



Grant Agreement:	247223
Project Title:	<b>Advanced Radio InTerface TechnologIes for 4G SysTems ARTIST4G</b>
Document Type:	PU (Public) (P/R/L/I)

Document Identifier:	<b>D1.4</b>
Document Title:	<b>Interference Avoidance Techniques and System Design</b>
Source Activity:	WP1
Editors:	David Gesbert and Tommy Svensson
Authors:	Valeria D'Amico, Bruno Melis, Hardy Halbauer, Stephan Saur, Nicolas Gresset, Mourad Khanfouci, Wolfgang Zirwas, David Gesbert, Paul de Kerret, Mikael Sternad, Rikke Apelfröjd, Maria Luz Pablo, Richard Fritzsche, Hajer Khanfir, Slim Ben Halima, Tommy Svensson, Tilak Rajesh Lakshmana, Jingya Li, Behrooz Makki, Thomas Eriksson
Status / Version:	1.1
Date Last changes:	15.07.12
File Name:	D1_4_v260.doc

Abstract:	<p>In this document we provide performance assessments of the most promising techniques that were studied within Work Package 1 (WP1) of the ARTIST4G project related to interference avoidance. The results are based on evolved techniques that were identified and classified in deliverable D1.1 and investigated in deliverable D1.2 and D1.3, as well as novel alternative techniques that are introduced and assessed in this document.</p> <p>Based on the insights from these performance results we provide a synthetic perspective over the most promising solutions for interference avoidance. We argue that these solutions can reach a satisfactory trade-off in terms of performance benefits vs complexity of implementation. Some of these techniques are also identified as complementary techniques towards an integrated interference avoidance concept.</p>
-----------	---

Keywords:	Interference avoidance, Coordinated Multi Point, Coordinated Scheduling, Coordinated Beamforming, Joint Processing, Inter Cell Interference Coordination, User Grouping, Clustering, Inter Cluster Interference Coordination, Single User MIMO, Multi User MIMO, Channel Estimation, Channel Prediction, Feedback, Robust design, Heterogeneous Networks, Game theory, Radio Access Network Architecture, Requirements, ARTIST4G.
-----------	---

Document History:	
30.06.2012	Version 1.1 with updated acronym list.
30.06.2012	Version 1.0 of document released.

## Table of Contents

<b>Table of Contents .....</b>	<b>3</b>
<b>Authors.....</b>	<b>5</b>
<b>1 Executive Summary .....</b>	<b>6</b>
<b>2 Introduction .....</b>	<b>7</b>
<b>3 Approaches and Techniques for Interference Avoidance System Design</b>	<b>8</b>
<b>4 Advanced Beamforming and Multi-cell Coordination .....</b>	<b>10</b>
4.1 Introduction .....	10
4.1.1 Dimensions for inter-cell coordination .....	11
4.1.2 Performance metrics .....	11
4.1.3 Types of information exchange between eNBs .....	11
4.2 Techniques for Advanced Beamforming and Multicell Coordination .....	13
4.2.1 Advanced Beamforming .....	13
4.2.2 Coordinated Beamforming .....	18
4.2.3 Coordinated Scheduling for Beam Coordination .....	22
4.3 Conclusions .....	29
<b>5 Advanced Joint Transmission Schemes for Multi cell cooperation .....</b>	<b>30</b>
5.1 Introduction .....	30
5.1.1 Introduction and Overview of the Joint Transmission Framework .....	30
5.1.2 Background Assumptions and Relations to Scenario 1 (Section 4) and to Scenario 3 (Section 6) .....	33
5.1.3 Joint Transmission CoMP: Promises and Challenges .....	34
5.2 The JT CoMP Framework with its Building Blocks .....	37
5.2.1 Clustering .....	39
5.2.2 Scheduling and User Grouping .....	47
5.2.3 Precoding .....	51
5.3 Practical Constraints and Enabling Technologies .....	59
5.3.1 Backhaul .....	59
5.3.2 Channel Estimation and Prediction .....	64
5.3.3 Feedback .....	72
5.4 Balancing the Joint Transmission Framework .....	77
5.4.1 Parameters and Design Variables .....	77
5.4.2 Performance in the presence of significant prediction errors .....	82
5.5 Conclusions and Discussion .....	85
<b>6 Advanced interference avoidance schemes for small cells deployments</b>	<b>89</b>
6.1 Introduction .....	89
6.2 Massive deployment of closed HeNB (with co-channel eNB) .....	91
6.2.1 Problem statement .....	91
6.2.2 Downlink HeNB/eNB ICIC with no direct cooperation capabilities .....	91
6.2.3 HeNB/eNB interference avoidance scheme with slow cooperation capabilities .....	94
6.3 The Femto campus use case .....	98
6.3.1 Problem statement .....	98
6.3.2 Centralized power control for femto campus .....	98
6.3.3 Coordinated scheduling for heterogeneous deployment .....	101
6.3.4 Resource allocation in slow fading interfering channels with partial knowledge of the channels .....	102
6.4 eNB/Pico/Relay Heterogeneous deployments .....	104
6.4.1 Problem statement .....	104
6.4.2 A Practical Iterative Algorithm for Joint Signal and Interference Alignment in Heterogeneous Networks .....	104

6.5 Conclusions .....	108
<b>7 Conclusions .....</b>	<b>109</b>
<b>Appendixes .....</b>	<b>111</b>
A1. Single-cell MU-MIMO scheme .....	111
A1-1 Transmit and receive filter design with limited signalling information .....	111
A2. Multi-cell MU-MIMO schemes .....	116
A2-1 Robust linear precoding with per-base-station power constraints .....	116
A2-2 An integrated design for downlink Joint Transmission CoMP .....	120
A2-3 Joint scheduling and power control with non-coherent transmission .....	139
A2-4 Waterfilling schemes for Zero-Forcing coordinated transmission .....	143
A2-5 Dynamic Partial Joint Processing .....	149
A2-6 Resource allocation for OFDMA Joint Processing CoMP .....	158
A2-7 Robust Precoding with Distributed Channel State Information .....	162
A2-8 Precoding optimization algorithm for coordinated beamforming .....	167
A2-9 Coordinated beamforming for interference rejection .....	170
A2-10 A Practical Iterative Algorithm for Joint Signal and Interference Alignment in Heterogenous Networks .....	175
A3. Advanced 3D Beamforming .....	177
A3-1 UE-specific horizontal and vertical beamsteering .....	177
A3-2 Distributed scheduling for beam coordination .....	184
A4. Enablers: channel estimation & feedback design .....	189
A4-1 Centralized/decentralized joint transmission with limited signalling information .....	189
A4-2 Kalman prediction of multi-site MIMO channels for CoMP .....	193
A4-3 Advanced channel prediction .....	204
A4-4 Feedback compression .....	211
A4-5 Advanced feedback compression schemes .....	215
A5. Clustering and user grouping .....	217
A5-1 Clustering and interference floor shaping based on partial CoMP .....	217
A5-2 Inter-cluster coordination with fractional frequency reuse .....	228
A6. Inter-Cell Interference Coordination .....	235
A6-1 Coverage control through non linear conjugate gradient optimization .....	235
A7. Coordinated Scheduling .....	252
A7-1 Performance evaluations of the interference management concept .....	252
A8. Scheduling for joint processing .....	255
A8-1 Impact of scheduling on the performance of downlink multicell processing ..	255
A8-2 Scheduling Aspects of Partial CoMP .....	259
<b>References .....</b>	<b>263</b>
<b>List of acronyms and abbreviations .....</b>	<b>271</b>

## Authors

Name	Beneficiary	E-mail address
Richard Fritzsche	TU Dresden	richard.fritzsche@ifn.et.tu-dresden.de
Tommy Svensson	Chalmers University of Technology	tommy.svensson@chalmers.se
Tilak Rajesh Lakshmana	Chalmers University of Technology	tilak@chalmers.se
Jingya Li	Chalmers University of Technology	jingya.li@chalmers.se
Behrooz Makki	Chalmers University of Technology	behrooz.makki@chalmers.se
Thomas Eriksson	Chalmers University of Technology	thomase@chalmers.se
Mikael Sternad	Uppsala University	mikael.sternad@signal.uu.se
Rikke Apelfröjd	Uppsala University	rikke.apelfrojld@signal.uu.se
Hardy Halbauer	Alcatel-Lucent Deutschland AG	Hardy.Halbauer@alcatel-lucent.com
Stephan Saur	Alcatel-Lucent Deutschland AG	Stephan.Saur@alcatel-lucent.com
Hajer Khanfir	Orange Labs	Hajer.khanfir@orange.com
Slim Bem Halima	Orange Labs	Slim.benhalima@orange.com
Nicolas Gresset	Mitsubishi Electric R&D Centre Europe	n.gresset@fr.mercede.mee.com
Mourad Khanfouci	Mitsubishi Electric R&D Centre Europe	m.khanfouci@fr.mercede.mee.com
Valeria D'Amico	Telecom Italia	valeria1.damico@telecomitalia.it
Bruno Melis	Telecom Italia	bruno1.melis@telecomitalia.it
Wolfgang Zirwas	Nokia Siemens Networks	wolfgang.zirwas@nsn.com
David Gesbert	EURECOM	gesbert@eurecom.fr
Paul de Kerret,	EURECOM	paul.dekerret@eurecom.fr
Maria Luz Pablo	Telefónica I+D	mlpg@tid.es

## 1 Executive Summary

Cooperation techniques at the lower layers between neighboring base stations offer a powerful tool for improving cellular wireless network performance, especially in ill-favored areas such as cell edge, black (shadowed) spots, etc. In particular, cooperation and coordination methods among interfering transmitters allow avoidance of interference before it is actually undergone by the receiving antenna at the user equipment side. Such techniques have been the focus of ARTIST4G's WP1, giving rise to a particularly rich set of research contributions. This document provides a synthetic perspective over the most promising ideas. It provides a classification of the proposed novel techniques, and highlights connections between them. Techniques are categorized according to the type of information exchange required between the base station and are labeled as interference coordination and multi-cell cooperation. It describes suitable scenarios of application in both homogeneous and heterogeneous networks and, where appropriate, illustrates the typical network performance gains.

Interactions with alternative or complementary interference-management innovations at other network components (such as UE side) or layers are briefly touched upon. Overall design considerations are also given in the document.

The results indicate the importance of interference avoidance as a way for upcoming wireless cellular network to deal with the problem of unfair radio quality distribution across the cell while meeting the new stringent constraints for overall cell traffic capacity.

## 2 Introduction

This document is the last deliverable for WP1 workpackage within the ARTIST4G project. The purpose of the document is to provide a synthetic perspective over some of the key contributions brought by the project consortium within the area of cooperative transmission methods for interference avoidance. As the last deliverable, the document describes the key lessons learned in terms of interference avoidance algorithms and formulates general system design guidelines. As the area is rich in recent proposed techniques from within and outside ARTIST4G, one challenge lies in the identification of which techniques actually makes sense for an application in 4G wireless networks, which methods require drastic evolutions of standards, and finally which technique reach a satisfactory trade-off in terms of performance benefits vs complexity of implementation to be worth considered. This document presents a selection of proposed methods, where the selection was carried out according to the above objective as much as possible.

The contributions are divided into three main sections, following a general panorama of techniques and challenges for interference avoidance in Section 3. In Section 4, the simplest form of interference avoidance methods is addressed, where the term “simplest” is to be understood in terms of a reduced need for multi-cell cooperation, reduced data exchange between eNBs, and often simplicity of implementation. The basic building blocks of interference avoidance through multi-cell coordination without exchange of user plane data are described there, namely coordinated beamforming and coordinated scheduling. In coordinated beamforming, mutually interfering eNBs coordinate the computation of single cell beams so as to minimize interference to each other while maximizing the received energy to the intended users. We show in particular how such approaches can benefit greatly from the addition of extra beamforming dimensions so as to render the beam design more discriminating (for instance using vertical dimension beamforming added to the conventional horizontal dimension). In Section 4, the gains of coordinating scheduling on interference reduction are shown, potentially used as a complement to coordinated beamforming.

In Section 5, more complex interference avoidance solutions are presented. These techniques rely on the assumption that user plane data can be shared by cooperating eNBs located within a cooperation cluster, thanks to a suitable backhaul architecture. Under this assumption, so-called Joint Precoding techniques across the cluster eNBs can be applied, which mimic the precoding methods used in multi-user MIMO systems and, in principle, intra-cluster interference can be fully eliminated. Section 5 addresses some of the key challenges to such approaches such as the design of cluster via suitable user and cell grouping algorithms, the management of inter-cluster interference via power control and robust beamforming methods, and finally the design of latency-robust feedback schemes exploiting channel predictions.

In Section 6, interference avoidance is investigated for the specific case of small cells and heterogeneous networks. There, practical constraints related to ease of implementation, low complexity, and distributed optimization to avoid heavy exchange of information between macro and femto cells are emphasized. In order to satisfy such constraints, techniques are proposed making use of clever power control protocols, together with coordinated scheduling and beamforming.

Finally, in Section 7 conclusions are given. Note that the proposed methods are only described in synthetic terms in the main sections of this document, while additional details for modelling, mathematical derivations and simulation results are presented in the Appendix chapters.

### 3 Approaches and Techniques for Interference Avoidance System Design

Operators of cellular networks want to provide a certain quality of experience to their users while the required cost for that purpose shall be minimized at the same time. Since future cellular networks will be mainly interference limited, a major interest of operators is the application of high-performing interference avoidance techniques. However, the effort for hardware and infrastructural upgrades should be as small as possible. A promising approach in this context is the application of coordinated multi-point (CoMP) techniques. In this section, basic aspects of interference avoidance techniques are discussed, focusing on the application of CoMP. On that base, a framework is presented targeting the design and optimization of a cooperative cellular system.

For the presented framework the degrees of freedom for optimizing a cellular system can be classified into three different parameter sets, which are related to certain influence quantities, namely environmental, traffic related and user specific properties. Each of these parameter sets is varying on a certain time scale.

Optimization related to environmental parameters is mainly static or varying on a time scale of month or years, according to changes like, e.g., the construction of buildings. Degrees of freedom relevant for optimization are mainly related to hardware or infrastructure. Regarding the deployment of base stations, small cells (micro or femto base stations) can be installed at cell edge areas to increase cell edge throughput, while interference to macro base stations can be handled by cooperation or by an appropriate frequency reuse scheme. Furthermore, base station locations (sites) are commonly divided into multiple sectors, while the amount of sectors per site and their orientation depends on environment aspects. Additional influence parameters which are optimized according to the environment are antenna patterns, mechanical tilting or the deployment of antenna-arrays in order to enable multiple-input multiple-output (MIMO) techniques. Note, that this optimization level also considers long term traffic density (over month or years). While currently deployed cellular networks avoid inter-cell interference by restricting the reuse of resources in adjacent cells, state of the art systems are focussing on reuse one networks. The resulting inter-cell interference is forced to avoid by cooperative techniques as well as an adequate optimization of the before mentioned parameters.

Data traffic within the network usually varies in the range of days or hours. Certain traffic behaviour can be identified over the day, depending on the day of the week. Such fluctuation can be handled by switching base stations on and off or adapting electrical down tilt or hand-over parameters. Since traffic related parameters can be adapted automatically it is often mentioned in the context of self-organizing networks. However, aspects which are basically related to user properties can also be mapped to the same optimization level, since they are handled in a time scale of hours, according to complexity aspects. Such parameters are referred to as semi-static and include frequency reuse, clustering and resource allocation among cooperation clusters. For low traffic situations a higher frequency reuse can be applied to save effort required for base station cooperation, while high traffic can be absorbed by an aggressive frequency reuse enabled by high-performing CoMP techniques. In this regard coordinated base stations form cooperation clusters. The size of a cooperation cluster can be semi-statically adapted to the traffic demand. Overlapping clusters can be used to ensure frequency reuse one, while resource allocation among clusters can be adapted to the traffic situation. Since reuse one systems basically suffer from a high interference floor, inter-cluster interference can be reduced by, e.g., electrical down tilt, and/or fractional frequency reuse

Based on the identified cooperation clusters and the allocated resources, interference within a cluster can be avoided by user specific optimization. Such optimization takes channel conditions or user positions into account. Regarding CoMP techniques, it is basically distinguished between *coordinated* and *cooperative* multi-cell transmission. For the former case only control plane information is exchanged between base stations to avoid inter-cell interference. Such control information is, e.g., SINR or channel state information. Cooperation refers to the additional exchange of user plane data, which enables joint precoding. While cooperative techniques provide theoretically better performance, coordinative approaches are typically less sensitive to practical impairments like outdated channel state information. Degrees of freedom for user specific optimization are: grouping users that share the same radio resource, allocating resources among user groups within a cooperation cluster, as well as the specific algorithm for spatial signal processing like coordinated beamforming or joint precoding.

The structure of the three level optimization framework outlined above is summarized in Table 3.1.

**Table 3.1: Structure of a three level optimization framework considering cooperative cellular networks in order to avoid interference.**

Influence Quantity	Time Scale	Degree of Freedom
<b>Environment</b>	Years/month	Deployment Sectorization Antenna properties Static frequency reuse
<b>Traffic</b>	Days/hours	Electrical down tilt Hand-over parameters Switching on/off base stations Semi static (fractional) frequency reuse Semi static clustering Semi static resource allocation (among overlapping clusters) Inter-cluster interference
<b>User property (position/channel state)</b>	Seconds/milliseconds	Scheduling/user grouping Dynamic resource allocation (within a cluster) Spatial signal processing algorithm

Interference avoidance aspects, focussing on coordinated multi-cell transmission are discussed in Section 4, while Section 5 covers techniques related to multi-cell cooperation. The discussed issues in Section 4 and 5 are basically analysed for homogeneous networks. Specific aspects which are focussing different cell sizes are discussed in Section 6.

## 4 Advanced Beamforming and Multi-cell Coordination

### 4.1 Introduction

Section 4 presents innovations working without the need for user plane cooperation. This comprises innovations without cooperation and with control plane exchange over the RAN (NO\_COOP and CP\_COOP as classified in [ARTD11]). These innovations were preferably investigated for application in the homogeneous macro cell deployment. They provide intra-cell or intercell interference avoidance capability at different levels of involvement of adjacent cells. Some of these innovations are applicable in each cell individually or require a predefined static configuration of all cells in the network. Interference avoidance is achieved through specific beam pattern shape and/or specific transmit and receive signal processing in each eNB. Other innovations of this section 4 go further towards exchange of control information with adjacent eNB within fixed or dynamically changing clusters, where the exchange of control information is expected to be either slow or fast. In this case, interference avoidance is achieved either through quasi-static configuration of resource usage or highly dynamic resource allocation in combination with appropriate transmit beam shaping. We even present some fundamental progress related to a smarter use of the spatial dimension offered by multiple antennas. This leads to an optimized use of the spatial degree of freedom, which in turn can be exploited in the context of multi-cell coordination.

Note that innovations which allow or rely on exchange of user data between eNBs (UP-COOP) are addressed in the following section 5. Heterogeneous networks (HetNets) comprising small cells operating on the same frequency resources are addressed in section 6.

Theoretical studies (e.g. [KFV06, FKV06]) indicate that the ultimate form of multi-cell cooperation (based on e.g. JP CoMP schemes) need to involve both the sharing of user plane data (data packets on the downlink and baseband received signals on the uplink) and of the channel state information (CSI) at all cooperating eNBs. Nevertheless, a significant gain through reduction of inter-cell interference can be achieved already in the case where eNBs are not informed of the user plane data originating from or intended to users located in interfering cells. The key underlying principle is that of interference diversity (in multi-antenna or multi-user domain). By this principle, interference signals are subject to the same type of random attenuation that otherwise affects the desired signals. This randomness can be exploited, especially when the pool of available users is large, by allocating to a user a slot of spectral resource that combines a good received desired signal level with an attenuated interference level. Importantly, multiple antenna beamforming at each eNB can be exploited to further accentuate the difference in level between the desired and the interfering signals. For a greater impact, the process of allocating resources can be coordinated across the interfering cells. Although this can be implemented with slow exchange of control information, additional performance is expected with a highly dynamic coordination reacting fast on changing resource allocation requirements in adjacent cells. Therefore, fast exchange of channel (resource control) information via backhaul link is desirable.

Since current LTE-A standard does not support UP-COOP yet, the NO-COOP and CP-COOP schemes addressed in section 4 are a possibility to improve performance already in short or medium time frame. When considering long-term deployments, it is expected that UP-COOP schemes can be applied, But there will remain difficulties and challenges especially for multi-vendor UP-COOP solutions, as pointed out in [ARTD43]. Therefore the CP-COOP schemes will remain relevant techniques even for long-term deployment scenarios, since they represent fallback solutions for situations where full UP-COOP is not applicable.

#### 4.1.1 Dimensions for inter-cell coordination

For the purpose of interference avoidance the available resources can be coordinated across several dimensions. Each of these dimensions leads to a different set of techniques. For some of these techniques also combinations of coordination dimensions are possible. The main coordination dimensions relevant for the innovations presented in section 4 are:

- Frequency slot allocation
- Time slot allocation
- Power allocation
- User group allocation
- Multiple antenna based beam allocation

#### 4.1.2 Performance metrics

Coordination methods in the domains of time, frequency, power, antenna pattern can be classified according to the type of performance metrics to be optimized. There are two leading approaches for this problem: First, the sum throughput over all users in the cooperating cells could be tried to be maximized, possibly with proportional fairness constraints. Second, resource (e.g. power) usage could be tried to minimize while achieving a minimal link quality target at each user. Both approaches have been considered in the innovations described in this section.

With respect to a main objective of the ARTIST4G project, the distribution of the user throughputs within the cell is of major interest. The most relevant metrics are the cell edge throughput measured as 5%-ile of the throughput CDF, and the average cell throughput. Improvements due to the innovative schemes can be covered by comparison of the individual metrics with a realistic baseline and are expressed in percent of the baseline performance. Also the Jain index is a figure for the improvement of the homogeneity of user throughputs across a cell. However, this metric is a relative measure and, if standing alone, does not reflect the absolute level of throughput improvements [ARTD51].

#### 4.1.3 Types of information exchange between eNBs

ENB coordination naturally goes at the expense of an overhead in the backhaul and over-the-air in terms of information acquisition and exchange between the coordinating cells. This information can be of several types described below and summarized in Table 4.1.

**Table 4.1: Types of information for coordination**

Rate of exchange	Instantaneous-channel related information	Statistical information
<b>Fast</b> ( $<$ fading coherence time)	Direct channel coefficient, interference channel coefficient, instantaneous SNR, SINR, SIR, power level, instantantaneous precoding coefficient, resource block assignment decision,	X
<b>Slow</b> ( $>$ fading coherence time)	X	Average SNR or SINR, Direct channel covariance, interference covariance, LOS component strength. NLOS probability.

The key differentiation is the time scale over which information is exchanged. We distinguish slow information exchange from fast information exchange. Slow data exchange relies on the computation of channel-related statistics, mainly correlation and average power based, which provide useful information about the interference created by a given user during a period of time related to the macroscopic fading (a region of tens to hundreds of wavelength around the initial position). Fast data exchange provides interference information on a time scale related to the fading coherence period (thus corresponding to a region of less than one wavelength around the initial user position). The key advantage of fast data exchange is to open up the possibility of computing instantaneous resource allocation and precoding solutions at the eNB transmitters based on the received interference information, allowing powerful interference cancelling solutions. It should be noted that for such schemes in addition to fast exchange of instantaneous information also statistical channel information might be useful. In contrast, coordination methods relying only on statistical information are limited to exploiting macroscopic diversity, for instance the fact that certain regions of the cells are shadowed from certain sources of interference thanks to distance-based power decay or obstruction from a hill or a construction. ARTIST4G builds up on both concepts to propose schemes offering a range of compromises between information overhead over the backhaul and over the wireless feedback channels, and interference mitigation performance.

## 4.2 Techniques for Advanced Beamforming and Multicell Coordination

Coordination in the spatial domain relies on the use of suitable multiple antenna combining and signal processing at the eNB side. Therefore, techniques which can improve the performance or robustness of smart antennas, even in a single cell context, are clearly of high interest. Better beamforming capability in a single cell scenario can lead, when exploited in a multi-cell context, to even more powerful coordination and interference mitigation capability.

ARTIST4G explores two basic mechanisms to achieve high performance multiple antenna combining. One mechanism targets at enhancements of multiple antenna combining methods using precoding schemes in the horizontal dimension to provide higher robustness with respect to errors and limitations in the channel state information feedback. The other mechanism basically expands the exploitation of only the horizontal dimension of beamforming by including additionally the elevation angle towards 3D beamforming. This reflects the fact that the properties of the considered macro cell deployment scenario can be better covered with a three-dimensional view. Both mechanisms finally are combined when for example using 3D beamforming in combination with elevation-related channel feedback information. This section 4.2 comprises solutions exploiting either one of these mechanisms or combinations of both. A common characteristic of all solutions provided in section 4.2 is that noexchange of UP information between eNBs is needed.

The addressed innovations can further be classified with respect to their involvement of adjacent eNBs, i.e. the amount of coordination among adjacent eNBs, and their complexity. Although for the lowest complexity solutions it can not be expected to get the same maximum performance improvement than for the more complex ones, they have their justification since they have the potential to provide reasonable gain already without or with only limited extension of the required capabilities of existing networks. Therefore they could be implemented at an early stage, whereas the more advanced solutions, which often require the extension of existing standards, can be applied in addition according to the increasing capabilities of the networks following the standards evolutions. In the following subsection 4.2.1 "Advanced Beamforming" the most promising solutions relying on beam pattern adaptation and related signal processing schemes are presented. There is no exchange of information between eNBs and no coordination on scheduling level assumed. Subsection 4.2.2 "Coordinated Beamforming" covers schemes which optimize the horizontal beam pattern, thus taking into account interference measurements from adjacent cells. Schedulers are not involved in the beam pattern optimization. Finally, subsection 4.2.3 "Coordinated Scheduling for Beam Coordination" involves scheduling as major means for interference avoidance through coordinated resource allocation making use of CP information exchange between eNBs and applying joint optimization of scheduling and beam pattern adaptation. Within each subsection the innovations are introduced in the order of increasing requirements and complexity.

### 4.2.1 Advanced Beamforming

In this subsection 4.2.1 "Advanced Beamforming" the most promising solutions relying on beam pattern adaptation are presented. These are pure beamforming schemes which do not use any coordination of radio resource allocation with adjacent cells. These schemes are working in a "single cell - like" operation mode. This means that each eNB applies individual algorithms for transmit signal processing, without taking into account feedback about the instantaneous pattern adaptation in adjacent cells. Of course these schemes are applicable and will operate usually in multi-cell scenarios. Gains are achieved in multi-cell scenarios as well.

In single cell operation there are two major effects influencing the performance. A first straightforward effect is the intra-cell interference in case of multi-user operation, which is caused by non-optimum antenna weights, unfavourable UE pairing or insufficient accuracy of channel estimation. Such non-optimum intra-cell operation leads to reduced MU-MIMO

performance, because it does not take into account all boundary conditions for improved cell edge and spectral efficiency maximization in the overall network. The innovations addressing these impacts exploit pairing optimization for improving capacity of MU-MIMO schemes for minimization of intra-cell interference.

A second effect is the impact of the antenna pattern itself, which impacts the capability to achieve full coverage. With antenna pattern optimization, up to individual per UE pattern adaptation a better coverage and cell edge throughput can be achieved. This individual per UE antenna pattern adaptation can be done by only maximizing the signal strength at the UE, or also trying to minimize signal strength at the same time in a direction where adjacent cell UEs are suffering from interference. With the approaches presented in this subsection this is achieved by limiting the minimum downtilt to a maximum value, which avoids adjacent cell interference when serving UEs close to cell border.

More specific, the innovations provided here are:

**Transmit and receive filter design with limited signalling:** This concept optimizes the MU-MIMO performance (sum throughput of multiple UEs sharing the same resources) in a single-cell context. It is based on an optimization of Rx filtering at the UE, taking into account the Tx beam pattern generated by precoding and the CSI. Additional transmission overhead due to signalling of the CSI to the eNB and the Rx filter coefficients to the UE is traded off against the achievable throughput increase. Only information available within the cell is needed, no coordination with and also no impact of other cells is taken into account.

**UE specific horizontal and vertical beamsteering:** According to the location of the UE, the vertical and the horizontal beam pattern are adapted dynamically. This is also done in a single-cell context. Due to more concentrated transmission of the energy, an increase in signal strength at the UE location and at the same time a reduction of the mean interference in adjacent cells when serving UEs close to their own eNB provides an increased overall spectral efficiency and an increased cell edge throughput.

In the following subsections these innovations are described in more detail.

#### 4.2.1.1 Transmit and receive filter design with limited signalling information

The application of MU-MIMO systems in cellular networks brings the possibility of performance gains compared to other multiple access strategies. Considering a single cell setup in the downlink, where the number of antennas placed at the eNB exceeds the number of antennas per UE, data streams to multiple UEs can be scheduled to the same transmission resource, while inter-stream interference is mitigated by linear precoding. If UEs are equipped with multiple antennas a part of the interference mitigation can be done at the receiver side.

With a joint transmit and receive filter design [ZWZ+05] higher rates can be achieved compared to systems where no receive filtering is applied at the UE [SM03]. However, the calculation of the receive filters requires knowledge of the precoded channel, which is assumed not to be a-priori available at the receivers. In order to find the most efficient signalling strategy, we do not restrict ourselves to existing pilot schemes of the LTE-A baseline system. In this contribution different strategies are investigated for signalling relevant information to make the receive filters available at the UEs. In order to compare the signalling strategies, a metric called net rate  $R_N$  is introduced, which is the rate  $R$  (in bit per channel use) for user plane transmission weighted by the amount of signalling required to achieve rate  $R$ . For each strategy first the net rate is maximized by finding the optimal amount of signalling.

The investigated strategies are basically distinguished between analog and discrete signalling. Analog signalling is based on reference signals (pilots) which are known at both, transmitter and receiver side. The reference signals are multiplied with the precoding matrix before they are transmitted. With this method the precoded channel is made available at the UEs, which are then able to calculate the receive filters on their own. Considering discrete signalling, quantized information is forwarded to the UEs. In this regard, two basic cases are distinguished. Either the

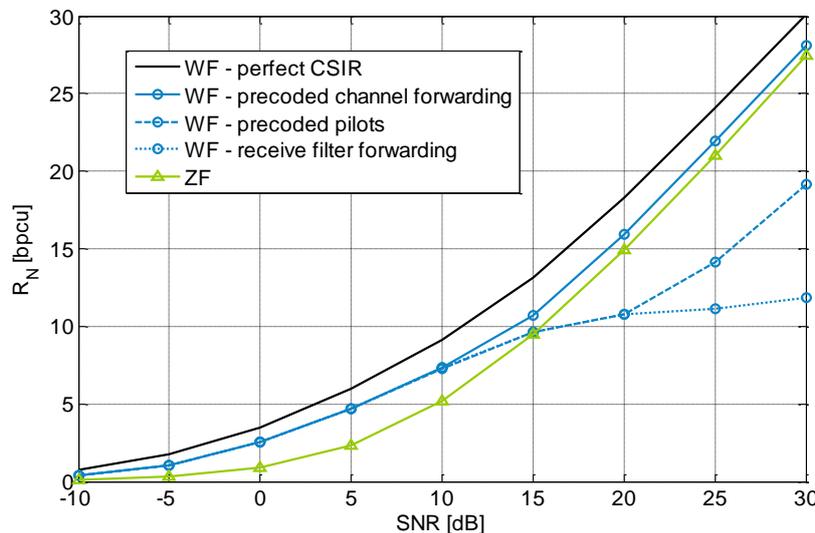
precoded channels or the already calculated receive filters are forwarded to the UEs. However, also in the latter case knowledge about the precoded channel needs to be available for detection. The scheme for considering imperfect channel state information at the receiver side (CSIR) is discussed in [MF11].

For simulations a setup is applied where one 4 antenna BS transmits 2 data streams to 2 UEs, respectively. Each UE is equipped with 2 antennas. Furthermore, perfect channel knowledge at the eNB is assumed which helps to identify the main effects of the different strategies.

Basic results can be seen in Figure 4.1. The net rate of the three signalling strategies with Wiener filtering (WF) is compared with zero-forcing (ZF), which requires no signalling since no interference remains at the receiver side. The upper bound refers to the case of perfect CSIR without spending any resources for signalling. For SNR smaller than 10 dB the signalling amount that maximizes the net rate is zero, and statistical knowledge of the precoded channel is used for receive filtering and detection. In the high SNR regime, it is preferable to forward the precoded channel, even if the block size is relatively small (the plot refers to a block size 25). Using precoded pilots, signalling becomes reasonable for high SNR. However, ZF outperforms this strategy. For receive filter forwarding it turned out, that signalling is not helpful at all. With this strategy it is preferable to perform receive filtering and detection based on statistical information (see further details in the Appendix A1).

With increasing block size, the resource consumption for signalling becomes less relevant and the performance of precoded channel forwarding and precoded pilots converges to the perfect CSIR case.

Applying the shown results to time varying channels, precoded pilots have an advantage compared to any forwarding strategy. The reason is that precoded pilots inherently include information of the current channel state, while pilot forwarding does only have access to a previous channel state. This aspect becomes especially important for FDD systems, where the delay between channel observation and actual data transmission includes the CSI feedback transmission.

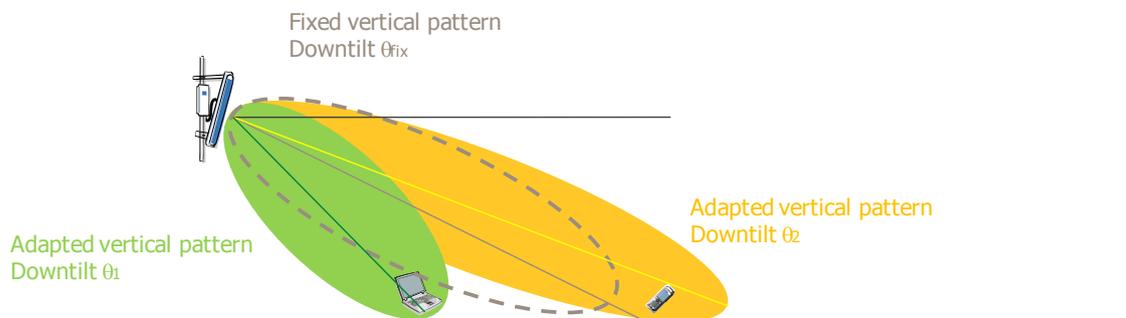


**Figure 4.1: The resulting net rate of the three investigated signalling strategies applying WF precoding and the comparison with ZF.**

#### 4.2.1.2 UE-specific horizontal and vertical beamsteering

Beamforming schemes, which concentrate the transmitted energy towards the UE, lead to a different statistic of interference characterized by lower mean interference power but increased

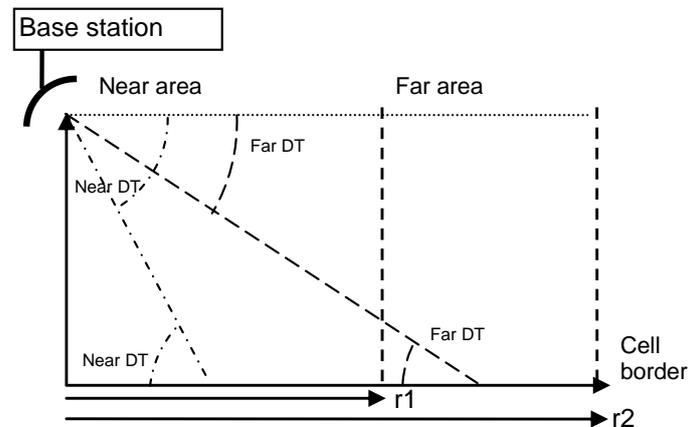
variance. Whereas in conventional systems the vertical antenna pattern was fixed and beam steering was applied only to the horizontal antenna pattern, the proposed new scheme of UE specific horizontal and vertical beam steering directs the main lobe of the horizontal *and the vertical* antenna pattern towards the location of the UE. Even without any coordination of radio resources among adjacent cells, this leads to an improvement of signal strength at the UE and further reduction of the interference in adjacent cells (especially when serving UEs close to their eNB). These two effects have the potential to increase the cell edge throughput and the spectral efficiency [ARTD12, SH11]. An example for adaptive vertical beam pattern pointing directly to different UEs, in contrast to a fixed vertical pattern (dashed line), is shown in Figure 4.2. It should be noted that we need not to steer the antenna pattern in a wideband fashion, but assume the possibility to steer the resources (i.e. subchannels) individually towards the UEs, to which they have been assigned. This can be achieved with new flexible antenna concepts enabling simultaneous individual beamsteering of parts of the resources to UEs in different locations.



