result of this procedure, the BS generates the (one or more) MUS group(s).

The following procedure shall take place for each MUS group, which is generated by step ③ above. When the *Tracking Mode Entrance Qualification Timer* is expired (a few tens of [ms]; system parameter) (④), the BS checks every MS in this MUS group if it would be eligible for the Tracking Mode (high channel temporal correlation; low mobility), based on its mobility (⑤). The BS can estimate each MS's mobility by itself or make use of the reported Doppler spread estimate f_{DoS}, if available. Here the BS might use the normalized Doppler spread estimate ($\gamma_D = \frac{|f_{DoS}|}{f_{D,max}}$, where $f_{D,max}$ is (Band-dependent) maximum Doppler shift) and its associated predefined threshold value ($\gamma_{D,threshold}$: Band-independent threshold value, taken from simulations/field trials) to remove the band dependency. A pseudo code of the operation at the expiration of the user grouping timer can be found below in Algorithm 1.

[4]Hereafter in this chapter circled numbers refer to the corresponding steps of the proposed scheduling algorithm which is depicted in Fig. 8.2

[5]The timer value can be tuned based on simulation results and/or lab/field test results.

```
Set the User Grouping Timer;
if User Grouping Timer is expired then
    Execute user grouping procedure to form MUS groups;
    for each MUS group do
        Set the Tracking Mode Entrance Qualification Timer;
        if Tracking Mode Entrance Qualification Timer is expired then
            for every MS in MUS group do
                if γ_D < γ_{D,threshold} then
                    Take this MS to TM set;
                end
            end
            Reset the Tracking Mode Entrance Qualification Timer;
        end
    end
    Reset the User Grouping Timer;
end
```

Algorithm 1: User grouping scheme

The following procedure shall take place for MSs operating in the Tracking Mode (TM). The BS commands MSs qualified for the TM to enter into the Tracking Mode operation (⑥), i.e., to report the high resolution channel information $\hat{\mathbf{H}}_i^h$ (with a differential or adaptive feedback encoding scheme) to make the best use of the MU-MIMO. The BS applies the MU-MIMO scheme for the MSs in the TM. As the high resolution codebook might be large and not required for the average MSs, this codebook can be downloaded from the BS to the MSs in the TM based on their needs, to save memory space at the MS. When the *Exit Qualification Timer* is expired (a few [ms]; system parameter) (⑦), the BS checks every MS in the TM if it satisfies the MUS group exit condition (⑧), e.g., both the MUS grouping condition being not satisfied and the block error rate (BLER) being higher than the predefined threshold value. If the MS is found to satisfy the exit condition, the BS removes these MS(s) from this MUS group.

Note that the removed MS(s) still belong to the Candidates set, therefore they are eligible for all the basic procedures (①-③). For the every remaining MS in the TM, the BS checks if it satisfies the Tracking Mode exit condition (⑨), e.g., both the normalized Doppler spread estimate being higher than the predefined threshold value and the block error rate being higher than the predefined

threshold value. If the MS is found to satisfy the TM exit condition, the BS removes these MS(s) from the TM operation. A pseudo code of the operation at the expiration of the exit qualification timer can be found below in Algorithm 2.

```
Set the Exit Qualification Timer;
if Exit Qualification Timer is expired then
    for every MS in TM do
        if (MUS grouping condition used in (③) not satisfied) AND (BLER
         > BLER_threshold) then
            Remove this MS from this MUS group;
        else
            if (γ_D > γ_D,threshold) AND (BLER > BLER_threshold) then
                Remove this MS from TM;
            else
                Keep this MS at TM;
            end
        end
    end
    Reset the Exit Qualification Timer;
end
```

Algorithm 2: User de-grouping scheme

Note that the removed MS(s) still belong to the associated MUS group set, therefore they are eligible for per MUS group procedures (④ onwards). Once a timer is expired, a new timer will be started automatically right after the required procedures are executed. Note that in Exit MUS Group Check (⑧), BLER check condition (BLER being higher than a predefined threshold) is AND-operated with the MUS grouping condition in (③) not being satisfied. Likewise, this is the case for the Exit Tracking Mode Check (⑨). In this case, the BLER check condition is AND-operated with the TM Entrance Qualification condition in (⑤) not being satisfied. As a result it should be okay to let the MS(s) remain in the MUS group/TM as long as its block error rate is kept low. This condition can take a more conservative notion by switching the AND to an OR operation. However, this might lead to too small number of MSs remaining in the TM (hence in MU-MIMO).

By setting the user grouping timer larger than the tracking mode entrance qualification timer which is larger than the exit qualification timer, the overall feedback overhead can be managed efficiently at the system level. the SU- and MU-MIMO grouping can be performed in a hierarchical

and systematic manner together with deployment of high resolution feedback, with a reasonable scheduling complexity at the BS. The rationale behind this is that the overall system complexity can be reduced by setting the timer values as above, since the size of the Candidate set is larger than the overall size of the MUS groups which is larger than the overall size of the TM sets. The user de-grouping scheme (Algorithm 2) needs to be executed more frequently than the user grouping scheme (Algorithm 1), as maintaining integrity of the MUS group or the TM set, rather than augmenting the group, is more critical to the system performance in terms of throughput.

Note here that the timer values can be set adaptively to the traffic situation (number of active MSs in the cell), e.g., different values for different times of the day (the higher the traffic, the shorter the timer values). The individual timer values and the exit qualification conditions are system parameters which can be used for the tuning of the overall system operation, which can serve the specific needs of the network operators, e.g., balancing between SU-/MU-MIMO user numbers, controlling the feedback overhead, etc.

8.4 Conclusion

In this chapter, the high granularity codebook operation scheme for MU-MIMO systems has been proposed. The proposed scheme facilitates a structured SU-/MU-MIMO user scheduling operation with a reasonable scheduling complexity at the network side. This can be achieved by incorporating a hierarchical user categorization scheme which is triggered by timer expiration events. The corresponding timer values can be defined as configurable system parameters. Finding an efficient way to configure the timer values might be an interesting area of investigation.

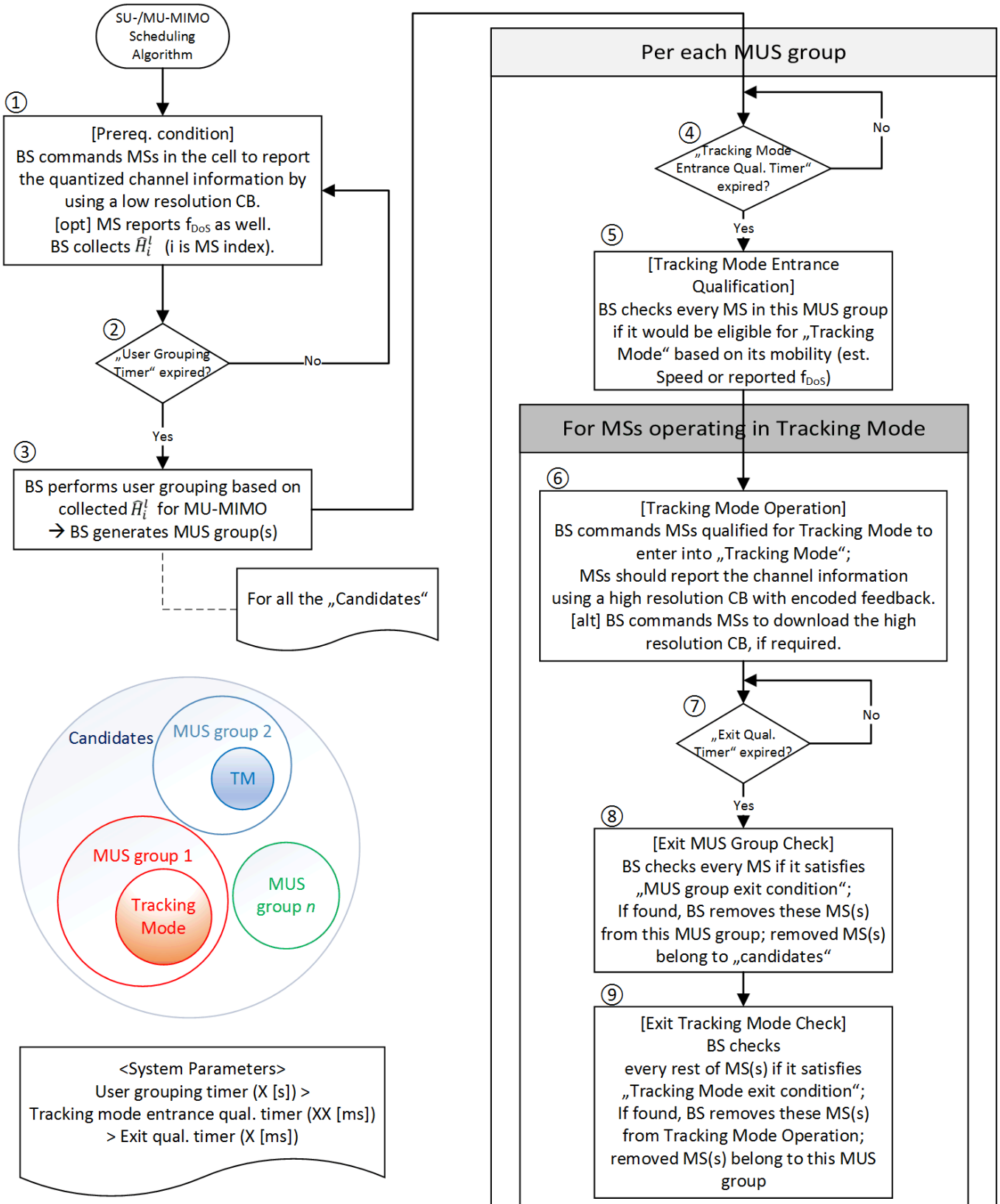


Figure 8.2: SU-/MU-MIMO scheduling algorithm at the network side

CHAPTER IX

Conclusion

The scope of this thesis is to develop efficient feedback techniques for multi-user MIMO systems, targeting distributed cooperative antenna systems in particular. As the associated BSs in distributed cooperative antenna systems are assumed not to have a means to promptly share downlink channel information among themselves, CSI feedback from the MSs should provide the BSs with all the necessary channel information over the air in case a coherent transmit filtering is adopted. One of the challenges for the distributed cooperative antenna concept to be employed in practical mobile communication systems, e.g., 3GPP LTE, is the fact that all the associated base stations should be able to acquire CSI from all the associated mobile stations. This would lead to a lot of feedback overhead considering the high resolution codebook required for interference mitigation and a large channel matrix due to the involvement of multiple BSs. This thesis has proposed efficient feedback methods which capture the essential channel information together with feedback compression methods which exploit the correlation of the channel in an effort to reduce the feedback overhead.

In Chapter II, the subspace based channel quantization method is proposed as a way of providing BSs with downlink channel state information in the presence of inter-user interference, which is applicable to the distributed cooperative antenna systems as well as MU-MIMO systems in general. The subspace based channel quantization improves the system performance significantly, compared to the analog pilot retransmission method with a relatively small feedback overhead. This is shown by simulations in which two or three BSs are cooperating to transmit data signals to two MSs through the same resources at the same time. Here both BSs and MSs are assumed to have a single antenna. This chapter also introduces an efficient codebook construction algorithm based on a well known LBG VQ scheme by adopting the chordal distance and modifying the optimality criteria accordingly. The codebooks generated by the proposed algorithm are shown to have better distance properties than Grassmannian codebooks that are currently available. Considering practical implementation

issues, we have also proposed a channel covariance matrix based nonfrequent CQ feedback method. It is shown that the proposed method is robust against the impacts of the noise, whereas it requires only a small amount of the uplink feedback overhead.

Chapter III introduces an alternative method of reducing the amount of feedback required for channel quantization which has been proposed in (*Roh and Rao*, 2004). This method is based on parameterization and quantization of channel parameters. The channel matrix can be decomposed to reveal its eigenvectors by using a SVD. As a result, we can acquire an unitary matrix which conveys channel spatial information. The geometrical structure of this unitary matrix is exploited to extract a set of Givens parameters that has a one-to-one mapping to the original matrix. The Givens parameter can be further quantized by using simple 1-bit Adaptive Delta Modulation (ADM) to allow further reduction of redundancy in time/frequency, which is intended to exploit the time/frequency correlation of the channel. Since an ADM provides a parameter tuning and tracking capability, Givens parameters can be quantized with only a small number of bits. The differential Givens (D-Givens) method, which is composed of Givens decomposition of the unitary matrix followed by adaptive delta modulation, is presented as a way of reducing the channel quantization feedback data rate. The author's contribution to this topic is the Givens parameter tracking feedback decision method which is based on the empirical normalized autocovariance function.