**Figure 4.2: Vertical beam steering**

To achieve vertical beam steering, an extension of the antenna system functionality is necessary. The vertical beam pattern of eNB antennas with a small half-power beamwidth (HPBW) can be realized with multiple vertically stacked closely spaced transmitting elements driven by the same signal with appropriate phase shift between the transmitting elements. For fixed downtilt this shift is achieved by a passive feeder network optimized for the targeted HPBW. To realize exact dynamic vertical downtilt adaptation to each UE location, ideally the antenna weights of each vertical element need to be controlled separately. Although such antenna concepts became meanwhile available [P12], also less complex alternative solutions are feasible.

The realization options investigated in ARTIST4G comprise exact main lobe steering to the UE without and with limitation of the minimum downtilt to reduce intercell interference. As less complex alternative also the selection of one out of two or three predefined fixed downtilts according to the location of the UE has been considered. In the case of two fixed downtilts, as indicated in Figure 4.3, UEs located in the near area up to a distance  $r_1$  from the eNB are served with the “near downtilt”, UEs located between  $r_1$  and  $r_2$  from the eNB use the “far downtilt”. Baseline is the case with one fixed downtilt.



**Figure 4.3: Selection among two fixed downtilts**

The location of the UE can be derived from the uplink transmission. Since the small HPBW used in vertical direction leads to a high correlation between the transmitting elements, the angle of arrival can be estimated from the signals of the vertical elements and is independent of the frequency within a reasonable bandwidth. So this method can be applied also in FDD systems.

The proposed scheme can basically be combined with almost all multi-antenna schemes dealing with horizontal beam pattern optimization. In this case the parameters of those schemes have an additional dimension depending on the selected downtilt. The downtilt adaptation itself is based on long-term statistical vertical channel information, which is possible due to the assumed strong correlation of the vertical antenna elements and helps reducing the estimation and signalling overhead compared to the MIMO algorithms applied in horizontal direction requiring instantaneous channel state information.

The UE specific horizontal and DT vertical beamsteering can operate completely independent from other adjacent cells and has shown benefits also in an interference free single cell scenario. The achievable performance depends on the UE distribution. The gains are due to statistical effects caused by the antenna pattern control. The main parameters like downtilt angle limitation in case of exact main lobe steering, or the two fixed downtilt angles for the approximating realization option depends mainly on the cell size and the eNB height and has to be statically configured.

Performance evaluations have been made for single and multi cell deployments with simulation assumptions according to the 3GPP simulation scenarios as described in [3GPP36814]. The mean spectral efficiency and the cell edge UE throughput improvements have been estimated for different values of HPBW and cell size in a 19 site / 57 cell scenario. UEs were randomly distributed. Pure vertical adaptation as well as combination of horizontal and vertical beam steering has been investigated. Detailed results have been presented in [ARTD12, SH11] and in the appendix A3 of this document.

A first verification of the behaviour of the vertical beam steering has been provided by WP6 with field measurements identifying the optimum downtilt for the maximum receive power level at the UE [ARTD61, ARTD62, KHS+12]. A summary is provided in Appendix A3-1. Additional measurements in the Dresden testbed are used to further investigate the radio channel behaviour in real propagation environment [DHG+12]. All these measurements revealed that the basic behaviour is as expected. However, in strong NLOS environments the deviation from this approach is visible. Depending on the height of the eNB above the rooftops and the number of reflections contributing to the receive signal, the optimum downtilt angle is smaller than expected from the pure geometrical analysis, as also theoretically predicted e.g. by [CL08].

## 4.2.2 Coordinated Beamforming

The schemes presented in this subsection 4.2.2 “Coordinated Beamforming” target at an optimization of the horizontal beam pattern to optimize performance within a cell. In contrast to the previous subsection, here also interference measurement information and scheduling decisions from adjacent cells are partially taken into account. This CP information is exchanged via the X2 interface. Based on this information each individual cell tries to optimize its horizontal beam pattern while minimizing the caused interference in the adjacent cells. Other than in the subsequent subsection, here this exchanged information is used solely to optimize each cell-individual beam pattern, without involving the schedulers and without coordinating the resource allocation of adjacent cells.

More specific, the innovations provided here are:

**Precoding optimization algorithm for coordinated beamforming:** This innovation focuses on a centralized optimization of the beam pattern of the full area of adjacent cells, where power control is considered as an additional parameter. The scheme relies on full CSI information, which is used for the optimization of precoding algorithms. Scheduling is not used as an optimization tool.

**Coordinated beamforming for interference rejection:** This scheme assumes the availability of the UEs to measure adjacent cell interference and the corresponding cell ID. This information is made available to the interfering cells, which shape their beam pattern in a way to optimize signal strength for the own UE while placing a minimum in the antenna pattern in direction of the interfered adjacent cell UEs. Although CP information is exchanged between the eNBs, the antenna pattern are generated in each cell individually, but based on this information.

In the following subsections these innovations are described in more detail.

### 4.2.2.1 Precoding optimization algorithm for coordinated beamforming

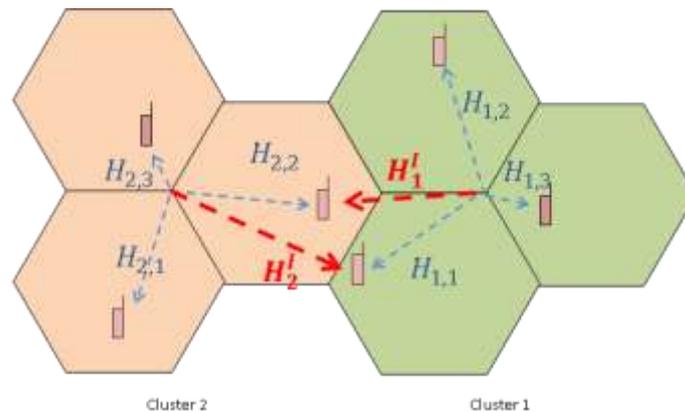
In this work, we consider the well known optimization problem where the sum-power of multiple transmitters is minimized under the constraint of quality of service (QoS) in multi-cell multiantenna cellular networks. It has been shown in [DY10] that this problem can be resolved by using Lagrange duality extended to the multi-cell case.

The algorithm proposed in [DY10] makes the assumption that all the users channels between all the transmit- and receive antennas have to be known at all the transmitters or the cells. To the best of our knowledge, two methods may be used to feed back the CSI (Channel state Information). In the first one, all the UEs feed back all their CSI to all the cells in the cooperation area [PHG08]. In the second one, each UE feeds back the CSI to the serving cell which in turn transmits it to the cells in the cooperation area. In fact, in the first step each Ue estimates the downlink channel from each cooperating cell to the UE and feeds back the CSI to its serving cell. At the second step each cell transmits all the received CSI to all the cells in the cooperation area.

In this case, the amount of exchanged information increases with the number of users and the number of coordinated cells in the cooperation area. Furthermore, the architecture has to provide low latency backhaul between all the coordinated cells which increases the cost of such deployment from the Operation and Maintenance (OAM) point of view. Some solutions based on clustering have been proposed in subsection 5.2.1, we consider the simplest one based on the 3GPP scenario1 [3GPP36.819] where each cluster is formed by B cells belonging to the same eNB and there are N cells in the entire network.  $N_c=N/B$  is then the number of clusters in the network. There are K users in each cell equipped with a single antenna.

However, the algorithm in [DY10] cannot directly be applied in a realistic network with clustering due to the inter-cluster interference. Since these algorithms only consider the inter-user and intercell interference cancellation, the interference from other clusters is not taken into account. In this case the inter-cell interference problem is turned into inter-cluster interference.

In this work, we propose an algorithm which transforms the transmit power minimization problem into transmit power and inter-cluster interference minimization under SINR constraints. Each cluster applies the proposed algorithm in order to minimize the transmit power of each cell and the interference created to the other clusters under the constraint of SINR user target. This optimization problem is resolved similarly to [DY10], since the downlink beamforming problem is equivalent to the uplink beamforming problem we can calculate iteratively the downlink precoding vectors and powers from the uplink ones. The algorithm in [DY10] has been modified to take into account to the inter-cluster interference into the calculation of the downlink powers and precoding vectors. The details of the algorithm are described in the Appendix A2-8.

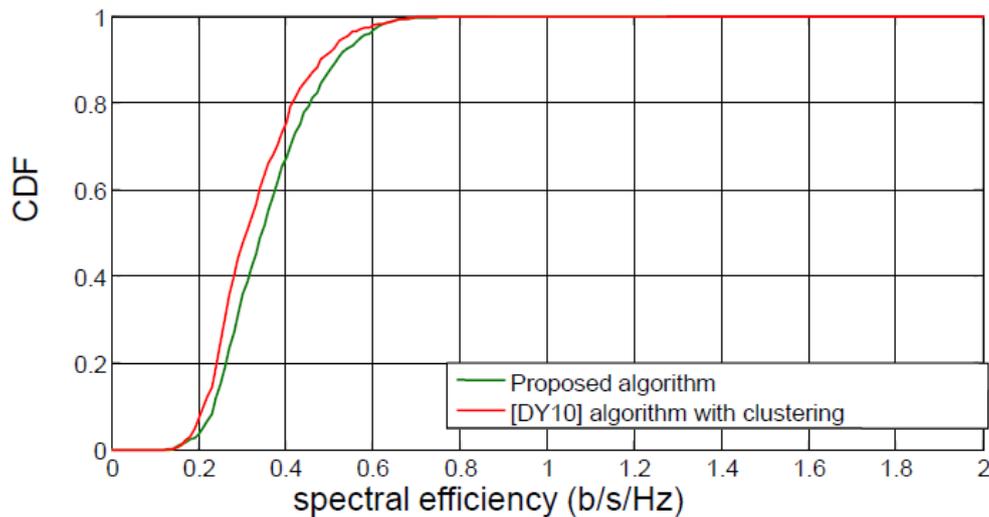


**Figure 4.4: A scenario with two clusters each one composed by three cells transmitting to three users**

Figure 4.4 represents two clusters with three cells serving three UEs. In the figure we represent only the UEs at the cell edge which receive the interference from the other cluster. The proposed algorithm is performed in each cluster and takes into account all the interferences that are created to the other UEs in the other clusters.

In addition, our algorithm avoids the fact that minimizing the multi-cell transmit powers under SINR constraint could lead to minimizing the power of some cells. In fact for the UEs which receive bad SINR because of the great interference coming from the other cells, in order to reach a better SINR our previous algorithm proposed in [D1.3] perform a power reduction on the interfering cells. This satisfies the SINR constraint of users which experience great interference at the price of the user which could be served rapidly and then leaves the network. This reduction of overall throughput has also been resolved by the 3D beamforming technique described in subsection 4.2.3.2. The other point of similarity with this technique is the use of limited coordination area to guarantee the limited exchange of information between the eNBs. However our proposed scheme allows the eNB to accept all the requests of interference reduction from the other eNBs without performing any prioritization cyclically shifted since the fairness among the cells is implicitly acquired.

The following figure shows the throughput performance gain of the proposed algorithm versus the algorithm proposed in [DY10].



**Figure 4.5: The cell spectral efficiency of the proposed algorithm compared to centralized scheme**

In order to have a fair comparison between both algorithms, the algorithm proposed in [DY10] is performed in clustering scheme. However, each cell exchanges the same information as our proposed algorithm but without applying the interference minimization.

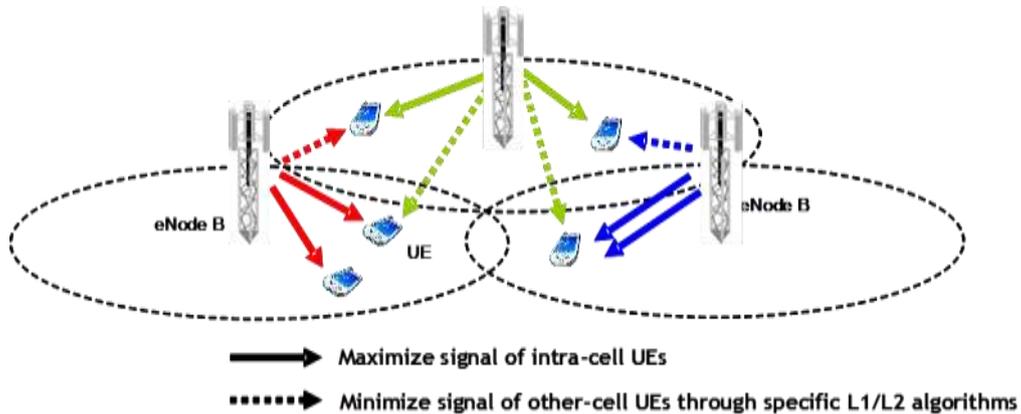
In fact the clusters exchange only  $(N/N_c) \cdot K$  interfered channels which reduces the exchanged information by  $N_c$ .

The simulation parameters, as well as the mathematical derivations, can be found in Appendix A8-2.

It is noticed that the overall throughput is improved by using our algorithm, but only 3% of the users at the cell Edge keep the same throughput as in [DY 10]. Since these users experience a very bad channel, reducing the interference by the other eNB and reducing the power by their own cluster lead to a sort of balance on their SINR. However, in 3GPP 5% of the CDF represents the cell Edge UEs whose performances are improved.

#### 4.2.2.2 Coordinated beamforming for interference rejection

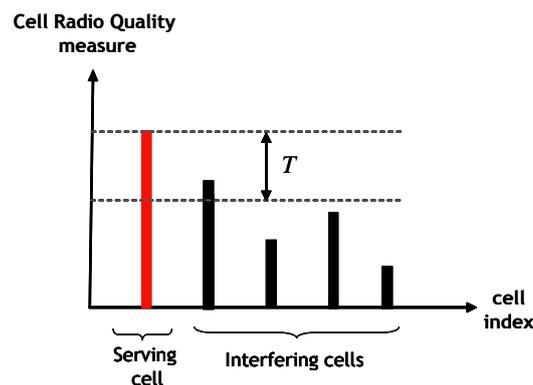
In this section it is described a downlink coordinated beamforming transmission scheme that requires only the exchange of instantaneous Channel State Information (CSI) and control over the backhaul links, without requiring any sharing of user data among the geographically separated transmission points. The user data is transmitted only from the serving cell with the advantage of limiting the latency/capacity requirements over the backhaul. The proposed Interference Rejection (IR) scheme is applicable on top of different Layer 1 MIMO transmission modes like for example Spatial Multiplexing (SM) SU-MIMO, Transmit Diversity (TxD) SU-MIMO, Single Cell MU-MIMO, etc. The scheme can be applied both to cells that belong to the same eNB (intra-eNB CoMP) or to cells that belong to different eNBs (inter-eNB CoMP). On top of this L1 interference rejection scheme, also coordinated L2 interference rejection mechanisms can be applied using a cross-layer design approach. The general principle of the multi-cell L1 interference control scheme is shown in the following Figure 4.6.



**Figure 4.6: The multi cell L1 interference control scheme.**

In particular, each UE monitors, through appropriate measurements, the neighbouring interfering cells identifying those that create the highest level of interference, by means of a threshold mechanism. The interference measure is performed at the UE by exploiting the Synchronization Signals (SS) or the Reference Signals (RS) transmitted by the different cells. Some possible examples of radio quality measures are the Reference Signal Received Quality (RSRQ), the Signal to Interference Ratio of Reference Signals (RS SINR) and the Reference Signals Received Power (RSRP).

The basic idea behind this mechanism is that the network sets, through downlink signalling, a dynamic threshold for each user. The threshold value may be a function of network parameters and/or specific user characteristics. The following Figure 4.7 shows the concept of dynamic threshold based on the cell radio quality indicator (e.g. based on RSRP, RS SINR, RSRQ, etc.). The threshold set by the network is denoted as  $T$ . The user terminal periodically measures the quality indicator of the different cells present in the network including the serving cell and the interfering cells. When the difference between the radio quality indicator of the serving cell and one interfering cell falls below the given threshold  $T$ , such interfering cell becomes a candidate cooperating cell, to be included in the cooperating set.



**Figure 4.7: Threshold mechanism applied in multi cell L1 interference rejection.**

The next step of the IR mechanism is that the interfered UE measures and feeds back to its serving cell the CSI information, together with the cell IDs of the interfering cell(s) that is (are) generating a high level of interference by exceeding the threshold  $T$ . This CSI is related to the radio link(s) between the interfering cell(s) and the UE. The serving cell then forwards this CSI together with the scheduling intentions (i.e. subbands and subframes index) for the interfered UE to the identified interfering cell(s) through the backhauling.

The CSI information and the scheduling intentions are then used by the interfering cell(s) in order to minimize the interference over the signalled resources (subbands and subframes) by shaping the radiation diagram. In particular the beamforming weights used by the interfering cell are calculated in order to maximize the signal quality for its served UE over the signalled resources and, simultaneously, to minimize the interference generated towards the selected other-cell UE by placing a minimum in the antenna radiation diagram. In order to limit the impact on the served users it is also assumed that each cell may activate the interference rejection mechanism over a given set of resources only towards a single other-cell UE. The selection of the specific other-cell UE toward which to place the minimum of the radiation diagram is determined by computing for each interfered UE the ratio between the power received from its serving cell and the power received from the  $j$ -th interfering cell, which has to activate the interference rejection mechanism. In formulas this ratio  $R^{(i)}$  for the  $i$ -th UE can be expressed as follows:

$$R^{(i)} = 10 \cdot \log_{10} \left( \frac{P_j^{(i)}}{P_{serv}^{(i)}} \right)$$

where  $P_{serv}^{(i)}$  is the power received by the  $i$ -th UE from its serving cell and  $P_j^{(i)}$  is the power received by the  $i$ -th UE from the  $j$ -th interfering cell. In the cases where over a given transmission resource there is more than one interfered UE for which the difference between the radio quality indicator of its serving cell and one interfering cell falls below the given threshold  $T$ , the  $j$ -th cell will activate the interference rejection mechanism towards the UE for which the ratio  $R^{(i)}$  is maximum.

Based on the exchanged CSI, each interfering cell then calculates the beamforming weights independently from the other cells. The method for the calculation of the beamforming weights is based on Multi-User Beamforming (MU-BF) [TSS05] with the maximization of the Signal to Leakage plus Noise Ratio (SLNR) [STS07]. The leakage for the user  $i$  caused by the  $j$ -th interfering cell, is defined as the power that this cell transmits to its served own-cell user with respect to the total power leaked from this cell to the user  $i$ .

### 4.2.3 Coordinated Scheduling for Beam Coordination

In this subsection innovative approaches are presented which rely on coordination of radio resource allocation among adjacent cells by making use of CP information exchange between eNBs. Exchange of CP information has less requirements on backhaul capacity than UP exchange. Therefore, the described innovations can provide reasonable performance improvements even if only low backhaul capacity will be available. The delay requirement can be relaxed if exchange of statistical or long-term information is sufficient, e.g. if the antenna elements are correlated. Even if enough backhaul capacity is available, there might be deployment scenarios where the ultimate performance gain versus processing complexity tradeoff is in favour of CP exchange solutions.

Coordination allows to explicitly avoid critical resource allocations causing strong interference. A prerequisite therefore is to identify interference sources in adjacent cells. Specific measurement procedures and feedback information is needed to apply scheduling in an optimized way.

The innovative schemes in this subsection are:

**Impact of coordinated scheduling on interference reduction:** This scheme takes into account the impact of coordination of the schedulers on interference, without considering specific beamforming techniques. It points out the capability of the scheduler itself with respect to interference reduction

**Distributed scheduling for beam coordination:** The scheduling and coordination schemes are extended to take into account also the interference impact depending on the vertical pattern adaptation. A distributed scheduling algorithm applies scheduling constraints to avoid high interference while scheduling UEs proportionally fair.

#### 4.2.3.1 Impact of coordinated scheduling on interference reduction

In this work, we investigate the performance of distributed scheduling based on opportunistic distributed scheduling. It has been shown in [GK11] that the scaling law for the achievable rates in terms of the number of UEs when using a distributed SINR maximizing scheduler, called *max-SINR scheduler*, is the same as the scaling law of an upper bound corresponding to a *max-SNR scheduler* with no interfering cell.

Yet, the analysis done in [GK11] dealt only with the scaling. In this work, we focus on the question to determine whether the rate of convergence is large enough to bring significant improvement at realistic number of users. Furthermore, we also investigate the performance when JP-CoMP is applied on the scheduled UEs. Indeed, one of the main goal of JP-CoMP is to manage interference and therefore it is not known to which extent it will improve the performance when it is applied after an opportunistic scheduler has reduced the amount of interference effectively received by the scheduled UEs.

We consider a multi-cell cellular network where each eNB is equipped with one antenna and transmits to only one UE. There are  $K$  UEs in each of the cells, also equipped with a single antenna. The noise at the UE is a zero mean AWGN and each eNB transmits with its maximal power. We consider a Rayleigh fading channel with a long term path loss effect where only the first ring of interferers is assumed to emit significant interference.

In a first step we discuss algorithms maximizing the performances without any consideration on fairness. We denote these algorithms as *Non-fair* in contrast to the alternative versions described later and denoted as *Fair* in which the Opportunistic Round Robin [KR03] algorithm is used so as to improve the fairness between the UEs.

#### Non-Fair Distributed Schedulers:

The first and main focus of our work is on the *max-SINR* scheduler, which consists in selecting distributively in each cell the UE with the maximal SINR. It is the most interesting distributed scheduler since it increases the gain of the direct link and reduces the interference at the same time. Once the user is selected, the rate can be computed directly by treating interference received as noise.

We also consider the performances of a less elaborate distributed scheduler which only selects the UE with the largest SNR without taking the interference into account and is denoted as the *max-SNR* scheduler.

Additionally, we discuss a *JP-CoMP ZF* scheme where a joint ZF precoder is applied on the eNBs inside a cooperation cluster *once* the users have been scheduled independently at the cooperating eNBs via the *max-SINR* scheduler. Thus, the interference emitted from the cooperating eNB to the scheduled UEs can be suppressed.

Finally, we consider a *no-interference upper bound* where the interference stemming from the eNBs inside the cooperation clusters have been suppressed for free.

#### Fair Schedulers:

A fairer alternative to these schedulers is called Opportunistic Round Robin (ORR) scheduler [KR03]. The principle of the ORR scheduler is to remove the UE from the set of UEs once it has been scheduled. The set of possible UEs is reduced by one and for the next time slot the scheduler is applied only on the remaining set of UEs. This continues until all the UEs have been scheduled and served once. It has for consequence that in  $K$  time slots, each UE is

scheduled one and only one time. Thus, the position of the UEs does not bring any diversity gain and the only multiuser diversity gain is obtained from the Rayleigh fading.

The principle of ORR scheduling does not state which figure of merit is used to select the UE, and we will in fact have an ORR version of each of the previously described schedulers (*max-SINR*, *max-SNR*, *JP-CoMP*, and *no-interference upper bound*).

In Figure 4.8, the average rate is plotted for  $K$  time slots, so that each UE can be served when the ORR algorithms are used. We observe that for both the unfair and the fair versions, JP-CoMP ZF achieves an average rate very close to the average no-intra-interference rate, while the schemes with single cell processing are characterized by larger losses. Still, distributed processing allows to achieve good performance with much lower requirements for the system.

We also note that the *max-SNR scheduler* introduces little loss compared to the *max-SINR scheduler* in both scenarios. However, the difference is much smaller in the case without fairness constraint. In that case, it seems that the two distributed schedulers have the same scaling in terms of the number of UEs. This is to put in relation with the fact that the scheduled UEs are then the UEs located very close to the eNBs, so that the interference power is very small. Consequently, the *max-SNR scheduler* and the *max-SINR scheduler* are likely to schedule the same UEs.

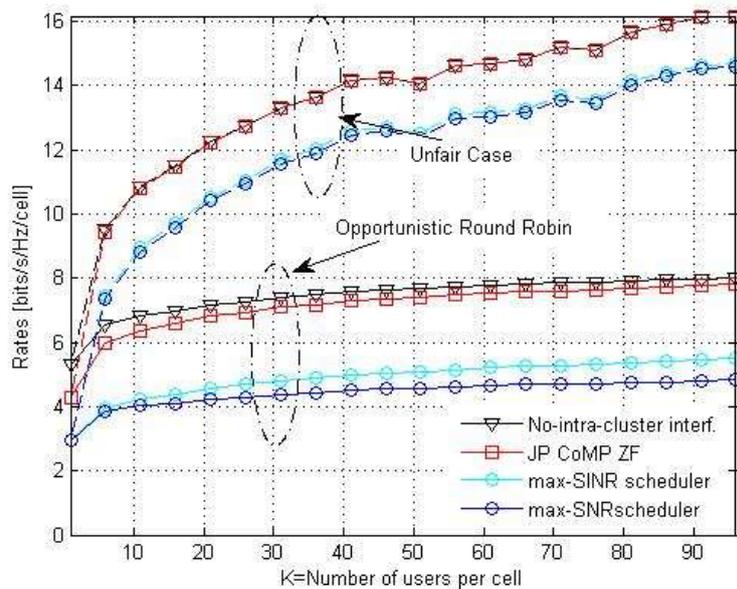


Figure 4.8: Average rate per cell and time slot.

With our simulations, we have analysed in a realistic environment the impact of the scheduler on the interference in a multi-cell scenario. We can see that distributed opportunistic scheduling leads to good performance compared to the ideal case without interference and is an efficient tool to manage interference without much requirement on the architecture. Yet, it cannot by itself manage interference and it has to be complemented by other methods reducing interference, like for example JP-CoMP. An analytical analysis of the rate of convergence and the rate loss due to distributed scheduling can be found in [KG11]. As a byproduct of this work, we also show that JP-CoMP is a useful scheme to manage interference, even when applied after an opportunistic scheduler.

#### 4.2.3.2 Distributed scheduling for beam coordination

Beam coordination aims at avoiding collisions of beams or precoded transmissions from adjacent cells to UEs using the same resources. For UEs especially at cell border this achieves significant interference reduction. Whereas the 3D beamforming presented in section 4.2.1.2

achieves this in a statistical way, this work exploits the combination of 3D beamforming with coordination to explicitly avoid beam collisions by putting appropriate constraints on the involved schedulers. While such constraints can reduce the interference from adjacent cells and improve the individual performance per scheduled UE, they might have also the negative impact to reduce the overall throughput if they restrict the number of simultaneously scheduled UEs too much. 3D beamforming is a means to overcome this by adding an additional degree of freedom for such constrained scheduling decisions to relax the restrictions for the UEs.

Scheduling with beam coordination optimizes system performance by assigning the same time and frequency radio resources to UEs in adjacent cells only if they can be sufficiently separated in space. For the space dimension horizontal and vertical beam parameters are adapted. Full reuse 1 is achieved through finding UEs for assignment of all available time and frequency resources in each cell, taking into account the spatial separation constraint.

The optimization ideally maximizes the overall cell throughput, while reducing the discrepancy between cell edge and mean cell throughput. It should also be possible to take into account a proportional fair constraint for the UEs.

The SINR of a scheduled serving cell UE (1 rx antenna assumed) is given by

$$SINR_i = \frac{s_i \cdot H_i^T \cdot P_{i,m}}{\sum_{\substack{j=1, \\ j \neq i}}^J (s_j \cdot H_{j,i}^T \cdot P_{j,m}) + N}$$

$s$  = transmitted symbol

$i, j$  = index of serving cell, interfering cells, respectively,  $J$  = number of coordinated interfering cells considered

$H_i^T$  = transposed channel vector between serving eNB and UE in serving cell

$H_{j,i}^T$  = transposed channel vector between interfering eNB and serving cell UE

$P_{i,m}$  = precoding vector  $m$  of eNB  $i$  out of  $M$  vectors comprising all combinations of horizontal precoding and available downtilts

$N$  = received noise power at serving cell UE

To maximize the cell edge throughput, an approach could be to maximize the minimum SINR for all UEs  $i$  within a coordination area for a specific scheduling decision by choosing suitable combinations  $P_{i,m}, P_{j,m}$ . Solving this problem is not straightforward and turns out to be computationally complex, especially since in proportional fair scheduling case dependencies between the selection of the  $P_{i,m}, P_{j,m}$  and the score of the UE occur. It will also cause intolerable delays.

Therefore in ARTIST4G more practical solutions and approximations like “implicit coordination” and “distributed horizontal and vertical beam coordination” have been analyzed, which either rely on pure statistical channel properties or approximate the target solution with reasonable performance and complexity.

### Implicit coordination:

This scheme works without any control information exchange between the eNBs. The radio resources are assigned according to the location of the UEs within a cell with the objective to minimize the interference. For this, orthogonal resources are assigned to terminals located close to the cell borders of adjacent cells. In this way coordination is achieved by a predefined configuration of a location-based resource assignment scheme. In a hexagonal cell scenario, this pre-definition area comprises 3 cells of a site and is repeated all across the network.

Each cell schedules its UEs individually based on location information of the UEs. Two specific resource assignment configuration schemes are shown in Figure 4.9 as examples. Each color

represents a subset of radio resources, which is preferentially assigned to UEs located in the respective area.

The first scheme, called “vertical sorting” (Figure 4.9 a), relies on a sorted list of UEs according to the vertical angle. The resources are assigned according to the vertical angles. So in this case the adaptation of the vertical beam pattern can coincide with the three areas of the different radio resources, but in general resource allocation and downtilt adaptation operate independently. For the second scheme called “horizontal sorting” (Figure 4.9 b), the UEs are sorted according to the horizontal angle, and the resource allocation is accordingly. The procedure for downtilt adaptation remains the same as for vertical sorting.

Performance was studied with a macro-cellular system scenario according to the 3GPP recommendation [3GPP36814] with 500 m inter-site distance and 12 randomly distributed UEs per cell. This scenario was extended with a UE-specific vertical downtilt adaptation and a vertical HPBW of 10°. Detailed descriptions and results can be found in the appendix section A3-2 and in [ARTD13]. Both downtilt adaptation and vertical or horizontal sorting reduce the interference, in particular for cell edge users, already on their own. The combination of both leads to additional gain, which is most significant for horizontal sorting in combination with 3D beamforming. With the specific simulation scenario used, up to about 40% additional coordination gain in cell edge throughput and 12% in spectral efficiency could be achieved on top of the pure stastical 3D beamforming gain.

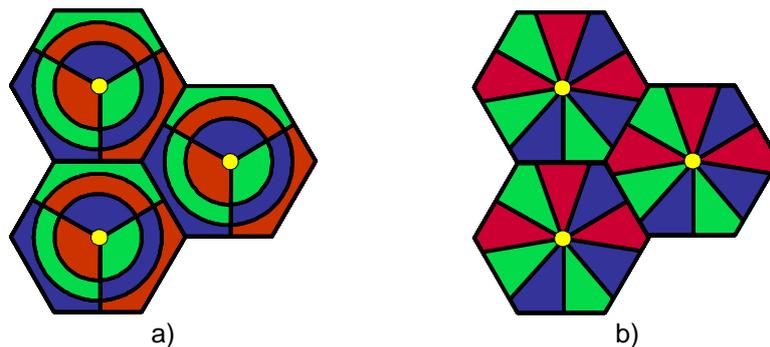


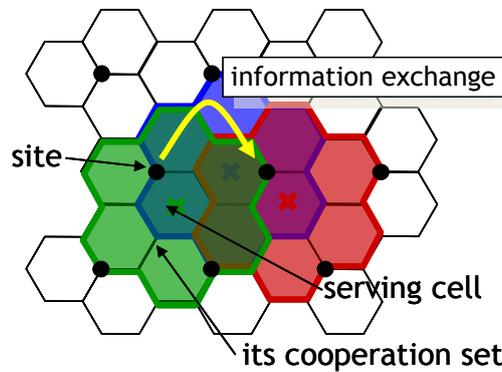
Figure 4.9: Implicit coordination schemes: a) vertical sorting, b) horizontal sorting

#### Distributed Horizontal and Vertical Beam Coordination:

The benefit of “Implicit Coordination” is exploited best if the UE distribution is assumed homogeneous across the cell area, so that all available resources can be used efficiently. If this is not the case, it appears that resources primarily reserved e.g. for the area close to the eNB have to be assigned also to farther UEs or vice versa, which may reduce the coordination gain. The following section describes a more advanced scheme aiming at dynamically optimizing the resource allocation depending on the UE locations and their scheduling priorities according to their score in a proportional fair scheduling scheme. So this is coming closer to the ideal optimization.

This scheme uses a distributed algorithm working on overlapping coordination areas around each cell represented by groups of seven cells (Figure 4.10) in the hexagonal scenario according to [3GPP36814]. It assumes the availability of explicit information on interference experienced by UEs in adjacent cells, for each beam pattern applied in the serving cell. Derived from an algorithm originally intended for pure horizontal interference coordination [MF11], this scheme has been extended towards 3D beamforming. Therefore the interference information is related to horizontal precoding as well as vertical downtilt adaptation. For example, if each UE would report the worst case interfering precoding matrix index (PMI) of each adjacent cell for each possible vertical antenna pattern to its own eNB, this information can be exchanged between the eNBs. The schedulers of the interfering cells, which are assumed to be weighted

proportional fair schedulers, can then accept some additional constraints to avoid unfavourable resource allocations causing high interference.



**Figure 4.10: Coordination areas**

The fact that each cell can be serving cell and part of another coordination area reflects the distributed nature of this scheme, but causes also conflicts between setting and accepting scheduling constraints. These conflicts can be resolved using a prioritization of the serving cells. Some of them are allowed to set constraints, whereas others have to accept constraints according to a cyclically shifted priority assignment guaranteeing fairness among the different cells.

The limited coordination area of 7 cells guarantees a limited exchange of information between eNBs. Each eNB corresponds with 6 surrounding eNBs. Based on this exchanged information, each eNB scheduler runs an individual algorithm optimizing the resource assignment of its cell. So this is a simplification towards a partial optimization over the network.

System simulations as shown in Figure 4.11, where the parameter  $\alpha$  is a fairness criterion, showed that with downtilt adaptation gains in spectral efficiency and cell edge throughput are achieved on top of pure coordination gain with fixed downtilt. Gain due to pure 3D beamforming without any coordination (blue curve) is already superior to the gain with optimized fixed downtilt and coordination with the optimum number of constraints (red curve). If to the 3D beamforming also coordination is applied (green curve), additional gains are achieved. This is the case even with the simplified model of two fixed downtilts instead of exact steering. Detailed system description and results have been partially shown in [ARTD13]. Additional results are presented in the Appendix A3-2.

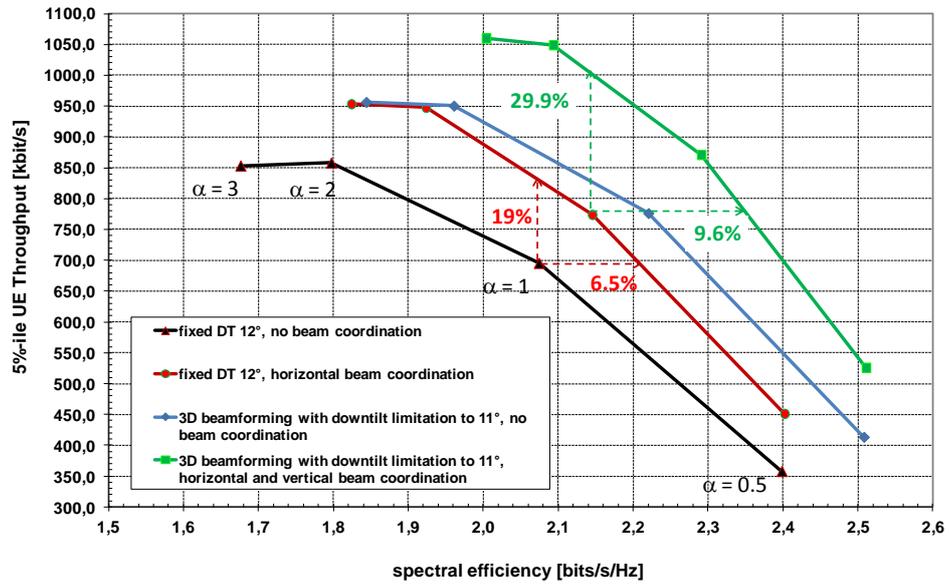


Figure 4.11: 3D Beamforming simulation results for direct steering with downtilt limitation of 11 °

### 4.3 Conclusions

This section 4 provides innovations which require either no or only a limited exchange of CP information between eNBs. This is in contrast to the JP CoMP schemes addressed in section 5, which depend on the capability of challenging UP data exchange over the RAN. We considered schemes, which involve different mechanisms and system parameters, and which have different levels of complexity and also different requirements on the architecture and deployment costs (see [ARTD43]).

The schemes grouped under the subsection 4.2.1 "Advanced Beamforming" achieve performance improvements through specific adaptation of the beam pattern, either with optimizing Tx and Rx filter design for a horizontal beam pattern ("transmit and receive filter design with limited signalling"), or with adaptation of the beam pattern also in vertical direction ("UE specific horizontal and vertical beamsteering").

In the next step addressed in subsection 4.2.2 "Coordinated Beamforming" schemes are considered, which optimize precoding of the horizontal antenna pattern. One solution ("Precoding optimization algorithm for coordinated beamforming") optimizes the antenna pattern centrally over the whole coverage area. A second solution ("Coordinated beamforming for interference rejection") requires measurements of adjacent cell interference and exchange of scheduling intentions to optimize the precoding vectors for interference minimization.

In a further step, also the schedulers have been involved. The investigation on "Impact of coordinated scheduling on interference reduction" shows that even without antenna pattern adaptation a reduction of interference can be achieved here based on opportunistic distributed scheduling. Finally, a distributed scheduling algorithm is combined with vertical beam pattern adaptation. This algorithm also requires adjacent cell interference measurements and exchange of scheduling constraints among eNBs. The scheduling constraints exchange can be envisioned for multi-vendor solutions, as an extension in future releases of the 3GPP LTE of the signalling exchanged over the X2 interface for ICIC.

With these innovations of relatively low complexity compared to JP CoMP, already reasonable gains in specific application scenarios have been shown and could be partially verified with field measurements. Some of these innovations are transparent to 3GPP standards (even Release 8 and 9) or require only minor enhancements. Therefore they will be applicable at an early stage. Others will require extensions of the standard, especially those requiring adjacent cell interference measurements. Additionally it can be pointed out that the schemes relying on multi-cell coordinated beamforming described in this section require multiple antenna elements or directional antennas at each eNB.

Although the innovations described in section 4 can not provide those ultimate performance gains as promised by the JP CoMP schemes addressed in section 5, they can be considered as entry schemes feasible with moderate network and system enhancements, even at a stage where the required advanced standard features and RAN capabilities needed for JP CoMP are not fully available.

## 5 Advanced Joint Transmission Schemes for Multi cell cooperation

### 5.1 Introduction

Section 4 has introduced novel schemes for interference avoidance by inter-cell coordination. We here explore a more challenging set of techniques: Joint multi-cell transmission of downlink messages to terminals in OFDM systems. Such Coordinated Multi-Point (CoMP) transmission techniques require joint processing of user plane data and are here denoted joint transmission CoMP, or JT-CoMP<sup>1</sup>

#### 5.1.1 Introduction and Overview of the Joint Transmission Framework

Joint transmission CoMP requires the user-plane data as well as jointly designed transmitters to be available at all participating nodes. It can thereby provide three potential benefits:

- **Macro-diversity gains:** Transmission over multiple independent fading channels can reduce the effect of fading by averaging. Macro-diversity transmission from different sites is particularly effective in reducing the effect of shadow fading.
- **Beamforming gains:** Coherent combining of transmissions from multiple antennas can increase the average received SNR.
- **Interference cancellation:** Interference can ideally be eliminated by joint coherent transmission among cooperating nodes: The phases and gains are adjusted so that the sums of interfering signals cancel at all receivers.

*Coherent* JT CoMP can give all three types of benefits also with a single antenna or a few antenna elements at each base station, but it is a challenging task: It requires synchronized base stations and a set of jointly designed beamformers. This, in turn, requires channel state information (CSI) at the transmitters of sufficiently high quality. It also requires the distribution of the jointly designed beamformers to all transmitters without undue delays.

Macro-diversity gains can also be harnessed by less powerful but also less demanding and more robust techniques named *non-coherent joint transmission*. In our integrated design, these techniques can be used as a fallback transmit mode for terminals with very uncertain channels.

Joint coherent downlink transmission has promised large gains in theory in multiple studies over the last years but, as discussed in the beginning of Subection 5.2 below, the gains obtained in more realistic system-level studies within 3GPP have been much more modest [3GPPTR1]. This has raised the question to what extent the theoretically motivated gains could be attained in a practical system. This question has led to an intense research effort within the ARTIST4G project. The work gradually revealed several key ingredients and mutually reinforcing design principles that *do* enable large performance gains. The final result is a design framework that is condensed into four points below. Under realistic conditions, this framework will still not attain the full gains promised by theory, but it does provide quite substantial gains.

For example,

75%-100% gains of sum throughput over cellular transmission seem attainable for 4 x 2 MIMO links (Subsection 5.4).

---

<sup>1</sup> A more general concept is *joint processing* CoMP, or JP CoMP. It includes coordinated scheduling (CS) or fast cell selection as well as joint transmission (JT) and other conceivable schemes, as well as uplink joint reception. Joint reception is outside of the scope of this document. Work on this topic has been performed within Workpackage 2 of ARTIST4G, with the final results reported in [ARTD25].

A sum throughput improvement of around 55% was measured for single-antenna transmissions over measured channels in a fully loaded 3-site CoMP setting in Appendix A2-2. The result was obtained with quite large channel impairments, 10 ms channel outdating at 5 km/h velocity at 2.66 GHz.

The JT CoMP framework that enables such gains is based on four mutually synergistic design principles for joint transmission in OFDM downlinks:

1. ***Use enlarged clusters of coordinated cells, with means for inter-cluster interference suppression (interference floor shaping).*** Due to backhaul limitations and control overhead, global joint optimization of downlink transmissions is unrealistic. Cooperation has to be limited to cooperation sets (clusters) of transmitters. But the use of intra-site CoMP or CoMP within small clusters can give only limited gains due to inter-cluster interference. In Subsection 5.2.1 below, we outline a clustering scheme that uses larger cluster sizes (typically 3 sites with 9 cells) and gives good control of inter-cluster interference. Only the strongest channels are measured and used, to limit the complexity of channel prediction and feedback schemes (“Partial CoMP”).
2. ***Scheduling and beamforming design is made independently at each cellular base station. This will directly generate a grouping of users at each frequency that have good channel properties for joint transmission.***

This “cellular user grouping” principle is outlined in Subsection 5.2.2 and in A2.2.5 in Appendix A2-2:

- Scheduling and beamforming is performed separately (but possibly coordinated) within each cell of the Master base station. We don't need new algorithms for this purpose, although advanced (possibly coordinated) algorithms are required for the highest performance. This facilitates a smooth introduction of JP CoMP in near-future systems. It will naturally place users that are closest to the same base station in *different* frequencies and/or beams.

- As a fortunate but very important side-effect, the groups of users within the cluster that will thereby share the same frequency resource will be well suited for JT CoMP: They will have well-conditioned channel matrices that become relatively easy to pre-invert by joint beamforming at the transmitters. This increases the attainable performance by coherent JT CoMP with linear precoders.

- The scheduling may be readjusted over the cluster to further enhance performance.

3. ***Use coherent robust linear precoding, with single-cell transmission or non-coherent joint transmission as fallback modes.***

A large number of precoders (for different resource blocks) need to be calculated for a wideband OFDM downlink. While nonlinear precoding is theoretically optimal, only linear precoding seems to provide feasible numerical complexity.

- Coherent joint transmission with linear precoders is proposed to be used wherever feasible. The precoders can be designed to optimize performance by averaging (marginalizing) the design criterion over uncertainties in the CSI due to prediction errors, feedback compression and other causes, see Subsection 5.2.3 below. This improves performance in non-ideal settings.

- However, the CSI will be outdated for high-velocity terminals and for large time delays even if prediction is used. Coherent JT CoMP then cannot be used. Non-coherent JT CoMP together with coordinated scheduling is here a higher-performance alternative scheme, while cellular transmission (with possible coordinated scheduling) would be the simplest alternative fallback mode.

4. **Use prediction of channel state information.** The coordination delays within a cluster of cooperating nodes may be significant. The channel state information at the transmitters, which needs to be quite accurate, is then degraded by outdated. The use of channel prediction based on noisy past and present channel samples, as discussed in Subsection 5.3.2 and A4-2 below, can predict the channel at the future instant when the downlink transmission will take place. This can compensate for backhaul time delays and extend the range of mobility for which JT CoMP can be used. Compared to using outdated estimates, prediction can almost triple the useful terminal velocity for a given delay, or delay for a given velocity. The interference floor shaping facilitates channel prediction.

An integrated design that is based on the above principles is presented in Appendix A2-2, and it is accompanied by a performance evaluation based on channel sounding measurements. The balancing of the parts of the design, and a discussion of the attainable performance gains, can be found in Subsection 5.4.

We have here focused on FDD downlinks in OFDM systems and also on homogenous deployments, in regular patterns of macro cells or micro cells. Indoor users have not been in focus as CoMP gains are small in noise-limited conditions. Noise-limited indoor users are better served by solutions that provide better indoor coverage such as indoor base stations.

Apart from the core framework, the work has produced a number of further important insights. In particular, we have studied the problem of decentralized design of coherent linear precoders, based on not fully consistent distributed channel state information, see Appendix A4-1. Robust MIMO precoding design schemes for this problem are presented in Subsection 5.3.1 and Appendix A2-7. We also discuss feedback requirements for FDD downlinks (Subsection 5.3.3, A2.2.8 in Appendix A2-2 and Appendices A4-4 and A4-5).

**The following appendices are directly related to Chapter 5:**

A2-1: Robust linear precoding with per-base-station power constraints.

A2-2: An integrated design for downlink Joint Transmission CoMP.

A2-3: Joint scheduling and power control with non-coherent transmission.

A2-4: Waterfilling schemes for Zero-Forcing coordinated transmission.

A2-5: Dynamic partial joint processing.

A2-6: Resource allocation for OFDMA Joint Processing CoMP.

A2-7: Robust precoding with distributed channel state information.

A4-1: Centralized/decentralized joint transmission with limited signalling information.

A4-2: Kalman prediction of multi-site MIMO channels for CoMP.

A4-3: Advanced channel prediction.

A4-4: Feedback compression.

A4-5: Advanced feedback compression schemes.

A5-1: Clustering based on partial CoMP.

A5-2: Inter-cluster coordination with fractional frequency reuse.

A8-2: Scheduling aspects of partial CoMP.

### 5.1.2 Background Assumptions and Relations to Scenario 1 (Section 4) and to Scenario 3 (Section 6)

JT CoMP requires fast user plane (UP) as well as control plane (CP) information exchange between the transmission points of a cooperation area (CA). The transmission points of the CAs will typically be spread over distributed locations of the network as this promises higher performance gains as compared to single site (intra-site) cooperation. The networks will here be assumed to be homogeneous and interference limited. Typically, an urban macro scenarios with similar parameters as the so called 3GPP LTE-Advanced case 1 [3GPPTR2] and inter-site distances (ISD) in the range of 500m is considered, but with an important difference. For ensuring an interference-limited scenario, the outdoor-to-indoor penetration loss has in simulations been set to 0 dB instead of 20 dB. In addition a carrier frequency  $f_c$  of 2.6 GHz instead of 2.0 GHz has been used.

The implementation of the fast exchange of UP and CP information is not the main focus of this study. It could be realized either by an appropriate X2 interface, relying e.g. on GB Ethernet or based on remote radio heads (RRH). For further details see Deliverable [ARTD43] from ARTST4G WP4 or solutions as being developed for different CoMP experimental systems [MF11].

The required CP information exchange might include scheduling requests, channel state information (CSI), rank indication (RI) and channel quality information (CQI). The required speed of exchange of UP and CP messages is related to the short term channel variations. It is therefore even for low mobility in the range of 1 to a few ms.<sup>2</sup>

In previous Artist4G deliverables [ARTD11], we distinguished between User-plane (UP) COOP, non UP COOP and Hetnet as being applied for Scenario 2, 1 and 3 respectively.

Contrary to Scenario 1 (non UP COOP) as described in Section 4, the here investigated Scenario 2 (UP COOP) assumes the presence of user data at multiple transmitters, so that joint precoding and transmission can be used. It focuses on finding and approaching the limits for interference mitigation, at least partly disregarding the resulting complexity for the time being. The goal is to find in principle implementable and realizable solutions.

Joint transmission CoMP promises higher gains – e.g. compared to Scenario 1 -, but it is at the same time the most challenging approach. For that reason it is seen as a mid- to long term research topic with the goal to approach the upper bound for interference mitigation. It should be noted that, in contrast to systems with network-wide precoding [BS06] this upper bound is not even known for clustered networks yet.

---

<sup>2</sup> These values refer to the X2 interface itself without accounting for processing and protocol delays. Such low values are realizable in many urban environments as routers and switches today can achieve processing delays of about or even below 100 $\mu$ s and fiber delays are just 3.3 $\mu$ s per one kilometre of fiber.

We see the potential to provide gains even in fully loaded networks as an important benefit of JT CoMP. Coordination schemes typically work best in medium to low load conditions<sup>3</sup>. Coordination schemes inherently require a degree of freedom for reallocation of resources.

Scenario 3 - as being described in Section 6 - concentrates on heterogeneous networks (Hetnet) scenarios. Partly different problems are then of relevance, such as coordination of closed subscriber group femto stations, or mutual interference between femto- and macro stations. It can be seen as an effort in parallel with UP COOP for boosting the capacity by adding many additional radio stations. An interesting next-step challenge is then the introduction of cooperative transmission in heterogeneous networks.

We now introduce the basic motivations and background assumptions in more detail, after which the elements of the framework are explained in Subsection 5.2 and Subsection 5.3.

### 5.1.3 Join Transmission CoMP: Promises and Challenges

Currently, mobile network operators (MNO) observe an exponential increase in user data traffic, generated by an increasing number of users, smart phones, tablets etc. Current technologies like UMTS HSPA and its evolution are still able to cope with the demand, but the traffic growth in the next years requires significantly higher network capabilities.

A simplified prediction/vision of this capacity expansion in [NSN11] is stating a factor of 1000 in 10 years. No single technology is likely to be able to provide such large gains economically. Several means have to be combined, in particular increased spectral efficiency, expanded spectrum and higher number of sites or cells per area – that in combination could attain the goal, with  $10 \times 10 \times 10 = 1000$  being one of many discussed proportions.

Each of these factors poses separate challenges.

An increase of the number of sites is closely related to local area offloading over Picocells, home eNBs or Femto stations. This leads to the heterogeneous network (Hetnet) scenarios as discussed in scenario 3, chapter 6. Deployment of large numbers of Pico and Femto-cells is not without costs, as it requires new sites and infrastructure, but the investment will typically be lower than that for a macro site.

Spectrum is mainly a regulatory issue. Additionally it requires large investments by the MNOs as can be seen from the latest spectrum auctions e.g. in Germany or Italy in 2011.

As a conclusion, in the longer run a significant leap in spectral efficiency for mobile radio networks is a desirable goal. Ideally, these gains in spectral efficiency should be achieved for the already deployed networks, by 'just' updating the eNB soft- and/or hardware, maybe accompanied by the installation of few further antenna elements (AE). Therefore the focus in the present chapter is on homogeneous macro cellular networks, on OFDM downlinks and on frequency division duplex (FDD) as the most widely deployed system.<sup>4</sup>

---

<sup>3</sup> Under special conditions the opposite will be true: Coordination will work best when there is a lot of multi-user diversity to exploit, and thus work best in high load conditions. This would be the case when the number of potential users increases with the load, when fairness between the users is unimportant and when delay constraints are unimportant. Then, coordinated scheduling and/or beamforming schemes would always be able to select compatible users for transmission, and this ability would increase with the number of potentially available users. However, in the scenarios we consider to be of most interest, these conditions will not be true. Even when the load is high, there will be relatively few (high-data rate) users that need to be served within rather short latency limits. Coordination schemes are extremely challenged under such high-load conditions, while cooperative transmission offers a potential solution.

<sup>4</sup> Time division duplex (TDD) will be getting more important in the future, but in the context of CoMP, TDD is the simpler variant because channel reciprocity makes it easier to achieve the transmitter side channel state information (CSIT). Otherwise we can apply more or less all means to be discussed here to FDD and TDD in similar ways. Regarding channel reciprocity in TDD, one should be aware that *interference* conditions at UE and eNB are different as there is a certain outdatedness of CSI information between UL and

A large increase in spectral efficiency is a real challenge, as HSPA as well as LTE have improved. For time-invariant channels, single link performance is already close to the well known Shannon bound, while multi-link and system spectral efficiencies have steadily improved over time. LTE as such profits for example from highly sophisticated schedulers, exploiting opportunistic multi-user scheduling gains with respect to frequency selective radio channels and interference conditions.

To improve the spectral efficiency, the two main tools are the link budget and control of the interference. Link budgets can be improved mainly by using small cells and by improving indoor coverage to avoid outdoor-to-indoor penetration losses.

Mutual interference in cellular radio systems in typical urban and macro, micro and pico cellular environments remains a main limiting factor for obtaining higher average as well as cell edge throughputs at high system loads. Theoretical results [KFV06],[FKV06] state that with perfect channel knowledge at the transmitter, full network wide cooperation between all transmitters of a network [SB05] enables the complete elimination of interference and even the exploitation of interference for useful data transmission [JJT+09]. Joint downlink transmission thereby promises to outperform conventional cellular networks by several hundreds of percent [TWS+09] with respect to spectral efficiency or coverage. Reduction of the interference could also boost the throughput gains attainable by MIMO processing.

Network-wide cooperation is typically infeasible in practical networks due to exploding overhead for reporting of channel components and extremely high backhaul requirements. Therefore, the network has to be partitioned into so called *cooperation areas (CA)* or *clusters*. This makes the reporting and backhaul problems manageable, at the price of introducing a separate problem: *inter-cooperation area interference*, which can easily spoil most of the cooperation gains. Therefore the main fundamental question motivating the here described research within the Artist4G project has been:

***'Is there a fundamental limit to interference mitigation in practical networks with clustered cooperation areas and how to approach this limit'.***

By an in-depth step-by-step analysis of typical conditions and effects in typical networks, an overall interference mitigation framework has emerged from the Artist4G collaboration, where joint precoding (JT CoMP) is only one – even so very important – piece of the overall puzzle.

In the project we have used a two-step approach.

Our first goal was to maximize system level gains by an advanced interference mitigation framework under ideal conditions, meaning ideal channel knowledge, ideal time and frequency synchronization, ideal backhaul with low delay and infinite capacity.

In parallel and based on the requirements found for the framework, we also worked on the main enablers, i.e. channel estimation and optimum reporting of channel components.<sup>5</sup>

In the end, both parts have to be combined to align achievable gains with in principle possible channel estimation accuracy and feedback overhead to come to an implementable real world concept.

Furthermore one has to keep in mind that such an interference mitigation framework will be introduced in real systems only in case relatively low-effort standardization strategies can be

---

DL transmission. Another problem for TDD and FDD is that transmission of sounding reference signals from many UEs might cause significant overhead and eventually drains the UE power supplies. Therefore, also TDD systems may gain from additive CSI reporting, as is being discussed in [SGS+09].

<sup>5</sup> Some of the implementation issues, in particular time- and frequency synchronization and the effect of phase noise have had a lower priority. These problems are known to be solvable, as verified by several real world 'CoMP' testbeds and demonstrations [MF11], even so potentially accompanied by some extra costs.

found, with good backward compatibility to e.g. LTE Release 8, 9, 10 or beyond as the currently most important mobile radio standard.

Transmitter-side processing is not the only tool for suppressing interference. Another natural option for cancellation of inter-cell interference would be to improve UE receiver capabilities. This can be done by increasing the number of receiver antennas and using advanced interference rejection combining (IRC) schemes. The work within this field in ARTIST4G is summarized in [ARTD25]. IRC is in the present investigation regarded as one piece of the puzzle, helpful for cancellation of far off interferers from outside of the cooperation areas, but it is not by itself adequate as the stand-alone solution.<sup>6</sup>

In our experimental results, we have assumed no use of IRC or other interference cancellation schemes at the receivers.

Still another way to mitigate interference might be to rely on multi-user scheduling gains: UEs can be scheduled over many cells in a coordinated manner, using coordinated scheduling (CS) so that the mutual interference will be low. It is known that asymptotically, the ergodic capacity then approaches that of an interference-free system in the limit of large numbers of potential users. For practical systems, it is important to understand how this limit is approached under typical radio conditions. As explained in Section 4 and in detail in [GK11], for practical numbers of UEs and reasonable time delays there remains a significant performance gap compared to interference-free or interference-exploiting systems.

---

<sup>6</sup> By adding antennas, the UEs would be able to cancel more interferers. The resulting performance will depend on their spatial degree of freedom and the spatial structure of the interferers. While this is in theory a powerful concept, having more than 2 UE antennas is technological challenging in small devices and is expensive as well. In the future we have to expect a variety of UEs where only a minority have more than 4 antenna elements. This makes it difficult to rely on this technique alone, since in macro cellular environments, the number of relevant interferers is typically large.

## 5.2 The JT CoMP Framework with its Building Blocks

The current status of cooperative networks – especially as being discussed in 3GPP with its CoMP study items and work items - is to some extent unsatisfying. In different studies of network wide cooperation [FKV06, JJT+09], large performance gains were found and predicted. In more realistic system level simulations, where cooperation is limited to cooperation areas of reasonable size, real schedulers are applied and/or more realistic loads as well as channel conditions are assumed, these purely theoretical results could not be verified so far. JT CoMP performance gains were quite low. Additionally, they would easily be spoiled by the introduction of realistic channel estimation and delayed feedback of channel information.

The report [3GPPTR1] contains the outcome of the extensive work done in 3GPP on CoMP performed during 2010-2011. Taking for example a 4x2 system from these results, with 4 Tx- and 2 Rx-antennas per cell and per UE, the MU-MIMO reference case achieves a maximum spectral efficiency (SE) of **3.1** bit/s/Hz/cell. The corresponding best-case JT CoMP system attains **4.0** bit/s/Hz/cell (including LTE physical layer overhead), indicating an average cell throughput gain of about 25-30%. At the same time, theoretical results for network wide precoding predict 15 to even 20 bit/s/Hz/cell (excluding overhead) for moderate signal to noise (SNR) ratios.<sup>7</sup>

A goal of Artist4G was to find the fundamental reasons for this large discrepancy and, more importantly, to overcome some of the identified shortcomings. The large performance gap has motivated us to concentrate on first order effects to identify the most relevant issues, while a further fine-tuning is left to future research.

It was found that not a single method or tool - e.g. like just enlarging the size of the cooperation clusters or defining a novel precoding scheme - will be sufficient. Instead, we gradually developed a whole framework and eventually obtained promising gains in the range of **75%-100%** over the above mentioned reference case. In a first assessment, a spectral efficiency of **6-7** bit/s/Hz/cell could be achieved under the assumption of ideal channel knowledge, full buffer traffic and a penetration loss of 0 dB.<sup>8</sup>

We will call the here developed and used advanced interference mitigation framework '**IMF-A**'.

### The Importance of the Signal-to-Noise Ratio

An important difference to the above mentioned 3GPP CoMP reference case (case 1) is the outdoor-to-indoor penetration loss (PL) of 20dB. Such a high PL is realistic, but according to our simulations it will for many UEs result in noise-limited instead of interference-limited conditions. As a result, a considerable part of the potentially attainable gains for the interference mitigation framework might be hidden by a noise floor.

To benefit from advanced interference mitigation, one has to ensure a sufficiently high signal to noise ratio (SNR). There are different means to approach this goal.

The easiest way is to limit advanced interference mitigation to outdoor UEs and to combine the overall system with indoor off-loading solutions like home eNBs or femtocells (see Section 6). Another direction is indicated in [3GPPTR1] Scenarios 3 or 4, where the macro eNB are supported by several low-power remote radio heads.

---

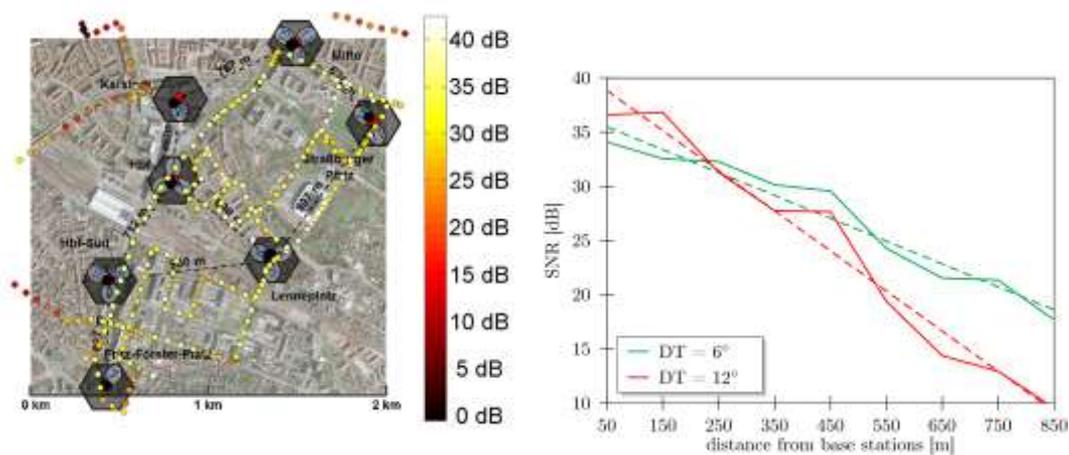
<sup>7</sup> The higher values would need nonlinear precoding techniques such as Dirty paper coding.

<sup>8</sup> JP CoMP is known to be sensitive to channel estimation errors [ZMS+09]. Nonetheless the above comparison with MU MIMO is fair as the reference case has been simulated with the same assumption of ideal channel knowledge. Nevertheless for a practical viable solution realizable channel estimation and reporting schemes with low to moderate overhead is one of the key research topics in Artist4G. These topics are discussed further in Subsections 5.3 and 5.4.

At the UE side, increasing the number of antenna elements and a maximum ratio combining (MRC) receiver would increase the Rx-power, i.e. improve the SNR at the cost of higher complexity.<sup>9 10</sup>

Figure 5.1 provides some real-world SNR results for an ISD of almost 1000m (left) showing the average mean SNR for a downtilt of 6° and 12°. Outdoor values of SNR > 30 to 35dB are quite common up to a distance of about 250m.

The system-level simulations for JT CoMP within the Artist4G project have been performed for a carrier frequency  $f_c$  of 2.6 GHz while 3GPP case 1 is for 2 GHz. Therefore, the Rx-power will be slightly higher for case 1 compared to our simulations. Furthermore, a decrease of the carrier frequency  $f_c$  to 800MHz would further benefit the SNR of the UEs. So TV white space might be an interesting scenario for IMF-A, where spectral efficiency is of great interest due to the limited available bandwidth.



**Figure 5.1: Left: SNR [dB] for strongest eNB in TUD tesbed; Right: average SNR [dB] versus distance from base stations, for two antenna downtilt (DT) values. Averages over results at 16 BSs. The dashed lines show linear fits.**

## Main Challenges

For achieving the mentioned high performance gains we had to develop advanced clustering schemes that ensure user-centric cooperation for almost all users. At the same time, the inter-cluster interference had to be minimized. Reducing the overall interference floor in the system with respect to intra- as well as inter-cluster interference was another high level goal. The use of appropriate types of user grouping and scheduling proved to be very important and powerful tools. Another task was to find precoders with high performance, low complexity and at the same time robustness against imperfections. Finally, as discussed above, interference mitigation by its nature provides gains only in interference limited scenarios. We assume that the system is working under such conditions.

<sup>9</sup> Increasing the eNB height is another way to reduce the typical pathloss values and might be an option for the rollout of new networks.

<sup>10</sup> Compared to case 1 in ITU, sometimes a few dB lower antenna cable loss is allowed, which is a typical implementation issue. Similarly helpful would be a reduced UE noise figure, where 9 dB has to be assumed in 3GPP evaluations.

## Main Enablers

The following main enablers will furthermore be needed to transfer the above mentioned ideal gains into gains of a real world system.

High accuracy multi-cell channel estimation and prediction will decide about the success or failure of the advanced interference mitigation schemes. This topic is discussed further in Subsection 5.3.2.

Channel prediction has to be supported by a feedback concept with low overhead due to the use of strong – but robust – compression techniques. We also need minimum possible feedback delay to minimize CSI outdated. See Subsections 5.3.3 and A2.2.8 in Appendix A2-2 for a further analysis of these aspects.

Backhaul links are required with sufficient capacity for exchange of UP and CP data with very short delays. While this can be seen to a large extent as an implementation issue, specific precoding solutions might help to relax the requirements on backhaul. This question is also related to the choice of centralized versus distributed calculation of joint precoders (beamformers) for the cooperation area. See Subsection 5.3.1.

Of special importance in this context are reliable channel prediction techniques to minimize the performance loss due to outdated of channel state information (CSI). State of the art Wiener and Kalman filtering is used to find the maximum user mobility, while still supporting the above mentioned large performance gains with acceptable degradations, see Appendix A4-2. Simultaneously, a novel powerful scheme, the so called model based channel prediction, has been studied, see Appendix A4-3. It might in the longer term help to stretch the possible mobility to larger speeds and beyond current state of the art with simultaneously low to very low feedback overhead in the future. Advanced channel prediction schemes for vehicles have also been developed, primarily for use in moving relay links [SGA+12], [ARTD35].

Now the Joint transmission framework – being called in the following IMF-A - will be introduced step by step. Beside the general ideas, their potential benefits and possible implementations will be outlined for a typical macro cellular radio network with inter-site distance of 500m and parameters according to Table 5.1, which are similar to 3GPP case 1 [3GPPTR1].

### 5.2.1 Clustering

Clustering is a very important first step for any JT CoMP scheme. It divides the whole network into cooperation areas of smaller, reasonable and realizable size. Going from network-wide to clustered cooperation means - independent on the size of the CAs – to change a potentially interference-free into an interference limited system. This is generally accompanied by a performance loss. Cooperation within the clusters might then cancel some of the interferers, but will leave residual inter-cluster interference.

There are many ways to cluster the network into cooperation areas:

#### **Single cell transmission – no cooperation**

- Single-cell transmission is the reference case without any cooperation.
- 3GPP LTE Release 8, 9 and 10 are frequency reuse 1 systems that require a proper handling of the inter cell interference. Sophisticated scheduling is the main means and has been proven to be very powerful and robust, at least in case of low to medium load. Inter-cell interference coordination (ICIC) has been specified already for LTE Rel 8 and is enhanced at LTE Release 10 [3GPP-R1105081], but it seems to provide only marginal further gains in macro cellular environments.

#### **Network-centric cooperation (Figure 5.2):**

- Network-centric cooperation means the use of a fixed pre-defined clustering, regardless of channel conditions or user locations.
- The most interesting case is that of **intra-site cooperation**, where all cells of a site form the cooperation areas, in our case with 3 cells or sectors per site. From an implementation point of view, intra-site cooperation is simple. It can be implemented in a single baseband unit and with a single local oscillator (LO). This avoids backhaul overhead, extra transmission delays and potential frequency offset or degradations due to LO phase noise.
- From a system level perspective, already a quite significant part of the potential gains are provided by reducing inter-sector interference.<sup>11</sup> Unfortunately, intra-site cooperation does not help the most interference-limited UEs at the cell borders between adjacent sites, which suffer from low Rx signal power in addition to interference.

#### **User-centric cooperation:**

- As the name states, for user-centric cooperation the cooperation areas are constructed around the UEs, so that they include e.g. the 3 strongest cells, thereby maximizing the overall CoMP gains. User-centric cooperation is an important prerequisite for larger CoMP gains, as cancelling of weaker interferers would provide only marginal gains in the presence of stronger interferers.
- The user-centric cooperation areas are typically set up semi-statically, based on so called reference signal received power (RSRP) measurements. (These are for a fixed Tx power equivalent to the estimated wideband pathloss to the different cells.) Moving UEs and fast scheduling between different UEs might then lead to a very dynamic setup of active cooperation areas.
- In macro cellular environments with a strong non-line-of-sight probability - i.e. strong shadowing – the strongest cells might be distributed over quite large areas. Due to wave guiding effects, there might be very far-off interferers.
- There will be an extremely high number of potential user-centric cooperation areas. The likelihood of finding more than one UE wishing to have one particular cooperation area will then be small. There might be hundred or more cell combinations, where most of them are desired only by a few percent of the UEs. JT CoMP requires that all UEs of a cooperation area are served simultaneously from the same cells. The probability of finding such e.g. 3 UEs having the same cooperation areas will be extremely small. The penetration rate (PR) of CoMP UEs is then extremely small.<sup>12</sup>
- Obviously the penetration rate of CoMP transmission must be high if large system level CoMP gains are to be attained. Typically, at this point one would find large CoMP gains for very few UEs (which just happen to be combinable with other UEs) and small to very small gains for the large remainder of UEs, since those have to be served from their second or even third best cell set.<sup>13</sup>

---

<sup>11</sup> In quite many cases – typically for about 40% of randomly distributed UEs - the cells of one site contain already the 2 or even 3 strongest interferers per cell. This can be explained by the relatively strong correlation of the signals originating from co-located cells of one site and nearby reflections, which is true especially near to the eNB.

<sup>12</sup> The penetration rate can be defined in different ways. Here it is the percentage of users being served by their n (often n=3) strongest cells. At least 'CoMP UEs' should benefit from the CoMP transmission so that e.g. their datarate with cooperation improves over single cell transmission.

<sup>13</sup> An obvious remedy is to apply well known optimization algorithms like those proposed in [BSX+10, PGH08] to find the best set of cooperation areas for a given user constellation. Unfortunately, this

### Network-wide cooperation:

- For network wide cooperation, there is only one single cooperation area.
- Network-wide cooperation in combination with optimal precoding represents our upper bound or benchmark, which we wish to approach as close as possible for our clustered network.

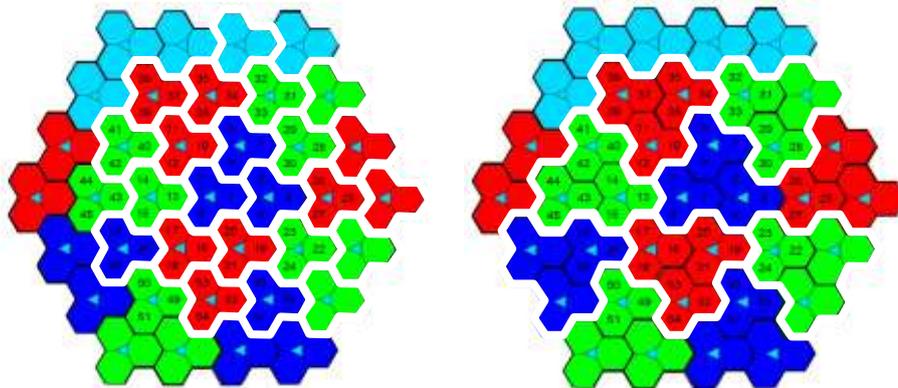


Figure 5.2: Homogeneous deployment. Triangles indicate sites with 3 sectors each. The white lines cluster the network into cooperation areas. Left: network centric/ intra-site cooperation of 3 cells; Right: network centric cooperation over 3 sites each having 3 sectors.

The Partial CoMP Clustering Scheme:

With **partial CoMP**, we introduce our first building block and first means for user-centric clustering. The scheme has three main aspects:

- **Expanded multi-site cooperation areas**, to facilitate user-centric cooperation.
- **Overlapping network-centric clusters**, here called **cover shifts**, that further facilitate user-centric cooperation and suppress inter-cluster interference, and
- **Partial reporting of channel components**, to reduce the feedback load. The last aspect has led to the term 'partial CoMP'.

#### 5.2.1.1 Expanded Cooperation Areas

A need for expanding or enlarging cooperation areas can be motivated in different ways. One goal is to increase the penetration rate of user-centric served CoMP UEs just based on combinatorial considerations. For example, the probability of finding 3 UEs having the same set of 3 strongest cells is very low. An enlarged cooperation area of e.g. 9 cells formed by 3 sites (see as example the right-hand part of Figure 5.2) will already cover 84 potential 3-cell clusters.