Chapter IV proposes the feedback overhead reduction methods for temporally-correlated channels. MU-MIMO systems require higher resolution codebooks than SU-MIMO systems to deal with interferences and it leads to an increase of feedback overhead. In this chapter three feedback compression methods, which are proposed by the author, are investigated exploiting the temporal correlation of the channel, i.e., combined codebook (cCB) design method, recursive codebook (rCB) design method and adaptive codebook (aCB) design method. The cCB design method introduces a hierarchical structure in existing subspace-based channel quantization codebooks. The combined codebook has a two-layered structure, which entails a two-layered feedback scheme, i.e., a coarse feedback and a fine feedback. The method for deciding feedback periods is proposed as well. The rCB design introduces the subspace tracking feedback in an effort to reduce the feedback overhead. The subspace tracking feedback encoding, which is also termed the recursive feedback, selects the channel matrix codes from a small size adaptive codebook, depending on the previous decisions. Regarding the operation scenario, the proposed feedback scheme operates in two modes, which are the subspace adjusting mode and the subspace tracking mode. The subspace tracking mode is operated by the recursive codebook, which is a subset of the complete codebook. The aCB design

introduces a transition probability based codeword sorting scheme in an effort to reduce the feedback overhead. The adaptive codebook sorts the codewords in the increasing order of the chordal distance with respect to the initial codeword, which is the codeword chosen at the previous time step. The adaptive codebook method in conjunction with the data compression scheme requires much less feedback overhead compared with other techniques at the cost of slightly increased memory requirements. It is worth to mention that a conceptually similar method to the author's proposals (combined, recursive, and adaptive codebook design schemes), i.e., to split the long-term and the short-term feedback or to exploit channel correlation over time for feedback overhead reduction, has been proposed and investigated for 3GPP LTE-Advanced (release 10 and beyond) standardization (*Lim et al.*, 2013). Even though this proposal was not adopted for LTE-Advanced in the end, future cellular systems, e.g., 5G, should support the high granularity/resolution codebook and its efficient deployment schemes to get the greatest benefit from MU-MIMO. According to the latest technical report on 5G New Radio Access Technology (*3GPP*, 2017), 3GPP has proposed a so called type II feedback for 5G. This new feedback type, which is an explicit channel feedback and/or codebook-based precoder feedback with a higher spatial resolution, is introduced to support MU-MIMO more effectively, compared with the legacy type I feedback which is a codebook-based precoder feedback with a normal spatial resolution. There have been active research works ongoing in this respect, on the development of the efficient explicit codebook design (*Ahmed et al.*, 2019a; *Mondal et al.*, 2019) and on the development of the feedback overhead reduction schemes (*Hindy et al.*, 2020; *Ahmed et al.*, 2019b). The authors of (*Onggosanusi et al.*, 2018) have provided a good overview of key features pertaining to CSI reporting and beam management for the 3GPP 5G New Radio (NR). In (*Mondal et al.*, 2019) researchers at Intel Corporation have provided MU-MIMO system performance comparison between LTE and NR focusing on higher order MU-MIMO and high resolution CSI feedback. Their results reconfirmed the importance of the feedback accuracy for MU-MIMO systems. In (*Hindy et al.*, 2020) the authors have proposed an efficient CSI feedback bit utilization scheme adopting a multi-stage quantization for the Release 16 Type-II codebook, whereas the authors of (*Ahmed et al.*, 2019b) have used a frequency domain compression method to reduce the Type-II CSI overhead. The authors of (*Ahmed et al.*, 2019a) have delivered a comparison of two practical explicit CSI feedback schemes, i.e., one to be based on time domain compression and the other to be based on principal component analysis (PCA). It has been shown that the PCA based method could achieve around 6 % performance gain at the cost of higher feedback overhead. It turns out that their PCA based method is effectively the same as the subspace-based channel quantization

method in this dissertation but with one important difference. They have directly quantized the coefficients after having performed a SVD, which leaves room for improvement when it comes to channel subspace aware vector quantization and efficient feedback compression. The techniques to cope with the high resolution feedback that are developed in this thesis can be useful and applicable for 5G in this regard.

In Chapter V, we extend the adaptive codebook design method into two-dimension, i.e., time and frequency. This is a logical extension of the method which relies on the correlation of the channel, as there exists correlation over adjacent subcarriers as well as subsequent time samples in multi-carrier transmission systems like OFDM. The chapter introduces the codeword interpolation method as well. Numerical results show that the two-dimensional adaptive encoding method with clustering of 6 adjacent subcarriers can greatly reduce the feedback overhead compared to the case without clustering (per-subcarrier encoding) in OFDM based systems. The clustering case achieves a performance that is 1.1 dB away from the reference case without clustering while using only 13.6 % of the feedback overhead.

In Chapter VI, the feedback methods introduced in the thesis are compared in terms of their performance, their required feedback overhead, and their computational complexity. They are the Givens rotation based channel quantization method (D-Givens) and the subspace based channel quantization method with adaptive codebook (aCB). The aCB method outperforms the D-Givens method in terms of the received SINR (1.7 dB to 4.8 dB gain, depending on the operational SINR) and the required feedback overhead (13.2 % of that of the D-Givens method). The D-Givens method achieves a reasonable performance with a moderate complexity.

In Chapter VII, the previous work on the subspace based channel quantization method is extended by introducing multiple antennas at the transmitter side (BS) and/or at the receiver side (MS) and by investigating the performance of the RBD precoding(/decoding) scheme as well as the ZF precoding scheme. The RBD precoding scheme requires the MS to compute the decoding matrix via a SVD in case of the multiple data stream transmission to each user. A cost-efficient decoding matrix quantization method is proposed which can avoid a complex computation at the MS while showing a slight SNR degradation for the users. A numerical analysis has been performed for cooperative antenna systems with 2 BSs and 1 or 2 MS(s) in which BSs and MS(s) can have a single antenna or two antennas as long as the total number of transmit antennas is equal to or greater than the total number of receive antennas, and two BSs cooperate to serve one or two MS(s) at the same time.

In Chapter VIII, a systematic scheduling mechanism at the network side is described which facilitates an efficient high resolution codebook operation.

As a future research direction, it would be interesting to investigate the feedback overhead reduction method further, as it can become a critical bottleneck in the interference-limited system since it would require a high resolution codebook. The feedback schemes can be combined with the existing interference mitigation methods, e.g., user selection/grouping. It can be also interesting to investigate a realistic feedback link which takes feedback delay and error into account, together with its impact on the overall system performance[1]. As massive MIMO and Full-Dimension MIMO systems have received much attention recently (*Marzetta*, 2010; *Ji et al.*, 2017), it is worthwhile to investigate how to design a limited feedback scheme for a system with a large number of transmit antennas. For example, 3GPP has addressed in release 13 a channel state information (CSI) feedback for a large number of transmit antenna elements by grouping them into two main categories: per-element[2] reporting (class A) and per-beam reporting (class B) (*Dahlman et al.*, 2016). CSI reporting class A is a direct extension of the CSI reporting scheme in earlier 3GPP releases to a larger number of antenna ports. Each antenna element transmits a unique CSI reference signal. As the amount of CSI reference signals would be determined by the number of transmit antenna elements, this approach is mainly suitable for a modest number of transmit antennas, e.g., up to 16. CSI reporting class B implies that each CSI reference signal is beam-formed using all the transmit antenna elements. The mobile device is supposed to measure the beam-formed CSI reference signals and to recommend a suitable beam among those beams, along with a preferred precoder matrix conditioned on the topmost beam being used for transmission. In this case, the number of simultaneous beams would determine the amount of CSI reference signals, the device channel estimator complexity and the associated feedback mechanism complexity. Hence it makes this approach primarily suitable for a large number of transmit antenna elements, e.g., 64. It is worthwhile to note that for both categories a codebook based feedback is still to be used to some extent. The feedback overhead reduction methods that have been proposed in this thesis can be applicable, in case 3GPP adopts a high resolution codebook. How to reduce the required memory and computational complexity at the mobile station can be an interesting topic as well. One of the research works in this direction can be found in (*Choi et al.*, 2015). It might be also interesting to see if recent advances on multidimensional parameter estimation techniques, e.g., R-D ESPRIT (*Haardt and Nossek*, 1998), Tensor-ESPRIT

[1] An analysis of the achievable data rates of the regularized block diagonalization (RBD) precoded multi-user MIMO broadcast channels with respect to the channel estimation error and feedback delay can be found in (*Song and Haardt*, 2009).

[2] per transmit antenna element

(*Haardt et al.*, 2008), and related analytical performance analyses (*Roemer et al.*, 2014; *Steinwandt et al.*, 2017), can be adopted for efficient feedback design.

APPENDICES

APPENDIX A

Symbols and Notations

The symbols and notations used in the thesis are as follows.

h	scalar
$\mathbf{h}$	vector
$\mathbf{H}$	matrix
$(\cdot)^{\mathcal{T}}$	transpose
$(\cdot)^{\mathcal{H}}$	Hermitian transpose
$(\cdot)^{-1}$	inverse
$\langle\cdot\rangle$	rounding operator
$\langle\mathbf{u},\mathbf{v}\rangle$	inner product, $\mathbf{u}^{\mathcal{H}}\mathbf{v}$
$\mathrm{tr}(\cdot)$	trace
$\mathrm{diag}(\mathbf{x})$	diagonal matrix with the elements of vector $\mathbf{x}$ on the main diagonal
$\lambda_{\max}(\cdot)$	largest eigenvalue
$\lvert\cdot\rvert$	magnitude of a scalar
$\lVert\cdot\rVert_2$	two-norm of a vector or a matrix, $\lVert\mathbf{a}\rVert_2 = \sqrt{\sum_i \lvert a_i\rvert^2} = \sqrt{\langle\mathbf{a},\mathbf{a}\rangle}$, $\lVert\mathbf{A}\rVert_2 = \sqrt{\lambda_{\max}(\mathbf{A}^{\mathcal{H}}\mathbf{A})}$
$\lVert\cdot\rVert_F$	Frobenius norm of a matrix, $\lVert\mathbf{A}\rVert_F = \sqrt{\sum_{i,j} \lvert a_{i,j}\rvert^2} = \sqrt{\mathrm{tr}(\mathbf{A}^{\mathcal{H}}\mathbf{A})} = \sqrt{\mathrm{tr}(\mathbf{A}\mathbf{A}^{\mathcal{H}})}$
$\mathbb{E}[\cdot]$	expectation
μ_X	mean of the random variable X
σ_X^2	variance of the random variable X
$\mathbf{R}_{\mathbf{x}}$	covariance matrix of the vector process $\mathbf{x}$, $\mathbb{E}[\mathbf{x}\mathbf{x}^{\mathcal{H}}]$
$c_{XX}(k)$	normalized autocovariance function of the real value process $X(t)$ with time difference k, $c_{XX}(k) = \frac{\mathbb{E}[(X(t)-\mu_X)(X(t+k)-\mu_X)]}{\sigma_X^2}$
$\mathbf{I}_N$	$N \times N$ identity matrix

$\mathbf{0}_{M\times N}$	all-zero matrix of size $M \times N$
$\mathbf{I}_{M\times N}$	defined as $\begin{bmatrix} \mathbf{I}_N \\ \mathbf{0}_{(M-N)\times N} \end{bmatrix}$ for $M > N$
$\mathbf{1}_k$	vector with k 1's
$[\mathbf{A}]_{i,j}$	$(i,j)^{th}$ entry of a matrix $\mathbf{A}$
$\lvert\mathcal{S}\rvert$	cardinality of a set $\mathcal{S}$
$d_c(\mathbf{T}_i, \mathbf{T}_j)$	chordal distance between $\mathbf{T}_i$, $\mathbf{T}_j$ which have orthonormal columns, $\frac{1}{\sqrt{2}}\|\mathbf{T}_i\mathbf{T}_i^{\mathcal{H}} - \mathbf{T}_j\mathbf{T}_j^{\mathcal{H}}\|_F$
$d_{c,\min}(\cdot)$	minimum chordal distance of the codebook
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	complex circular symmetric Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$
$\mathbb{C}^{m\times n}$	set of $m \times n$ complex matrices
N_{BS}	number of cooperating BSs in cooperative area (CA)
N_{MS}	number of MSs in CA
N_t	number of transmit antennas at BS
N_r	number of receive antennas at MS
N_{tr}	total number of receive antennas in CA, $N_{tr} = N_{\mathrm{MS}}N_r$
N_{tt}	total number of transmit antennas in CA, $N_{tt} = N_{\mathrm{BS}}N_t$
n_s	number of data streams for each MS, $n_s \leq N_r$
$\mathbf{s}$	$\mathbb{C}^{N_{tr}\times 1}$ data symbol vector, $\mathbf{s} = [s_1, \ldots, s_{N_{tr}}]^{\mathcal{T}}$
$\mathbf{W}$	$\mathbb{C}^{N_{tt}\times N_{tr}}$ precoding matrix
$\mathbf{H}$	$\mathbb{C}^{N_{tr}\times N_{tt}}$ channel matrix
g	normalization factor imposed by the transmit power constraint
i_{BS}	BS index
i_{MS}	MS index
i_t	transmit antenna index at BS
i_r	receive antenna index at MS
i	row index of the channel matrix $\mathbf{H}$, $i = N_r(i_{\mathrm{MS}} - 1) + i_r$
j	column index of the channel matrix $\mathbf{H}$, $j = N_t(i_{\mathrm{BS}} - 1) + i_t$
$h_{i,j}$	channel coefficient between the i_rth receive antenna of the i_{MS}th MS and the i_tth transmit antenna of the i_{BS}th BS, $h_{i,j} = [\mathbf{H}]_{i,j}$
σ_N^2	noise variance
$\mathbf{n}$	$\mathcal{CN}(0, \sigma_N^2\mathbf{I})$ additive noise sample vector
$\hat{\mathbf{s}}$	$\mathbb{C}^{N_{tr}\times 1}$ data symbol vector estimate

$\mathbf{R_n}$	$\mathbb{C}^{N_{tr}\times N_{tr}}$ covariance matrix of the noise, $\mathbf{R_n} = \mathbb{E}[\mathbf{n}\mathbf{n}^{\mathcal{H}}]$
$\mathbf{R_s}$	$\mathbb{C}^{N_{tr}\times N_{tr}}$ covariance matrix of the data symbols, $\mathbf{R_s} = \mathbb{E}[\mathbf{s}\mathbf{s}^{\mathcal{H}}]$
P_{tx}	maximum transmit power in CA
$\mathbf{H}_j$	$\mathbb{C}^{N_{tt}\times N_r}$ transpose of the channel matrix for MS j
$\mathbf{U}_j$	$\mathbb{C}^{N_{tt}\times N_{tt}}$ left singular vectors obtained by SVD of the channel matrix for MS j $\mathbf{H}_j$
$\mathbf{V}_j$	$\mathbb{C}^{N_r\times N_r}$ right singular vectors obtained by SVD of the channel matrix for MS j $\mathbf{H}_j$
$\mathbf{U}_j^{(\mathrm{S})}$	$\mathbb{C}^{N_{tt}\times N_r}$ basis for the signal subspace of the channel matrix for MS j $\mathbf{H}_j$
$\mathbf{U}_j^{(0)}$	$\mathbb{C}^{N_{tt}\times(N_{tt}-N_r)}$ basis for the left nullspace of the channel matrix for MS j $\mathbf{H}_j$
$\hat{\mathbf{U}}_j^{(\mathrm{S})}$	$\mathbb{C}^{N_{tt}\times N_r}$ quantized version of the column space basis vectors $\mathbf{U}_j^{(\mathrm{S})}$
$\mathbf{C}_i$	$\mathbb{C}^{N_{tt}\times N_r}$ code of index i
$\mathcal{C}$	codebook of size N_{cb}, set of all codes $\mathbf{C}_i$
n_{CB}	codebook size in bits, $2^{n_{\mathrm{CB}}} = \lvert\mathcal{C}\rvert$
N_{cb}	codebook size, $N_{cb} = 2^{n_{\mathrm{CB}}}$
$\hat{\mathbf{H}}_j^d$	$\mathbb{C}^{N_{tt}\times N_r}$ quantized version of the channel matrix for MS j $\mathbf{H}_j$, $\hat{\mathbf{H}}_j^d = \hat{\mathbf{U}}_j^{(\mathrm{S})}$
$\mathcal{G}(m,n)$	set of all n-dimensional subspaces of the space $\mathbb{C}^m$ (Grassmannian space)
$\mathcal{T}$	training sequence of size M which consists of basis for the left nullspace of the channel matrix samples, $\mathcal{T} = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_M\}$, where $\mathbf{X}_m \in \mathbb{C}^{N_{tt}\times N_r}$ is a sample of $\mathbf{U}_j^{(\mathrm{S})}$
$\mathcal{R}_n$	encoding region (Voronoi cell)
$\mathcal{P}$	partition of the space, $\mathcal{P} = \{\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_{N_{cb}}\}$
$centroid(\mathcal{S})$	optimum code calculated for a given encoding region $\mathcal{S}$
T_c	coherence time
T_s	OFDM symbol duration, $T_s = 71.37\mu s$
$\mathbf{R}_{\mathbf{H}_j}$	channel covariance matrix for MS j, $\mathbf{R}_{\mathbf{H}_j} = \mathbb{E}\left[\mathbf{H}_j\mathbf{H}_j^{\mathcal{H}}\right]$
N_w	window size used for estimation of the channel covariance matrix (unit: number of OFDM symbols)
f_D	maximum Doppler frequency
v_{MS}	mobile speed
c	radio propagation speed
$\mathbf{D}_k$	$\mathbb{C}^{N_{tt}\times N_{tt}}$ kth diagonal matrix obtained by Givens rotation based decomposition of $\mathbf{U}_j^{(\mathrm{S})}$, $\mathbf{D}_k(\phi_{k,k}, \cdots, \phi_{k,N_{tt}}) = \mathrm{diag}(\mathbf{1}_{k-1}, e^{j\phi_{k,k}}, \cdots, e^{j\phi_{k,N_{tt}}})$
$\phi_{k,l}$	$(l-k+1)$th phase for kth diagonal matrix $\mathbf{D}_k$
$\mathbf{G}_{p-1,p}(\theta_{k,l})$	$\mathbb{R}^{N_{tt}\times N_{tt}}$ Givens matrix which operates in the $(p-1,p)$ coordinate plane

$\theta_{k,l}$	lth rotational angle for kth diagonal matrix $\mathbf{D}_k$
$\phi[k]$	kth sample of phase ϕ
$\hat{\phi}[k]$	kth phase constructed by adaptive delta modulation (ADM)
$\Delta[k]$	ADM step-size for kth sample
$c[k]$	ADM encoded bit for kth sample, $c[k] \in \{-1, +1\}$
α	ADM step-size scaling factor, $\alpha > 1$
$[\tau]$	unit of the feedback period which is the number of OFDM symbols, $[\tau] = \frac{\tau}{T_s}$
N_{GP}	number of Givens parameters
τ_{ref}	Givens parameter feedback period of the reference case with mobile speed v_{MS}
$\hat{c}_{X_n X_n}^{(v_{\mathrm{MS}})}([\tau_n])$	normalized autocovariance function of the nth Givens parameter with mobile speed v_{MS} at temporal separation τ_n
$\mathcal{C}^c$	coarse codebook of size N_c
i_c	coarse codebook index
n_c	coarse codebook size in bits, $2^{n_c} = \lvert\mathcal{C}^c\rvert$
N_c	coarse codebook size, $N_c = 2^{n_c}$
$\mathbf{C}_{i_f}^{i_c}$	$\mathbb{C}^{N_{tt} \times N_r}$ fine code of index i_f which belongs to the i_cth fine codebook $\mathcal{C}_{i_c}^c$, $\mathbf{C}_{i_f}^{i_c} = \mathbf{C}_{(i_c-1)N_f+i_f}$
$\mathcal{C}_{i_c}^c$	fine codebook of index i_c, $\mathcal{C}_{i_c}^c = \{\mathbf{C}_1^{i_c}, \mathbf{C}_2^{i_c}, \cdots, \mathbf{C}_{i_f}^{i_c}, \cdots, \mathbf{C}_{N_f}^{i_c}\}$
i_f	fine code index
n_f	fine codebook size in bits, $2^{n_f} = \lvert\mathcal{C}_{i_c}^c\rvert$
N_f	fine codebook size, $N_f = 2^{n_f}$
n_t	combined codebook size in bits, $n_t = n_{\mathrm{CB}} = n_c + n_f$
τ_c	coarse feedback period
τ_f	fine feedback period
$\mathcal{C}^r$	recursive codebook of size N_{rc}
n_r	recursive codebook size in bits, $2^{n_r} = \lvert\mathcal{C}^r\rvert$
N_{rc}	recursive codebook size, $N_{rc} = 2^{n_r}$
J_p	channel state (quantized channel index) at time p
P_i	stationary probability of channel state being i at time p, $P_i = \Pr\{J_p = i\}$
$P_{i,j}$	transition probability of going from channel state j at time $p-1$ to channel state i at time p, $P_{i,j} = \Pr\{J_p = i \lvert J_{p-1} = j\}$
$\mathcal{N}_i(\epsilon)$	ϵ-neighborhood of the Markov state i at time p
$\mathcal{N}_j'(\epsilon)$	next step ϵ-neighborhood of the Markov state j at time $p-1$

τ_a	subspace adjusting feedback period
τ_t	subspace tracking feedback period
$\mathcal{C}^a$	adaptive codebook of size N_{cb}
$\mathbf{C}_i^a$	$\mathbb{C}^{N_{tt} \times N_r}$ ith member of the adaptive codebook $\mathcal{C}^a$
$\mathcal{C}^a[J_p]$	adaptive codebook for channel state J_p at time p
K	cluster size in number of subcarriers
K'	interpolation area size in number of subcarriers
N	total number of subcarriers
k	subcarrier index in an interpolation area, $1 \leq k \leq K'$
l	interpolation area index, $0 \leq l \leq N/K' - 1$
θ_l	phase rotation parameter for interpolation area l
$\widehat{\mathbf{h}}_{j,l}$	$\mathbb{C}^{N_{tt} \times 1}$ quantized version of the channel vector for MS j at interpolation area l, $\widehat{\mathbf{h}}_{j,l} = \mathcal{Q}(\mathbf{H}_j(lK' + 1))$
$\widetilde{\mathbf{h}}_j(lK' + k; \theta_l)$	$\mathbb{C}^{N_{tt} \times 1}$ interpolated channel codeword for kth subcarrier in $(l+1)$th interpolation area
c_k	linear weight value for codeword interpolation
N_{sym}	feedback period in number of OFDM symbols

APPENDIX B

Channel Quantization Criterion for the MISO Case

Proof of Equation (2.14)

Proof: First, we provide the formula which explains how the chordal distance is related with the inner product, when it is used for two unit norm vectors $\mathbf{v}_i, \mathbf{v}_j$ where $\mathbf{v}_i^{\mathcal{H}}\mathbf{v}_i = \mathbf{v}_j^{\mathcal{H}}\mathbf{v}_j = 1$.

$$\begin{aligned} d_c^2(\mathbf{v}_i, \mathbf{v}_j) &= \left(\frac{1}{\sqrt{2}} \|\mathbf{v}_i\mathbf{v}_i^{\mathcal{H}} - \mathbf{v}_j\mathbf{v}_j^{\mathcal{H}}\|_F \right)^2 \\ &= \frac{1}{2}\mathrm{tr}\left((\mathbf{v}_i\mathbf{v}_i^{\mathcal{H}} - \mathbf{v}_j\mathbf{v}_j^{\mathcal{H}})(\mathbf{v}_i\mathbf{v}_i^{\mathcal{H}} - \mathbf{v}_j\mathbf{v}_j^{\mathcal{H}})^{\mathcal{H}} \right) \\ &= \frac{1}{2}\mathrm{tr}\left(\mathbf{v}_j^{\mathcal{H}}(\mathbf{v}_i\mathbf{v}_i^{\mathcal{H}} - \mathbf{v}_j\mathbf{v}_j^{\mathcal{H}})(\mathbf{v}_i\mathbf{v}_i^{\mathcal{H}} - \mathbf{v}_j\mathbf{v}_j^{\mathcal{H}})^{\mathcal{H}}\mathbf{v}_j \right) \\ &= \frac{1}{2}\left(1 - \mathbf{v}_j^{\mathcal{H}}\mathbf{v}_i\mathbf{v}_i^{\mathcal{H}}\mathbf{v}_j \right) \\ &= \frac{1}{2}\left(1 - |\langle \mathbf{v}_i, \mathbf{v}_j \rangle|^2 \right), \end{aligned} \tag{B.1}$$

where $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \mathbf{v}_i^{\mathcal{H}}\mathbf{v}_j$. The following property is used from the first to the second line: $\|\mathbf{A}\|_F = \sqrt{\mathrm{tr}\left(\mathbf{A}\mathbf{A}^{\mathcal{H}}\right)}$, and from the third to the fourth line the $\mathrm{tr}(\cdot)$ operation is omitted since its argument has a scalar value. From (B.1), the decision criterion in terms of the chordal distance can be formulated as follows.

$$\arg\min d_c(\mathbf{v}_i, \mathbf{v}_j) = \arg\max |\langle \mathbf{v}_i, \mathbf{v}_j \rangle| \tag{B.2}$$

On the other hand, the column space basis vector $\mathbf{u}_{\mathrm{S}} \in \mathbb{C}^{N_{tt} \times 1}$ of the channel vector $\mathbf{h} \in \mathbb{C}^{N_{tt} \times 1}$ can be found as (the user index j is omitted for brevity):

$$\mathbf{h} = \mathbf{U}_h \mathbf{\Sigma}_h \mathbf{V}_h^{\mathcal{H}} = v_h \sigma_h \mathbf{u}_{\mathrm{S}}, \tag{B.3}$$

where v_h and σ_h have scalar values and $\mathbf{u}_{\mathrm{S}} \in \mathbb{C}^{N_{tt} \times 1}$. The rightmost form is an "economy size" version of the SVD, where v_h is in effect same as $\mathbf{V}_h$ (size: 1×1), and σ_h is the only non-zero

singular value in $\mathbf{\Sigma}_h$. The followings hold: $\sigma_h = \|\mathbf{h}\|_2$ and $v_h \in \{+1, -1\}$, since $\mathbf{h}$ is a vector. Therefore, $\mathbf{u}_{\mathrm{S}}$ can be expressed in terms of the directional vector of the channel as follows.

$$\mathbf{u}_{\mathrm{S}} = v_h \frac{\mathbf{h}}{\|\mathbf{h}\|_2} = v_h \mathbf{v}, \tag{B.4}$$

where $\mathbf{v} = \mathbf{h}/\|\mathbf{h}\|_2$ is the directional vector of the channel. Since v_h decides the sign only, the following criterion holds:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{c}_i \in \mathcal{C}} d_c(\mathbf{u}_{\mathrm{S}}, \mathbf{c}_i) = \arg \max_{\mathbf{c}_i \in \mathcal{C}} |\langle \mathbf{v}, \mathbf{c}_i \rangle| \tag{B.5}$$

Therefore, the chordal distance based channel quantization criterion (2.11) can be simplified to the inner product based criterion (2.14). ■

APPENDIX C

On Decision of the Feedback Period

This section is an appendix to Section 4.1. This section investigates how to determine the feedback periods τ_c and τ_f to guarantee the target performance for a given profile, e.g., the dimension of the quantized downlink channel $\mathbf{H}_j \in \mathbb{C}^{N_{tt} \times N_r}$, the carrier frequency f_c, the speed of MS v_{MS}, and the number of feedback bits n_c, n_f.