Any UE having its 3 strongest downlink channels from the 9 cells within this cooperation area can therefore be served user-centric. For this reason, use of clusters of more than 3 cells is a

---

typically provides only moderate to minor gains. Optimization will just find less bad cell sets without solving the issue fundamentally, i.e. finding a clustering where all or nearly all users are served user-centric from their  $n$  (e.g.  $n=3$ ) strongest cell.

first step to combine the benefits of static network-centric clustering with the gains of user-centric cooperation.<sup>14</sup>

In the main scenario used in the present chapter, the cooperation areas have been expanded to *3 sites comprising 9 cells*. From a practical point of view, this is an interesting choice. Use of intra-site cooperation only is not sufficient to exploit the interesting CoMP gains for cell edge UEs, located between two different sites and suffering from low received power. Connecting to the nearest adjacent two sites is then the natural first step to inter-site cooperation. As soon as one cell of an adjacent site is included into the cooperation area, the other cells of that site can be incorporated into the cooperation area easily, at least from a backhaul point of view. It is furthermore of potential practical advantage that intra-cluster coordination will only use backhaul connections between nearest neighbouring sites.

### 5.2.1.2 Cover Shifts – Overlapping Cooperation Areas

Enlarging cooperation areas solves the problem only partly. Even a more aggressive expansion will not really help, as there will be always UEs at the cooperation area borders. The border areas are large as compared to the center area due to geometry.

As illustrated in Figure 5.3 for “Cover shift 1”, the red UEs having their legs in two adjacent cooperation areas will suffer from strong inter-cooperation area interference.

For that reason we propose overlapping cooperation areas – here called cover shifts. Cooperation between different sites is set up differently in different orthogonal resources like frequency bands or time slots. This allows the eNBs to schedule UEs into their best fitting cover shift, being the one with the highest number of strongest cells for this UE.

It can be shown that with 6 cover shifts, all CA edge UEs in one particular cover shift can be re-scheduled into the cooperation area center of another cover shift. *This concept therefore allows the system to serve almost all UEs user-centric*, i.e. from cooperation areas that include at least their 3 strongest cells.<sup>15</sup> For the LTE Advanced case 1 in [3GPPTR1], a very promising CoMP penetration rate of about 90% can then be achieved.

### Partial Reporting

Enlarging cooperation areas is not for free. It will increase the number of downlink channels, here denoted channel components (CC), per cooperation area. This generates quite some overhead for CSI estimation and reporting. For that reason, it is proposed to restrict reporting/prediction to the  $N_p$  strongest out of  $N$  CCs in the cooperation area or, alternatively, to report only CCs with powers above a threshold, as in Appendix A2-5 and Appendix A5-1. The coordinating unit then sets the unreported CCs to zero as the best guess, but in a robust precoder design it still uses an uncertainty estimate for missing CCs, see Appendix A2-2.

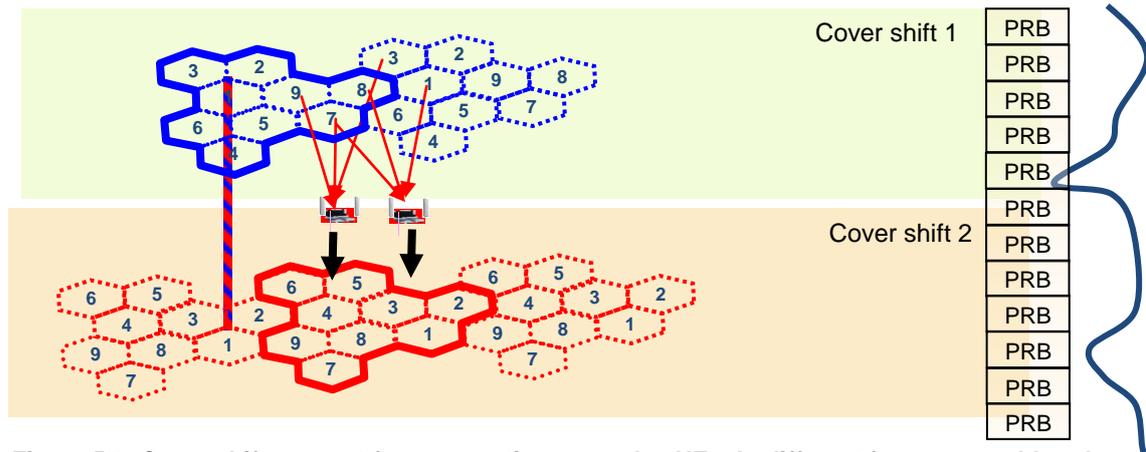
As a result of neglecting channel components, there will be precoding errors leading to some intra CA interference, but this can be kept small as the unreported CCs are per definition small

---

<sup>14</sup> From a theoretical point of view, one can analyze the optimum cooperation area size based on the useful cooperation range  $R_u$  per UE. Here,  $R_u$  is defined as the radius of an area around the UE containing cells that – if being part of the cooperation area - potentially contribute to the UE performance while the further gain by adding cells outside  $R_u$  will with probability  $P_u$  be  $\leq \varepsilon_u$ . As we are interested in first order effects,  $P_u$  and  $\varepsilon_u$  might be chosen as 0.9 and 0.05. This parameter depends on many side conditions and might vary between 500m to few km for typical SCMe channel conditions. For a 500m inter-site distance, 9-cell clusters should allow the exploitation of at least the first order CoMP gains.

<sup>15</sup> This would limit the bandwidth per UE, if the UE is associated with only one cover shift. If a UE needs a larger instantaneous bandwidth, it must be associated with multiple cover shifts comprising wider time-frequency bands, and thus to multiple cooperation areas. Since these other shifts may not contain the strongest transmitters, the attainable CoMP gain will then be reduced, in comparison with cases when a UE is associated with only one cooperation area. However single-cell scheduling gains can always be exploited in case these are larger than the CoMP gains due to the full scheduler flexibility.

or at least smaller than the reported CCs. The threshold and the number of relevant CCs depends on the overall scenario, the intended performance and also to a large extent on the precoding as the precoder will convert CSI errors into intra-cooperation area interference. The number  $N_p$  of reported CCs per UE typically varies between 3 and 5 CCs, but might in cells with multi-antenna transmitters increase up to 10 and even more reported CCs per UE. This would still represent a significant saving, for example, under the assumption of 2 to 4 Tx antennas per cell, i.e. overall 18 to 36 channel components within a single CA comprising of 9 cells.



**Figure 5.3: Cover shift concept for cooperation area edge UEs: In different frequency subbands (sets of physical resource blocks, PRB), cooperation is performed between different sets of sites.**

### 5.2.1.3 Interference Floor Shaping

As mentioned above, the distances to strong transmitters as seen from an individual UE might in some cases be very large, when distant transmitters have almost line-of-sight channels to an UE. Countering such effects by forming cooperation areas over 7 or more sites seems to be unrealistic and ineffective. Therefore, these far-off interferers will generate an interference floor that may significantly limit the potential CoMP gains. This is illustrated by the 2-user rate region in Figure 5.4. Assuming an interference limited scenario with a very low noise floor, the rate region without cooperation will be limited by the inter-cell interference (blue area). The optimum rate region with CoMP is (or might be) just limited by the maximum modulation and coding schemes (MCS) for example 64QAM9/10 for LTE. This assumes that that the interference from other clusters is below the noise floor. In case there is a noise floor generated by the out-of-cluster cells, the rate region will be reduced significantly to the red area. (A similar effect would happen in case of a significant noise floor, but this would violate our main assumption of being in an interference-limited scenario.)

Interference floor shaping as such is difficult to achieve in conventional networks as there is typically a trade-off between reduced interference at the outside of the cell versus signal quality within the cell, leaving relatively little room for optimization. From the beginning of cellular radio systems, antenna tilting has been used to localize signal power to the vicinity of the base stations. It is commonly used with good results for network planning [WLS+00].<sup>16</sup>

<sup>16</sup> For the LTE Advanced case 1, increasing and optimizing the antenna tilts for all eNBs from 7 to 15 degree provided a substantial gain in spectral efficiency helping to fulfil the target requirements of e.g. 2.6 bit/s/Hz/cell for a DL 4x2 antenna configuration.

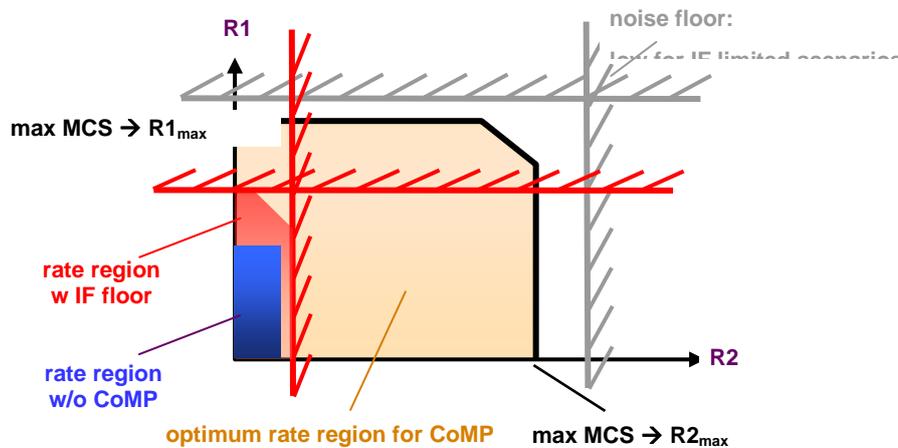


Figure 5.4: Illustration of two-user rate region with and without interference floor.

The use of JP CoMP in combination with the above introduced enlarged cooperation areas over for example 9 cells opens the door for much more powerful means for inter-cluster interference suppression.

One possibility is the use of advanced vertical beamsteering antennas, as discussed in Chapter 4 above. For example, by applying *low vertical tilt angles of e.g. 7° to the inbound wideband (WB) transmit beams* of a CA, a high received signal strength is achieved for the CA center area. Oppositely *for outbound transmit beams a strong tilt of e.g. 15°* can be used, ensuring a fast decline versus distance of interference power leakage into other CAs of the same cover shift. This effect can be further enhanced by also using power allocation with strong / weak power for inbound / outbound beams. Interference floor shaping for cooperation clusters by such means has been denoted the **“Tortoise concept”**.

Figure 5.5 left illustrates the basic Tortoise principle for a 3 site / 9 cell scenario and for 2 (fixed) beams per cell. The use of JT CoMP is an essential part of the concept, as the weakly tilted inbound beams generate very strong intra-CA interference. Without JT CoMP, this interference would result in very low SINRs for the served UEs, while JT CoMP can leverage these strong signals for an improved CA center coverage.

Assuming an LTE Advanced system supporting 64QAM as maximum modulation and coding scheme (MCS), SINRs in the range of 17 to 20dB are required. Accordingly the inter-CA interference leaking from one CA to another should be below -20dB as well. In the right-hand part of Figure 5.5, the tortoise concept has been investigated by a ray tracing simulation for a typical set of 3 sites in Munich downtown, verifying the intended localization of interference power to the vicinity of the CA center. Within a distance of 500m around the center area of the CA, the interference power has fallen below -20dB, indicated by the green coloured area. Only few waveguiding effects along long streets will need special attention, for example by the use of IRC receivers.

The proposed interference floor shaping is basically simple and robust. In combination with the Cover shift concept, a subband-wise adaptation of antenna tilts will be needed. This suggests a need for active antenna systems (AAS), with frequency-wise steerable vertical beamforming, similarly as the technologies investigated for scenario 1 in Section 4.

If active antenna systems are not deployed, cover-shift specific antenna tilting cannot be used. Then, a fractional frequency reuse scheme, implemented for the 9-cell cooperation areas can be used as a less powerful, but much simpler, interference floor shaping strategy. Fractional frequency reuse strategies used for this purpose are investigated in Appendix A5-2.

In Appendix A5-2, fractional frequency reuse is evaluated on 3-cell/sector clusters, with single-antenna base stations and one sector per site. The scheme FFR-1 is shown to provide a significant rate (and thus SINR) improvement as compared to frequency reuse 1. An evaluation

in terms of SINR CDFs in larger (9-cell) clusters would be required to fully evaluate its efficiency.

A further relevant (but optional) aspect of the CoMP concept is the use of wideband beams within cells, which might be set up statically or semi statically. The linear CoMP precoders will then be applied on top of these static wideband beams (see Appendix A2-2). Therefore, detrimental flashlight effects due to adaptive beam steering can be avoided. This enables more accurate interference estimation in the rest of the network.

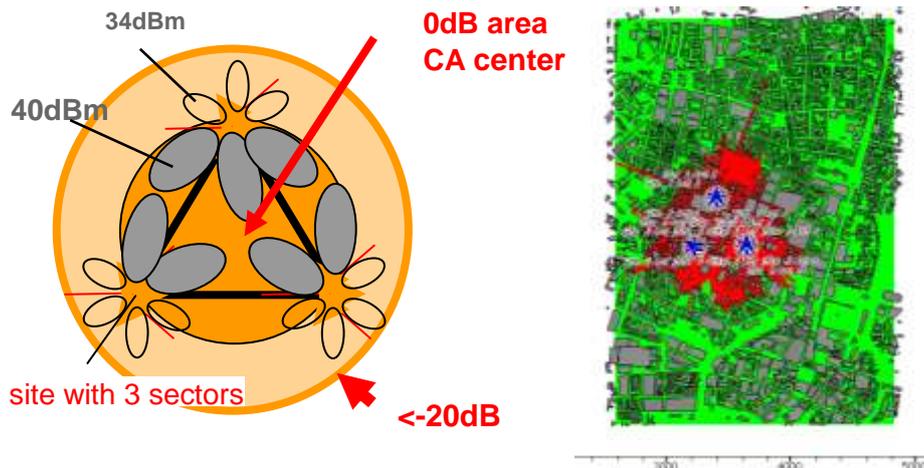


Figure 5.5: Left: Illustration of interference floor shaping based on the Tortoise concept. The grey transmit beams have a low antenna tilt of 7° in combination with strong Tx power, while outbound beams use 15° downtilts and lower power in the transmit resources (cover shift) used within the cooperation area. Right: Ray-tracing simulation verifying localization of interference to single cooperation area; green: interference power -20dB below that of CA center.

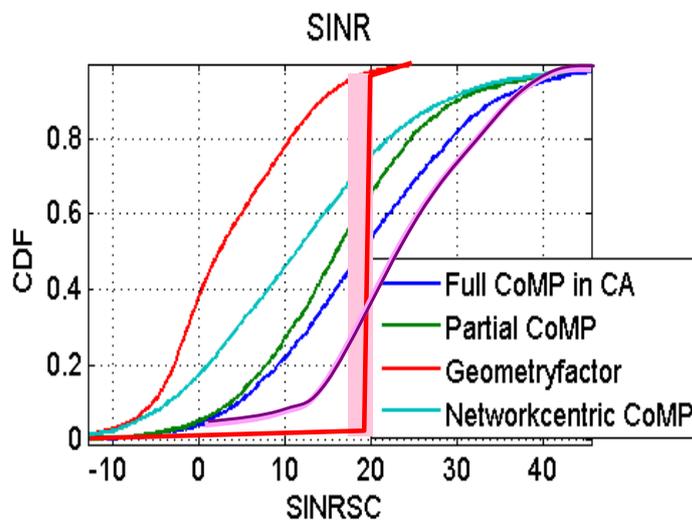


Figure 5.6: Optimum CDF for an 'interference free' system (red) and ideally achieved performance with interference floor shaping (magenta). Thin lines are for the other proposed CoMP schemes.

#### 5.2.1.4 Some SINR Results

Combining the proposed pieces of the puzzle - i.e. *enlarged cooperation areas*, *partial CoMP* in combination with *cover shifts* and the *Tortoise concept* for interference floor shaping- significant SINR gains were achieved as can be concluded from an exemplary simulation in Figure 5.6.

The figure contains the SINR cumulative distribution functions (CDF) of different CoMP schemes, using the simulation parameters of Table 5.1.

The reference case is the Geometry factor as defined in [3GPPTR2]. The label 'Network centric cooperation' refers to the case of cooperation over 3 cells of a site, i.e. intra-site cooperation. The blue line is for 'full cooperation' over 9 cells and includes the cover shift concept together with full reporting of all channel components. For 'partial CoMP' the reporting of CCs has been limited to the ones being above a power threshold  $TH_{CC}$ .

The thin lines represent results in the absence of additional interference floor shaping. An artificial ideal CDF has been added for comparison as a broad vertical pink-red line. It represents an SINR of about 20dB for all UEs of the network. This would lead to the maximum spectral efficiency in an LTE downlink together with best possible user fairness, i.e. all UEs are served by using the maximum possible modulation and coding scheme.

The *thick magenta curve* in Figure 5.6 is the best so far achieved result for the overall interference mitigation framework, including interference floor shaping by varying antenna downtilt and transmit power control. It demonstrates the large gains due to interference floor shaping, especially important for cell edge users. For the overall concept, about 70% of the users achieve an excellent SINR of at least 20dB, thereby verifying the potential of the interference control part of the overall JT-CoMP framework.

While promising these results are only an intermediate step as each cell here serves just one single UE with a single data stream. A specific CoMP scheduler designed to maximize the number of UEs per cell and per CA without sacrificing the SINR will be introduced in the next Subsection.

**Table 5.1: Main simulation parameters**

Number of cells	57
Number of sites	19
Cells (Sectors) per site	3
Sector width	120 deg
Number of PRB/subframe	32
Bandwidth per PRB	180 kHz
Transmit antennas per cell	4
Receive antennas per UE	1 / 2
Joint precoding	Zero forcing
Channel modell	SCME
Inter-site distance	500 m
Antenna tilting $T_{IB}/T_{OB}$	10 / 15 deg
Indoor penetration loss	No
CSI	Ideal (If reported)

## 5.2.2 Scheduling and User Grouping

Large SINR gains for single UEs per cell are only the first part of the story. These SINR gains have to be converted into higher spectral efficiency and user fairness by multi-user scheduling and CoMP grouping. The main tasks are to find suitable user groups served simultaneously and cooperatively in the CA for each physical resource block (PRB). Important aspects are to simultaneously exploit CoMP gains as well as frequency diversity gains to the largest possible extent.

Current channel prediction techniques, as outlined in Subsection 5.3.2 below, have a limited prediction horizon. In combination with the high accuracy requirements and high number of channel components for advanced cooperation schemes, only slow to nomadic users should therefore be served by a JT COMP transmission mode. Generally, this is expected to be a minor restriction as typical power users will be static. Nonetheless as a first scheduling decision, all fast moving UEs should be precluded from CoMP transmission.

### 5.2.2.1 Ideal Scheduler

The goal of an ideal scheduler would be to serve as many UEs as possible for a given set of spatial degree of freedom, with the highest possible throughput. This means being able to use a high-rate modulation and coding scheme (MCS), without too high packet error rate. In case of LTE Advanced, the highest modulation scheme is 64QAM. For evaluation of some main scheduling effects, we will below use a 4x2 MIMO system, i.e. each cell has 4 Tx- and each UE 2 Rx-antennas. Otherwise the simulation parameters are as in Table 5.1.

Optimization of CoMP scheduling represents – similar as for MU-MIMO scheduling - a multi dimensional non-convex optimization problem. The goal is, in general, to maximize the spectral efficiency while keeping a predefined degree of fairness or, even better, to improve the degree of fairness as this is one of the main ARTIST4G goals. This is equivalent to maximizing the sum rate over all  $N_{DS}$  simultaneously served data streams per CA, constrained by the intended degree of fairness.

For JP CoMP, the following aspects have to be taken into account:

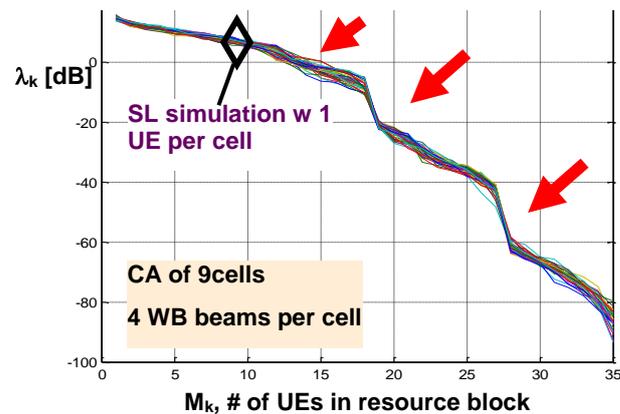
- Setup of cooperation areas that allow a high penetration rate of user-centric served UEs
- User grouping that provides a high mutual orthogonality for users who share each subcarrier or physical resource block  $PRB_i$ . This means, in particular, that different users should have their strongest channels to different transmit antennas or beams.
- Allocation of numbers of streams per UE, per cell and in the end per CA. This is a trade-off because when increasing the number of streams per cooperation area,  $N_{DS}$ , the goodput (GP) per stream will typically decrease. (Note that GP counts the throughput of correctly received retransmission blocks and therefore represents the throughput excluding the overhead for hybrid ARQ retransmissions.)
- For each potential user group, the optimum cooperation-area wide precoder matrix  $\mathbf{W}$  and UE individual beamformers or filters  $\mathbf{F}_k \in \mathbb{C}^{N_{UE} \times 1}$  – including phase and power - have to be derived, which in general is a non convex optimization problem by itself.
- Scheduling of UE sets to PRBs or subcarriers, to maximize overall capacity by exploiting multi-user scheduling gains, while ensuring at least proportional fairness for all UEs.

Taking all these scheduling dimensions into account, one can formulate the overall optimization problem for maximum spectral efficiency  $SE_{max}$  as follows:

$$SE_{max} = \arg(\max(\sum_{n_{DS}} \sum_{N_{PRB}} GP(CA_b^{opt}, UE_{1..K}^{opt}, \mathbf{W}^{opt}, \mathbf{F}_{1..K}^{opt}, PRB_i) / N_{cell})) \quad (5.2.1)$$

### 5.2.2.2 User Grouping

Enlarged cooperation areas have been introduced above to obtain sufficiently large penetration rates. Furthermore, the tortoise concept has helped to decouple adjacent CAs. Therefore, the analysis of the user grouping can be restricted here to a single CA comprising 9 cells - i.e. 3 sites a' 3 cells - for an exemplary 4x2 MIMO system according to [3GPPTR1]. The maximum number of servable data streams per physical resource block  $N_{DS,max}$ , is  $N_{DS,max} \leq \min(N_{tx}, MN_{rx})$ , where  $M$  is the number of UEs, so each cell might serve up to 4 data streams. This corresponds to  $4 \cdot 9 = 36$  potential UEs (or UE antennas) per CA simultaneously served on each PRB.



**Figure 5.7: Singular values of a channel matrix of dimension  $36 \times M_k$ , with 36 transmitters with fixed beams, as a function of the number of UEs  $M_k$ , that are randomly selected to co-exist in one resource block.**

Under these conditions, Figure 5.7 illustrates the typical singular value distribution of the total channel matrix from all 36 transmitters to  $M$  users, as being observed for an increasingly loaded system with  $M$  increasing. Serving more than  $k=13..18$  UEs seems to be quite unreasonable due to the very high spread of singular values, with some singular values being extremely small. Additionally, the variation over frequency is small as indicated by the differently colored lines.

SVD-based joint transmit- and receive beamforming would be the best transmission strategy (although it can not be implemented for the here considered non-cooperative multi-receiver case). Its performance would be severely limited by a large singularvalue spread of the total channel matrix. Beamformer design would be sensitive and difficult and only a minority of the singular values could contribute meaningfully to the sum throughput. Given these results, it could be concluded that attainable gains for JP CoMP indeed seems to have a fundamental limit.

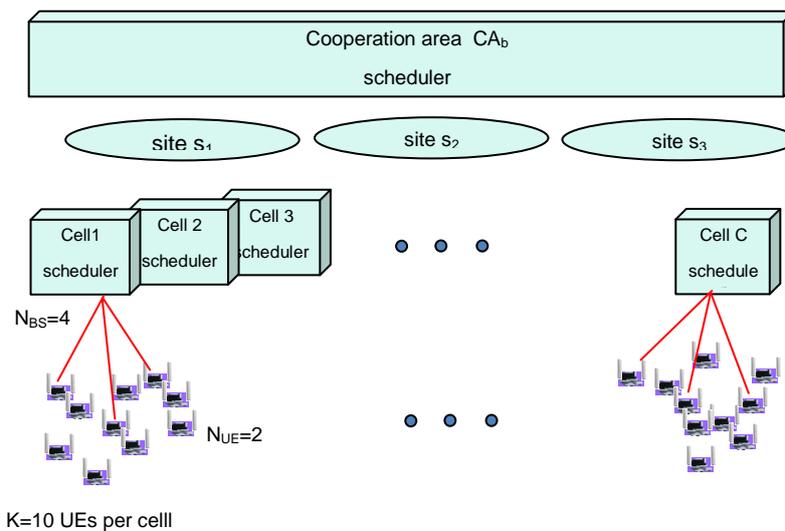
However, there is already a useful hint in Figure 5.7. Serving at most 2 instead of 3, or at most 3 instead of 4 UEs per cell would be accompanied by about 10-20dB lower spread of singular values. The number of served users per 9 cells would then drop from above 18 to max 18 and from above 27 to max 27 UEs. This would result in large reductions of the singular value spreads of the channel matrices.

The conclusion is that loading one single extra UE into one single cell might already spoil the overall channel conditions for JT CoMP. A main reason for this is the rather poor orthogonality of the fixed beams within cells, due to a significant beam overlap. A further reason for large eigenvalue spreads in highly loaded systems with random user grouping or positioning, is that some users will be located far from all transmitters. In a highly loaded system, we then in effect have more users than the number of useful transmitters. This effect is clearly evident also in the evaluations of Appendix A2-2 in the "random user grouping" cases.

Radio channels from different cells and sites are most of the time uncorrelated. Therefore, mutual orthogonality is typically spoiled between the UEs of a *single* cell and not between UEs of *different* cells. This observation is very helpful. It points to a strategy for reducing the CA-wide untractable grouping problem for e.g. 90 UEs (9 cells each scheduling 10 active UEs) to 9 low-complexity *cell specific schedulings*:

In a first stage, cell-specific schedulers would place all users *within cells* on orthogonal or close to orthogonal frequency-spatial transmission resources. The users in different cells that are thereby allocated to a resource block will then mostly have *differing* (instantaneously) strongest eNB's or beams. The resulting channel matrices therefore tend to be well-conditioned, with reduced spread of singular values as compared to a random user grouping. The resulting joint precoder design problem for these user groups therefore becomes easier to solve.

In a second stage, sets of linear JT CoMP precoders are then designed for the whole cooperation area, one per PRB. In each PRB, the sets of UEs are those that have been scheduled in each PRB by the cell-specific scheduler. The joint precoder designs are performed for channel matrices that include the cell-specific beamformers, as outlined in Section A2.2.1 in Appendix A2-2. As investigated there, such pre-selected and pre-beamformed sets of users have much better channel properties for JT CoMP as compared to random user grouping followed by a joint CA-wide precoder design.



**Figure 5.8: Two-stage scheduling, with initial scheduling per cell and subsequent CA-wide CoMP precoder designs.**

In each PRB, the schedulers therefore just have to find the best groups of e.g. 3 UEs out of overall e.g. 10 UEs, which are to be served by e.g. 4 beams within one particular cell. This is precisely what multi-user MIMO schedulers are being designed to do today in LTE and LTE-A systems, even so the scheduling itself will be more complex as it has to be adapted to the IMF-A framework.

As a further benefit, this leads to a smooth scheduler evolution by reusing well known - potentially enhanced - MU MIMO schedulers. The difference now is that the second stage JP CoMP processor is running over the whole CA. It uses the per-cell scheduling decisions plus the accordingly reported CSI information to perform a joint precoding design per PRB, resulting in a precoding matrix  $\mathbf{W}$  over all cells of the CA. The design of this CA-wide joint precoding matrix is discussed in Subsection 5.2.3 below. As compared to cellular MU-MIMO transmission via separate cellular downlink beamformers, most of the UEs will then experience a further rank enhancement and only by chance will some of the UEs see bad channel conditions, requiring a final fine tuning of the CA-wide scheduler.

In the evaluations of Appendix A2-2, this user grouping strategy has been found to be the most important effect that generates and explains the obtained CoMP gains.

### 5.2.2.3 An Overbooking Strategy to Maximize the Attainable Load

What is described in this subsection is an optional part of the JT CoMP concept. It is at present under study as a way of maximizing the attainable load and thus the attainable spectral efficiency and throughput at high loads.

Already mentioned above was that per cell sets of e.g. 3 UEs in a 4-antenna cell should be found. This we call an “optimistic overbooking” strategy. With four antenna elements per cell, a maximum of 4 data streams per cell would be possible, but it is well known at most 3 users should in practice be served. In single-cell designs, it is hard to serve 3 users per PRB in 4-antenna cells successfully. The “optimistic overbooking strategy” strives to attain this high number of users per cell in a JT CoMP setting.

For uniform linear arrays (ULA) with  $\lambda/2$  antenna spacing or similarly with  $\lambda$ -spacing, relatively broad wideband beams will be generated, that have a strong mutual overlap. In case more than two UEs should be served simultaneously, there will then be a badly conditioned channel matrix. Therefore, when trying to serve 3 UEs per cell simultaneously one has to be optimistic, and the hope is that the CA-wide JT CoMP precoding will result in a cooperation area-wide channel matrix with acceptable eigenvalue spread. Relative to what is normally possible by single cell MU-MIMO transmission, the here proposed system is thus ‘overbooked’ with 3 UEs per cell, all being in CoMP mode. Only in the end some fine tuning is done where needed. This aggressive cell loading strategy guarantees that all potential COMP gains obtainable for a given number of spatial degrees of freedom are exploited and that the penetration rate of CoMP UEs becomes near to 100%.<sup>17</sup>

This overbooking strategy at present works only in combination with an extensive per-cell optimization, including an exhaustive search for all possible sets of 3 user groups (e.g. 120 potential groups in case of 10 UEs), adapting the UE filters and finding the best allocation of UE sets to PRBs. For the last task, a per-set instead of per-UE Score-based scheduler has been found to be useful as it will lead to a good degree of fairness.

An important property that limits the performance of joint linear precoders in too highly loaded cells or cooperation areas is the power normalization loss problem: With a fixed number of transmit antennas and an increasing number of UEs per PRB, linear zero-forcing precoders, which perform channel inversion, will tend to get larger matrix elements. The whole matrix has to be rescaled to satisfy per-antenna or per transmitter power constraint. This rescaling reduces the total utilized transmit power, and the reduction increases when the singular value spread of the channel grows. This is called the power normalization loss. This does not affect the intra-cluster SIR, since the inversion still works, but it will worsen the SNR and the inter-cluster SIR. To prevent such effects, linear precoders should not exclusively be designed to take intra-cluster interference into account, but should also consider noise and out-of cluster interference. In Appendix A2-2, a multi-step approach to solve the power normalization loss problem is presented. Fortunately, use of the “cellular” scheduling and user grouping strategy outlined above reduces the power normalization loss problem significantly, as shown in A2-2.

### 5.2.2.4 Scheduler Performance

The SINR values are measured at the output of the cooperation area scheduler for all resource blocks and all UEs. Then, by using the SINR to throughput mapping from [KPK+09] as described in Appendix A8-2, summing up all contributions and dividing the throughput by the number of cells, the average cell throughput without overhead is obtained. To obtain the resulting spectral efficiency (SE) one has to multiply with  $(1-0.43)=0.57$ , which includes the

---

<sup>17</sup> This is an important difference to many known schedulers that use very restrictive selection criteria for CoMP UEs, based on small threshold values.

overhead for synchronization channels, guard intervals and bands, reference signals and control channels as known from LTE Release 8.

For the example according to results from Appendix A8-2, the attained spectral efficiency was:

$$SE = \frac{0.57}{9} \sum_K \sum_{N_{PRE}} Bits(SINR(PRB_i, UE_k)) = 6.5 \text{ bits/s/Hz/cell.}$$

This corresponds to a brutto data rate  $TP_{brutto}$  (excluding all overhead) of the transmission system of  $TP_{brutto} = 6.5/0.57 = 11.4$  bits/s/Hz/cell.

Compared to the values reported in [3GPPTTR1], this would be an improvement of a factor of 2.1 over Rel10 MU MIMO and of 1.6 over the best JT CoMP results in [3GPPTTR1] so far. However, it should be noted that due to the limited statistics of these system level simulations, further validation of the results are needed. In addition, further optimizations are possible so that the final performance might differ slightly from the preliminary one presented here.

### 5.2.3 Precoding

For systems at low load, joint scheduling is a sufficient tool for avoiding strong mutual interference. Users that would cause high mutual interference will then be served on orthogonal resources in time, space and/or frequency. However, as the load increases, user traffic must be allocated to overlapping resources, causing a decreasing Joint scheduling (JS) gain. Here coherent Joint Transmission (JT) CoMP is a possible remedy. This precoding technique both increases the received power and cancels intra-cluster interference, by transmitting messages from different antennas. The general principle is to let useful components arrive with additive phases at receiver antennas, while the phases of interferences are adjusted so that they cancel. Linear joint precoders can be designed by the Zero-forcing (ZF) criterion to fulfill this objective.

It has however been shown for MISO systems that the quality of such joint linear precoding is highly depended of the quality of CSI at the transmitter side, see e.g. [ZPO08]. The problem for precoders used in JT CoMP would be even more difficult: The need for sharing CSI between eNBs results in outdated CSI, due to e.g. backhaul transmission.

Within the JT CoMP framework, we use several tools to handle these difficulties:

- **Channel prediction:** The accuracy of the CSI can be improved by the prediction techniques which have been examined within the project, see Subsection 5.3.2.
- **Robust precoding for coherent transmission.** If the CSI is of medium quality, e.g. for pedestrian users with high to medium high SINR, then linear precoding by zero forcing often proves to be insufficient [ASA12]. Here it can be beneficial to use robust precoding schemes that take the statistical knowledge of the uncertainty of the CSI into account. Such robust precoding schemes have previously been investigated for MISO systems, see e.g. [HDJ+04], and MIMO systems, see e.g. [SD08],[ZPO08]. We here propose and evaluate two Robust Linear Precoding (RLP) schemes for JP CoMP that have been developed within the project [FF12],[ASA12]. These schemes have the benefit that they take uncertainty estimates (variances and covariances) of the channel estimates into account when designing the precoding matrix, at the cost of a somewhat higher computational complexity than traditional ZF. In addition, a stochastic optimization of linear precoders based on Particle Swarm Optimization (PSO), has been evaluated within the project, see [LBS12a] and [LBS12b]. This algorithm stochastically arrives at a stable or equilibrium solution for a given criterion or objective function, which can be designed to take uncertainties into account. It significantly outperforms the other schemes at a cost of high computational complexity.

- **The “cellular” user grouping** described above results in easier channel inversion problems, and thus in precoders that are much less sensitive to model uncertainties. This effect is illustrated rather strikingly in Appendix A2-2.
- **Mode-switching to Non-coherent joint transmission when appropriate:** Both ZF and RLP assume coherent joint transmission. When we have very low quality CSI, e.g. for fast moving users or users with very low SINR, coherent JT cannot handle the intercluster interference very well, since it relies on reasonably accurate prediction of the phases of the signals that will be received [LSB+11]. For these users we have investigated a fallback mode using non-coherent JT together with joint scheduling. The performance of noncoherent JT is lower than that of coherent transmission, based on perfect CSI but it is also degraded much less by channel prediction errors, since it does not require phase information. Therefore, it also requires less CSI exchange between base stations. It offers significant performance gains at the cell borders as compared to uncoordinated cellular transmission.

We assume per base station power constraints. The schemes mentioned above design the precoding matrices for each physical resource block separately. One option is then to also divide the total power equally over all subcarriers. This simplifies the design significantly, by enabling separate and independent optimization of the JP precoder matrices that act within each physical resource block. However, to improve throughput further, a waterfilling model for allocating different powers to different subcarriers is examined within the project. This scheme also makes it possible to redistribute the power over the different transmit antennas at a base station, when this is allowed.

### 5.2.3.1 Model of the Linear Multi-Site Transmission System

We focus on Orthogonal Frequency Division Multiplexing (OFDM) downlink transmission where a cluster of  $N$  eNBs, each with  $t$  antennas, transmit jointly to  $M$  user, each with  $r$  antennas. Channel models are obtained for sets of subcarriers (physical resource blocks) and are described by a complex channel matrix  $H$ , of dimensions  $(Mr-Nt)$ , which can be written as the sum of the estimated channel matrix  $\hat{H}$  and the estimation error matrix  $\Delta H$

$$H = \hat{H} + \Delta H.$$

The system model, including the transmit- and receive filters (beamforming matrices)  $B$  and  $U$ , is shown in Figure 5.9. The  $t$  first rows of the transmit filter  $B$  constitutes the precoding matrix for the first base station, the following  $t$  rows are specific for the second base station and so on. These rows can be designed either centralized in a CU and then distributed to their corresponding eNB or decentralized in each eNB. The receive filter  $U$  has a block diagonal structure, where each  $r \times r$  block is specific for each user. If some combinations of transmit antennas cannot use coordinated transmission, then the corresponding elements of  $B$  will be set to zero. A general discussion on backhaul limitations can be found in Subsection 5.3.1. For comments on backhaul limitation for specific precoding schemes, please see Appendix A.2.

In the Figure 5.9, the column vector  $s$  represents the payload symbols intended for all users at one subcarrier within an OFDM symbol (a resource element). The columns of the precoder matrix  $B$  distribute these symbols for joint transmission over some or all of the  $Nt$  transmitters. The column vector  $y$  contains the corresponding received complex-valued signals at all  $Mr$  receiver antennas for all users. The vector  $n$  is the sum of noise and intercluster interference, which we for simplicity will refer to as noise in this section.

The target matrix  $D$  is a diagonal matrix symbolizing the goal of intra-cluster-interference free transmission. The diagonal elements of  $D$  should be set to realistic values. A discussion on the choice of diagonal elements of  $D$  is included in Appendix A2-2. When the system coincides with the target system ( $UHB=D$ ), each element in the estimated symbol vector  $\hat{s}$  on the receiver side will be a scaled version of the corresponding symbol in the message vector, i.e.  $\hat{s} = z$ .

Assuming that the noise and intercluster interference floor is satisfactorily suppressed, by e.g. using the clustering principles introduced in Subsection 5.2.1, we have an intracluster interference-limited system. The design of the linear precoding systems can therefore focus on minimizing the intracluster interference. However, caution must be taken to ensure that this does not lead to a solution where the strongest eNBs, i.e. those with highest channel gain to the users in the group, are silenced when using a design objective that takes only intra-cluster interference into account such as the zero-forcing criterion. Such a solution would cause a low signal power at the receiver in addition to a low interference power and is hence not desirable. Therefore, the noise level has to be considered in the design, and this is done in different ways for the considered precoding schemes.

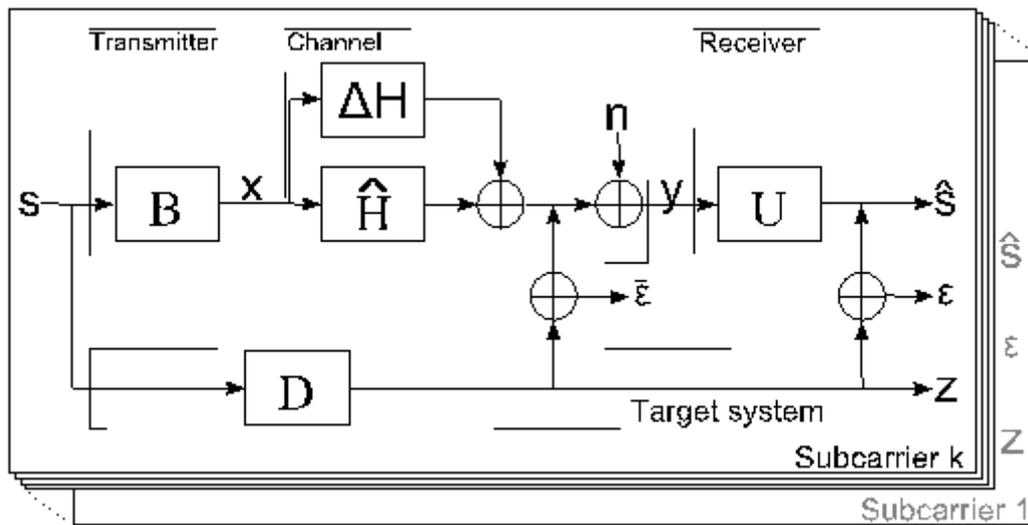


Figure 5.9: System considered for precoding design.

A brief overview of each scheme is described in the following. For more detailed descriptions, please see Appendix A2, sections A2-1 to A2-5.

### 5.2.3.2 Linear Precoder Design Schemes

#### Robust linear precoding (RLP) for coherent joint transmission

Two RLP schemes have been evaluated that use consistent (but uncertain) CSI for a centralized precoder design. Both of them utilize statistical information of the CSI error when designing the complex-valued precoding matrix  $B$ , to be used within cooperation clusters. The two design methods differ in their choice of minimization criteria.

The first scheme is a RLP Minimum Mean Square Error (RLP-MMSE) solution, presented and evaluated in Appendix A2-1. It minimizes the MSE at the receiver side (the power of  $\epsilon$  in Figure 5.9.) by jointly designing the transmit filters and the receive filters, based on channel state information available at the transmitter, under a specified per-resource-block power constraint. As the noise and the maximum transmit power is incorporated into the minimization criteria, the problem of creating solutions with low transmit power from the strongest eNBs is avoided.

The second scheme is an RLP solution based on the Automatic Control robust Feed-forward (RLP-ACFF) filter developed in [ÖAS95] and [SA93]. This robust precoder is used in a downlink JP-CoMP concept outlined in Appendix A2-2. It is there used to iteratively minimize an objective represented by a function of the expected interference power, noise power, and the expected signal power. The powers are averaged (marginalized) over model uncertainties. The objective function could, for example, represent the (local) sum-rate in the resource block. It is optimized by iteratively adjusting a scalar parameter, or a small set of parameters, when solving the RLP

design problem [ASA12]. This scheme has the benefit that it is able to focus on criteria closely related to the user satisfaction and it also effectively handles the power scaling problem. It is in Appendix A2-2 used in combination with a very simple yet effective user grouping scheme that forms groups of users with well-conditioned channel matrices. In contrast to RLP\_MMSE, the RLP-ACFF scheme designs the transmit filter only. This algorithm is faster than the RLP-MMSE algorithm.

These two schemes thus strive to adjust linear precoder matrices of the same type as those used by the zero forcing algorithms. The designs are however more elaborate in that they do not target perfect removal of intra-cluster interference. Instead, their aim is to balance the effects of the three major causes of performance degradation: Remaining intra-cluster interference, interference from outside of the cluster, and the extra interference that is caused by mismatched precoders due to channel model errors.

In Appendix A2-7, a different type of problem is also solved by performing a robust design. The problem here is that the backhaul links might sometimes have so high delays and/or so limited capacity that a centralized precoder design becomes impractical. Here, a decentralized design is to be performed of coherent linear precoders. This design is based on not fully consistent distributed channel state information. This problem is discussed in Subsection 5.3.1 below.

### **Stochastic optimization of linear precoding for coherent joint transmission**

The Particle Swarm Optimization (PSO) scheme [KE95] uses a stochastic optimization method sprung from studies on how swarms of animals, such as a flocks of birds or shoals of fish, moves when e.g. escaping from an enemy. The modelling of this movement, which can be used for optimization, was simplified by [KE95]. We here use the algorithm developed in [Eng05] which is described in Appendix A2-5. It is designed to adjust the complex values of all elements of the precoding matrix  $B$  to optimize the criterion function under per-antenna transmit power constraints. This is a nonconvex optimization problem with respect to the precoder gains, and the PSO scheme can here be applied as a search tool for the optimization. To make a fair comparison, the evaluations in this section use the estimated sum-rate function (also used for the RLP-ACFF algorithm) as the optimization criterion. The receive filter  $U$  is set to a unit matrix. The basic PSO, used in this work, finds a local optimum that outperforms the RLP solutions. It has a high computational complexity, but as the swarm optimization is carried out at the CU the complexity seems acceptable in the case of small precoding matrices. The complexity is discussed in [LBS12b].

### **Non-coherent joint transmission and scheduling**

For fast-moving users, we have no reliable CSI at the transmitters even when using channel prediction. Without phase information, coherent joint transmission cannot be used. A possible fallback transmit scheme could then be coordinated scheduling and beamforming as described in Chapter 4. We have explored an alternative JT scheme that can provide improved multi-cell diversity, as compared to schemes that do not distribute the user data to multiple cells.

The proposed fallback mode uses noncoherent joint transmission (NCJT) within cooperation clusters. Coordinated transmit points of a user are then assumed to have the same data and signalling information to provide joint transmission using the same modulation and coding scheme (MCS). Different non-coherent transmission solutions for the single cell case can be extended for CoMP systems, e.g., Space-Time Block Coding (STBC)/Space-Frequency Block Coding (SFBC), Cyclic Delay Diversity (CDD). The received multi-path data is non-coherently combined at the receiver side (the received powers from multiple transmit points are summed) to achieve a diversity gain.

With the objective of maximizing the system sum rate under per-point power constraints, binary power control combined with greedy user selection turns out to be an efficient suboptimal resource allocation algorithm for non-coherent joint transmission scenarios [LSB11]. This low-complexity semi-distributed algorithm can offer a good balance between joint transmission and

interference coordination and it provides a substantial cell-edge performance improvement as compared to cellular transmission. For more details please see Appendix A.2-3 and [LPA+12].

### Waterfilling

In a separate problem solved in Appendix A2-4, we consider the optimization of the power assigned to the user streams in a coordinated base station downlink environment with OFDM. Here, the channels are assumed precisely known. In this scenario the base stations perform distributed cooperative processing with e.g. a ZF scheme to remove interference among users. To that end, a precoding matrix is calculated so that the weight vector associated to each user has to be orthogonal to the subspace spanned by other users' channels. Under this ZF-based CoMP strategy it can be observed that, for perfect CSI, the overall system is then a set of parallel noninterfering channels.

In order to optimize the power assignment and so maximize a user-defined weighted<sup>18</sup> sum of the rates for the set of users, two schemes based on the waterfilling technique are proposed and compared to the optimal solution, which can be obtained numerically by using convex optimization. A constraint on the maximum available power for transmission from each base station is considered in all cases. First, a technique (*Modified Waterfilling, MWF*) is outlined, where the most stringent of the constraints on the maximum available power for transmission from each base station is considered. The problem is reduced to an "equivalent" base station having for each symbol transmitted to each user the precoding weights whose sum of squared values is maximal among all the BSs.

In order to further simplify this solution we may consider the fact that in a practical realization, the values of these new precoding weights are close to each other for all users, receiver antennas and subcarriers. Under this assumption, the MWF solution can be reduced to a classical power optimization by the waterfilling algorithm, where the water level is modified only by the user priorities. For more details about both waterfilling techniques see Appendix A.2-4.

#### 5.2.3.3 Performance Comparisons of the JT CoMP Precoders in a Simple Scenario

To compare the precoding techniques, a simple FDD downlink experiment has been setup with  $N=2$  single antenna eNBs at a distance of 500 m and  $M=2$  single antenna users, both moving at a speed of 5 km/h, i.e. at a fast pedestrian speed. We assume the path loss model for urban environments from [3GPP25996b]. In different experiments, the users are localized in three different ways, denoted Setup 1-3, respectively, below.

1. Both users are in their respective cells, close to the cell border.
2. Both users are at a distance of 125 m from their Master eNB.
3. One user located at the cell edge while one user is 125 m from its Master eNB.

For the channel error model, we assume that the channel is predicted 8 ms<sup>19</sup> into the future using a Wiener filter<sup>20</sup> with a perfect knowledge of the channel statistics. The predicted downlink channels are then fed back over the uplink, using a 5-bit quantization scheme for the phase of each complex channel. In the high SNR range, the quantization error will then be the dominant channel error and in the low SNR range, the channel prediction error will dominate.

Figure 5.10 - Figure 5.12 below show the performance in setups 1, 2 and 3 for NCJT, JT RLP (both RLP-MMSE and RLP-ACFF) and PSO-optimized precoding, both for the channel errors as described above and for perfect CSI. To see the gains achieved from taking channel knowledge

<sup>18</sup> See Appendix A2-4 for details on the weights.

<sup>19</sup> This corresponds to the feedback control loop, including processing time for channel prediction, feedback and backhaul delays, processing time for precoding design and any additional delays before actual transmission.

<sup>20</sup> This provides worse predictions than if a Kalman predictor would have been used.

into account, a ZF scheme, denoted as NTxZF, is also investigated. In this scheme we have chosen the target matrix  $D$  as described in [PBG11] (see Appendix A2-2 for a discussion on this choice).

For comparison, we also investigate coordinated scheduling (CS), implemented as a max rate scheduler, which gives the whole bandwidth to the user with the instantaneous highest channel gain from one transmitter. Transmission is performed over that transmitter, while the other one is silenced.

Finally, we compare to a FDMA scheme where both eNBs have access to half the bandwidth for their respective user, so no interference is generated. The performance is given in terms of average rate of the two UEs and is plotted against cell edge SNR.<sup>21</sup> To better emphasize the differences between the schemes, Table 5.2 shows the per-user performance at a cell-edge SNR of 15 dB in experiment Setup 3, where the user positions are asymmetric. Using a proper out-of-cluster interference suppression scheme, 15 dB is a reasonable cell edge SNR.

Both RLP schemes outperform the ZF scheme at imperfect CSI. However, using perfect CSI, the RLP-MMSE scheme is outperformed by ZF, especially when UEs are closer to the eNBs. This is due to the different choices of the target matrix  $D$  in these two designs. The RLP-ACFF scheme (that uses a simple choice of  $D$  discussed in Appendix A2-2) performs marginally better than ZF also with perfect CSI in the high-SNR regions.

The figures clearly show that the best precoding algorithm to choose when we have channel errors is the *PSO-based optimization of the precoding matrix*. The PSO scheme being stochastic in nature has the quality that it can search many options and therefore, PSO has the best performance for all SNR regions and all setups. Only when the channel quality is very high and the channel is perfectly known, is PSO outperformed by RLP schemes. This can be explained by the fact that the PSO finds the best of a number of investigated optimums, but not necessarily the global optimum. However, when very little noise is present, the optimal solution is to suppress all interference, which is what the linear schemes have as part of their basic design. They hence perform better in these extreme scenarios. Thus, if the computational complexity of the PSO scheme can be supported, we recommend that it is used.

When a system is subjected to a more limited computational capacity, so that the PSO solution is infeasible, then a mode switching algorithm offers an attractive solution. In these setups, if the cell-edge SNR is below 20 dB, the coherent JT schemes are the better mode to choose in terms of sum-rate, as compared to coordinated scheduling in all the setups (compare CS to RLP schemes with imperfect CSI). Coherent JT CoMP attains a better sum rate performance in all investigated conditions when the CSI is of high quality. (Compare CS with perfect CSI to RLP or ZF with perfect CSI.)

Comparing the two robust coherent JT schemes in case of imperfect CSI, the RLP-ACFF performs marginally better than the RLP-MMSE. For this reason, and since it is also the faster of the two algorithms, we propose that, under limited computational capacity, and at reasonable channel errors the RLP-ACFF is a sensible choice of precoder. It can be noted that in this comparison, we have focused on relatively easy problems, when both UEs have different serving eNBs. In Appendix A2-2 it is shown that the RLP-ACFF also performs well for harder precoding design problems.

It is clear that the impact of the quantization error is a limiting factor in the high cell-edge SNR region  $>25$  dB SNR for the coherent JT schemes. For setup 1 and 3, the coherent joint transmission schemes are here outperformed by coordinated scheduling, due to their reliance on coarsely quantized CSI. This shows that the fall-back mode that uses either coordinated scheduling or NCJT is an important aspect of a system. A similar behaviour will also occur in lower SNR regions when the channel predictions are of much poorer quality than the ones used

---

<sup>21</sup> Note that the SNR of the UEs located at a distance of 125 from their master eNBs will be higher than the one given in the figures.

here, due to e.g. high UE mobility or the need of a higher prediction horizon due to slow backhaul. For more discussions on when to switch to NCJT or coordinated scheduling as a function of the prediction horizon, please see Appendix A2-3.

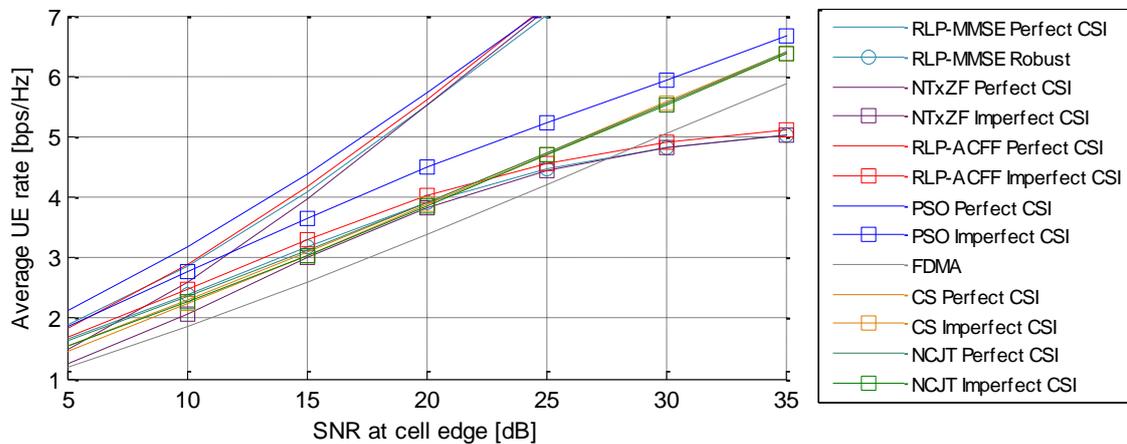


Figure 5.10: Performance of the different schemes for setup 1.

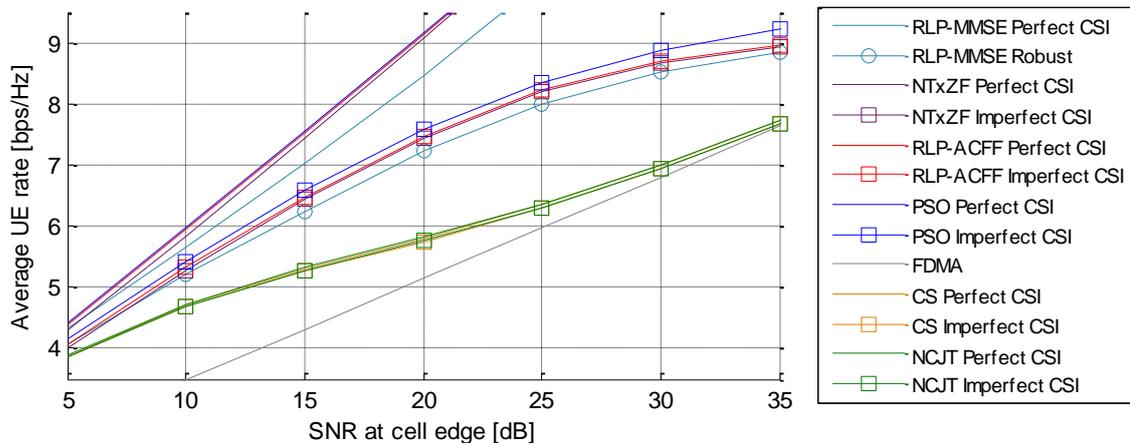
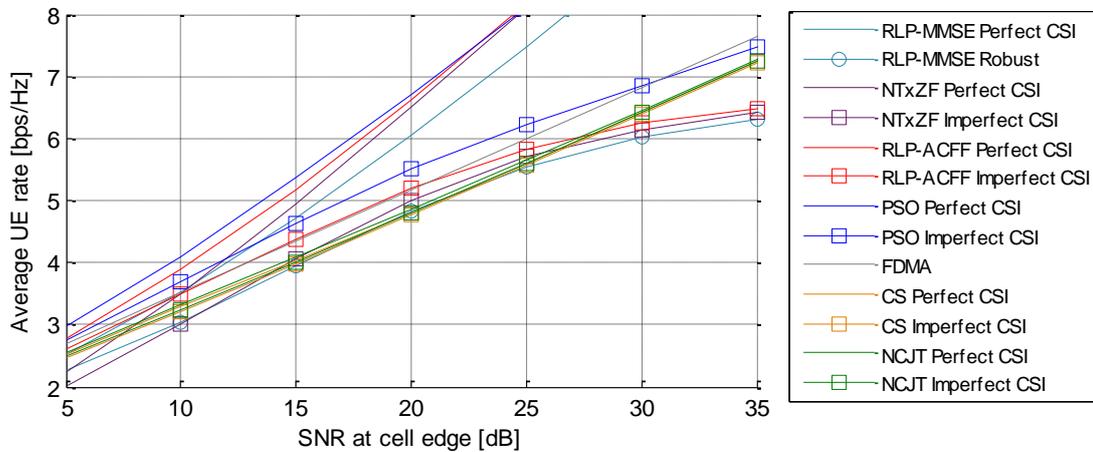


Figure 5.11: Performance of the different schemes for setup 2

In these simple experiments we have very little gain between NCJT and coordinated scheduling. Looking closely we see that NCJT always performs as good as coordinated scheduling or better. In other studies with 3 BS and 3UE [LPA+12] we have seen a larger gain from applying NCJT compared to coordinated scheduling. The difference is that in this simulation the users are allocated to a BS based on instant channel quality. This would correspond to the use of fast cell selection. In the [LPA+12] study, they are instead allocated based on location, in which case the shadow fading can cause the other BS to be stronger and in these situations NCJT provides a larger gain. Also, the simulation presented here has fewer users. This creates fewer options for the NCJT scheme.



**Figure 5.12: Performance of the different schemes for setup 3.**

The distribution of the sum throughput over the two users when they have different channel conditions can be studied in Setup 3. (It should be noted that fairness between users was nowhere used as a criterion or side constraint in the investigated designs.) Table 5.2 shows that the “fairest” scheme, in the sense of making user capacities least unequal, is then the RLP-MMSE followed by ZF. This can be explained by considering the objectives of the schemes. ZF is designed to remove interference for all users equally and RLP-MMSE is designed to minimize the residual equally for all users. The other schemes (RLP-ACFF, PSO and CS) in these simple experiments all optimize with respect to sum-rate and will therefore rather throw out a user if transmission to it does not benefit the sum-rate. To make these schemes fairer, a different objective function can easily be implemented in all of them.

**Table 5.2: Average user rate performance for the different schemes at a cell edge SNR of 15 dB for setup 3.**

Precoding Scheme	Perfect CSI UE 1	Perfect CSI UE 2	Imperfect CSI UE 1	Imperfect CSI UE 2
PSO	7.0682	3.6519	6.7709	2.5054
RLP-ACFF	6.7352	3.6101	6.0839	2.6761
RLP-MMSE	5.2846	4.1438	4.4611	3.4315
NTxZF	5.9488	3.9278	5.1440	2.9763
Coordinated Scheduling	8.0807	0.7832	8.0159	0.7480

For more discussions on when to switch to non-coherent JT or coordinated scheduling as a function of the prediction horizon, please see Appendix A2-3.

### 5.3 Practical Constraints and Enabling Technologies

This section discusses three important properties of a transmission system that constrain the performance of joint transmission CoMP, and it outlines strategies for best overcoming these such constraints:

The capacity and delay properties of the fixed backhaul links within the cooperation cluster are crucial constraints, and best ways of using them are important enablers. This topic is discussed in Subsection 5.3.1. This section also discusses the use of distributed calculation of joint transmission strategies to partly overcome backhaul limitations.

The possibilities for the prediction of channels from multiple sites are discussed in Subsection 5.3.2. The use of Kalman prediction strategies is presented here, and also in Appendix A4-2. The discussion of channel prediction also covers the closely related topics on which known reference signals to use, the reference signal overhead, where to place the predictors for FDD downlinks (in the UEs or on the network side) and the consequences for the resulting feedback reporting overhead. Some more futuristic prediction schemes that have been studied within the project are also described.

Subsection 5.3.3 finally discusses the uplink reporting overhead that is required by downlink JT CoMP schemes. This overhead aspect has often been seen as a key show-stopper for CoMP transmission schemes.

Some other constraints and enabling technologies have not been studied within the project, in particular requirements and techniques for time and frequency synchronization. The main reason for this is that these problems seem solvable with present techniques. It can also be noted that the cooperation clusters discussed in the JP CoMP framework consist of nearest neighbouring macro or micro sites. Therefore, the relative delay differences, and the corresponding synchronization problems and possible inter-symbol interference problems, should be rather small.

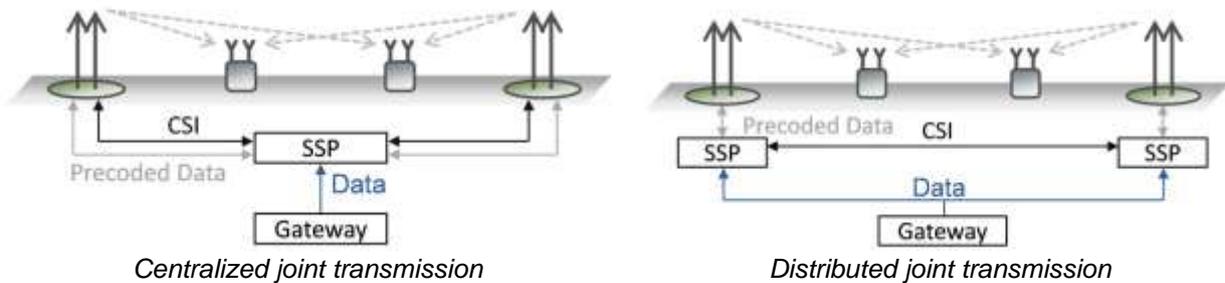
As in all OFDM systems that propose to use high-order modulation extensively to improve spectral efficiency, phase noise will be a challenge. The solution here lies in the construction of better local oscillators.

Note, in [JWS+08] [JTS+08] a more detailed analysis regarding the quite stringent frequency requirements between cooperating eNBs can be found. Ideally the frequency deviation  $f_{\Delta}$  should be in the order of less than a few Hz. Exemplary for a carrier frequency  $f_c$  of 2.6 GHz it means  $f_{\Delta} \ll 0.5$  ppb and is therefore more than a factor lower than typical LTE requirements of 5 ppb. However, in combination with channel prediction  $f_{\Delta}$  will be inherently predicted and canceled, i.e. we can expect a certain relaxation for our IMF-A framework. The generally unpredictable LO phase noise is another issue demanding high quality LOs.

#### 5.3.1 Backhaul

A precondition in JT CoMP is the exchange of CSI between collaborating network nodes, which are cooperating eNBs or a central unit (CU). The nodes are connected via a gateway to the core network. Based on CSI and user data of all cooperatively served UEs, spatial signal processing (SSP) is performed by precoding the data signals in order to avoid interference, as described in the previous section. Based on the locality of the SSP, centralized and distributed setups can be distinguished, as shown at the left and at the right of Figure 5.13 respectively. In the former case a central unit collects CSI from the eNBs as well as the respective user data from the core network. After SSP the precoded data vector is split into multiple partitions, each of them corresponding to one eNB. In contrast, for distributed JT the central unit is eliminated and CSI is exchanged directly between the eNBs. However, data of all cooperatively served users need to be available at every eNB. Note that the distinction between centralized and distributed setups refers to the logical structure and does not necessarily reflect the physical topology. Basically, both strategies can be applied to the same physical backhaul structure. However, as long as the backhaul is not affected by any limitations both schemes result in the same performance.

As shown in [J05], coherent super-positioning of data signals as it is forced for JT CoMP is fairly sensitive towards inaccuracies in CSI. In a practical system CSI accuracy is mainly influenced by lossy compression and latency in combination with time-variant channels. In FDD systems both effects suffer from CSI feedback transmission. Regarding currently deployed mobile networks in the context of JT, additional CSI inaccuracy is introduced by backhaul limitations in rate and latency.

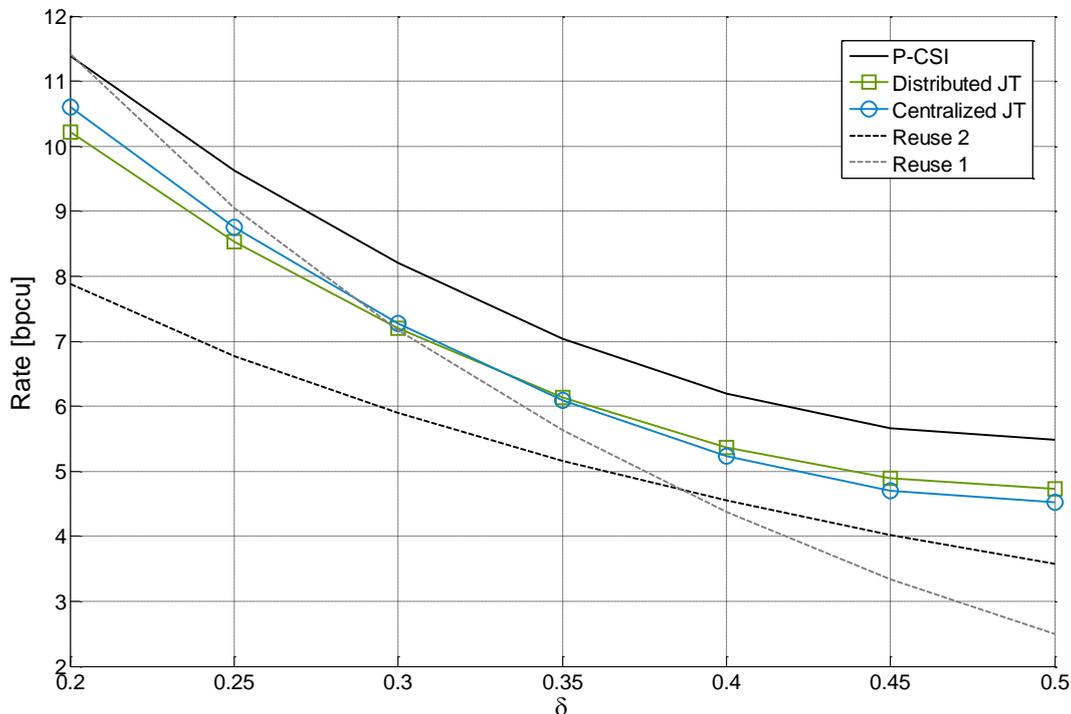


**Figure 5.13: Centralized versus distributed joint transmission**

This section discusses the impact to the performance of JT under backhaul limitations. A model for CSI impairments considering a certain quantization granularity as well as outdated CSI is introduced in Appendix A4-1. Performance gains achieved by using distributed instead of centralized processing is discussed. Since current backhaul structures are typically affected by high latencies (in the range of 10 ms) mainly due to routing issues, the deployment of low latency connections is discussed. Such connections can be realized by, e.g. directive wireless connections or prioritization within the protocol. However, the capacity of low latency connections might be fairly limited. The tradeoff between capacity and latency of the backhaul is analyzed in this section.

Cell edge performance gains can be observed by distributing the SSP to the eNBs, since only CSI from other cell UEs is affected by backhaul limitations while CSI of the inner cell UEs can directly be used for SSP. Basic performance results in terms of rate are shown in Figure 5.14.

At the cell edge ( $\Delta$  close to 0.5) performance gains can be achieved with JT compared to non-cooperative reuse 2 transmission (black dashed line) where resources are exclusively used by one UE. At the cell center ( $\Delta$  about 0.25) JT with imperfect CSI is outperformed by non-cooperative reuse 1 transmission. If the UEs are very close to its eNBs even JT with perfect CSI is outperformed by non-cooperative transmission, since it ensures per-BS-power constraints and an eNB would scale its transmit power below the maximum if another eNB is achieving the limit. Comparing centralized and decentralized JT, the latter scheme shows small performance gains in the CoMP relevant region close to the cell edge ( $\Delta$  close to 0.5).



**Figure 5.14:** Rate in bits per channel use (bpcu) achieved with different transmission strategies in a two-cell setup as illustrated in Figure 5.13. The abscissa is the distance of the UE from its BS relative to the inter-site distance.

### Robust Precoding With Distributed Channel State Information

In the distributed joint transmission setting described on the right part of Figure 5.13, the CSI is exchanged between the eNBs and each eNB runs locally the algorithm for JP-CoMP. With perfect cooperation between the eNBs, this leads exactly to the implementation of the transmission scheme designed. Yet, in practical systems, the CSI is likely to be degraded by the exchange through the backhaul links between the eNBs. Especially, the sharing of the CSI introduces inherently some delay so that the CSI received at every cooperating eNB will be aged differently. Additionally, the CSI could be also quantized further before the CSIT exchange to reduce the amount of data that is sent over the backhaul.

Therefore, it is practically relevant to study the impact of having inconsistent CSI versions at the different eNBs on the performance of distributed JP-CoMP. This problem can be seen as a “Team Decision Problem” [ZG10] since the cooperating eNBs transmit jointly the user’s data symbols but do not have access to the same channel state information.

In Appendix A2-7, we study the losses induced by the inconsistencies between the CSI estimates at the different eNBs and we develop transmission schemes being more robust to this imperfect sharing of the CSI between the eNBs.

Since the main purpose of the JP-CoMP scheme is to reduce the interference level, we consider in our analysis the high SNR regime. As a first step, we focus on the pre-log factor of the sum rate achieved (a.k.a. the number of degree-of-freedom or pre-log factor) which is the most important figure of merit when the SNR increases.

To model the imperfect knowledge of the channel at the eNBs, we consider that the channel estimates come from a digital quantization with a finite number of bits. At eNB  $j$ , the estimate of the channel  $\mathbf{h}_i$ , which represents the links between all the eNBs and UE  $i$ , is then denoted by  $\mathbf{h}_i^{(j)}$  and comes from a quantization scheme using  $B_i^{(j)}$  bits. We consider that the channel estimates

are obtained at each eNB from a different codebook to model the inconsistency between the estimates at the different eNBs.

In a word, this means that even if the number of bits quantizing the CSI is the same at different eNBs, the estimates at the two eNBs are of the same quality but are not equal. This corresponds to a setting where the two eNBs have the same average quality of the CSI but where the degradation is still unique to each eNB, be it due to the CSI-aging or due to an imperfect CSI feedback.

In that setting, we show that applying the conventional transmission schemes to manage the interference (ZF at high SNR) leads to a strong degradation of the sum rate achieved by the cooperation cluster. Indeed, it is the worst accuracy across the eNBs and the channel estimates, which dictates the pre-log factor achieved.

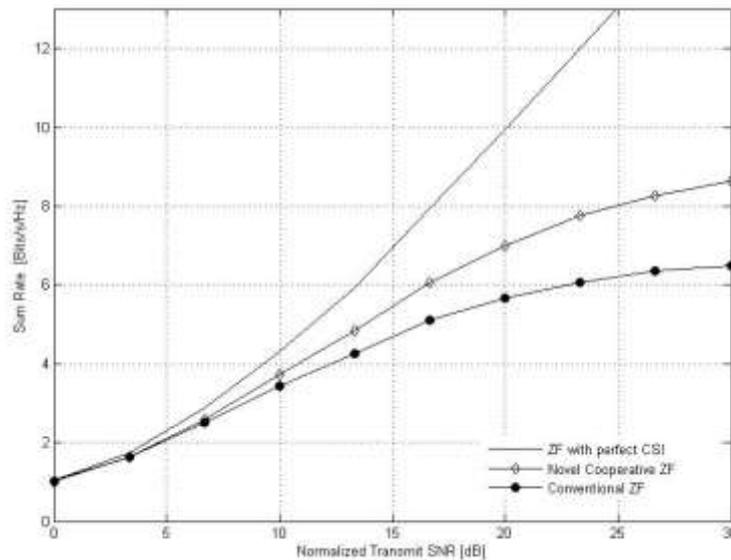
As a consequence, we develop a new JP-CoMP scheme, called **cooperative ZF**. This scheme is more adapted to this distributed CSI configuration and is thereby more robust to the unequal quality of the CSI at the eNBs. Our scheme consists in letting the eNB with the most accurate CSIT adapt its transmit coefficient to the transmission realized by the other eNB which has the less accurate CSIT. Particularly, we show [KG11b] that our scheme improves significantly the pre-log factor by avoiding that the minimal accuracy imposes the pre-log factor. Furthermore, we demonstrate that the pre-log factor achieved corresponds to the one achieved with the most accurate CSI available across the eNBs, so that our precoding scheme is optimal in terms of the maximization of the pre-log factor. More details on these results can be found in Appendix A2-7.