C.1 System Requirements

There are constraints imposed by the system implementation issue:

- *The coarse feedback period should be an integer multiple of the fine feedback period.*

$$\tau_c = n\tau_f, \ n \in \mathbb{N}, \tag{C.1}$$

where $\mathbb{N}$ denotes the set of all natural numbers, i.e., $\mathbb{N} = \{1, 2, 3, \cdots\}$. This constraint enables both coarse and fine feedback operations to work in a synchronized way, and prevents redundant usage of the feedback resource. For example, if $\tau_c = 2.5\tau_f$ provides us with the target performance, it leaves the MS two inefficient options. One is to send the fine feedback message 3 times more frequently as the coarse feedback message so as to guarantee the target performance, and another is to send the fine feedback 2 times more frequently as the coarse feedback in order to use less uplink resource for feedback. The former is inefficient since it consumes uplink resource more than required. The latter is also inefficient in that it cannot achieve the target performance since it sends fine feedback less frequently than required.

- *The fine feedback period should be an integer multiple of the symbol duration.*

$$\tau_f = mT_s,\ m \in \mathbb{N}, \tag{C.2}$$

where T_s denotes the symbol duration time. This constraint simply reflects the fact that the symbol is the basic unit of the physical layer operation. The physical layer frame consists of several symbols in time, and the basic unit of channel estimation, data transmission, and control signaling including feedback operations, is a symbol. Therefore, the fine feedback period is rounded off to be an integer multiple of the symbol duration time.

C.2 Semi-analytic Approach to Feedback Period Decision

As briefly mentioned in the previous section, finding an optimal feedback period is a challenging task, since the cost function is dependent on many factors, including the dimension of the channel matrix, the number of channel quantization bits, the channel quantization algorithm and the codebook construction method as well as the channel characteristics. Among them, the numbers of feedback bits n_c, n_f are parameters we want to best utilize, and the speed of MS v_{MS} is a variable which is difficult to control in the system design. However, the system related parameters, which are the dimension of the channel matrix, the channel quantization algorithm, the codebook construction method and the carrier frequency, are static or can be forced to be static during the feedback operation since they are system design parameters. Thus, the feedback period decision problem can be stated as follows which is addressed in the following proposed method: Given the dimension of the channel matrix (N_{tt}, N_r), the channel quantization algorithm ($\mathcal{Q}(\ )$), the codebook construction method (mLBG( )) and the carrier frequency f_c, find the best resource utilizing τ_c, τ_f with respect to v_{MS} such that the target performance is achieved.

- CEAC based Feedback Period Decision Method

  The Complementary Envelope Autocorrelation Coefficient (CEAC) based method is the semi-analytic scheme which finds the solution with the help of the channel envelope autocorrelation function. This method is applicable to the uncorrelated Rayleigh fading channel case. This method works as follows:

  1. Find the reference target performance-achieving feedback period τ_c for given n_c and v_{MS}, by simulations.

This will serve as a reference target performance-achieving feedback period. Even though there is a rule of thumb to estimate the coherence time of the channel, $T_c = \frac{0.423}{f_m}$, where $f_m = \frac{v_{\mathrm{MS}}}{c} f_c$, this leads us to much relaxed criteria. For instance, we can achieve less than 0.5 dB degradation in terms of 50 % outage SINR compared with the $\tau_c = T_s$ case, when applying $\tau_c = 10\ T_s$ (T_s is the OFDM symbol duration for LTE, 71.37 μs) for a $n_c = 3$ bit codebook with $v_{\mathrm{MS}} = 10$ m/s. This is a tolerable loss which satisfies our target performance criteria. On the other hand, the rule of thumb gives us $T_c \approx 4.881$ ms, which is almost 7 times of τ_c. This brings about more than 2 dB degradation.

2. Based on the reference feedback period τ_c found empirically in step 1, find the feedback period of our interest for given n_f and v_{MS} by using the CEAC function.

   The CEAC function, which is denoted as $p(\Delta t)$, is defined as follows:

$$p(\Delta t) := \rho(0) - \rho(\Delta t), \text{for } 0 \leq \Delta t \leq \tau_0, \tag{C.3}$$

   where $\rho(\Delta t) = J_0^2\left(2\pi v_{\mathrm{MS}} \Delta t/\lambda\right)$ is the Envelope Autocorrelation Coefficient (EAC) function (*Jakes*, 1974) and τ_0 is the first zero crossing point of $\rho(\Delta t)$ which satisfies $\tau_0 \in \{\tau_i : \rho(\tau_i) = 0, \tau_i \leq \tau_{i'}, \forall i < i', i = 0, 1, 2, \cdots\}$. $J_0()$ is a zeroth-order Bessel function of the first kind[1]. $\rho(\Delta t)$ represents the correlation between the channel's responses to sinusoids received with a time separation equal to Δt, where v_{MS} is the mobile speed, and λ is the wavelength of the propagating signal. Fig. C.1 depicts the envelope autocorrelation coefficient as a function of the relative time delay Δt. The decrease of the correlation for a longer time separation is clearly identified in this figure. We can regard the CEAC as a measure of how uncorrelated the received signals are, for a time separation Δt. The CEAC function $p(\Delta t)$ is a monotonically increasing function from 0 (completely correlated) to 1 (completely uncorrelated) in the interval $[0, \tau_0]$.

   We should elaborate the time delay Δt as a function of the number of feedback bits n (n_c or n_f) to take the channel quantization resolution into account. The time delay is meaningful only when it leads to a channel quantization output transition, since a feedback message is not necessary when the output does not change over time. Hence we define $\Delta t(n)$ as a minimum recognizable time delay, which is the average of minimum time delays such that the output of the n bit channel quantization changes. By inserting $\Delta t(n)$

[1]The derivation of $\rho(\Delta t)$ is based on the assumption of a uniform scattering environment without a dominant Line Of Sight (LOS) component. The proposed method, hence, is applicable to the uncorrelated Rayleigh fading case.

into (C.3), we can acquire $p(\Delta t(n))$ which is a measure of uncorrelated-ness corresponding to the minimum recognizable time delay for a n bit channel quantization.

It is not a trivial task to come up with a completely analytic formula $p(\Delta t(n))$ in terms of n. However, the relative formula can be drawn as follows. Our empirical research shows that $p(\Delta t(n))$ is inversely proportional to the resolution of the n bit channel quantization which corresponds to the number of encoding regions (2^n). It is formulated as

$$p(\Delta t(n)) \propto \frac{1}{2^n}. \tag{C.4}$$

Once we have a reference case ($\Delta t(n')$ for n') which is acquired by the empirical approach, we can find $\Delta t(n)$ for a given n by exploiting (C.4) as follows. Let us assume that we have a reference recognizable time delay $\Delta t(n')$ which is empirically proven to achieve the target performance, e.g., less than 0.5 dB degradation from the ideal case, for the given number of feedback bits n'. The CEAC value $p(\Delta t(n))$ for n bit feedback can be calculated as the following:

$$p(\Delta t(n)) = \frac{p(\Delta t(n'))}{2^{n-n'}} \tag{C.5}$$

This means that the level of uncorrelated-ness for n feedback bits should be reduced to 1 over $2^{n-n'}$ of that for n' feedback bits. In other words, the level of correlation should be increased $2^{n-n'}$ times, since one encoding region is divided into $2^{n-n'}$ encoding regions, and this increase of the channel quantization resolution should be reflected in p. The linear relationship in (C.5) comes from (C.4). Equation (C.5) allows us to calculate the CEAC value, once the reference value is given. Now we have a p value of our interest n, so we can find a $v_{\mathrm{MS}}\Delta t/\lambda$ value by using the inverse function of (C.3) for $0 \leq \Delta t \leq \tau_0$. The whole process can be summarized as follows:

(a) Calculate the CEAC value for the target performance-achieving reference case.

$$p(\Delta t(n')) = 1 - \rho(\Delta t(n')) = 1 - J_0^2\left(2\pi\frac{v_{\mathrm{MS}}\Delta t(n')}{\lambda}\right) \tag{C.6}$$

(b) Calculate the CEAC value for the case of our interest by (C.5).

$$p(\Delta t(n)) = \frac{p(\Delta t(n'))}{2^{n-n'}} \tag{C.7}$$

(c) Calculate the normalized delay value $v_{\mathrm{MS}}\Delta t(n)/\lambda$ to get the feedback period of our

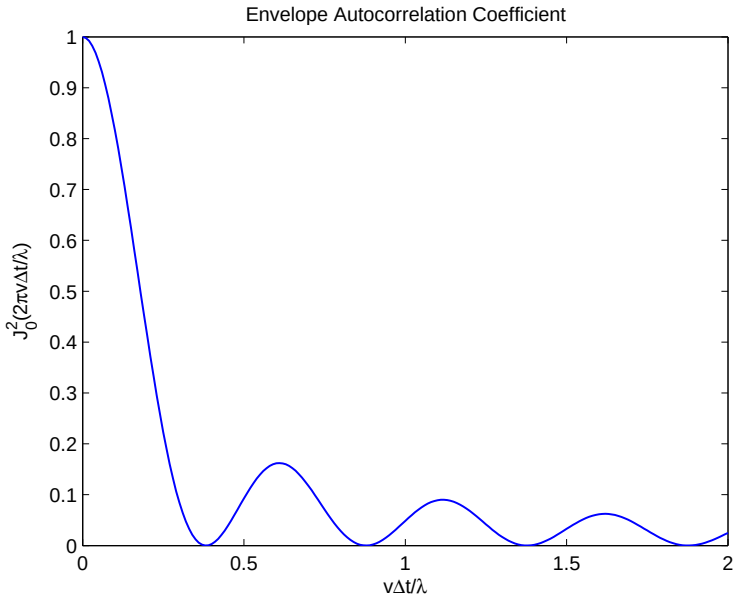


Figure C.1: Envelope autocorrelation coefficient function $\rho(\Delta t)$ over the relative time delay Δt

interest by using the inverse CEAC function.

$$\frac{v_{\mathrm{MS}}\Delta t(n)}{\lambda} = \frac{1}{2\pi}\left(J_0^2(1-p(\Delta t(n)))\right)^{-1}, \text{ for } 0 \le p \le 1 \tag{C.8}$$

Equation (C.8) is depicted in Fig. C.2. We should note that we can acquire $\Delta t(n)$ for a given v_{MS} by (C.8).

Example C.1

Consider the 2 BSs-2 MSs cooperative antenna system with single antenna BS and MS. We want to find a feedback period for the 5 bit combined codebook, of which 3 bits are allocated to the coarse feedback and 2 bits are to the fine codebook. Simulation results show that by setting $\tau_c = 10\ T_s (T_s = 71.37\mu s)$ for $n_c = 3$ bit coarse feedback, we can achieve a less than 0.5 dB degradation compared with the $\tau_c = T_s$ case. Simulations have been performed under the following system parameters: $v_{\mathrm{MS}} = 10$ m/s, Urban Macro channel, $\mathbf{H}_j \in \mathbb{C}^{2\times 1}$(the dimension of the channel matrix to quantize is 2×1). Decide the fine feedback period τ_f for $n_f = 2$ which satisfies the CQ resolution requirement imposed by the reference target performance-achieving

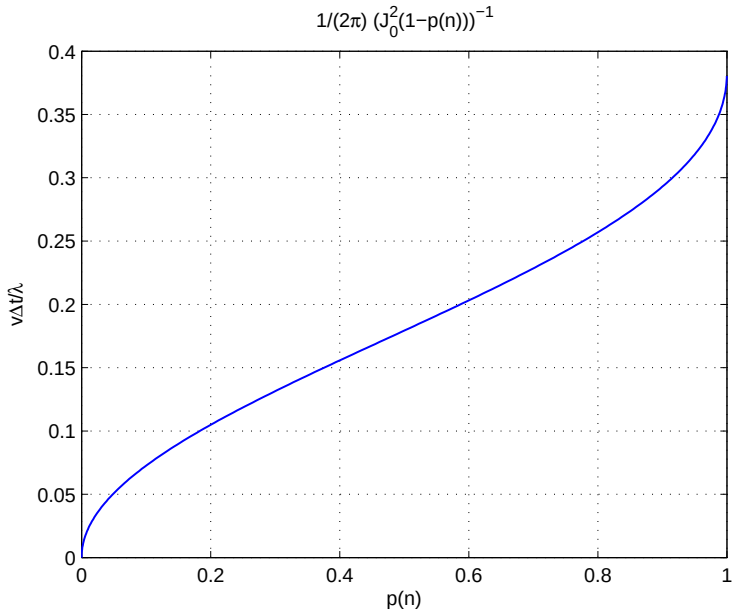


Figure C.2: Inverse CEAC function for $0 \leq p \leq 1$

feedback period τ_c. What happens if the mobile speed v_{MS} is set for 5 m/s? Compare this result with the case $n_f = 1$.

Solution We follow the steps explained above.