In Figure 5.15, we show the sum rate which is achieved in a Rayleigh fading channel when two eNBs cooperate to transmit to two UEs, each equipped with one antenna. The channel matrix is composed of two channel vectors  $\mathbf{h}_1^H$  and  $\mathbf{h}_2^H$ , each of dimensions 2 by 1, representing the links from the two eNBs to UE 1 and UE 2, respectively, and have to be known at each eNB. We use for the simulations  $B_1^{(1)}=9$  and  $B_2^{(1)}=6$  for the channel estimate at eNB1, and  $B_1^{(2)}=6$  and  $B_2^{(2)}=9$  for the channel estimate at eNB 2. This cooperative model where eNB 1 has a more accurate knowledge of the channel to UE 1 and eNB 2 knows more accurately the channel to UE 2 is well adapted to the cellular setting where the eNB obtains easily an accurate CSI relative to the UE inside its cell but has more difficulty to obtain the CSI relative to the users in the neighboring cells.

We can observe in the figure that the derived robust JP-CoMP scheme significantly improves the performance compared to the conventional JP-CoMP schemes. Furthermore, the improvement becomes even more important when the accuracy of the CSI becomes more unequal across the eNBs.

From this figure, it is possible to observe how CSI inconsistencies between the eNBs can lead to significant losses. Therefore, there is the need to enforce more consistency between the signals transmitted by the eNBs. This can be done by designing novel precoding schemes taking into account the accuracy of the CSI at the different eNBs, e.g., it could be more efficient to use a less accurate CSI which is available at all the eNBs, instead of letting each eNB use its most accurate (or recent) estimate.

More details on the design of schemes which are robust to the CSI discrepancies, can be found in Appendix A2-7.



**Figure 5.15: Sum rate in terms of the SNR with the number of feedback bits  $[B_1^{(1)}, B_1^{(2)}] = [6,3]$  and  $[B_2^{(1)}, B_2^{(2)}] = [3,6]$ .**

We have shown that inconsistencies between the channel estimates at the cooperating eNBs can lead to large degradations of the performance. To improve over this aspect, we have provided a precoding scheme being more robust in that distributed joint precoding scenario. In this paragraph we have described our results for the case of two cooperating eNBs serving two UEs. These results have been extended to the case of more cooperating eNBs in [KG2011c]. To gain robustness with more cooperating eNBs, we also show in [KG2011c] that a feedback based on hierarchical codebook can be very helpful. Finally, these works have demonstrated that not only the accuracy of the channel estimates is of interest but also the consistency between the estimates at the cooperating eNBs and that it is possible to improve over the transmission schemes conventionally used.

### 5.3.2 Channel Estimation and Prediction

High quality Channel State Information (CSI) at the transmitter side is crucial for coherent downlink JT CoMP for multiple users. It determines the attainable interference suppression. In this section, we introduce the increased challenges with respect to quantity, quality and potential outdatedness of the CSI required for JT CoMP. Kalman prediction helps to solve the outdatedness problem for an important class of mobility scenarios. It is described here, with focus on system design aspects. A more detailed exposition is found in Appendix A4-2. We also present two novel concepts intended to push the performance limits for vehicular users: In one concept that has recently been verified by measurements, a “predictor antenna” is placed in front of the transmit/receive antenna(s) on the roof of a vehicle. In another futuristic concept, the idea is to pre-calculate the propagation environment from a ray tracing simulation map, and to use pre-calculated values instead of channel predictions.

#### 5.3.2.1 CSI at the Transmitter for Coherent Joint Transmission CoMP: Three Challenges

If we compare the CSI requirements for coherent JT CoMP with the CSI that is used at present in LTE for link adaptation, scheduling and beamforming, we first note a difference in **quantity**. At present, CQI (powers or gains) and phase information with very limited resolution in the form of the best beam (codebook entry) is reported, for one eNB. For JT CoMP we will need phase and amplitude information for the channels from all transmitters that potentially take part in the joint transmission.

The required frequency granularity of this information is up to one set per resource block bandwidth. The required time granularity is on the order of 5 ms.

We here inevitably confront the need of a large increase in feedback bandwidth in FDD systems. This increase in uplink control signalling has to be balanced against the potential gains unlocked by applying JT CoMP in downlinks. These tradeoffs are discussed in Subsection 5.4. TDD systems confront no less severe system design challenges in introducing JT downlink CoMP.<sup>22</sup>

*Second*, the challenge is not only one of quantity but of **quality**. Interference cancellation will demand higher accuracy of estimated complex channels, and it is important to understand why. Main-lobe beamforming and non-coherent CoMP both work by increasing signal power additively. This is a relatively insensitive operation. Uncorrelated and zero-mean 10% (or -10 dB) relative channel power estimation errors in each channel then end up causing a 10% error in the beamformed sum power. In contrast, coherent JT CoMP strives to also reduce interference subtractively: Interference channels are cancelled, with opposite phase. Subtraction is a sensitive operation. For example, a relative error of -10 dB implies that if the channel is used for coherent JT, then interference due to that channel can on average be suppressed by at most 10 dB by cancellation. The problem can be further amplified if the precoder design is sensitive. This is a problem for linear Zero-forcing precoders that are based on ill-conditioned channel matrix inverses. Appendix A2-1 and A2-2 explore the effects of different levels of CSI inaccuracy on the performance of linear joint precoders.

---

<sup>22</sup> If we would use uplink-downlink channel reciprocity to estimate the downlink channels and thereby avoid a need for feedback, then large pools of potential TDD downlink users would have to transmit sounding reference signals with adequate density in time and frequency in uplinks. This is to enable all the potentially involved eNB antennas to receive and predict the uplink channels over wide bandwidths. Adequate uplink sounding reference signals for this purpose have not been defined, and the RS overhead problem could become quite serious. The problem of detecting and predicting the uplink channels at far-off eNB could become quite severe, since UE transmitters have limited power. Furthermore, all the involved antennas, at all eNB, would need to be calibrated to utilize channel reciprocity. On balance, it seems better to estimate TDD downlink channels using downlink RS as in the FDD case, even taking the resulting feedback load into account.

*Third*, we have the problem of **CSI outdateding**. The channel varies in time, with the fastest variations being due to the short-term fading, on a millisecond scale. Even if estimates of all relevant complex channels are delivered with sufficient accuracy, they might already be irrelevant on arrival. The problem here is the sum of time delays in the transmission control loop. In an FDD downlink, reference signals are transmitted in the downlink, and the channel is then estimated at the time-frequency positions of these reference signals. This estimation takes some time. The estimates are then transmitted to the network side on an uplink control channel. The transmit strategy (joint precoder, link, adaptation, scheduling) then has to be determined for a future subframe, and distributed to all involved nodes in time for actual transmission in that subframe. Finally, the transmission takes place. The original channel estimate should then still be of adequate quality.

The downlink transmission control loop in LTE and LTE-advanced requires the channel to be known 5 ms ahead in time. CoMP processing will add to the delays, and thus to the required channel prediction horizon. The added delay will be at least 1 ms, and may be 5-20 ms in systems that use the present X2 interface for eNB-eNB communication. Faster alternatives for control signalling and coordination are studied within the project [ARTD43].

When studying the application of downlink JT CoMP for LTE systems, the following velocity ranges will be of interest:

- **Static** (non-moving) users: The majority of future data users will likely be at static locations. However, it is usual to assume a velocity of 3 km/h even in the static case, to handle mobility in the environment. This ensures that the solution has some robustness to nonstationarities.
- **Pedestrian (or nomadic) users**. Here, a common simulation assumption is to let users all have equal velocity of 5 km/h, but with differing directions and locations.
- **Vehicular users**, for which velocities of 50 km/h and above are relevant.

A transmission loop delay requiring a latency  $L$  in time corresponds to a lag of

$$\frac{Lv}{\lambda} = \frac{Lv f_c}{c} \quad (5.3.1)$$

wavelengths in space, where  $v$  is the relative velocity and  $\lambda$  is the wavelength at frequency  $f_c$ .

As shown in Appendix A4-2, the extrapolation of the present channel to a transmission time 5 ms later is adequate only for static and pedestrian users at 2.66 GHz carriers. At pedestrian and low vehicular velocities, a large performance improvement can be obtained by using **predictors** to counteract the latency: The time-frequency channel is then extrapolated by  $L$  in time based on its statistical properties and past channel measurements.

Introduction of prediction almost triples the range in space over which channel estimates remain useful. At 5 km/h at 2.66 GHz, it extends the latencies that can be handled from 5 ms to 15 ms. This is the range of delays that would be encountered if inter-site CoMP schemes were introduced in present systems, without assuming direct fiber-optic connections between sites. Use of optimal prediction essentially solves the CSI outdateding problem for users up to fast pedestrian velocities at carriers above 2 GHz, and it extends the range to vehicular velocities at 700 MHz.

At vehicular velocities of  $> 50$  km/h, prediction by at least 5 ms would according to the relation (5.3.1) correspond to a spatial prediction of half a wavelength or more at carrier frequencies above 2 GHz. As illustrated by Figure 5.16 in the following subsection, reliable prediction of the short-term fading of radio channels cannot be performed over such long horizons, based on noisy past channel estimates only [Ekm02], [Aro11]. Coherent joint transmission CoMP schemes that rely on prediction of channel phase properties then cannot be used with presently available prediction accuracies. Such users need to be served by other means. This is one motivation for the need of mode switching, a concept introduced and used in Subsection 5.2.3.

### 5.3.2.2 CSI Prediction based on Past and Present Channel Estimates

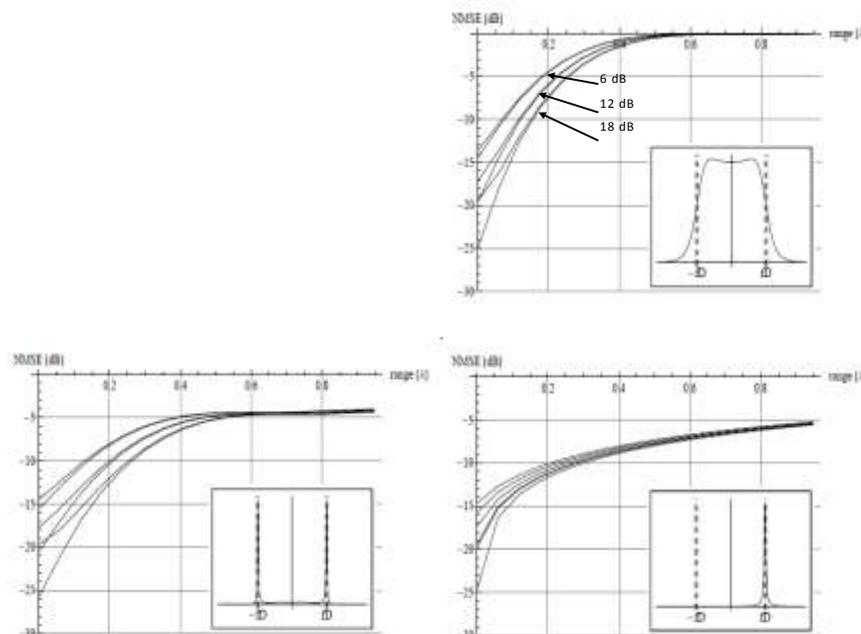
In Appendix A4-2 we present and evaluate OFDM channel prediction algorithms that enable JT CoMP to work reliably for terminals with high pedestrian velocities at 2.66 GHz, or for vehicular velocities at 700 MHz. We mainly discuss FDD downlink prediction, but the method is relevant also for TDD and for uplinks. For FDD systems, predictions of downlinks must be based on downlink reference signals.

We use a Kalman state-space channel prediction algorithm, making extensive use of recent results from [Aro11] and [ASA12]. Kalman predictors provide two advantages:

First, they provide optimal linear predictions of multiple channels that are based on measurements over an unlimited past time window, and a selectable frequency window. Second, at no extra cost they also provide the second order statistics of the prediction accuracy. This statistical information can then be included into a robust precoder design in different ways (see Subsection 5.2.3 and Appendix A2-2). The result is less intra-cluster interferences when robust precoders are applied on the true, unknown, channels.

The Kalman predictor is based on a linear state space model. This model can be based on Auto-regressive (AR) models for each estimated channel component. The state space model represents the second order statistics of the variability over time and frequency of all OFDM channel components of interest. A vector of state estimates is updated every time a Reference Signal (RS) is received, using channel measurements at the RS time-frequency locations. From the state estimate, the predicted CSI can then be calculated for a given prediction horizon.

#### Kalman prediction performance: Theoretical observations.



**Figure 5.16: Normalized mean square prediction error (NMSE) of the complex channel gain, in dB, versus the prediction horizon  $\lambda$  expressed in carrier wavelengths, for non-line-of-sight propagation with a flat Doppler spectrum (upper) for an 4th order AR approximation of a Jakes/Clakes spectrum (lower left) and for a Doppler spectrum with one main peak (lower right). The curves correspond, in pairs, to SNR levels of 6, 12 and 18 dB, respectively. In each pair, the upper and lower curves show the NMSE performance for frequency selective and flat fading channels, respectively. (Figures 6.8, 6.9 and 6.10 in [Aro11].)**

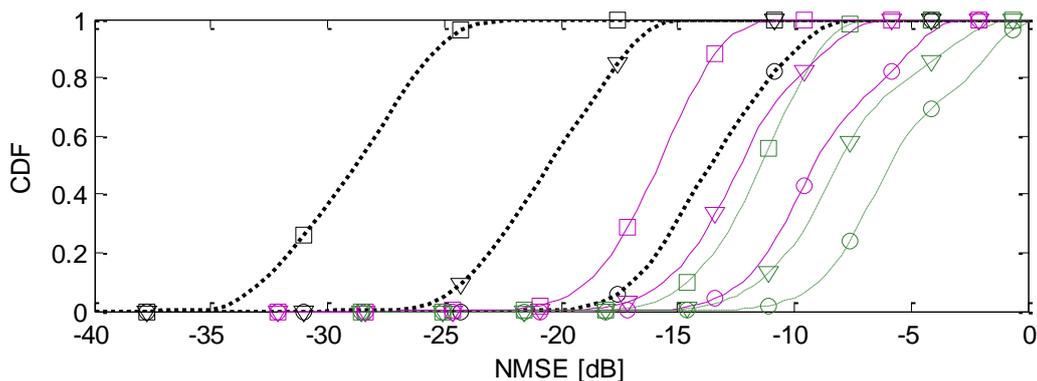
The figures above are theoretically calculated stationary Kalman prediction accuracies from [Aro11]. They refer to the simultaneous estimation of four independently fading complex channels (e.g. OFDM channels from different eNB), all having the same SNR at the reference signal positions. These results indicate the typical Kalman predictor performance as a function of the prediction range in wavelengths, expressed as the normalized mean square prediction error (NMSE) of time-varying complex OFDM channel coefficients. A very important channel property that influences the predictability is the Doppler spectrum. The upper plot illustrates results for the difficult case of a flat Doppler spectrum. The lower left plot shows the more commonly discussed Jakes, or Clarke spectrum. This is obtained when a very large number of isotropically distributed local scatterers are all localized in the horizontal plane. As shown, the performance is also affected by the SNR.

In general, it is not possible to predict non-line-of-sight fading channels with an NMSE accuracy better than -8 dB more than half a wavelength ahead, at reasonable SNRs. Useful prediction ranges are often in the interval 0.15-0.30 wavelengths. Fading processes with flat Doppler spectra are harder to predict than fading processes with large single peaks (lower right-hand plot in Figure 5.16). The predictability of the Jakes Doppler spectrum is intermediate between these cases. The presence of line-of-sight components would improve the predictability.

The Doppler spectrum itself varies on a time-scale related to that of the shadow fading which causes propagation paths to change in strength, appear and disappear. For pedestrian velocities, this occurs on a time scale on the order of a few seconds. A channel predictor must be appropriately tuned to the fading statistics and must therefore be re-tuned on this timescale.

#### Kalman prediction performance: Measurement-based evaluations.

The prediction performance is in Appendix A4-2 evaluated by using channel sounding measurements from an urban environment which covers a considerable range of power delay profiles and Doppler spectra. On average, the resulting measured performance for channel sounding measurements falls in-between the theoretical results presented above for the cases of flat Doppler spectra and Clarke Doppler spectra.



**Figure 5.17: Prediction NMSE for filter estimates (black dotted),  $L=10$  ms or  $0.123\lambda$  in space (purple solid lines) and  $L=23$  ms or  $0.28\lambda$  (green dashed lines) at 5 km/h at 2.66 GHz, using an orthogonal RS pattern. Cdf's for RS SNR intervals of [0, 10] dB (circles), [10, 20] dB (downwards triangles) and [20, 30] dB (squares).**

Figure 5.17 illustrates the results. It shows the normalized mean square prediction error for complex frequency-domain OFDM downlink channel coefficients. Three downlink OFDM channels at 2.66 GHz from three transmitters at different locations are estimated simultaneously using orthogonal reference signals (RS). Results are for filter estimates (performing no prediction), for 8 prediction steps, corresponding to  $L=10$  ms at 5 km/h (or  $0.123\lambda$  by eq. (5.3.1) and for 18 prediction steps, corresponding to  $L=23$  ms or  $0.28\lambda$  at 5 km/h. The results are

shown as cdfs for different intervals of the SNR at the reference signal locations. Reference signals from each transmitter are here placed on every third subcarrier (of 15 kHz width), spaced 1.3 ms apart in time.

The *interpolation gain* of a channel estimator is discussed in Subsection 5.4 as an important parameter that influences the performance of JT-CoMP schemes. It represents the estimation accuracy improvement relative to the ambient average SNR, and is given by  $-(\text{SNR (dB)} + \text{NMSE (dB)})$ . For example, if the SNR is 20 dB and the average prediction NMSE is -25 dB, the interpolation gain is said to be +5 dB. It is seen in the figure above that the interpolation gain will be positive for filter estimates. This is due to the noise averaging effect of the estimator. It will however be zero or negative for predictors at the considered prediction horizons.

#### **Kalman predictor performance and consequences for CoMP system design:**

**Performance.** Use of predicted channels as opposed to outdated CSI is crucial for prediction horizons of 5 ms or more even for slow moving UEs (up to 10 km/h). At reasonable noise and out-of-cluster interference levels at the RS positions, Kalman predictors can provide sufficient channel prediction accuracy for these “pedestrian” UEs. In our setting, the use of prediction instead of using outdated CSI almost triples the prediction horizon for which acceptable accuracy is obtained on average. The useful range e.g. corresponds to prediction horizons (feedback loop latencies) of 15 ms for 5km/h users at 2.66 GHz and to horizons of 5 ms for velocities of 57 km/h at 700 MHz.

**Effects of noise and delays.** A well designed suppression of the inter-cluster interference and the use of fast backhaul links that limit the required prediction horizons are both important for enabling accurate predictions. If these properties are improved, for example by reducing the basic feedback latency of LTE from 5 ms to perhaps 3 ms, in combination with maximum coordination latencies within (fibre-connected) cooperation clusters of 1 ms, then useful predictions could be obtained also at higher velocities.

**Appropriate reference signals.** Prediction of multiple channels that have very different channel gains benefits from orthogonal RS patterns, so we primarily recommend the use of orthogonal reference signals. The CSI-RS defined in LTE release 10 provide adequate orthogonal reference signal locations for use in 9-cell clusters. To limit overhead CSI-RS would be transmitted infrequently, every 5 ms, a channel sampling rate adequate for fast pedestrian users at 2 GHz and for vehicular users at 700 MHz.

**Proposed reference signal pattern.** For clusters comprising 9 cells at 3 sites, each with up to 4 antenna ports/beams, we propose the use of CSI-RS that are repeated at most every 5 ms. Each cell/eNB is allocated 4 time-frequency symbols so 36 out of 40 available positions are used. The 4 RS per cell may use either orthogonal signals or four quasi-orthogonal (code-orthogonal) sequences, to distinguish up to four antenna ports per eNB. Predictors for channels from 9 eNBs can then be implemented as 9 separate Kalman predictors, which decreased computational complexity.

#### **5.3.2.3 Physical Location of Downlink Predictors: In UEs or on the Network Side?**

Predictors for downlink channels can either be located in the UEs or in the fixed network.

- If the predictors for the downlinks are *located at UE's*, then their predictions are transmitted on an uplink control channel to the strongest eNB of each UE, denoted the *Master eNB*. This feedback is performed on the short term fading time scale and needs to have low latency. The prediction error statistics must also be reported, but this can be performed on a much slower time-scale, related to that of the shadow fading.
- If predictors are *located at the network side*, then each UE sends (compressed) measurement reports over an uplink control channel, that are used for the prediction.

The choice between these alternatives is influenced by the reporting overhead they incur, by complexity and by other considerations.

The required uplink feedback overhead *per UE* seems for these two alternatives to be rather similar. Both predictions and measurement reports would need to be sampled and updated equally densely. A frequency- and temporal resolution equal to that of the reference signals would typically be used in both cases.

As a rough estimate, this would, as discussed in Subsection 5.3.3 below, with the above suggested reference signal pattern, generate complex numbers at up to 36 reference signal locations per resource block for LTE Rel 10 CSI reference signals, to be reported every 5 ms. We here exemplify by assuming that on average complex numbers for the 5 strongest received channels are reported. Insignificant quantization losses would require 10 bits per reported complex coefficient, corresponding to a quantization error floor at -30 dB.<sup>23</sup> The corresponding maximal reporting overhead would then be  $5 \times 10 \text{ bits} \times 200 \text{ reports/s} = 10 \text{ kbits/s} = 10 \text{ bits/ms}$  per physical resource block bandwidth per reporting UE, in both of the considered alternatives. The quantization granularity could be reduced. Furthermore, feedback information can be compressed by lossy compression before being transmitted over the uplink, see e.g. Appendix A4-4 and A4-5 below and [EO07].

When the *total* uplink feedback load is considered, placement of the predictors at the UEs becomes clearly more attractive than placement on the network side. As outlined in Section A2.2.8 of Appendix A2-2, and further discussed in Subsection 5.3.3 below, placement of the predictors in UEs makes it possible to use a scheme that can drastically reduce the required feedback rate.

It was shown in Chapter 8.6 of [Aro11] that the computational complexity for prediction over a whole 20 MHz LTE band are well within reasonable limits for placing predictors for downlinks in the UEs. For example prediction of 4 frequency-selective channels spanning 20 MHz requires  $10^6$  real operations per filter update. With 5 ms RS spacing in time, which corresponds to 200 updates/s, this would result in 200Mops/s. This result assumes parallel sets of Kalman filters for different frequency blocks, and use of orthogonal RS. It also utilizes the fact that the Kalman filter recursions stabilize quickly.

However, placing predictors in UEs would have three drawbacks:

- First, the predictors implemented in terminals may vary widely in quality, depending on receiver type (computer, mobile phone, etc.), brand and model.
- Second, if a large number of adjacent RS-bearing subcarriers are predicted jointly, then the accuracy of the Kalman filters, but also their computational complexity, increases. The eNBs are better equipped than UEs to handle the resulting higher computational complexity.
- Third, as discussed in Appendix A2-1, channel predictions may be required for different prediction horizons. Predictors at UEs would increase the uplink feedback overhead if they reported predictions for multiple horizons. Predictors on the network side cause no such problems.

Prediction of channels with very different received powers using quasi-orthogonal RS pattern could require measurements to be sampled and fed back with very high dynamic range, leading to a large feedback load for a solution with predictors at the network side. In the above

---

<sup>23</sup> If predictors are located at the UEs, quantization error of the channel predictions that are fed back would act as a noise floor in the prediction accuracy, or a lower limit of the NMSE. A level of -30 dB would then have insignificant contributions on the precoder performance. If predictors are located at the network side, then quantization of the fed-back measurements on which the predictions would be based would act as added noise in the measurements. Also here, a -30 dB quantization noise level would lead to insignificant performance losses at realistic SNR levels.

proposed RS pattern, where different eNB in the cluster use orthogonal RS positions, this problem vanishes.

On balance, no definitive recommendation can be given at the moment about the over-all best location of predictors. Location at the UEs seems favored due to its ability to reduce the feedback overhead. This is an important topic for further study.

#### 5.3.2.4 Use of “Predictor Antennas” on Vehicles

Channel prediction based on past measurement will in general be inadequate at vehicular velocities at carrier frequencies above 1 GHz. For such users, a new solution has been investigated within the project: Placing a “predictor” antenna in front of the main antenna, both placed at the roof of a vehicle. This innovation is described in [ARTD35] and in [SGA+12]. Such antennas can be part of an array, and be used for spatial combining as well as for prediction.

Recent measurements have shown the signal in such a predictor antenna, placed 0.5-3 wavelengths ahead in the direction of travel, can be made highly correlated to the complex channel that will later be perceived when the transmit/receive antenna moves to the same position. This extends the useful prediction horizons (providing NMSE -8 dB or better) by an order of magnitude, as compared to the horizons shown in Figure 5.16. It represents the first really efficient technique for enabling adaptive modulation, opportunistic frequency-domain scheduling, adaptive beamforming and coherent JT downlink CoMP for vehicular users in LTE systems. It is of particular interest for relay links to moving relay nodes in busses or trams, which constitute natural hot-spots.

#### 5.3.2.5 Model-Based Channel Prediction

Measurement-based updating of radio propagation maps are used today to help network planning and optimization. This is done on a *tens-of-meters resolution* in space.

Use of GPS positioning could improve this resolution to *meter scales*. One could then for example learn the shadow fading pattern along streets. This knowledge could be used to pre-adjust resource allocation and handover algorithms that work on the shadow-fading time-scale. Such concepts have been studied over the last decade by several research consortia, for example the EU WHERE project.

Within ARTIST4G, we have investigated the use of model-based tools on an even smaller spatial resolution: The *decimetre scale*. If successful, mapping of the radio propagation environment on a decimetre scale could help to predict the short-term fading. This is the scale that is needed for the pre-adjustment of coherent joint processing schemes.

The general concept is here to pre-calculate the propagation environment in large detail from a ray tracing simulation map, and to use pre-calculated values instead of channel predictions. When the user travels through the environment, its positioning is determined by GPS and may then be fine-tuned by pattern comparison with the propagation map. Given sufficiently detailed knowledge of the propagation environment, a 3D location determined with centimetre precision, and an accurate velocity vector, the phases and amplitudes of the received signals from different transmitters could then be predicted by map-based ray tracing pre-calculation for arbitrary future locations. This concept is described in more detail in Appendix A4-3.

To realize this concept, very high accuracy would be required in three aspects of the scheme:

- *Map accuracy*. To be able to reproduce phase properties of the radio propagation, the object map of the environment needs to describe all objects in the environment that may cause significant scattering, of size one decimetre or larger. Their positions need to be determined to within a few centimetres. Their reflectivity and scattering properties and coarseness of their textures also need to be known and described on this scale.
- *Path calculation accuracy*. If a total sum of propagation paths to a user is to be calculated with a total relative accuracy of -10 dB, then the sum of inaccuracies in each path, plus the inaccuracy caused by neglected paths, cannot be allowed to be larger.

- *Positioning accuracy.* Positioning errors by more than a few centimetres would result in more than 10% phase errors at carrier frequencies above 2 GHz.

These are obviously quite challenging demands. Strategies and possibilities for meeting them are discussed in Appendix A4-3.

While the accuracy of model-based methods would today not be adequate as stand-alone channel estimators, these methods can still have an important role as support algorithms for Kalman or Wiener signal-based channel predictors. Such predictors need to periodically readjust their filter parameters, which are adjusted to the Doppler statistics of the fading channels. A map-based background processing of the user trajectory could alert the predictors to immediate changes in the propagation environments, such as the appearance/disappearance of important reflexes. This could be used to trigger timelier re-estimation of autoregressive fading models, as compared to continuing with outdated models until the next regularly scheduled model updating event.

### 5.3.3 Feedback

In mobile radio systems a reasonably small feedback (FB) rate is of greatest importance for any new transmission control scheme. For the control of downlink transmissions, the feedback has to be transmitted by the UEs in uplink (UL), which is generally limited by smaller UE Tx-power, strong and varying UL interference conditions, and battery limitations of the UEs. Additionally UE vendors are strongly in favour of simple algorithms with few processing steps and low memory requirements. Ideally processing should be limited to subframes or to frames.

The IMF-A framework for JT-CoMP that is outlined in Subsection 5.2 promises large performance gains under ideal conditions. At the same time the feedback requirements increase compared to less powerful schemes. The main challenge is due to the enlarged cooperation areas (CA) with an accordingly larger number of channel components (CC) in combination with high accuracy requirements. These accuracy requirements are partly a direct consequence of the high number of CCs within a CA – i.e. due to summing up of many small error components – together with the generally increased targeted performance of IMF-A.

System level simulations have been performed for IMF-A with varying levels of simple IQ-quantization of the CCs with different number of quantization bits  $N_Q \in \mathbb{N}$  to assess the sensitivity of the downlink capacity to feedback quantization errors. These results are in Section 5.4 used for balancing of the overall IMF-A framework, using a suitable accuracy level for the reported channel state information (CSI) for all CCs.

A related question is how the CSI information should be reported. There is a long research history in this area, which led to many different proposals, mainly with the goal to reduce the feedback rate by suitable compression techniques.

Such schemes can generate a wide variety of feedback data rates. For example, 3GPP LTE Release 8 relies on feedback of so called precoding matrix index (PMI) information, which for SU- and MU-MIMO attains a very strong compression ratio, since it is just an index into a relative small size unitary precoding matrices [HT11]. Codebook based reporting with small sized codebooks is an implicit reporting scheme, which loses information about the CSI of the CCs. For example, in the case of several Tx-antennas there will be only a combined feedback for all Tx- Rx-links.

At the other end of the complexity scale would be full explicit reporting of all CCs with a prescribed quantization without use of compression. The CCs might then be quantized as I-Q-, phase- and amplitude- or vector quantization values. Vector quantization [HZH08] is known to provide a useful reduction of the required number of bits for a given accuracy level.

Different compression schemes for quantized CCs have been proposed, where in particular lossless compression (similar to the well known ZIP compression of e.g. word documents) is interesting as it often reduces the feedback rate significantly without any loss in accuracy.

Similarly for low mobility UEs – as being anyway assumed for JP COMP – many different tracking proposals exist [HZH08]. These exploit the large coherence time of the radio channels by reporting only changes or model-based residuals (delta reporting). A survey of feedback schemes for the purpose of link adaptation is given in [EO07] and a comparison of some schemes was presented in [ARTD12].

While powerful, lossless compression and tracking solutions might have issues with robustness against feedback errors: Lossless compression will lose all CSI information in case of a single error. Tracking solutions have the well known problem of error propagation. Therefore, it will be necessary to harden the feedback link by adding sufficient redundancy. This redundancy eats up parts of the compression gains. A suitable balancing results in a net FB rate reduction.

To increase robustness, some degree of self containment of the CSI feedback is helpful. While adding some extra overhead, it would allow the receiver to reconstruct the CSI from a single CSI report at least to some prescribed accuracy level. This known accuracy level might be

taken into account by the robust precoder (see above) to achieve the best possible performance under the given circumstances. Furthermore, a suitable combination with hybrid ARQ (HARQ) will help to recover fast from failure conditions [ZMS+09].

Due to the large variety of options and their potential combinations no specific feedback proposal will be given here. Any realistic proposal for a feedback scheme would have to be a strongly influenced by standardization trends, which are typically difficult to foresee.

Instead we will assess first order effects and provide an upper bound on the feedback rate that would result in marginal performance losses, and that uses no compression. Several indications are then given on that this maximal rate could be reduced.

### Partial Reporting

The IMF-A framework as such is not limited to specific scenarios or antenna constellations. As a high end solution we investigated a deployment with four Tx- and two RX-antennas per cell and per UE. From a feedback point of view this leads to CAs comprising overall of 4 times 9 = 36 WB beams to two Rx-antennas per cell, corresponding to 72 CCs. Assuming a suitable Rx combiner at the UEs this can be reduced to 36 CCs.

Helpful is that due to shadowing effects many CCs are very small with marginal effect on the total Rx signal of a certain UE. In case of the partial (reporting) CoMP concept [MZ11] only the  $n$  strongest CCs or only those CCs above a certain threshold  $TH_{CC}$  are being reported. The unreported CCs will result in intra-CA interference, which has to be kept sufficiently low.

After a series of SL simulations for the above described scenario it was found that a threshold of  $TH_{CC} = 0.05$  equal to -26dB with respect to the power of the strongest CC seen by a certain UE will lead to negligible intra-CA interference as compared to the case of full and ideal CSI information if this limited set of reported channels is used in linear zero-forcing precoders. Further analysis for other scenarios can be found in [MZ12].

Depending on the UE location, the number of CCs being above a threshold  $TH_{CC}$  of about 0.05 varies between three and more than 20 CCs. Typical mean values are around  $N_{CC,mean} = 12$  CCs.

It is proposed to use an adaptive reporting - i.e. adaptive number of reported CCs - based on wideband pathloss measurements as being defined by LTE Release 8 for the reference signal received power (RSRP) values. On average the partial reporting can save about  $36/12 = 300\%$  for this specific scenario with almost no performance degradation.

### Quantization Level

For the IMF-A framework system level simulations have been performed, assessing the performance for different simple IQ-quantization levels. For a reasonable performance loss an IQ quantization of  $N_Q=5 + 5=10$  bits seems to be required.<sup>24</sup>

The reported CCs would have varying average powers that would vary slowly. It would therefore be feasible to report the average power of each CC slowly and adjust the quantization of each CC accordingly, so that the whole dynamic range of the quantizer can be used for each CC. This method is assumed to be used in the following reasoning, and equal quantization levels are assumed to be used for all CCs, strong as well as weak ones.

---

<sup>24</sup> In these specific simulations, this was true for the strongest as well as the weakest CCs. This might seem somewhat counterintuitive a one might expect only small sensitivity of the precoder performance to estimation errors in weak CCs. This effect can be partly explained by the higher number of weak compared to strong CCs. Furthermore it might be related to the higher precoding phase sensitivity compared to that of amplitudes. In case of a weak CC being 10 to 20 dB weaker than the strongest CC the achievable phase resolution would be very poor. Further investigations are planned to analyze either suitable phase- and amplitude quantization or vector quantization schemes. For the time being we assume 10 bit per CC for calculation of a worst case upper bound.

### Upper Bound: explicit feedback of channel components without compression

Given a maximum UE mobility, one can define the necessary update rate for the CSI of the CCs or conversely, one can define a suitable feedback repetition rate and limit the maximum allowed mobility to ensure sufficient correlation over time. State of the art channel prediction is possible for about  $0.2\lambda$  as has been explained in more detail in the previous Subsection. A prediction horizon of 10 ms for 5 km/h nomadic users at 2.66 GHz carriers would correspond to  $0.123\lambda$  by eq. (5.3.1). Good prediction accuracies are attainable for such prediction horizons, see Appendix A4-2. A maximal reporting interval of 10ms, corresponding to a (minimal) feedback rate  $FB_{rate} = 100$  CSI reports per second can therefore be assumed. This 10 ms interval equals the LTE framelength.

To limit inter-cluster interference, each cooperation cluster would be allocated only a fraction of the total time-frequency transmission resources. Therefore, each UE would need to report the channel at most for this fraction of resources. We here use the fraction 1/6, discussed in the “cover shift” concept of Subsection 5.2.1.2. Each UE would then have to report at most for a single cover shift spanning over  $1/6^{th}$  of the full frequency band being in case of LTE typically 20MHz or 100PRB, i.e. for about 18PRBs. For proper oversampling of the channel transfer function, there should be at least  $N_{RS,PRB} = 2$  RSs per PRB.

The CA is here assumed to consist of 9 cells. We assume that there are on average  $K_{cell}$  active UEs per cell, where we below set  $K_{cell} = 60$  for illustration. Each cell would participate in 6 overlapping cooperation areas (cover shifts). Within each cover shift, each cell would then have  $K_{CS} = K_{cell}/6 = 10$  active UEs on average.

Taking the above mentioned numerical values for the relevant parameters one can calculate the upper bound feedback rate  $FB_{UB}$  per UE, per cell and per CA:

$$FB_{UB}(UE) = N_{CC,mean} * N_Q * N_{PRB} * N_{RS,PRB} * FB_{rate} = 12 * 10 * 18 * 2 * 100 = 432kb/s$$

$$FB_{UB}(cell) = FB_{UB}(UE) * K_{cell} = 26Mb/s$$

$$FB_{UB}(CA) = FB_{UB}(cell) * 9 = 233Mb/s$$

Here,  $N_{CC,mean}$  represents the average number of channel components that needs to be reported by each UE while  $N_Q$  is the quantization level, in bits per complex-valued channel component. The quantity  $N_{PRB}$  is the number of physical resource blocks predicted per UE.  $N_{RS,PRB}$  stands for the number of reference signal positions (and corresponding estimated CCs) allocated per transmit antenna per physical resource block. Finally,  $FB_{rate}$  is the assumed feedback rate per second.

These upper-bound figures would represent an unacceptably high control feedback overhead load in the uplinks of each cell.

### Influence of the placement of the channel predictors.

As discussed in Subsection 5.3.2.3 above, channel predictors for FDD downlinks may be physically located at the UEs or on the network side. In the former case, their predictions are fed back to the network over the uplink control channels. In the latter case, measurements need to be fed back, and the predictors operate based on these measurements.

For these two alternatives, the feedback load per predicted resource block per user becomes roughly equal. However, when the *total* uplink feedback load is considered, placement of the predictors at the UEs becomes clearly more attractive than placement on the network side.

The reason is that there in general are  $M = 9K_{CS} = \text{UEs}$  per CA that could be scheduled within a subframe, of which  $M_s \ll M$  are actually scheduled in a particular subframe. The scheduling and link adaptation decision is today based on quite coarse CQI that is fed back to the network side. In LTE Release 8, PMIs are reported with 3 / 4 bits in case of 2 / 4 Tx-antennas, i.e. using a very coarse feedback which generates low overhead. The frequency-dependent CSI required per user for successful coherent CoMP is significantly larger, as discussed here and in Section 5.3.2.1.

The following scheme is proposed in Section A2.2.8 of Appendix A2-2: If predictors are located at the UEs, and if a small extra reporting delay is acceptable between scheduling and transmission, then the UEs may report complex channel gains only for the  $M_s$  UEs/data streams that have actually been scheduled. A similar concept has been proposed in [SVH+11].

If instead the predictor were placed on the network side, measurements would need to be reported from *all*  $M$  UEs that may *potentially* be scheduled and for *all* resource blocks that may *potentially* be used. This is because Kalman and (and Wiener) predictors need a time-history of past measurements on which to base their predictions. They cannot be immediately cold-started, accuracy during their initial transient is worse than after a few samples, when the accuracy has converged to its steady-state value.

Thus, if the above proposed scheme for feedback with predictors located in the UEs is used, we could in the example for the maximal feedback load per cell substitute the number of *actually scheduled* users for the number of potentially scheduled users. Instead of 10 users per cover shift per cell we would schedule at most 3 per cell, using the optimistic overbooking strategy. This reduces  $K_{\text{cell}}$  from  $6 \times 10 = 60$  UEs per cell to  $6 \times 3 = 18$  UEs per cell. Furthermore, feedback needs to be reported only for the *resource blocks allocated to that user*. This reduces the feedback load further by a large factor. With this modification, the feedback load per cell in the example above can be reduced from 26 Mbits/s to a few Mbit/s over a 20 MHz bandwidth.

We are now down at much more reasonable figures for the feedback load, without having introduced any compression schemes, or having made compromises with respect to the JT-CoMP performance.

There are many other aspects that can be used to further reduce the feedback load. They are discussed briefly below.

### General Feedback Considerations

In a long-term perspective, the use of model-based channel prediction as discussed in Appendix A4-3 could be used to reduce the required feedback load, as outlined in Appendix A4-5: If (parts of) the channel quality can be derived from knowledge of the location and of the environment, then the need for explicit measurements and feedback could be reduced, perhaps drastically reduced. Since accurate model-based prediction of the small-scale fading is challenging, this remains a speculative long-term possibility.

In the mid-term, simple state of the art compression techniques might be applied. From a technical point of view compression ratios of 5 to 10 should be possible by suitable combination of e.g. vector quantization, tracking schemes or codebooks [LH04][HT11], tensor based feedback schemes [RHG06], lossless compression techniques [HZH08], etc. resulting in feedback rates of about 50 to 100kbit/s/UE. Generation of high-precision feedback only for scheduled users and their allocated PRBs, as proposed above, would reduce the average feedback rate by a further large factor. We are then down at uplink feedback rates that can be considered reasonable, in relation to the large potential downlink gains of JT-CoMP.

Robust precoding solutions might help to take the over the time decreasing CSI reliability of typical tracking solutions into account. Infrequently occurring complete feedback failures might be tackled by novel HARQ solutions [ZMS+09]. This is an interesting area of further research.

For the uplink direction reciprocal relative performance improvements can be expected as from the IMF-A DL scheme. Hence the increasing feedback overhead will be partly compensated by the higher UL performance, due to the use of uplink CoMP.

A further consideration is with respect to the typical use case of JP CoMP. One can expect that a mobile network operator (MNO) will put UEs into CoMP mode only in case this is really needed, i.e. only in case a cell is highly loaded or even overloaded. UEs in low loaded cells can fall back to a low size reporting mode. To cope with varying load conditions one can apply a traffic light model, i.e. in 'green' mode - broadcasted by the serving eNB - the UEs are in low FB mode, in case the eNB indicates a 'yellow' light for medium load the UEs prepare for accurate CSI reporting, but do not transmit this information yet. Only in case of a 'red' light the UEs have to really transmit with a high feedback rate. This concept helps to save UL capacity and UE battery power as far as possible.

MNOs typically never exploit the network capacity to 100 per cent, but only up to e.g. 30 to 50% of the maximum capacity to ensure a stable operation. As a result the percentage of time the UEs are really in high speed reporting mode might be quite small.

From a physical or technical point of view the feedback should be very flexible and adaptable even so this might make standardization as well as system operation more complex. Suitable resource allocation algorithms should hide this complexity to the operator. A suitable adaptation will help to keep reported CSI accuracy in balance with the channel correlations in time and frequency, the UE noise conditions, the visibility of relevant channel components being very different for different UEs as well as the potential scenario dependend system-level CoMP gains. Ideally it avoids any irrelevant feedback overhead not leading to any performance gains, but provides sufficient information to get close to the performance with ideal CSI knowledge.

As a last topic we strongly recommend to enhance the feedback by a suitable CSI reliability information. This needs to be updated at a low rate for each UE, corresponding to the shadow fading time-scale. As outlined in Section 5.2.3 and discussed in more detail in Appendix A2-2, this information will be of great importance for finding the best robust precoding solution and robustness is one of the main challenges in the context of the IMF-A framework.

## 5.4 Balancing the Joint Transmission Framework

In this subsection, the goal is to assess the performance attainable when using the proposed advanced interference mitigation framework (IMF-A) for JT downlink CoMP and to find some suitable operation points where the relevant system design parameters are well harmonized and balanced. For illustration, think of CSI reporting, where a very fine granular quantization would be useful only in case of correspondingly accurate channel estimation, as the feedback bits would otherwise be wasted.

### 5.4.1 Parameters and Design Variables

The most influential system properties and design variables are illustrated together with their main interrelations in Figure 5.18. The noise floor  $NF$  - which is basically the SNR distribution over all UEs - is a convenient starting point. It is determined by the over-all scenario which includes the inter-site distance, the RF carrier frequency and the statistics of the outdoor-to-indoor penetration loss (PL) for as function of the locations. The  $NF$  can be affected for example by the maximum Tx power and the UE noise figure. A black arrow indicates that the  $NF$  directly affects the attainable spectral efficiency (SE) as it limits the achievable rate region in a similar way as the interference floor discussed in Subsection 5.1.3. More importantly, a low SNR will also degrade the channel estimation accuracy, which is indicated by a red arrow. This, in turn, may lead to large errors in the elements of the precoding matrix.

The quantization of the channel components and the use of a potentially lossy feedback compression will to some extent degrade the precoder, and thus reduce the spectral efficiency.

As countermeasures, there are some options to reduce the precoder sensitivity against CSI errors, represented by the two red boxes. This includes the precoder design such as the use of robust precoding as discussed in Subsection 5.2.3. The precoder sensitivity can also be reduced by appropriate user scheduling and grouping. In addition the use of CAs of smaller size typically leads to more robust precoders than the use of large CAs.

Linear precoding for a given set of users will result in a power normalization loss (PNL) when the design is scaled to fulfil the per-antenna power constraints. A given set of algorithms of scheduling, user grouping and precoder design will result in a statistical distribution of PNLs. The power normalization loss is relevant for the CSI sensitivity. It also directly reduces the spectral efficiency (SE) that can be attained for ideal CSI by coherent linear joint transmission via the resulting reduced SINRs of the UEs after power normalization.

Of equal importance as the noise floor is the interference floor. It can be significantly improved by the previously discussed tools for inter-cluster interference control. By variation of the interference 'I' in the SINR, it directly influences the spectral efficiency SE (black arrow). The channel estimation will benefit in a similar manner from a low interference at the the CSI reference signals and the robust precoder would benefit from the resulting lower uncertainty of the reported CCs (blue arrows).

Channel estimator performance and the resulting CSI accuracy are very important for JT CoMP due to the high sensitivity of interference cancellation to CSI errors. As discussed in Subsection 5.2.3 above, cooperation should be limited to low mobility UEs, using a mode switching scheme. The SNR, the UE mobility, the required prediction horizon and the carrier frequency together leads to a certain channel prediction performance and CSI accuracy, as discussed in Appendix A4-2. Schemes for partly countering large backhaul delays by decentralized precoder design have been discussed in Subsection 5.3.1 above.

Helpful for reducing the channel prediction error are

- the already mentioned interference floor shaping
- a proper CSI reference signal design, including a suitable selection of the number of RSs per physical resource block and per channel component,

- the potential use of CSI RS power boosting and optimum muting patterns [ARTD25].

The channel prediction/estimation accuracy can be represented either in the form of a normalized mean square error, or in terms of an interpolation gain (IPG) for the channel estimation, as discussed in Subsection 5.3.2.2. (The interpolation gain can be seen in Figure 5.17 in Subsection 5.3.2.2, by comparing e.g. the median NMSE to the median of the corresponding RS SNR used for generating the CDF.)

In the following, the high level model from Figure 5.18 will be used together with some system level simulation results to estimate the adjustments of the design parameters that would be reasonable for attaining a desired target performance.

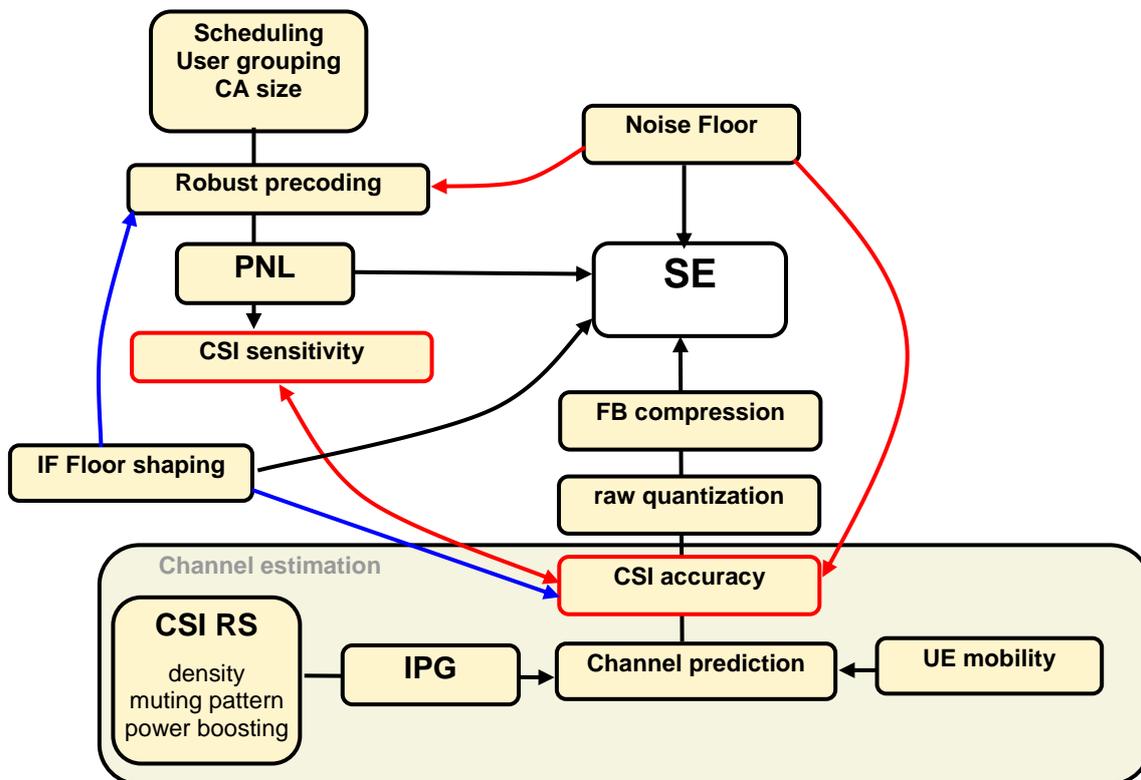


Figure 5.18: Mutual interaction between certain design parameters for the IMF-A framework.

Some important parameters are first discussed separately to obtain a high level feeling for suitable parameter selections, using system level simulations for 9-cell clusters with 4 transmit antennas per cell.

**Reporting Threshold  $TH_{cc}$ :** As IMF-A assumes a partial reporting of CCs, a suitable threshold  $TH_{cc}$  has to be found. All CCs with an Rx power  $TH_{cc}$  dB below that of the strongest CC will not be reported to save feedback overhead and to limit the effort for channel estimation. The unreported CCs are set to zero in the CA-wide precoding matrix. This leads to some intra-CA interference which reduces the SINR for the affected UEs. By system-level simulations, has been found that a  $TH_{cc}$  of about -26dB ensures good performance and -32dB close to perfect performance. For 4 x 2 systems (four antennas per base station), the number of CCs above  $TH_{cc}$  has been found to vary depending on the UE location between three and almost 20 CCs - with on average of 12-13 CCs. This clearly suggests the need for an adaptive reporting scheme.

Figure 5.19 contains the SINR CDFs from system-level simulation results for one CA for the settings  $TH_{cc} = -26$  and  $-32$  dB. Also illustrated - for  $TH_{cc} = -32$  dB – is the strong variability of the effective SINR between different UEs and different PRBs.

**Quantization of CCs:** The number of quantization bits  $N_Q$  has been varied in system-level simulations and in Figure 5.20, the SINR CDFs for different quantization levels have been plotted. The simplest IQ quantization scheme without additional compression is being used. For  $N_Q = 10$  bit in combination with CC-power dependent scaling, more or less all UEs have an SINR larger than 20 dB. With  $N_Q = 9$  there is a minor degradation. Therefore, a 5+5=10 bit (I+Q) quantization is concluded to be close to the upper bound on the required quantization granularity.

**Noise limited UEs:** Based on the system level simulations, Figure 5.21 contains the attained spectral efficiency (including physical level overhead) as a function of the noise floor level (NF), which is varied from  $-15$  to  $25$  dB. Also, different artificial interpolation gains IPG of 3, 12, 30 and 50 dB are assumed to illustrate the effect of the channel estimation/prediction accuracy.

- An IPG of 50 dB is unrealistic, but provides the SE close to that for ideal CSI knowledge. With respect to the parameters of 3GPP case 1, including an UE noise figure  $NF_{UE}$  of 9 dB, a PL of 20 dB and a  $f_c = 2.0$  GHz the corresponding overall noise floor (NF) would be about 25 dB. The resulting SE is approximately 3 bit/s/Hz/cell.
- An outdoor-to-indoor power loss (PL) of 0 dB would correspond to a NF of 5 dB, leading to a spectral efficiency of around 6.5 bit/s/Hz/cell. SNR values below 5 dB (PL=0) have been added to the figure to see to which extend the system is still noise limited and to assess potential further gains by a further reduction of NF. For example, a UE noise figure of 4 dB instead of 9 dB (NF=0 dB instead of 5 dB) would increase the SE from 6.5 to 7 b/s/Hz/cell. For lower NFs the system starts to saturate.

**Interpolation Gain IPG, channel estimation accuracy and channel prediction accuracy:** The interpolation gain mentioned above is defined as the normalized mean square error (NMSE) of the CSI estimate versus the SNR at the UE receiver of  $UE_k$ , i.e.

$$IPG_k \text{ [dB]} = -(NMSE_k(CSI) + SNR_k(N_s)).$$

In particular for radio channels with a large coherence time or coherence bandwidth, the achievable IPG for channel *estimation* might be in the range of several dBs [Aro11],[WMM+05]. In the future this might be enlarged for example by power boosting of RSs, increasing the number of RSs per CC or other by more advanced schemes.

Note, however, that interpolation gains that are discussed mostly refer to channel estimation, not channel *prediction*. For larger prediction horizons, the interpolation gain is rapidly reduced and becomes negative, as is evident in in Figure 5.17 in in Subsection 5.3.2. It is important to take estimates of the mobility statistics of the assumed user distribution into account, and convert it to prediction “interpolation” gains for the assumed channel prediction schemes. This can be done using the performance results of Appendix A4-2.

In Figure 5.21, the spectral efficiency has been simulated for different IPGs of 3, 12, 30 and 50 dB to see the sensitivity of the overall system against CSI estimation errors. With an IPG of 3 dB, which is quite easily achieved for static users and non-bursty (full load) traffic, the SE would fall back from ideally 6.5 to 4 b/s/Hz/cell (assuming a NF=5dB which gives a PL= 0 dB). A SE of 6 b/s/Hz/cell would be attainable by reducing the UE noise figure from 9 to 5dB (NF=0dB) plus improving the IPG e.g. by power boosting and/or by increasing the number of REs for CSI RSs from 3 to 12dB. It should be noted that these SL simulations have been conducted with a ZF precoder. A robust RLP precoder is expected to handle CSI errors much more effectively.

It can be concluded that advanced cooperation systems will transfer any improvements in channel estimation and reporting into a higher performance.

### Harmonization:

For harmonization of the IMF-A framework, the effects of different degradations due to the reporting threshold  $TH_{CC}$ , the quantization of CSI feedback, various SNR limitations and the channel estimation errors due to a low IPG should not be too dissimilar. To obtain the highest performance, they should furthermore all be as low as practically possible. The realistic parameter settings will vary with the assumed deployment scenario and system setup.

Figure 5.22 is basically the same as Figure 5.18 but in addition, parameter values for the most advanced system setup are given. We are here targeting a spectral efficiency (SE) of up to 5 to 6 b/s/Hz/cell while using sophisticated - but realistic - channel estimation and reporting schemes, and by using 9 cooperating base stations per cluster, each having 4 antenna elements.

This set of parameter values represents a set that is deemed both desirable and attainable. It has been selected based on all so far available results, including the simulation results provided in this subsection.<sup>25</sup>

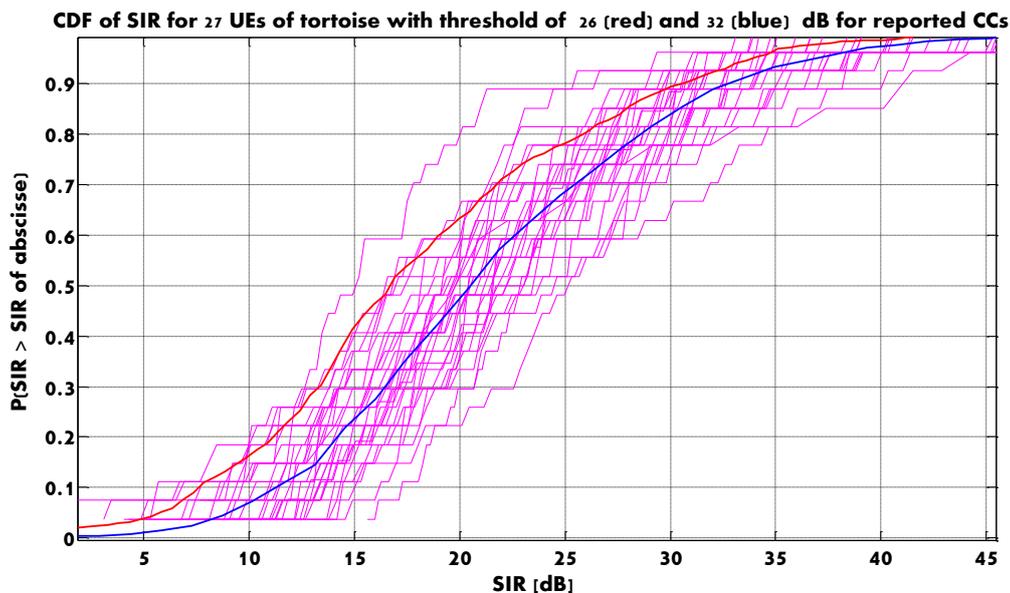


Figure 5.19: CDF of SIR for single CA for a reporting threshold of CCs of  $TH = -26$  (red) and  $-32$ dB (blue); magenta: SIR per PRB for  $TH = -32$ dB.

<sup>25</sup> At the end of the ARTIST4G project, we have found numerous avenues to further significantly improve the resulting performance by further fine tuning, e.g. by replacing the simple ZF by a RLP precoder or by novel schedulers that better take care of the SNR conditions of UEs. However, the performance impact of these modifications could not be investigated by system-level simulation for 4 x 2 MIMO transmission within the time-limits of the project.

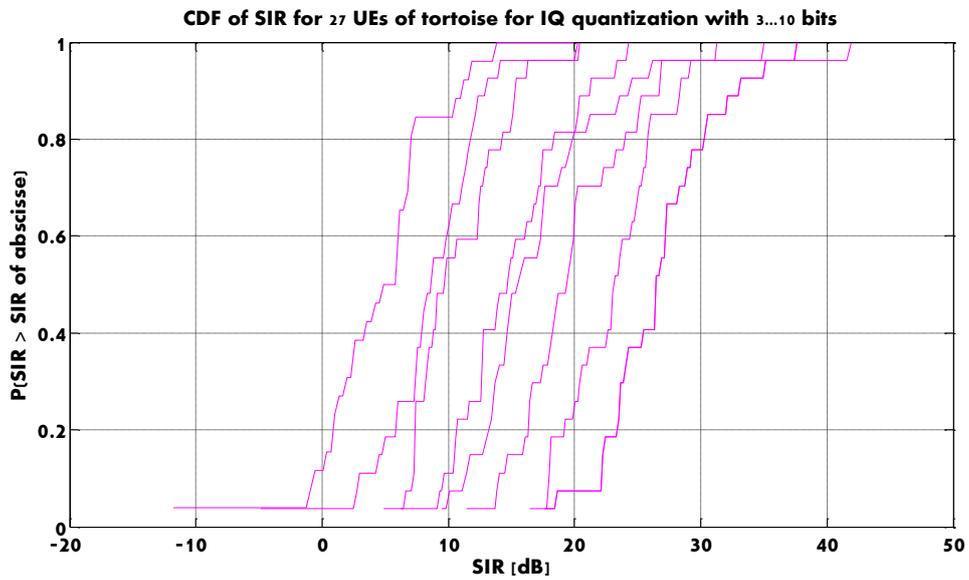


Figure 5.20: CDF of SIR per CA for IQ-quantization with 3 to 10 bit per I- and Q-component.

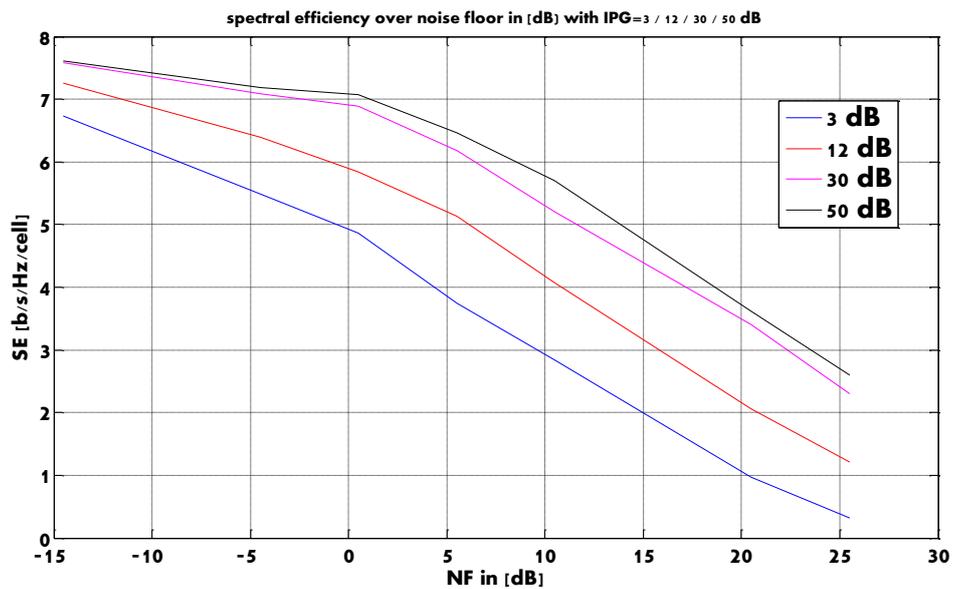


Figure 5.21: System level simulation results of spectral efficiency SE for variation of noise floor NF with IPG as parameter.

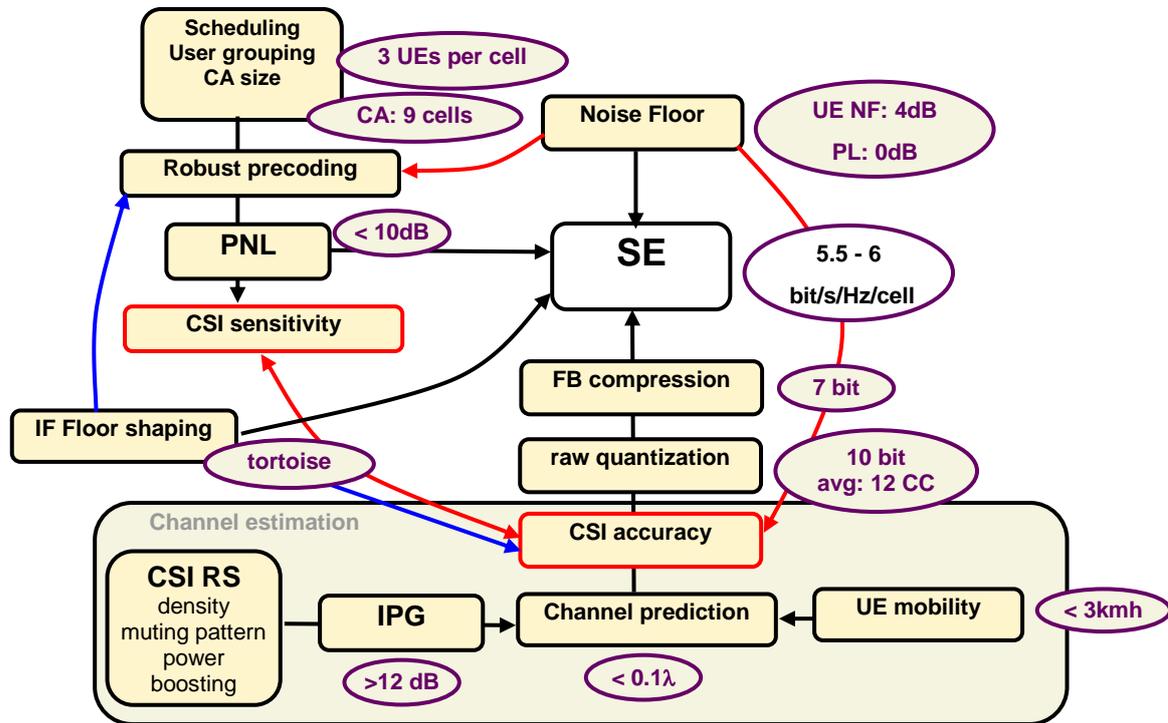


Figure 5.22: Parameter selection for IMF-A targeting highest performance.

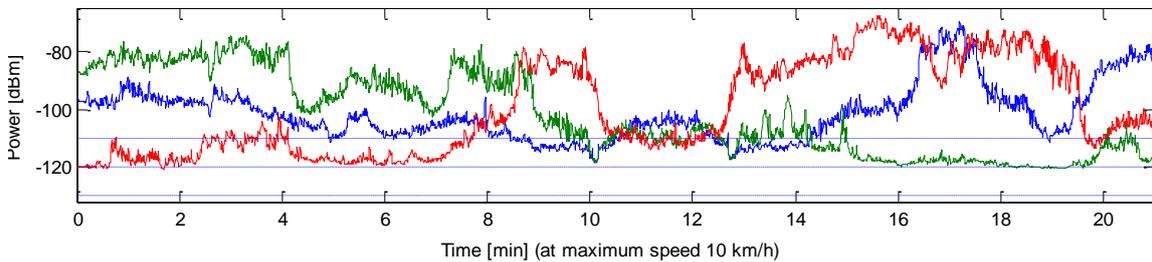
#### 5.4.2 Performance in the presence of significant prediction errors

The section above illustrated that the channel prediction accuracy (represented by the interpolation gain) has a large influence on the attainable performance. The highest performance is therefore obtained for static users. Interestingly, quite large gains can still be obtained for fast pedestrian users, for which the channel predictors provide only small or even negative interpolation gains. These results have been obtained by designing and evaluating robust and ZF linear precoders on measured channel data, using realistic prediction errors that were obtained from Kalman predictors. These investigations have been performed in Appendix A2-2 and are summarized here. The setting is partly different from the one above: Three sites with one single-antenna transmitter each are used (since measurements were available for such a setting) and the performance is measured in terms of the Shannon rate for the resulting precoded channel, not for the actual data rate in an LTE-like system simulation. Still, the results provide quite useful insights.

Here, channel samples have been obtained from sounding measurements at 2.66 GHz over a 20 MHz bandwidth, received from three transmitters in an urban environment<sup>26</sup>. The measurements have been collected by a vehicle. Noisy received signals from the three transmitters at the UE positions are generated for appropriate simulated UE velocities, based on these measured channels, see Appendix A4-2 for details. The received signal powers at different positions along the measurement route are shown in Figure 5.23 below. They can be used to represent signals from three single-antenna transmitters (cells), located at three sites, with 360-600 m inter-site distance.

<sup>26</sup> The data set was within ARTIST4G used by Uppsala University. It was previously collected by Ericsson Research in December 2008 in the Stockholm suburb Kista. See also Subsection 4.4.1 in [ARTD12].

In total 9 single-antenna UEs are placed randomly along the measurement route. At most 3 of these 9 active UEs are served within each frequency resource block by the three base stations. Noise and out-of-cluster interference is simulated by adding constant noise levels, which are indicated by three dashed horizontal lines in the Figure 5.23 below. On average over these three levels, the signal-to-out-of-cluster interference distribution is close to that which was generated by system simulation in a clustered 3-site environment when the “Tortoise” interference floor shaping method was applied, as illustrated by Figure 5.6. Differences in performance between different transmission schemes that are averaged over these interference levels should therefore indicate the relative performances of different schemes in an interference environment generated by the suggested clustering scenario.



**Figure 5.23. Signal strengths of received signals from the three omnidirectional single antenna transmitters (full lines) and the three power levels of noise added in the simulations (dotted lines).**

We let the UEs move at a maximum velocity of 5 km/h along the measurement route for 128 ms. The performance of different transmit schemes is evaluated in terms of the (Shannon) sum-rate, without any overhead, averaged over the 128 ms interval, and then also averaged over 1000 randomly selected 9-UE position sets.

The assumed total transmission control latency, including precoder calculation and backhaul delays, is here assumed to be 10 ms, i.e. double the LTE feedback latency of 5 ms. Frequency-domain Kalman predictors use measurements at orthogonal reference signals that are transmitted by the three base stations with 1.2 ms intervals to predict the downlink channels from the three transmitters as outlined in Appendix A4-2. The estimates are reported to a central controller. For each subframe in time and each physical resource block in frequency, robust linear precoders are designed based on the estimated channels and also their covariance matrices, which indicate the channel uncertainties.

**Table 5.3: Average and 5 % percentile of the *sum*-Shannon rates per cell, in bits/s/Hz/Cell, when 9 UEs are served by 3 single-antenna sites. Use of 10 ms predictions, 5 km/h at 2.66 GHz.**

	JT CoMP using Robust Linear Precoding		Single-cell transmission	
	Ave	5% perc.	Ave	5% perc.
Random grouping with RR	5.5	2.4	-	-
Cellular grouping with RR	7.7	4.0	4.9	2.3
Cellular grouping with SB	8.5	5.1	5.5	3.5

Table 5.3 exemplifies one result obtained in Appendix A2-2. It compares the average sum Shannon rate per cell obtained for five different strategies for scheduling and user grouping:

- **Single-cell transmission with Round Robin (RR) scheduling:** Users are served by transmitter with the highest average channel power. UEs allocated to same transmitter are given different frequency resources, with equal total allocation per UE.
- **Single-cell transmission with Score-based (SB) scheduling** [Bon04]. The opportunistic SB scheduler is similar to proportional fair scheduling. It uses predictions of the frequency-dependent channel gains to allocate frequency resources to UEs.
- **JT CoMP with random user grouping and RR scheduling:** All users are served jointly by three transmitters. Each of 9 users is allocated 1/3 of the frequency resources.
- **JT CoMP with cellular user grouping and RR scheduling.** Out of the 9 users, the ones with the same closest base station are allocated different frequency resources in a RR manner (e.g. 4 users are each given  $\frac{1}{4}$  of the resources). For each physical frequency resource block, the (up to 3) users within the cooperation area who share that particular block will be served by JT CoMP.
- **JT CoMP with cellular user grouping and SB scheduling.** Same as above, but using Score-based opportunistic scheduling based on predicted frequency-selective channels.

We see that JT CoMP provides significant gains over single-cell transmission with respect to the average sum rate in these experiments at 2.66 GHz, with 5 km/h user mobility in combination with 10 ms transmission latency. In this setting, with a 10 ms prediction horizon, the channel predictions produce accuracies illustrated by purple solid curves for the NMSE prediction performance in Figure 5.17. In these situations, the interpolation gain of the estimator is negative. Still, JT CoMP attains large performance gains relative to single-cell transmission.

The obtained gains depend to a striking extent on the utilized user grouping scheme. With random user grouping, JT CoMP would outperform the single-cell average sum rate (with RR scheduling) by only 12%. With cellular user grouping, the gains grow to 54% when using SB scheduling and to 57% when using RR scheduling.

Also, these results illustrate that quite large CoMP gains can be obtained when single-antenna base-stations cooperate, although higher gains are expected in the highly loaded 4 x 2 MIMO multi-user MIMO cases discussed in the previous subsection and in Subsection 5.2.2.3.

## 5.5 Conclusions and Discussion

In this subsection, a starting point has been the maximization of the system level gains under ideal conditions, in particular assuming ideal CSI. Some first steps were taken to understand the limitations and requirements regarding channel estimation as well as for CSI feedback. Without suitable enhancements, a significant part of the system level gains under ideal conditions are easily lost, but for some well justified enhancements for channel estimation, feedback, robust precoding in combination with fine tuning of the CoMP scheduler, a significant part of the gains is achievable even under real world conditions.

The here developed IMF-A framework for JT CoMP was introduced and elaborated on in Subsection 5.1.1, with the design principles:

1. use of enlarged clusters of coordinated cells, with means for inter-cluster interference suppression (interference floor shaping),
2. scheduling and beamforming design is made independently at each cellular base station,
3. use of coherent robust linear precoding, with single-cell transmission or non-coherent joint transmission as fallback modes, and
4. use of prediction of channel state information.

A general conclusion is that this framework enables a significant step ahead with respect to the state of the art performance as known for example from [3GPPTTR1].

However, it has to be emphasized that only first order effects have been analysed within the timeframe of the project and that the obtained results are preliminary. Given the complexity of JT-CoMP system level simulations – even by taking greatest care - there is always the risk of introducing modelling errors, which might easily affect the given performance results. The complete IMF-A framework could be implemented so far only by a single project partner. While the presented numerical results are likely to be subject to revision, the qualitative conclusions about relative benefits of the introduced features of the IMF-A framework are expected to hold.

To summarise some of the main points of Subsection 5.3, the key enablers for ensuring a practical high JT CoMP performance in FDD downlinks are:

- mode-switching between different transmission modes, which utilizes a strong restriction of the maximum user mobility for users participating in coherent JT-CoMP.
- state of the art channel prediction to overcome CSI outdated due to backhaul latencies, and
- an enhanced feedback scheme that should be combined with a yet unspecified compression scheme to obtain a reasonable uplink feedback overhead.

In addition, new methods were presented in Subsection 5.3.1 for decentralized design of coherent linear precoders, based on not fully consistent distributed channel state information. Robust MIMO precoding design schemes for this problem are presented in Appendix A2-7.