1. Calculate the CEAC value for the target performance-achieving reference case ($\tau_c = 10\,T_s$ for $n_c = 3$):
$$p(\Delta t(n_c)) = 1 - J_0^2\left(2\pi \frac{10 \cdot 10T_s}{\frac{c}{f_c}}\right) = 0.08$$

2. Calculate the CEAC value for the case of our interest, $n_f = 2$:
$$p(\Delta t(5)) = \frac{p(\Delta t(3))}{2^{5-3}} = 0.02$$

3. Calculate the normalized delay value $v_{\text{MS}}\Delta t(5)/\lambda$ to get the fine feedback period, τ_f:
$$\frac{v\Delta t(5)}{\lambda} = \frac{1}{2\pi}\left(J_0^2\left(1 - p(\Delta t(5))\right)\right)^{-1} = \frac{1}{2\pi}\left(J_0^2(1 - 0.02)\right)^{-1} = 0.032$$

   The graph of the inverse CEAC function is shown in Fig.C.3. The normalized delay for the reference case $n_c = 3$ is 0.065, which is about two times of that for the fine feedback case. Therefore, we can set $\tau_f \approx \frac{\tau_c}{2} = 5\,T_s$, provided that other parameters, v_{MS} and λ,

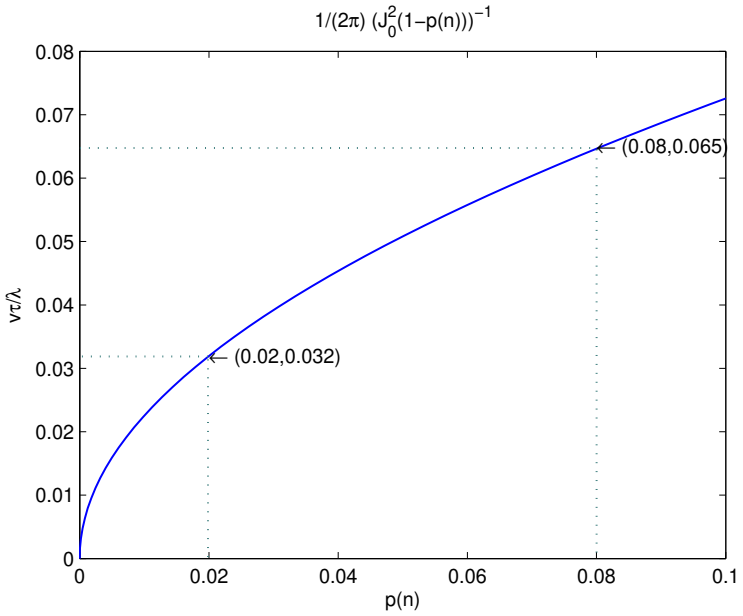


Figure C.3: Inverse CEAC function (in detail)

remain unchanged. We can get $\tau_c = 20\ T_s$ and $\tau_f = 10\ T_s$ for $v_{\mathrm{MS}} = 5$ m/s.

What happens if 1 bit is allocated to the fine feedback? With $n_f = 1$ we get $\tau_f = 7.27\ T_s$, which doesn't allow τ_c to be an integer multiple of τ_f. This makes the coordination of operations of the combined codebook tricky. With the combination of feedback bits being denoted as (n_c+n_fb), (3+2b) case is more efficient than (3+1b) case in terms of synchronization of the feedback periods, which is directly associated with the efficient usage of the feedback resources.

C.3 Numerical Results

The performance of the combined codebook is shown in Fig. 4.2. The 3+2 bit combined codebook with the feedback period pair ($[\tau_c]$,$[\tau_f]$)=(10,5), (3+2bCQ$_1$), is less than 0.1 dB away from the performance of the 5bCQ case. The following cases have also been tested - the 3+2bCQ case with various other feedback periods, i.e., (10,10) and (20,5), and each case is denoted by the subscript 2 and 3, respectively. None of these cases outperforms the case (10,5), and the performance degradations are 0.7 dB and 0.8 dB for the longer fine feedback period case and the longer coarse feedback period case, respectively. We should note that the found feedback period pair guarantees the target

performance. In this case, ($[\tau_c]$,$[\tau_f]$)=(10,5) is the best feedback period pair, and any other case with a longer period results in a performance degradation.

Additional simulation results for the case in which 3 bits are allocated for the fine feedback are shown in Table C.1. The CEAC value for $n_f = 3$ ($n_{cb} = 6$) is $p(\Delta t(6)) = p(\Delta t(3))/2^{6-3} = 0.01$, and the corresponding normalized delay value $v_{\mathrm{MS}}\tau_f/\lambda$ is 0.0225 which leads to the fine feedback period $\tau_f \approx 3.64\ T_s$. The τ_f is conservatively set as 3 T_s. For simplicity, τ_c is set as 9 T_s instead of 10 T_s, since it is an integer multiple of 3 T_s. Simulation results show that the (9,3) case (3+3bCQ$_1$) is only 0.3 dB away from the reference case (6bCQ) in terms of 50 % outage SINR. (9,1) case (3+3bCQ$_2$), which triples the fine feedback frequency, cannot guarantee a better performance. Any attempt to reduce either the fine feedback frequency (3+3bCQ$_3$) or the coarse feedback frequency (3+3bCQ$_4$) brings about performance degradation.

Table C.1: Combined codebook performance and feedback overhead for (n_c+n_fb)=(3+3b) case, 2 BSs - 2 MSs with v_{MS}=10 m/s

Profile	50 % outage SINR [dB]	Feedback overhead [bits/sym/subcarrier]
6bCQ	20.7	6.00
3+3bCQ$_1$ (9,3)	20.4	1.33
3+3bCQ$_2$ (9,1)	20.4	3.33
3+3bCQ$_3$ (10,5)	20.0	0.90
3+3bCQ$_4$ (18,3)	19.8	1.17

In this section, the feedback period decision method is introduced for given number of feedback bits and the mobile speed, based on the reference case. The simulation work is required only once at the beginning to search for the target performance-achieving point, and the rest of the process relies on the analytic approach. The method herein provides a tool which can be used for design of the combined codebook, resource allocation for the coarse/fine codebook, and decision of the feedback periods.

APPENDIX D

Computational Complexity Comparison of the D-Givens and the aCB based Feedback Schemes

This section introduces a computational complexity comparison of the differential Givens (D-Givens) and the adaptive codebook (aCB) based feedback schemes. The following notations are used in this section.

$d_c(\mathbf{T}_i, \mathbf{T}_j)$	chordal distance between $\mathbf{T}_i$, $\mathbf{T}_j$ which have orthonormal columns, $\frac{1}{\sqrt{2}}\|\mathbf{T}_i\mathbf{T}_i^{\mathcal{H}} - \mathbf{T}_j\mathbf{T}_j^{\mathcal{H}}\|_F$
N_{BS}	number of cooperating BSs in cooperative area (CA)
N_t	number of transmit antennas at BS
N_r	number of receive antennas at MS
N_{tt}	total number of transmit antennas in CA, $N_{tt} = N_{\mathrm{BS}}N_t$
n_s	number of data streams for each MS, $n_s \leq N_r$
$\mathbf{H}_j$	$\mathbb{C}^{N_{tt}\times N_r}$ transpose of the channel matrix for MS j
$\mathbf{U}_j^{(\mathrm{S})}$	$\mathbb{C}^{N_{tt}\times N_r}$ basis for the signal subspace of the channel matrix for MS j $\mathbf{H}_j$
$\mathbf{C}_i$	$\mathbb{C}^{N_{tt}\times N_r}$ code of index i
n_{CB}	codebook size in bits
n_a	adaptive codebook search space size in bits, i.e., number of the required bits to encode one node
K	cluster size in number of subcarriers
K'	interpolation area size in number of subcarriers
N	total number of subcarriers
N_{sym}	feedback period in number of OFDM symbols

Note that with respect to LTE, 6 resource block assignment (N6 BW[1]) with a feedback period of

[1]Each resource block occupies 12 subcarriers, and N6 BW corresponds to 6 resource blocks.

1 subframe, i.e., $N = 72, N_{sym} = 14$, has been assumed when calculating the complexity of the proposed methods in the following subsections.

D.1 Complexity of Givens Rotation based Channel Decomposition Method with ADM

The computational complexity of the D-Givens (Givens decomposition + ADM) method at the MS is summarized in Table D.1[2]. In Step 1, the Givens decomposition of the channel matrix is performed per node[3], which is followed by the ADM operation per each Givens parameter (Step 2). This D-Givens operation (Step 1 and Step 2) per node should be repeated for the number of nodes per N6 BW ($N = 72$) in Step 3. In Step 4, a feedback period adjustment constant is multiplied for a fair comparison of the simulation results with the aCB based method, which takes into account the fact that D-Givens simulation results have been acquired with a parameter tracking interval of 10 symbols, unlike the case of the aCB simulations (14 symbols).

Table D.1: Computational complexity of the D-Givens method at MS

Steps	Note	Number of real multiplications
1	Givens decomposition of $\mathbf{H}_j$ per node	$S_1 = \sum_{i=1}^{n_s} 12(N_{tt} - i)(n_s - i)$
2	No. of ADM op. per node	$S_2 = (2N_{tt} - 1)n_s - n_s^2$
3	No. of D-Givens op. per N6 BW	$S_3 = 72$
4	feedback period adjustment constant	$S_4 = \frac{14}{10}$
Total	$(S_1 + S_2)S_3S_4$	$100.8\left\{\sum_{i=1}^{n_s} 12(N_{tt} - i)(n_s - i) + 2n_sN_{tt} - (n_s^2 + n_s)\right\}$

D.2 Complexity of Subspace based Channel Quantization Method with aCB

The computational complexity of the aCB method at the MS is summarized in Table D.2[4]. In Step 1, the basis for the signal subspace of the channel matrix is acquired per node. This can be done by a Givens decomposition of the channel matrix. In Step 2 and 3, the chordal distance is calculated for the basis with respect to every code in the adaptive codebook search space of size 2^{n_a}. In principle the whole adaptive codebook of size $2^{n_{\mathrm{CB}}}$ shall be searched, but the search space needs to be restricted for practical applications. In Chapter V, the average number of bits to encode one note is about 2.2 bits in case of $v_{\mathrm{MS}} = 3$ m/s, $N = 72$, $K' = 12$, $N_{tt} = 2$, $n_s = N_r = 1$, and

[2]Division operation, which is very costly in implementation, is not required.
[3]Complexity of the Givens decomposition can be found in (*Ansari et al.*, 2004).
[4]Division operation is not required.

$n_{\mathrm{CB}} = 10$, which can be used as a guide for selection of n_a. This aCB operation (Step 1, 2 and 3) per node should be repeated for the number of nodes per N6 BW ($\frac{2N}{K'} = \frac{2 \cdot 72}{12} = 12$) in Step 4.

Table D.2: Computational complexity of the aCB method at MS

Steps	Note	Number of real multiplications
1	Finding $\mathbf{U}_j^{(\mathrm{S})}$ per node	$S_1 = \sum_{i=1}^{n_s} 12(N_{tt} - i)(n_s - i)$
2	$d_c(\mathbf{U}_j^{(\mathrm{S})}, \mathbf{C}_i)$ per code	$S_2 = \left\{2(N_{tt}^2 - N_{tt})4n_s + 4N_{tt}\right\} + 2(N_{tt}^2 - N_{tt}) + N_{tt}$
3	No. of d_c op. per node	$S_3 = 2^{n_a}$
4	No. of aCB op. per N6 BW	$S_4 = 12$
Total	$(S_1 + S_2 S_3)S_4$	$12\left\{\sum_{i=1}^{n_s} 12(N_{tt} - i)(n_s - i) + 2^{n_a}\left\{(8n_s + 2)N_{tt}^2 + (3 - 8n_s)N_{tt}\right\}\right\}$

D.3 Complexity Comparison

Table D.3 and Fig. D.1 show the computational complexity comparison of the D-Givens and the aCB based methods for the various cases of (N_{tt}, n_s)[5].

Table D.3: Computational complexity of the D-Givens and the aCB methods at MS (real multiplications)

(N_{tt}, n_s)	D-Givens	aCB	Ratio (aCB/D-Givens)
(2,1)	202	1654	8.21
(3,1)	403	4135	10.26
(4,1)	605	7719	12.76
(2,2)	1411	2680	1.90
(3,2)	3024	7070	2.34
(4,2)	4637	13444	2.90
(6,2)	7862	32148	4.09
(8,2)	11088	58792	5.30

D.4 Conclusion

For the cases of our interest which can be configured by the total number of transmit antenna(s) and the number of data streams per MS, the aCB based feedback scheme commands higher computational complexity at the MS side, when being compared with the D-Givens based method.

[5]In case of $n_s > 1$, the straightforward extension of the proposed method has been assumed without considering feedback of the channel magnitude information. In case of the aCB method, $n_a = 2.2$ is assumed for simplicity, but it should be noted that the selection of n_a can be a function of the following factors, e.g., the channel condition, N, K', N_{tt}, and n_s.

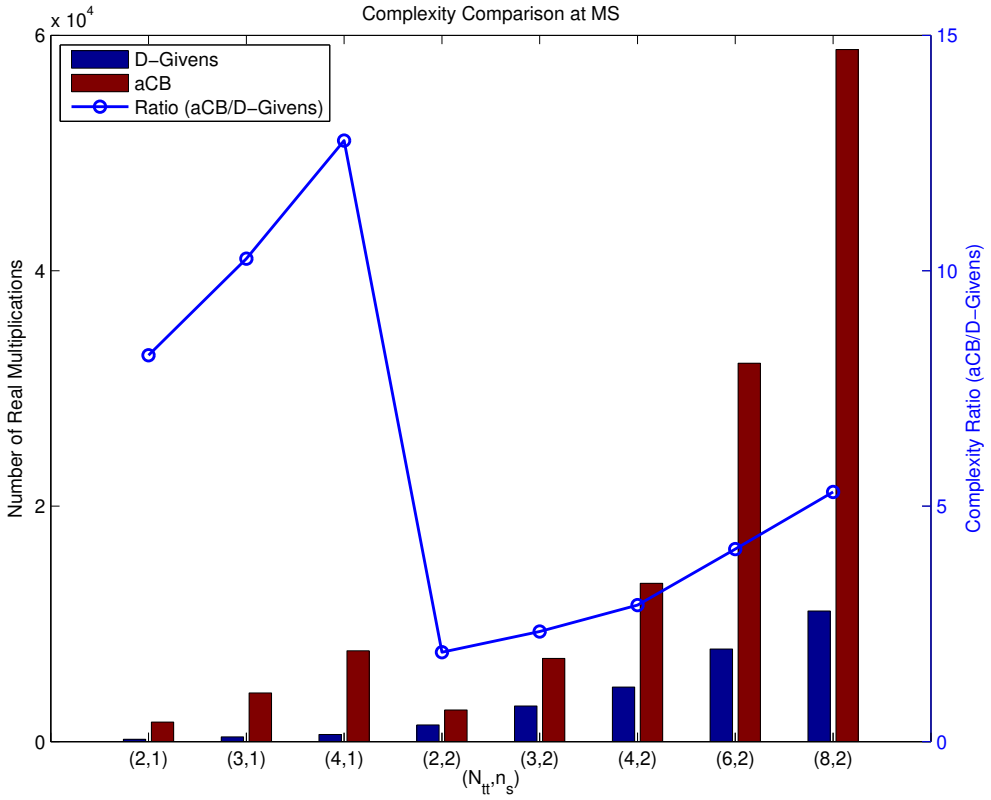


Figure D.1: Complexity Comparison at MS for $N = 72$ and $N_{sym} = 14$ (N:total number of subcarriers, N_{sym}:feedback period in number of OFDM symbols, N_{tt}:total number of transmit antennas in CA, n_s:number of data streams for each MS)

BIBLIOGRAPHY

BIBLIOGRAPHY

3GPP (2007), Spacial channel model for Multiple Input Multiple Output (MIMO) simulations, *TR 25.996*, 3rd Generation Partnership Project (3GPP).

3GPP (2008), Technical specification group radio access network; requirements for evolved UTRA (E-UTRA) and evolved UTRAN (E-UTRAN) (release 8), *TR 25.913*, 3rd Generation Partnership Project (3GPP).

3GPP (2009), Technical Specification Group Radio Access Network; Further Advancements for E-UTRA Physical Layer Aspects (Release 9), *TS 36.814*, 3rd Generation Partnership Project (3GPP).

3GPP (2012), Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 10), *TS 36.211*, 3rd Generation Partnership Project (3GPP).

3GPP (2017), Study on New Radio (NR) Access Technology; Physical Layer Aspects (Release 14), *TR 38.802 v2.0.0*, 3rd Generation Partnership Project (3GPP).

Ahmed, R., K. Jayasinghe, and T. Wild (2019a), Comparison of Explicit CSI Feedback Schemes for 5G New Radio, in *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pp. 1–5, doi:10.1109/VTCSpring.2019.8746469.

Ahmed, R., F. Tosato, and M. Maso (2019b), Overhead Reduction of NR type II CSI for NR Release 16, in *WSA 2019; 23rd International ITG Workshop on Smart Antennas*, pp. 1–5.

Akoum, S., and R. W. Heath, Jr. (2009), Limited feedback for temporally correlated MIMO channels with other cell interference, *CoRR*, *abs/0912.2378*, informal publication.

Aktas, D., M. N. Bacha, J. S. Evans, and S. V. Hanly (2006), Scaling results on the sum capacity of cellular networks with MIMO links, *IEEE Transactions on Information Theory*, *52*(7), 3264–3274.

Alcatel Lucent (2009), R1-083110, MIMO and CQI-techniques for LTE-advanced, *Tech. rep.*, 3GPP TSG RAN WG1 54bis.

Aldajani, M., and A. Sayed (2001), A stable adaptive structure for delta modulation with improved performance, in *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'01)*, vol. 4, pp. 2621–2624, Salt Lake City, UT, USA.

Alexiou, A., and M. Haardt (2004), Smart antenna technologies for future wireless systems: trends and challenges, *IEEE Communications Magazine*, *42*(9), 90–97, doi: 10.1109/MCOM.2004.1336725.

Andrews, J. G., A. Ghosh, and R. Muhamed (2007), *Fundamentals of WiMAX: understanding broadband wireless networking*, Pearson Education, New Jersey, USA.

Ansari, M., et al. (2004), Unified MIMO Pre-Coding based on Givens Rotation, *TR IEEE C802.16e-04/516*, IEEE 802.16 Broadband Wireless Access Working Group.

ANSI/IEEE (2003), IEEE 802.11 Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, *IEEE Standard 802.11*.

Astely, D., E. Dahlman, P. K. Frenger, R. Ludwig, M. Meyer, S. Parkvall, P. Skillermark, and N. Wiberg (2006), A future radio-access framework, *IEEE Journal on Selected Areas in Communications*, *24*(3), 693–706.

Baum, D., G. Del Galdo, J. Salo, P. Kyösti, T. Rautiainen, M. Milojevic, and J. Hansen (2005), An interim channel model for beyond-3G systems: extending the 3GPP spatial channel model (SCM), in *Proceedings IEEE 61st Vehicular Technology Conference (VTC 2005 Spring)*, Stockholm, Sweden.

Bazzi, S., G. Dietl, and W. Utschick (2014), Subspace precoding with limited feedback for the massive MIMO interference channel, in *Proceedings IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM 2014)*, Coruña, Spain.

Biglieri, E., J. Proakis, and S. Shamai (1998), Fading channels: information-theoretic and communications aspects, *IEEE Transactions on Information Theory*, *44*(6), 2619–2692, doi: 10.1109/18.720551.

Bingham, J. A. C. (1990), Multicarrier modulation for data transmission: An idea whose time has come, *IEEE Communications Magazine*, *28*(5), 5–14.

Buzuverov, A., H. Al-Shatri, and A. Klein (2015), On the degrees of freedom of the three-user MIMO interference channel with delayed CSIT, in *Proceedings IEEE Information Theory Workshop (ITW 2015)*, pp. 239–243, Jeju Island, Korea.

Buzuverov, A., H. Al-Shatri, and A. Klein (2016), On the achievable degrees of freedom of the MIMO X-channel with delayed CSIT, in *Proceedings IEEE Information Theory Workshop (ITW 2016)*, pp. 464–468, Cambridge, United Kingdom.

Caire, G., and S. Shamai (2003), On the achievable throughput of a multiantenna Gaussian broadcast channel, *IEEE Transactions on Information Theory*, *49*(7), 1691–1706.

Castañeda, M., A. Chathoth, and J. A. Nossek (2009), Achievable rates in the SIMO-uplink/MISO downlink of an FDD system with imperfect CSI, in *Proceedings IEEE 6th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 121–125, Siena, Italy.

Chang, R. W. (1966), Synthesis of Band-Limited orthogonal signals for multichannel data transmission, *Bell Systems Technical Journal*, *45*, 1775–1796.

Cheng, Y., S. Li, J. Zhang, F. Roemer, B. Song, M. Haardt, Y. Zhou, and M. Dong (2014), An efficient and flexible transmission strategy for the multi-carrier multi-user MIMO downlink, *IEEE Transactions on Vehicular Technology*, *63*, 628–642.

Choi, J., and R. W. Heath, Jr. (2005), Interpolation based transmit beamforming for MIMO-OFDM with limited feedback, *IEEE Transactions on Signal Processing*, *53*(11), 4125–4135.

Choi, J., D. J. Love, and T. Kim (2015), Trellis-extended codebooks and successive phase adjustment: A path from LTE-Advanced to FDD massive MIMO systems, *IEEE Transactions on Wireless Communications*, *14*(4), 2007–2016.

Choi, W., and J. G. Andrews (2007), Downlink performance and capacity of distributed antenna systems in a multicell environment, *IEEE Transactions on Wireless Communications*, *6*(1), 69–73.

Chong, J. H., C. K. Ng, N. K. Noordin, and B. M. Ali (2014), A low computational complexity V-BLAST/STBC detection mechanism in MIMO system, *Human-centric Computing and Information Sciences*, *4*(1), 1–28.

Dahlman, E., H. Ekström, A. Furuskar, Y. Jading, J. Karlsson, M. Lundevall, and S. Parkvall (2006), The 3G Long-Term Evolution - radio interface concepts and performance evaluation, in *Proceedings 63rd IEEE Vehicular Technology Conference (VTC 2006 Spring)*, pp. 137–141.

Dahlman, E., S. Parkvall, and J. Skold (2016), *4G, LTE-Advanced Pro and The Road to 5G*, Academic Press.

Daniels, R. C., K. Mandke, K. T. Truong, S. Nettles, and R. W. Heath, Jr. (2008), Throughput/delay measurements of limited feedback beamforming in indoor wireless networks, in *Proceedings IEEE Global Communications Conference (Globecomm 2008)*, pp. 4593–4598, New Orleans, LA, USA.

Dawod, N. H., I. D. Marsland, and R. H. M. Hafez (2006), Improved transmit steering for MIMO-OFDM downlinks with distributed base station antenna arrays, *IEEE Journal on Selected Areas in Communications*, *24*(3), 419–426.

Dharamdial, N., and R. Adve (2005), Efficient feedback for precoder design in single- and multi-user MIMO systems, in *Proceedings 39th Conference on Information Sciences and Systems (CISS'05)*, Baltimore, MD, USA.

Dunn, P. F. (2010), *Measurement and Data Analysis for Engineering and Science*, CRC Press, London, UK.

Ericsson Mobility Report (2016), Ericsson Mobility Report: on the pulse of the networked society, *Tech. Rep. EAB-16:006659 Uen, Revision A*, Ericsson AB.

Ericsson Mobility Report (2020), Ericsson Mobility Report, *Tech. Rep. EAB-20:004467 Uen*, Ericsson AB.

Foschini, G. J., and M. J. Gans (1998), On limits of wireless communications in a fading environment when using multiple antennas, *Wireless Personal Communications*, *6*(3), 311–335, doi: 10.1023/A:1008889222784.

Foschini, G. J., D. Chizhik, M. J. Gans, C. Papadias, and R. A. Valenzuela (2003), Analysis and performance of some basic space-time architectures, *IEEE Journal on Selected Areas in Communications*, *21*(3), 303–320.

Fuchs, M., G. D. Galdo, and M. Haardt (2007), Low-complexity space-time-frequency scheduling for MIMO systems with SDMA, *IEEE Transactions on Vehicular Technology*, *56*, 2775–2784.

Furht, B., and S. A. Ahson (2016), *Long Term Evolution: 3GPP LTE radio and cellular technology*, CRC Press.

Gersho, A., and R. M. Gray (1992), *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Norwell, MA, USA.

Godara, L. C. (1997a), Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations, *Proceedings of the IEEE*, *85*(8), 1195–1245.

Godara, L. C. (1997b), Applications of antenna arrays to mobile communications. I. Performance improvement, feasibility, and system considerations, *Proceedings of the IEEE*, *85*(7), 1031–1060.

Goldsmith, A. (2005), *Wireless Communications*, Cambridge University Press.

Goldsmith, A., S. A. Jafar, N. Jindal, and S. Vishwanath (2003), Capacity limits of MIMO channels, *IEEE Journal on Selected Areas in Communications*, *21*(5), 684–702.

Golub, G. H., and C. F. Van Loan (2012), *Matrix computations*, vol. 3, The Johns Hopkins University Press.

Grünheid, R., E. Bolinth, and H. Rohling (2001), A blockwise loading algorithm for the adaptive modulation technique in OFDM systems, in *Proceedings 54th IEEE Vehicular Technology Conference (VTC 2001 Fall)*, vol. 2, pp. 948–951, Atlantic City, NJ, USA.

Haardt, M., and J. A. Nossek (1998), Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems, *IEEE Transactions on signal processing*, *46*(1), 161–169.

Haardt, M., F. Roemer, and G. D. Galdo (2008), Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems, *IEEE Transactions on signal processing*, *56*(7), 3198–3213.

Hahn, J., M. Meurer, P. W. Baier, and W. Zirwas (2006), Spread-spectrum based low-cost provision of downlink channel state information to the access points of FDD OFDM mobile radio systems, in *Proceedings IEEE 9th International Symposium on Spread Spectrum Techniques and Applications (ISSSTA 2006)*, pp. 10–15, Manaus, Brazil.

Han, Y., W. Shin, and J. Lee (2014), Projection based feedback compression for FDD massive MIMO systems, in *Proceedings IEEE Globecom Workshops (GC Wkshps 2014)*, pp. 364–369, Austin, TX, USA.

Hanzo, L., C. H. Wong, and M. S. Yee (2002), *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalised and Space-Time Coded TDMA, CDMA and OFDM Systems*, John Wiley & Sons.

Harashima, H., and H. Miyakawa (1972), Matched-transmission technique for channels with intersymbol interference, *IEEE Transactions on Communications*, *20*(4), 774–780.

Hassibi, B., and H. Vikalo (2005), On the sphere-decoding algorithm I. Expected complexity, *IEEE Transactions on Signal Processing*, pp. 2806–2818.

Hata, M. (1980), Empirical formula for propagation loss in land mobile radio services, *IEEE Transactions on Vehicular Technology*, *29*(3), 317–325.

Haustein, T., J. H. Kim, and W. Zirwas (2009), Method of and device for transmission of data in a communication network, Patent Application, EP 2009/2114029 A1.

Haykin, S. (2009), *Communication Systems*, 5th ed., Wiley Publishing, New York, NY, USA.

Heath, R. W., Jr., T. Wu, and A. C. K. Soong (2009), Progressive refinement of beamforming vectors for high-resolution limited feedback, *EURASIP Journal on Advances in Signal Processing*, *2009*, doi:http://dx.doi.org/10.1155/2009/463823.

Hindy, A., U. Mittal, and T. Brown (2020), CSI Feedback Overhead Reduction for 5G Massive MIMO Systems, in *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0116–0120, doi:10.1109/CCWC47524.2020.9031236.

Hochwald, B. M., T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. L. Urbanke (2000), Systematic design of unitary space-time constellations, *IEEE Transactions on Information Theory*, *46*(6), 1962–1973.

Hoeher, P., S. Kaiser, and P. Robertson (1997), Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering, in *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97)*, vol. 3, pp. 1845–1848, Munich, Germany.

Holma, H., and A. Toskala (2006), *HSDPA/HSUPA for UMTS: High Speed Radio Access for Mobile Communications*, John Wiley & Sons.

Hsu, C.-Y., R.-J. Chen, and F.-S. Tseng (2013), A bit-adaptive PMI feedback mechanism, in *Proceedings 78th IEEE Vehicular Technology Conference (VTC 2013 Fall)*, Las Vegas, NV, USA.

Hu, H., Y. Zhang, and J. Luo (2007), *Distributed Antenna Systems: Open Architecture for Future Wireless Communications*, Auerbach Publications, Boston, MA, USA.

Huang, K., B. Mondal, R. W. Heath, Jr., and J. G. Andrews (2006a), Effect of feedback delay on multi-antenna limited feedback for temporally-correlated channels, in *Proceedings IEEE Global Communications Conference (Globecomm 2006)*, San Francisco, CA, USA.

Huang, K., B. Mondal, R. W. Heath, Jr., and J. G. Andrews (2006b), Limited feedback for temporally-correlated channels: Feedback rate and delay, *CoRR*, *abs/cs/0606022*, informal publication.

Huang, K., R. W. Heath, and J. G. Andrews (2009), Limited feedback beamforming over temporally-correlated channels, *IEEE Transactions on Signal Processing*, *57*(5), 1959–1975, doi: 10.1109/TSP.2009.2014272.

Huawei (2010), R1-100252, Adaptive codebook designs and simulation results, *Tech. rep.*, 3GPP TSG RAN WG1 #59.

IEEE (2004), IEEE 802.16 Standard for local and metropolitan area networks part 16: Air interface for fixed broadband wireless access systems, *IEEE Standard 802.16*.

IEEE (2005), IEEE 802.16e Mobile WirelessMAN Standard: Physical and medium access control layers for combined fixed and mobile operation in licensed bands, *IEEE Standard 802.16e*.

Isukapalli, Y., and B. D. Rao (2007), Finite rate feedback for spatially and temporally correlated MISO channels in the presence of estimation errors and feedback delay, in *Proceedings IEEE Global Communications Conference (Globecomm 2007)*, pp. 2791–2795, Washington, DC, USA.

Jafar, S. A., S. Vishwanath, and A. Goldsmith (2001), Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback, in *Proceedings IEEE International Conference on Communications (ICC 2001)*, vol. 7, pp. 2266–2270, Helsinki, Finland.

Jakes, W. C., Jr. (1974), *Microwave Mobile Communications*, Wiley, New York, NY, USA.

Ji, H., Y. Kim, J. Lee, E. Onggosanusi, Y. Nam, J. Zhang, B. Lee, and B. Shim (2017), Overview of Full-Dimension MIMO in LTE-Advanced Pro, *IEEE Communications Magazine*, (accepted for publication).

Jindal, N. (2006), MIMO broadcast channels with finite rate feedback, *IEEE Transactions on Information Theory*, *52*(11), 5045–5059.

Joham, M., W. Utschick, and J. A. Nossek (2005), Linear transmit processing in MIMO communications systems, *IEEE Transactions on Signal Processing*, *53*(8-1), 2700–2712.

Jungnickel, V., T. Wirth, M. Schellmann, T. Haustein, and W. Zirwas (2008), Synchronization of cooperative base stations, in *Proceedings IEEE International Symposium on Wireless Communication Systems (ISWCS'08)*, Reijavik, Island.

Juntti, M. J., K. Hooli, K. Kiiskila, and J. Ylioinas (2007), Space-time equalizers for MIMO high speed WCDMA downlinks, *IEEE Transactions on Wireless Communications*, *6*(7), 2582–2592.

Karakayali, M. K., G. J. Foschini, and R. A. Valenzuela (2006), Network coordination for spectrally efficient communications in cellular systems, *IEEE Wireless Communications Magazine*, *3*(14), 56–61.

Kim, J. H. (2016), Communication network component and method for requesting channel information, Patent Application, WO 2018063519 (A1), filed September 28, 2016, and issued April 5, 2018.

Kim, J. H., and W. Zirwas (2009), Method and device for data processing and communication system with data processing device, Patent Application, US 2009/0122851 A1, filed May 14, 2009, and issued February 19, 2013.

Kim, J. H., W. Zirwas, and M. Haardt (2007), Efficient feedback via subspace based channel quantization for downlink transmission in distributed cooperative antenna systems, in *Proceedings International ITG/IEEE Workshop on Smart Antennas (WSA 2007)*, Vienna, Austria.

Kim, J. H., E. Schulz, and W. Zirwas (2008a), Adaptive codebook with compressed feedback, European Patent, EP 2159978 (A1).

Kim, J. H., E. Schulz, and W. Zirwas (2008b), Method for providing channel information in a radio communications system and mobile station thereof, Patent Application, EP 2008/1962539 A1.

Kim, J. H., W. Zirwas, and M. Haardt (2008c), Efficient feedback via subspace-based channel quantization for distributed cooperative antenna systems with temporally correlated channels, *EURASIP Journal on Advances in Signal Processing*, *2008*, 1–13, doi:http://dx.doi.org/10.1155/2008/84729.

Kim, J. H., W. Zirwas, and M. Haardt (2008d), Feedback reduction methods for distributed cooperative antenna systems with temporally-correlated channels, in *Proceedings 16th European Signal Processing Conference (EUSIPCO 2008)*, Lausanne, Switzerland.

Kim, J. H., W. Zirwas, and M. Haardt (2008e), Adaptive codebooks for efficient feedback reduction in cooperative antenna systems, in *Proceedings IEEE Global Communications Conference (Globecomm 2008)*, pp. 3380–3384, New Orleans, LA, USA.

Kim, J. H., E. Schulz, and W. Zirwas (2009), Method for data transmission in a radio communication system as well as radio station and radio communications system, Patent Application, US 2009/0221315 A1.

Kim, J. H., E. Schulz, and W. Zirwas (2010), Adaptive codebook with compressed feedback, European Patent, EP 2159978 (A1).

Kobayashi, K., T. Ohtsuki, and T. Kaneko (2008), MIMO systems in the presence of feedback delay, *IEICE Transactions*, *91-B*(3), 829–836.

Kotterman, W. A. T., C. Schirmer, M. Landmann, and G. Del Galdo (2014), On arranging dual-polarised antennas in 3D wave field synthesis, in *Proceedings IEEE 8th European Conference on Antennas and Propagation (EUCAP 2014)*, pp. 3406–3410, The Hague, Netherlands.

Kotterman, W. A. T., C. Schirmer, M. H. Landmann, and G. Del Galdo (2017), New challenges in over-the-air testing, in *Proceedings IEEE 11th European Conference on Antennas and Propagation (EUCAP 2017)*, pp. 3676–3678, Paris, France.

Kuehne, A., and A. Klein (2008), Throughput analysis of multi-user OFDMA-systems using imperfect CQI feedback and diversity techniques, *IEEE Journal on Selected Areas in Communications*, *26*(8), 1440–1450.

Kuehne, A., and A. Klein (2012), Hybrid adaptive/non-adaptive multi-user OFDMA systems in the presence of user-specific imperfect channel knowledge, *EURASIP Journal on Wireless Communications and Networking*, *2012*(1), 1–16.

Lee, B., J. Choi, J.-Y. Seol, D. J. Love, and B. Shim (2015), Antenna grouping based feedback compression for FDD-based massive MIMO systems, *IEEE Transactions on Communications*, *63*(9), 3261–74.

Lim, C., T. Yoo, B. Clerckx, B. Lee, and B. Shim (2013), Recent trend of multiuser MIMO in LTE-Advanced, *IEEE Communications Magazine*, *51*(3), 127–135.

Linde, Y., A. Buzo, and R. M. Gray (1980), An algorithm for vector quantizer design, *IEEE Transactions on Communications*, *COM-28*(1), 84–95.

Love, D., R. W. Heath, Jr., W. Santipach, and M. Honig (2004), What is the value of limited feedback for MIMO channels?, *IEEE Communications Magazine*, *42*(10), 54–59.

Love, D. J., and R. W. Heath, Jr. (2004), Grassmannian beamforming on correlated MIMO channels, in *Proceedings IEEE Global Communications Conference (Globecomm 2004)*, vol. 1, pp. 106–110, Dallas, TX, USA.

Love, D. J., and R. W. Heath, Jr. (2005), Limited feedback unitary precoding for spatial multiplexing systems, *IEEE Transactions on Information Theory*, *51*(8), 2967–2976.

Love, D. J., R. W. Heath, Jr., and T. Strohmer (2003), Grassmannian beamforming for multiple-input multiple-output wireless systems, *IEEE Transactions on Information Theory*, *49*(10), 2735–2747.

Love, D. J., R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews (2008), An overview of limited feedback in wireless communication systems, *IEEE Journal on Selected Areas in Communications*, *26*(8), 1341–1365.

M. Soleymani, D., A. Puschmann, E. Roth-Mandutz, J. Mueckenheim, and A. Mitschele-Thiel (2016), A hierarchical radio resource management scheme for next generation cellular networks, in *Proceedings IEEE WCNC Workshop on Device to Device communications for 5G Networks*, Doha, Qatar.

Manolakis, K., C. Oberli, and V. Jungnickel (2010), Synchronization Requirements for OFDM-based Cellular Networks with Coordinated Base Stations: Preliminary Results, in *Proceedings 15th International OFDM-Workshop 2010 (InOWo'10)*, Hamburg, Germany.

Marqués, A. G., F. F. Digham, and G. B. Giannakis (2006), Optimizing power efficiency of OFDM using quantized channel state information, *IEEE Journal on Selected Areas in Communications*, *24*(8), 1581–1592.

Marzetta, T. L. (2010), Noncooperative cellular wireless with unlimited numbers of base station antennas, *IEEE Transactions on Wireless Communications*, *9*(11), 3590–3600.

Marzetta, T. L., and B. M. Hochwald (2006), Fast transfer of channel state information in wireless systems, *IEEE Transactions on Signal Processing*, *54*(10), 1268–1278.

Mohr, W. (2007), The WINNER (Wireless World Initiative New Radio) project - development of a radio interface for systems beyond 3G, *International Journal of Wireless Information Networks*, *14*(2), 67–78.

Mondal, B., S. Dutta, and R. W. Heath (2007), Quantization on the Grassmann Manifold, *IEEE Transactions on Signal Processing*, *55*(8), 4208–4216, doi:10.1109/TSP.2007.896112.

Mondal, B., V. Sergeev, A. Sengupta, G. Ermolaev, A. Davydov, E. Kwon, S. Han, and A. Papathanassiou (2019), MU-MIMO and CSI Feedback Performance of NR/LTE, in *2019 53rd Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, doi: 10.1109/CISS.2019.8692922.

Motorola (2010a), R1-100190, Low-Overhead Feedback of Spatial Covariance Matrix, *Tech. rep.*, 3GPP TSG RAN WG1 #59.

Motorola (2010b), R1-100189, Importance of Higher Order MU-MIMO in LTE-A, *Tech. rep.*, 3GPP TSG RAN WG1 #59.

Motorola (2010c), R1-101129, On Extensions to Rel-8 PMI Feedback, *Tech. rep.*, 3GPP TSG RAN WG1 #60.

Mueckenheim, J., A. Puschmann, D. M. Soleymani, E. Roth-Mandutz, A. M. Waswa, and A. Mitschele-Thiel (2016), On D2D-communication for resource efficient data transmission of delay-critical services, *ITG-Fachbericht-Mobilkommunikation–Technologien und Anwendungen.*

Mukkavilli, K. K., A. Sabharwal, E. Erkip, and B. Aazhang (2003), On beamforming with finite rate feedback in multiple-antenna systems, *IEEE Transactions on Information Theory*, *49*(10), 2562–2579.

Mwanje, S., G. Decarreau, C. Mannweiler, M. Naseer-ul Islam, and L. C. Schmelz (2016a), Network management automation in 5G: Challenges and opportunities, in *Proceedings IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC 2016)*, pp. 1–6, Valencia, Spain.

Mwanje, S. S., and A. Mitschele-Thiel (2014), Distributed cooperative Q-learning for mobility-sensitive handover optimization in LTE SON, in *Proceedings IEEE Symposium on Computers and Communication (ISCC 2014)*, pp. 1–6, Madeira, Portugal.

Mwanje, S. S., and A. Mitschele-Thiel (2015), Concurrent cooperative games for coordinating SON functions in cognitive cellular networks, in *Proceedings IFIP/IEEE International Symposium on Integrated Network Management (IM 2015)*, pp. 1298–1303, Ottawa, Canada.

Mwanje, S. S., L. C. Schmelz, and A. Mitschele-Thiel (2016b), Cognitive cellular networks: A Q-learning framework for self-organizing networks, *IEEE Transactions on Network and Service Management*, *13*(1), 85–98.

Narula, A., M. J. Lopez, M. D. Trott, and G. W. Wornell (2002), Efficient use of side information in multiple-antenna data transmission over fading channels, *IEEE Journal on Selected Areas in Communications*, *16*(8), 1423–1436, doi:10.1109/49.730451.

Onggosanusi, E., et al. (2018), Modular and High-Resolution Channel State Information and Beam Management for 5G New Radio, *IEEE Communications Magazine*, *56*(3), 48–55, doi: 10.1109/MCOM.2018.1700761.

Palomar, D. P., M. Bengtsson, and B. E. Ottersten (2005), Minimum BER linear transceivers for MIMO channels via primal decomposition, *IEEE Transactions on Signal Processing*, *53*(8-1), 2866–2882.

Parkvall, S., E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wänstedt, and K. C. Zangi (2008), LTE-advanced - evolving LTE towards IMT-advanced, in *Proceedings IEEE Vehicular Technology Conference (VTC 2008 Fall)*, pp. 1–5, Calgary, AB, Canada.

Paulraj, A. J., D. A. Gore, R. U. Nabar, and H. Bolcskei (2004), An overview of MIMO communications - a key to gigabit wireless, *Proceedings of the IEEE*, *92*(2), 198–218, doi: 10.1109/JPROC.2003.821915.

Proakis, J. (2000), *Digital Communications*, 4 ed., McGraw-Hill.

Qualcomm Inc. (2010), R1-100684, Extending the UE feedback for efficient MU-MIMO and CoMP, *Tech. rep.*, 3GPP TSG RAN WG1 #59.

Raleigh, G. G., and J. M. Cioffi (1998), Spatio-temporal coding for wireless communication, *IEEE Transactions on Communications*, *46*(3), 357–366.

Ravindran, N., and N. Jindal (2007), MIMO broadcast channels with block diagonalization and finite rate feedback, in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007)*, vol. 3, pp. III–13, Honolulu, HI, USA.

Ravindran, N., and N. Jindal (2008), Multi-user diversity vs. accurate channel feedback for MIMO broadcast channels, in *Proceedings IEEE International Conference on Communications (ICC'08)*, pp. 3684–3688, Beijing, China.

Redl, S., M. Weber, and M. W. Oliphant (1998), *GSM and Personal Communications Handbook*, Artech House, Norwood, Massachusetts.

Roemer, F., M. Haardt, and G. D. Galdo (2014), Analytical performance assessment of multi-dimensional matrix-and tensor-based ESPRIT-type algorithms, *IEEE Transactions on signal processing*, *62*(10), 2611–2625.

Roh, J. C., and B. D. Rao (2004), An efficient feedback method for MIMO systems with slowly time-varying channels, in *Proceedings IEEE Wireless Communications and Networking Conference (WCNC 2004)*, vol. 2, pp. 760–764, Atlanta, GA, USA.

Roh, J. C., and B. D. Rao (2006a), Transmit beamforming in multiple-antenna systems with finite rate feedback: a VQ-based approach, *IEEE Transactions on Information Theory*, *52*(3), 1101–1112.

Roh, J. C., and B. D. Rao (2006b), Design and analysis of MIMO spatial multiplexing systems with quantized feedback, *IEEE Transactions on Signal Processing*, *54*(8), 2874–2886.

Roh, J. C., and B. D. Rao (2007), Efficient feedback methods for MIMO channels based on parameterization, *IEEE Transactions on Wireless Communications*, *6*(1), 282–292.

Sallakh, U., S. S. Mwanje, and A. Mitschele-Thiel (2014), Multi-parameter Q-learning for downlink inter-cell interference coordination in LTE SON, in *Proceedings IEEE Symposium on Computers and Communication (ISCC 2014)*, pp. 1–6, Madeira, Portugal.

Salo, J., G. Del Galdo, J. Salmi, P. Kyösti, M. Milojevic, D. Laselva, and C. Schneider (2005), MATLAB implementation of the 3GPP Spatial Channel Model (3GPP TR 25.996), *Tech. rep.*, Helsinki University of Technology.

Sampath, H., S. Talwar, J. Tellado, V. Erceg, and A. Paulraj (2002), A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results, *IEEE Communications Magazine*, *40*(9), 143–149.

Samsung Electronics Co. Ltd. (2009), R1-082886, Inter-cell Interference mitigation through limited coordination, *Tech. rep.*, 3GPP TSG RAN WG1 #54bis.

Samsung Electronics Co. Ltd. (2010a), R1-100116, Comparisons and performance evaluation of differential feedback proposals for Rel. 8 PMI enhancements, *Tech. rep.*, 3GPP TSG RAN WG1 #59.

Samsung Electronics Co. Ltd. (2010b), R1-100117, Further discussions on feedback framework for SU/MU-MIMO, *Tech. rep.*, 3GPP TSG RAN WG1 #59.

Scaglione, A., P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath (2002), Optimal designs for space-time linear precoders and decoders, *IEEE Transactions on Signal Processing*, *50*(5), 1051–1064.

Schirmer, C., M. Landmann, W. Kotterman, M. Hein, R. Thomä, G. D. Galdo, and A. Heuberger (2014), 3D Wave-field synthesis for testing of radio devices, in *Proceedings IEEE 8th European Conference on Antennas and Propagation (EUCAP 2014)*, pp. 3394–3398, The Hague, Netherlands.

Schulz, P., et al. (2017), Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture, *IEEE Communications Magazine*, *55*(2), 70–78.

Sharif, M., and B. Hassibi (2005), On the capacity of MIMO broadcast channels with partial side information, *IEEE Transactions on Information Theory*, *51*(2), 506–522.

Sharif, M., and B. Hassibi (2007), A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users, *IEEE Transactions on Communications*, *55*(1), 11–15.

Simon, C., and G. Leus (2008a), Feedback quantization for linear precoded spatial multiplexing, *EURASIP Journal on Advances in Signal Processing*, *2008*, 1–13, doi: http://dx.doi.org/10.1155/2008/683030.

Simon, C., and G. Leus (2008b), Feedback quantization for linear precoded spatial multiplexing, *EURASIP Journal on Advances in Signal Processing*, *2008*, doi: http://dx.doi.org/10.1155/2008/683030.

Sklar, B. (1997), Rayleigh fading channels in mobile digital communication systems part I: Characterization, *IEEE Communications Magazine*, *35*(9), 136–146.

Somekh, O., O. Simeone, Y. Bar-Ness, and A. M. Haimovich (2006), Distributed multi-cell zero-forcing beamforming in cellular downlink channels, in *Proceedings IEEE Global Communications Conference (Globecomm 2006)*, San Francisco, CA, USA.

Song, B., and M. Haardt (2009), Effects of imperfect channel state information on achievable rates of precoded multi-user MIMO broadcast channels with limited feedback, in *Proceedings IEEE International Conference on Communications (ICC'09)*, pp. 1–5, Dresden, Germany.

Song, B., F. Roemer, and M. Haardt (2008), Efficient channel quantization scheme for multi-user MIMO broadcast channels with RBD precoding, in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)*, pp. 2389–2392, Las Vegas, NV, USA.

Spencer, Q. H., C. B. Peel, A. L. Swindlehurst, and M. Haardt (2004a), An introduction to the multi-user MIMO downlink, *IEEE Communications Magazine*, *42*(10), 60–67.

Spencer, Q. H., A. L. Swindlehurst, and M. Haardt (2004b), Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels, *IEEE Transactions on Signal Processing*, *52*, 461–471.

Stankovic, V., and M. Haardt (2004), Multi-user MIMO Downlink Precoding for Users with Multiple Antennas, in *the 12th Meeting of the Wireless World Research Forum (WWRF)*.

Stankovic, V., and M. Haardt (2008), Generalized design of Multi-User MIMO precoding matrices, *IEEE Transactions on Wireless Communications*, *7*(3), 953–961.

Steinwandt, J., F. Roemer, M. Haardt, and G. D. Galdo (2017), Performance analysis of multi-dimensional ESPRIT-type algorithms for arbitrary and strictly non-circular sources with spatial smoothing, *IEEE Transactions on signal processing*, *65*(9), 2262–2276.

Su, Q., and S. Schwartz (2001), Effects of imperfect channel information on adaptive loading gain of OFDM, in *Proceedings 54th IEEE Vehicular Technology Conference (VTC 2001 Fall)*, vol. 1, pp. 475–478.

Telatar, I. E. (1999), Capacity of multi-antenna Gaussian channels, *European Transactions on Telecommunications*, *10*, 585–595.

Thiele, L., T. Wirth, T. Haustein, V. Jungnickel, E. Schulz, and W. Zirwas (2009), A unified feedback scheme for distributed interference management in cellular systems: Benefits and challenges for real-time implementation, in *Proceedings 17th European Signal Processing Conference (EUSIPCO 2009)*, pp. 1489–1493, Glasgow, Scotland.

Thomas, T. A., K. L. Baum, and P. Sartori (2005), Obtaining Channel Knowledge for Closed-loop Multi-stream Broadband MIMO-OFDM Communications Using Direct Channel Feedback, in *Proceedings IEEE Global Communications Conference (Globecomm 2005)*, pp. 3907–3911, St. Louis, USA.

Tomlinson, M. (1971), New automatic equaliser employing modulo arithmetic, *Electronics Letters*, *7*(5), 138–139.

Tse, D., and P. Viswanath (2005), *Fundamentals of Wireless Communication*, Cambridge University Press.

Vikalo, H., and B. Hassibi (2005), On the sphere decoding algorithm II. Generalizations, second-order statistics, and applications to communications, *IEEE Transactions on Signal Processing*, pp. 2819–2834.

Vishwanath, P., and D. N. C. Tse (2003), Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality, *IEEE Transactions on Information Theory*, *49*(8), 1912–1921, doi: 10.1109/TIT.2003.814483.

Vishwanath, S., N. Jindal, and A. J. Goldsmith (2003), Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels, *IEEE Transactions on Information Theory*, *49*(10), 2658–2668.

Visotsky, E., and U. Madhow (2001), Space-time transmit precoding with imperfect feedback, *IEEE Transactions on Information Theory*, *47*(6), 2632–2629.

Vojčić, B. R., and W. M. Jang (1998), Transmitter precoding in synchronous multiuser communications, *IEEE Transactions on Communications*, *46*(10), 1346 – 1355, doi:10.1109/26.725312.

Wade, G. (1994), *Signal Coding and Processing*, Cambridge University Press, Shaftesbury Road, Cambridge, United Kingdom.

Wang, H., and N. Moayeri (1995), Finite-state Markov channel: A useful model for radio communication channels., *IEEE Transactions on Vehicular Technology, 44(1):163–171.*

Wang, Z., and G. B. Giannakis (2000), Wireless multicarrier communications, *IEEE Signal Processing Magazine*, *17*(3), 29–48.

Weingarten, H., Y. Steinberg, and S. Shamai (2006), The capacity region of the Gaussian multiple-input multiple-output broadcast channel, *IEEE Transactions on Information Theory*, *52*(9), 3936–3964.

Weinstein, S., and P. Ebert (1971), Data transmission by Frequency-Division multiplexing using the discrete fourier transform, *IEEE Transactions on Communications*, *19*(5), 628–634, doi: 10.1109/tcom.1971.1090705.

Wyner, A. D. (1994), Shannon-theoretic approach to a Gaussian cellular multiple-access channel, *IEEE Transactions on Information Theory*, *40*(6), 1713–1727.

Xia, P., S. Zhou, and G. B. Giannakis (2004), Adaptive MIMO-OFDM based on partial channel state information, *IEEE Transactions on signal processing*, *52*(1), 202–213.

Ye, S., and R. S. Blum (2003), Optimized signaling for MIMO interference systems with feedback, *IEEE Transactions on Signal Processing*, *51*(11), 2839–2848.

Ye, S., R. S. Blum, and L. J. Cimini (2006), Adaptive OFDM systems with imperfect channel state information, *IEEE Transactions on Wireless Communications*, *5*(11), 3255–3265.

Yoo, T., and A. J. Goldsmith (2006), On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming, *IEEE Journal on Selected Areas in Communications*, *24*(3), 528–541.

Yoo, T., N. Jindal, and A. J. Goldsmith (2007), Multi-antenna downlink channels with limited feedback and user selection, *IEEE Journal on Selected Areas in Communications*, *25*(7), 1478–1491.

Yu, W., and J. M. Cioffi (2004), Sum capacity of Gaussian vector broadcast channels, *IEEE Transactions on Information Theory*, *50*(9), 1875–1892.

Yu, X.-H., G. Chen, M. Chen, and X. Gao (2005), The FuTURE Project in China, *IEEE Communications Magazine*, *43*(1), 70–75.

Zacarías, E., B, S. Werner, and R. Wichman (2009), Limited feedback multiuser MIMO techniques for time-correlated channels, *EURASIP Journal on Advances in Signal Processing*, *2009*, doi: http://dx.doi.org/10.1155/2009/104950.

Zhang, H., and H. Dai (2004), Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks, *EURASIP Journal on Wireless Communications and Networking*, *2004*(2), 222–235, doi:http://dx.doi.org/10.1155/S1687147204406148.

Zhang, J., R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath (2009), Networked MIMO with clustered linear precoding, *IEEE Transactions on Wireless Communications*, *8*(4), 1910–1921, doi:http://dx.doi.org/10.1109/TWC.2009.080180.

Zhang, P., X. Tao, J. Zhang, Y. Wang, L. Li, and Y. Wang (2005), A vision from the future: beyond 3G TDD, *IEEE Communications Magazine*, *43*(1), 38–44.

Zhou, S., and G. B. Giannakis (2004), How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?, *IEEE Transactions on Wireless Communications*, *3*(4), 1285–1294.

Zhou, S., and B. Li (2006), BER criterion and codebook construction for finite-rate precoded spatial multiplexing with linear receivers, *IEEE Transactions on Signal Processing*, *54*(5), 1653–1665.

Zirwas, W., E. Schulz, J. H. Kim, V. Jungnickel, and M. Schubert (2006), Distributed organization of cooperative antenna systems, in *Proceedings European Wireless Conference 2006*, Athens, Greece.

Zirwas, W., J. H. Kim, V. Jungnickel, M. Schubert, T. Weber, A. Ahrens, and M. Haardt (2007), *Distributed Antenna Systems: Open Architecture for Future Wireless Communications*, chap. 10. Distributed organization of cooperative antenna systems, Wireless Networks and Mobile Communications Series, Auerbach Publications, Boston, MA, USA.