### **The attainable performance gains when using JT CoMP in FDD downlinks**

As one main conclusion relevant CoMP gains close to 100% over state of the art MU-MIMO seems to be possible not only theoretically but even under real world conditions. Importantly, many of our results have been backed up by real world measurements, including e.g. SNR distributions, vertical antenna tilting effects, the tortoise concept for interference floor shaping, WB beamforming patterns, attainable prediction accuracies and the effects of prediction errors.

The current status is that with moderate enhancements of the enabling technologies, a performance with a spectral efficiency of about 5 bit/s/Hz/cell – equivalently to about 70% gain over a 4x2 MU-MIMO system (3.1 bit/s/Hz/cell [3GPPTR1]) - seems to be possible.<sup>27</sup>

Midterm, a spectral efficiency of around 6 bit/s/Hz/cell seems to be in reach by quite realistic means, although a price will have to be paid by using a significant uplink feedback overhead and downlink reference signal overhead.

Ultimately the limit seems to be – at least for the most advanced here investigated 4x2 scenario and under ideal assumptions – in the range of 7 to 7.5 bit/s/Hz/cell. This is close to a factor of 2.5 improvement from the single-cell MU-MIMO case.

A simpler 3 cell measurement-based case study with single antennas was summarized in Subsection 5.4.2 above is described in detail in Appendix A2-2 and A4-2. Here, sum throughput gains of around 55% with respect to single-cell transmission were attained, in situations with 5 km/h (pedestrian) user mobility and 10 ms feedback control latency. This shows that significant gains are attainable also when simple single-antenna sites cooperate, and these gains are attainable at significant user mobilities and large downlink transmission control latencies.

Any progress in channel estimation in combination with corresponding feedback improvements will be directly converted into further performance gains. Generally the CSI accuracy should be increased as far as possible, and combined with a robust precoding scheme to make the best out of the available information. For that purpose it is very important to make the reliability of the CSI information available to the control unit(s) that designs the precoder, for example via a low rate feedback channel.

An important further aspect is the noise level of the system. The IMF-A framework is able to effectively combat inter-cell interference, but the network will benefit only in case of really interference limited scenarios. For an inter-site distance of 500m some UEs are still noise limited and therefore need special attention by the mode-switching scheme. Especially under the assumption of a strong outdoor-to-indoor penetration loss (PL) of e.g. 20 dB, the benefits of the IMF-A framework are noticeably reduced. In this respect, the trend towards decreasing cell sizes will tend to increase the attainable JT-CoMP gains.

### **The main challenges for attaining JT CoMP gains: Realizability and standardization impact**

- An important part of the concept is the use of **user centric setup of cooperation areas** with enlarged (at least 9-cell, 3-site) CAs together with the cover shift concept. In urban areas and for fiber connected sites together with modern routing or switching equipment typical backhaul delays of about 1ms are possible and even common. Some further investigations are here motivated to assess the required number of RSRP measurements as compared to LTE Release 8.
- **Interference floor shaping** - e.g. by means of active antennas - should be manageable as well given that active antennas are available and attracts increasing interest. Alternatively the use of fractional frequency reuse represents an interesting fallback solution.
- Very important has been the **per cell optimization of the user grouping** including specific CoMP schedulers and harmonized Rx-filters for the scheduled UEs. For multi-

---

<sup>27</sup> The here mentioned spectral efficiencies have been calculated taking into account a typical overhead of  $\text{OH}_{\text{LTE}}=47\%$  (plus an effective SINR degradation of 1.5dB) including the frequency guard band (10%), the OFDM symbol guard interval ( $\approx 10\%$ ), overhead for CRSs (5%) and for 3 OFDM symbols (21%) for the physical downlink control channel (PDCCH). A similar overhead has been assumed for the cellular MU MIMO case in 3GPP studies. More realistic analysis would furthermore have to take into account DMRS overhead, which can be expected to be the same for CoMP as for single-cell MU MIMO.

antenna base-stations, the optimum MU-MIMO scheduling and beamforming that is required to obtain the highest spectral efficiency results might be difficult to implement: It requires more CSI information per UE than is presently available and this will have standardization implications. For further study is to find alternative means like nonlinear precoding techniques or UE-individual vertical antenna tilting (see chapter 4) providing equivalent effects, i.e. a high orthogonality of simultaneously scheduled UEs per cell.

- The overhead for CSI RSs and CSI reporting as well as performance of state of the art CSI prediction has been discussed above. These challenges seem to be at least manageable. An important extension that has some standardization impact will be the calculation and feedback of reliability information regarding all CCs, which is required for the robust precoder calculation.

### Areas for future study

We conclude by briefly commenting on some related areas, potentially competing techniques and some areas for future study that are important to assess in relation to the here discussed methods for JT downlink CoMP.

**Interference alignment (IFA):** Interference alignment [PH09] is currently a research topic that generates considerable interest. Using IFA, precoding can be performed without exchange of user plane data. According to our terminology it is thus a CP COOP instead a UP COOP scheme. While need for CP COOP only is an interesting benefit, it has to be paid for by a typically reduced performance compared to a JP CoMP transmission, as one degree of freedom is lost for the interference terms. In the here presented investigation, we are interested in upper performance bounds and thus see IFA more as a backup solution in certain scenarios, e.g. in case of missing backhaul links. The main above mentioned challenges like accurate CSI estimation and its feedback are expected to be similar challenging for IFA as for JP CoMP. In case a fast backhaul connection is available, JP CoMP is therefore the preferable solution.

**Non linear (NL) Precoding:** Non linear precoding [C83] is known to provide – at least theoretically – further performance gains over linear coherent JP CoMP in cases when linear JP CoMP suffers from spatially incompatible UEs which results in large power normalization losses. Indeed in the future NL precoding techniques might be introduced on top of IFM-A. In our present investigation, NL precoding has not been considered for IMF-A as we have been interested in first order effects. With an optimum user grouping ensuring good user orthogonality, we do not expect large NL precoding gains but this is an important aspect to investigate further. Additionally, NL precoding techniques are expected to be even more challenging than linear precoder designs with respect to CSI accuracy requirements.

**Complexity:** Complexity of the required algorithms as well the implementation of the different algorithms will be very important for the realization in the future, and is a natural topic for further investigations. This aspect has been only partly tackled here, again as the focus had been so far on first order effects as well as on the upper performance bounds. Nonetheless, all components of the IMF-A framework are expected to be implementable with the expected processing capabilities in few years from now.

**System Load: How to utilize the improved performance:** Mobile network operators (MNO) today operate their systems typically quite far from the maximum theoretical possible load, just to ensure a continued stable operation of the network in the presence of fluctuations in the capacity requirements. This could be for example at an offered load of 50% as indicated in Figure 5.24 by the vertical dashed line.

System concepts that provide performance gains mainly for low to medium loads and only provide marginal gains for full load are represented by the blue line in Figure 5.24. From our current understanding, the benefit of such proposals seems to be limited from a capacity point

of view. For a stable operation that continues to ensure 50% headroom to full load, the improved system would have to work at the new, only slightly increased, 50% capacity point indicated by the blue horizontal line. Otherwise, if one would try to exploit the throughput gains at the old 50% offered load point by increasing the average load correspondingly, then the system would easily get into overload due to small fluctuations in the capacity requirements. In contrast, the IMF-A framework has been optimized for the full load case. Therefore, the 50% operation point scales linearly with the attainable maximal load. Thus, a MNO that introduces this transmission scheme will be able to serve more UEs while still preserving a 50% margin to the (new) full load case (red horizontal line). From an absolute margin point of view, the MNO might even choose to go beyond the 50% offered load point as the 100% load is now larger in absolute terms. The MNO will be provided with considerably wider latitude to re-distribute resources to users with poor connections. By lifting the maximal capacity limit, the here outlined JT-CoMP scheme can be used as a tool for obtaining a more even distribution of resources among users.

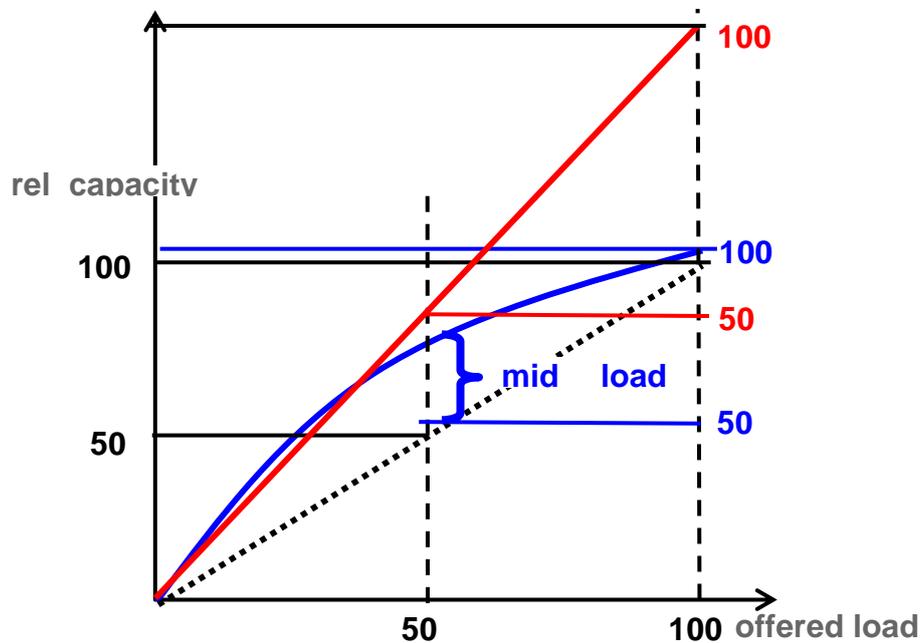
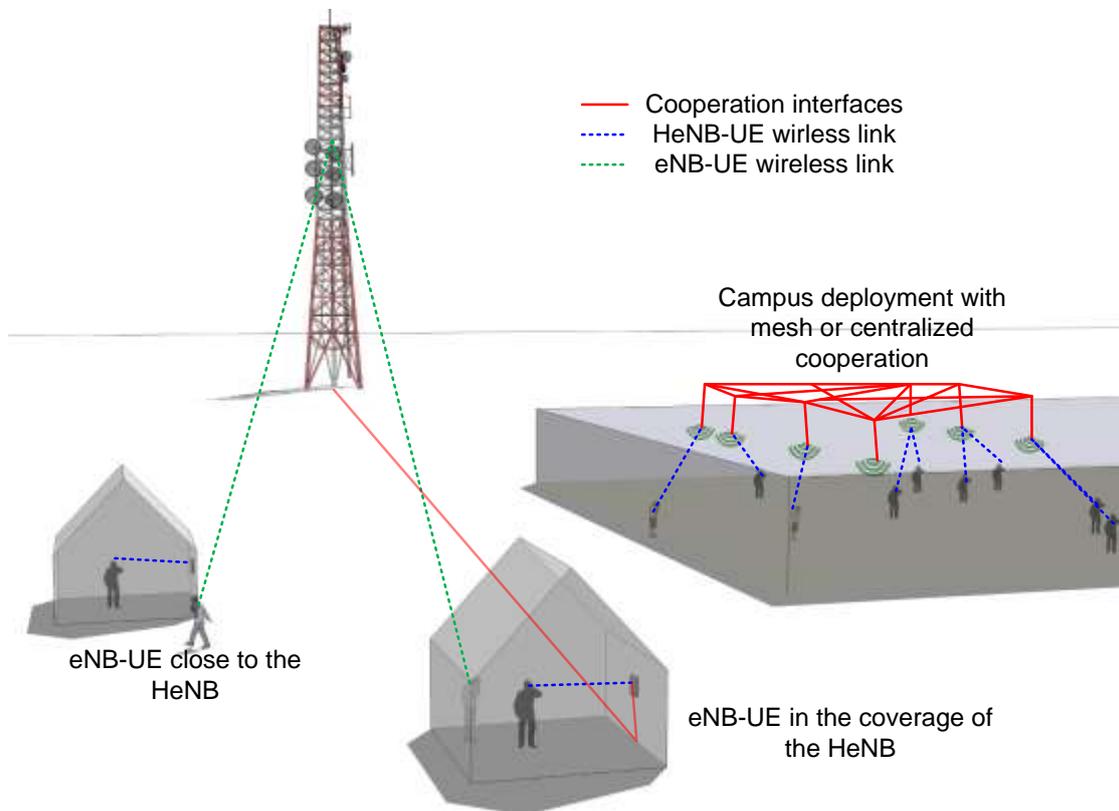


Figure 5.24: capacity versus load for IMF-A (red) versus system with mainly mid load gains (blue)

## 6 Advanced interference avoidance schemes for small cells deployments

### 6.1 Introduction

Besides the cooperation techniques between eNBs of a macro cellular network, as presented in sections 4 and 5, another solution to provide a ubiquitous user experience is the densification of the network with the benefit of dividing the number of UEs per cell and improving the propagation conditions, at the price of increasing the deployment cost. The heterogeneous network concept goes one step forward by considering multiple overlays of deployment each with different cell sizes, and allows the offloading of larger cell traffic to smaller cells. Also, different types of cells are considered, and classified with respect to the way they are connected to the core network, or to their deployment characteristics.



**Figure 6.1: Different femtocell use cases addressed with interference avoidance schemes**

According to [VENU12], “In 2016, traffic will grow 18 times (from 2011). 80% of mobile device usage is in home, office or indoors”. Thus, the indoor deployment of femtocells is the main target of this section, with different use cases illustrated in Figure 6.1.

The first important use case we focus on addresses Home eNBs (HeNBs). The idea of deploying small cells inside home comes from the observation that a significant amount of traffic is generated for serving users at home. However, the deployment of HeNBs must be low-cost, plug and play, and must not degrade the existing macro cellular network. Thus, advanced interference avoidance techniques have been developed for taking the best benefit from the offloading capabilities provided by HeNBs with the low cooperation level they can allow. These techniques are discussed in Subsection 6.2.

Another use case addresses specific dense deployments of small cells, such as campus, malls, large indoor public places, or extremely dense urban environment. These deployments are by essence random, but cooperation is made possible by a gateway-based architecture (see [ARTD42]). In Subsection 6.3, we address the interference avoidance schemes considering radio resource management techniques comprising power control and resource allocation.

Finally, for longer term deployments, we also have to consider that higher spectral efficiencies must be provided outdoor by the use of cooperation between layers of the heterogeneous networks. The non regularity of heterogeneous networks precludes the use of the specific designs developed in Sections 4 and 5. However, advanced distributed MIMO techniques are particularly well adapted to interference avoidance in HetNets, where the smaller cell is mainly interfered only by the larger cell. Thus, we investigate the performance of interference alignment schemes in Subsection 6.4. Interference alignment consists in confining interfering signals in a reduced dimension subspace leaving the remaining dimensions for useful signal transmission.

## 6.2 Massive deployment of closed HeNB (with co-channel eNB)

### 6.2.1 Problem statement

A home base station (HeNB) is by definition deployed indoor for covering customers' homes to provide high offloading capabilities to the operator and better user experience and additional services to the customer. The connectivity with the core network is relying on the fixed internet access of the customer, and the deployment appears to be random from a radio point of view.

In near future, when the extensive use of all available spectrum will be necessary to support the forecasted traffic demand growth, we can't afford the luxury of deploying HeNBs on orthogonal resource to the macro networks'. When addressing a massive deployment of HeNBs on the same channel as the already deployed macro cellular network, cross-interference between the two layers exists. It can be limited with interference avoidance techniques requiring different levels of cooperation and targeting different deployment steps, from short term to long term.

Furthermore, in the baseline 3GPP LTE-A system, a macro eNodeB can cooperate with only tens of radio neighbours through an X2 interface. This limitation is the result of the complexity for a node to setup and maintain logical X2 interfaces at the architecture level, and might be less limiting in longer term deployments.

In the downlink, a UE attached to the eNB, and not belonging to the Closed Subscriber Group (CSG) of a neighboring HeNB can suffer from high interference from it without any hand-over possibilities. For short-term deployments without any possible cooperation between the eNBs and HeNBs, we have developed HeNB power setting functions optimized so as to provide a protection for the outdoor eNB-UEs close to HeNBs. The results are shown to achieve a good trade-off between the HeNB and eNB cell-edge throughput, and protect most of the eNB users from the HeNB interference. The main achievements are presented in Subsection 6.2.2. However, a power setting strategy cannot solve cases where eNB-UEs are located inside the building covered by a closed-CSG HeNB. We address this issue via the coordination of the time scheduling and beamforming of the eNB so as to limit at maximum the interference generated by the eNB on the closest HeNBs. This approach has been previously detailed in [ARTD13] and the achievements will be summarized in Subsection 6.2.3.2.

When considering a large number of HeNBs under the coverage of an eNB, interference is also observed in uplink. We present in Subsection 6.2.3.1 the main achievements of UE power control strategies with long-term centralized cooperation between the eNB and the HeNBs that allows to achieve a good trade-off between the eNB and HeNB cell-edge throughput.

### 6.2.2 Downlink HeNB/eNB ICIC with no direct cooperation capabilities

In [ARTD13], we have presented a power setting strategy for HeNB deployments in order to protect the macro UE located outside and close to a building of a HeNB. We define a high interference zone (HIZ) as the locations where eNB UEs experience a SINR below a given threshold. In our scenario, the SINR mainly depends on the useful signal power received by the eNB, on the interference received by the HeNB, and also on the interference from the neighboring macro stations. We neglect the impact of the other HeNBs by observing that the coverage of a HeNB is limited, which induces that the probability to be interfered by more than one HeNB is low. By consequence, when all HeNBs are set with the same transmit power, the HIZ area varies with the position of the HeNB. For HeNBs deployed at the cell center, the HIZ is small because the eNB useful power is large with respect to the interference from the other nodes and from the HeNB. For HeNBs deployed at the cell edge, the HIZ is large since the impact of the HeNB on the SINR is non-negligible for eNB-UEs located around the HeNB.

Thus, a first strategy is to try to equalize the size of the HIZ so that each HeNB has an equivalent impact on the eNB UEs, and to tune the power control strategy so as to reach a target trade-off between the eNB and HeNB network cell edge throughput.

As introduced in [ARTD13] and further developed in [GUI11-1], we first compute a function  $g()$  as the macro performance degradation in the HIZ

$$g(P_M, I, P_{t,F}^{sol}) = SINR_M(P_{t,F}) / SINR_M(0).$$

where

$$SINR_M(P_{t,F}) = \frac{P_M}{N + I_o + P_{t,F}G_F},$$

and where  $P_{t,F}$  is the transmit power of the most interfering HeNB for a given eNB-UE,  $G_F$  the path gain from this HeNB to the eNB-UE,  $P_M$  is the received power from the eNB,  $N$  the constant AWGN power level at the UEs, and  $I_o$  the interference power level received from other eNBs. As a remark,  $SINR_M(0)$  denotes the SINR of the eNB-UE when the HeNB is turned off.

Then, we compute the HeNB transmit power for a target macro performance degradation  $g_{th}$ .

$$P_{t,F} = \frac{g^{-1}(P_M, I, g_{th})}{G_F} = \frac{(1/g_{th} - 1)I}{G_F}.$$

First, we assume that the HIZ is sufficiently small to approximate  $P_M$  and  $I_o$  as constant in the HIZ.

Thus, in order to compute its transmit power, each HeNB collects long term SINR measurements from its HeNB-UEs, and computes the distribution of  $SINR_M(P_{t,i})$  from these measurements. An evaluation of the HeNB building wall attenuation is necessary. An identification of the measurements associated to HeNB-UEs inside or outside the building is performed by a simple analysis of discontinuities in the c.d.f. of the SINRS measured collected at the HeNB on the long term from all HeNB-UEs. The measurements of HeNB-UEs must be corrected by the wall attenuation, so as to virtually place them as if they were outside the HeNB building, in the HIZ. Finally, the path gain  $G_F$  can be modeled with a log-normal distribution and a path loss model, or the exact distribution be obtained with HeNB-UEs measurements

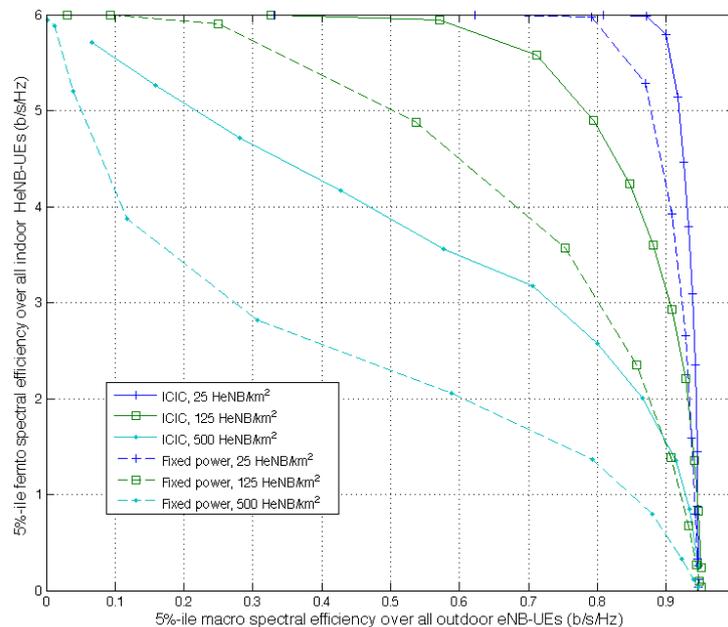
In conclusion, the power of each HeNB is computed from the measurements it obtained from its HeNB-UEs, without any cooperation with the eNB. However, the parameter  $g_{th}$  allows to adjust the trade-off between the HeNB network and the eNB performance, and must be provided to all HeNBs, with RRC signalling for example.

In Figure 6.2, we show the tradeoff between the 5%-tile of the eNB-UEs (MMT) and the HeNB-UEs (FMT) for different densities of deployments (25 125 or 500 HeNBs/Km<sup>2</sup>) with the system level simulation parameters of Table 6.1. We observe that with a constant power setting for all HeNBs, the eNB cell-edge throughput is highly degraded when the density of HeNB increases while the proposed power setting strategy highly limits the impact of the HeNB deployment on the eNB cell-edge throughput. For example, for a target 1 b/s/Hz cell edge throughput for the HeNB network, only 3% degradation is observed on the eNB cell edge throughput when a very high density of 500 HeNB/km<sup>2</sup> is deployed, while 10% degradation is observed without power setting-based ICIC.

Thus, the goal of protection the eNB-UEs close to HeNBs is achieved, without any cooperation between the eNB and the HeNBs (CP\_NOOCOOP in [ARTD11]). This innovation can be applied for current LTE deployments, and can also be applied in the future as cooperation between HeNBs and eNBs on top of more advanced cooperation schemes.

**Table 6.1: System level simulation parameters**

<b>Simulated propagation model for the eNB</b>	
Total eNB transmit power	43 dBm
Distance dependent mean path loss (dB)	$128.1 + 37.6 \log_{10}(d)$ , $d$ in km
eNB Antenna type	directional (for 3-sectorized sites) with vertical selectivity
eNB Antenna gain	14 dB
Shadowing standard deviation for eNB	8 dB
Shadowing correlation for two eNBs	0.5
Shadowing correlation distance	50 m
Wall penetration loss $A_w$	10 or 20 dB
Small-scale channel	Rayleigh ITU-TU6
<b>Simulated propagation model for the HeNBs.</b>	
Distance dependent mean path loss for indoor (dB)	$37 + 30 \log_{10}(d)$ , $d$ in m
Distance dependent mean path loss for outdoor (dB) $r$ is the radius of the femto building	$37 - 6.7 \log_{10}(r) + A_w - 36.7 \log_{10}(d)$ , $d$ in m
FBS antenna type	Omni-directional
FBS antenna gain	5 dB
Shadowing standard deviation for eNBs	10 dB
Shadowing correlation distance	Uncorrelated
Wall penetration loss $A_w$	10 dB
Small-scale channel	Rayleigh ITU-InH NLOS
Simulated 3GPP-LTE physical layer.	
Modulation waveform	OFDM
Bandwidth	5 Mhz
FFT size	512
Useful sub-carriers	300
Sub-carrier spacing	15 kHz
MMT/FMT allocation size	12 sub-carriers
Maximum spectral efficiency	6 b/s/Hz (SIMO, 64-QAM, coderate 1)



**Figure 6.2: Tradeoff between the eNB cell edge throughput and the HeNB network cell-edge throughput with and without the proposed power setting strategy and for different HeNB deployment densities**

### 6.2.3 HeNB/eNB interference avoidance scheme with slow cooperation capabilities

In this section, we now assume that a cooperation interface is available between the macro and the HeNB network for slow cooperation (CP\_COOP(S) in [ARTD11]). In Subsection 6.2.3.1, we first investigate the uplink scenario, and show that a minimum level of cooperation is needed. We focus on keeping this cooperation level as minimal as possible, by the use of a central unit that aggregate all HeNBs and allows to optimize them as a whole from the eNB perspective, and consider only optimization based on long term statistics. In Subsection 6.2.3.2, we target scenarios where an eNB-UE is close to a CSG-HeNB while not belonging to its CSG. The occurrence of such a scenario is of small but non-null probability, but might lead to an outage for the eNB-UE. Thus, for these specific situations, we assume that a cooperation link can be set up between the eNB and the HeNB so as to perform a coordinated beamforming and scheduling, and we have focused on schemes requiring slow control plane exchange.

#### 6.2.3.1 Semi-centralized power setting for UL ICIC in co-channel eNB/HeNB deployments

The interference from HeNB-UEs on eNB-UEs observed in the uplink at the eNB is to our knowledge not addressed yet in the literature. We have focused on this topic after observing that each HeNB-UE generates a negligible interference level at the eNB, but their potential high number transmitting on the same resource adds up into a resulting non-negligible interference level. The high number of HeNB under the coverage of the eNB makes it impossible in a mid-term deployment to rely on advanced interference avoidance scheme involving an eNB/HeNB cooperation. Instead, we proposed in [ARTD13] and [GUI11-2] to address the interference from the HeNB network as a unique interferer from the eNB point of view, which requires three features: an optimization algorithm, an architecture solution and physical layer enablers for relaxing even more the architecture constraints.

The optimal power control and interference avoidance scheme would gather measurements from all eNB-UEs and HeNB-UEs in the central unit. This would require periodic measurements updates so as to at least track the long-term SINR changes due to mobility, or even to optimize the transmit power according to the fast fading channel variations. As we are targeting mid-term

deployments for which cooperation is highly constrained by the architecture, our goal is to find the best trade-off between performance and signaling overhead. Thus, we are targeting solutions with very slow control plane exchange, i.e., with exchanges not faster than every minute.

The algorithm is performed in a centralized fashion, in a central unit that computes the optimization for each cell of the eNB network independently (resulting in a semi-centralized solution). Thus, we propose a sub-optimal strategy where power control is performed in a semi-static fashion, and in two steps:

First, each HeNB and eNB performs a power control algorithm without taking into account any impact on the other cells (as in the baseline system). Of course, the power control strategy of each HeNB has an impact on the performance since it shapes the interference distribution from the given HeNB to the eNB and also impacts the performance of the HeNB uplink. We have evaluated 5 different power setting methods at the HeNB that will be detailed in the following.

Then, a correction is applied, and common to all HeNBs-UEs so as to finely tune the trade-off between the macro and HeNB network performance. The correction is a semi-static parameter computed from long-term statistics collected at the central unit. More precisely, the statistics are computed from the measurements performed at the HeNB-UE from the eNB and HeNB downlinks (e.g., the distribution of the path gains containing information of multiple positions of HeNB-UEs in the HeNB coverage). Thus, the semi-static correction parameter is not computed for an instantaneous snapshot of HeNB-UEs, but for a statistical distribution of the HeNB-UEs instead, which drastically reduces the signaling overhead with the central unit, at the price of a sub-optimality of the power setting. The correction factor is broadcasted to all HeNBs and shifts the resulting distribution of interference from all HeNBs in order to reach a target degradation at the eNB when comparing the performance with and without HeNB deployment (see [GUI11-2] for more details).

We have evaluated 5 methods for the power control at HeNB, leading to different correction parameter computation and signaling requirements, also taking into account the actual load of each HeNB. The three first methods are used as reference extremum strategy so as to introduce the naïve solution of method 4 which provides the essence of our strategy, which is improved in method 5. In the following, we detail the five methods and provide their advantage and drawback, as well as performance assessments as illustrated in Figure 6.3.

**Method 1-No power control:** All HeNB-UEs transmit with the same power. In a degraded version, no information is exchanged between nodes, but the performance is improved by exchanging the HeNB load information, or an window-average version of it; with the central unit.

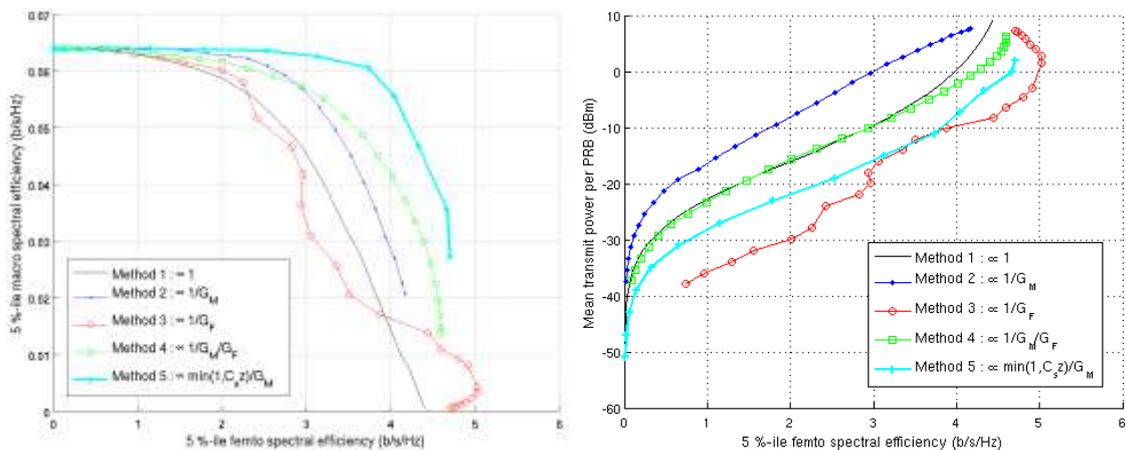
**Method 2-Macro protection:** The HeNB-UEs transmit power is set according to the inverse of its path gain to the eNB, which maximizes the eNB performance but degrades the HeNB performance. As for Method 1, only the HeNB load feedback is needed. We observe on Figure 6.3 that the cell-edge tradeoff between the eNB and HeNB network is improved at the price of larger power consumption.

**Method 3-Femto capacity maximization:** The HeNB-UEs transmit power is set to maximize the HeNB performance, i.e., by inverting the path gain to the HeNB. In order to control the eNB performance degradation, the average ratio of the path gain from the eNB and the HeNB, measured by the HeNB-UEs on a long term basis, is provided to the central unit. The cell-edge throughput tradeoff is generally not improved, but the power consumption is decreased.

**Method 4- Combination of Method 2&Method3:** The HeNB-UEs transmit power is set to invert the product of the path gains between the target HeNB-UE and the eNB and HeNB. This makes a tradeoff between the useful signal received at the HeNB and the interference received at the eNB. Indeed, when the HeNB-UE is close to the HeNB, there is no need to transmit at high power or when the HeNB-UE is close to the eNB, its transmit power should be limited in order to control the interference generated on the eNB-UEs. Only the expectation of the inverse of the path gain to the HeNB needs to be provided to the central unit for the computation of the

correction factor. We observe on the system level simulations that the tradeoff between the cell-edge throughputs of the HeNB and eNB network is improved, for a fixed power consumption with respect to Method 1.

**Method 5-improvement of Method 4:** The HeNB-UEs transmit power is set as the minimum of the inverse of the path gain between the target HeNB-UE and the eNB and the weighted path gain between the target the target HeNB-UE and the HeNB. The weighting parameter makes a tradeoff between the Method 2 and Method 3, and is computed at the central unit from the long term statistics provided by all HeNBs (see [GUI11-2] for more details). In that case, each HeNB computes the mean and variance of the path gains between the HeNB-UE and the eNB or the HeNB, and feed it back to the central unit that computes the correction and weighting parameters and broadcast them back the the HeNBs. This Method allows for both improving the tradeoff between the cell-edge throughputs of the HeNB and eNB, and exhibiting a low HeNB-UE power consumption.



**Figure 6.3: Trade-off between the cell-edge throughputs of the HeNB and eNB (left), and the HeNB-UE power consumption (right)**

In conclusion, we have shown that we can achieve a good protection of the eNB network, while guaranteeing good performance for the HeNB users. The signaling exchange between the eNB, the HeNBs and a central unit is limited. Also, the central unit broadcast semi-static parameters to whole HeNBs for tuning the uplink power control of the HeNB UEs.

In terms of architectures, two enablers are provided in [ARTD4.2] and are the results of the cooperation between WP1 and WP4. The location of the central entity is discussed and the HeNB-GW architecture seems to be promising for several reasons. First, if the HeNB-GW supports the optimization functionalities, a unique cooperation channel is established between the eNB and the HeNB-GW, which circumvent the limitation at the eNB of direct cooperation with a large number of radio neighbours.

In general, a Physical Cell Identity (PCI) is selected for each eNB or HeNB. A PCI is mapped to a synchronization signal, which is broadcasted by the node at the physical layer. This allows for one node to make measurements on its radio neighbours via its UE. From the signaling, a node can also obtain information for contacting the radio neighbours and set up cooperation interface at the core network level. Thus, a Group Physical Cell Identity GPCI is introduced to allow for the eNB making measurements via the eNB-UEs of the HeNBs, as if they were a single interferer. More precisely, the HeNB transmit, on top of (or orthogonal to) their own synchronization signal mapped to their Physical Cell Identity (PCI), a signal mapped to a PCI common to the whole HeNB fleet and addressing the HeNB-Gateway (HeNB-GW) for setting up a cooperation interface.

In conclusion, the eNB only sees one Grouped-HeNBs interferer at the radio and architecture level, the HeNB-GW cooperate with the eNB and all HeNBs, collect measurements on the long term basis, and broadcast semi-static parameters that allows to achieve a good tradeoff between the eNB and HeNB cell edge performance.

### 6.2.3.2 Coordinated scheduling based on restriction requests

The power setting strategy detailed in previous subsection cannot solve cases where eNB-UEs are located inside the building covered by a closed-HeNB (i.e., where the access control is limited to a closed subscriber group CSG), as illustrated in Figure 6.1. Indeed, the path gain of the HeNB is not weighted by the wall attenuation while the useful path gain to the eNB is, which creates at least 20dB penalty on the SINR when compared to the case where the eNB-UE is outside the building. As a result, the HeNB must be turned off to enable the communication between the eNB and the eNB-UE. Thus, more advanced schemes are required in that case relying on the orthogonalization of the resource use. The carrier aggregation (CA) is one of the most important features of the 3GPP LTE-A, and can be used for the frequency domain orthogonalization of the small cell and Macro UEs. Sub-frame muting via ABS (almost blank sub-frame) is another alternative studied in 3GPP, which allows for a partial orthogonalization in the time domain. We propose an alternative via the coordination of the time scheduling and beamforming of the eNB so as to limit the interference generated by the eNB on the closest HeNBs, while not losing spectrum usage at the HeNB. This has been proposed in [ARTD13] and the achievements will be summarized here.

The innovation is based on the 3GPP LTE Rel.8 precoding matrices codebook, and a first step of scheduling and link adaptation is performed at each node. Then, based on the measurements performed at the UEs, a selection of highest interference generating precoding matrices is done at the interfered node, and a restriction request of the codebook is sent to the interfering node. This innovation is particularly relevant for unbalanced interference setting when the most significant part of the interference is generated by one node on the other only. Indeed, in that case, the restriction requests are only mono-directional from the interfered node on the other, and no decision making is needed at the nodes. In our scenario, the eNB send restriction request to the HeNB who is interfering only on one eNB due to its small coverage area. From the architecture point of view, the number of active eNB-UEs in such situations that require the use of the innovation is low at a given time, but stays for a long time. Thus, we can imagine setting up few X2 interfaces with the identified interfering HeNB. The signaling for applying the restriction request is not supported yet by the baseline system, and must be included in future standards to enable the use of the innovation. System level simulations have shown that a gain of 40% of throughput can be achieved for the eNB-UEs in the same building as the HeNB, while not degrading significantly the HeNB performance.

## 6.3 The Femto campus use case

### 6.3.1 Problem statement

In this section, we address the scenario where femto-cells are used to cover indoor spaces such as companies, malls, hall stations, etc... The main difference to the massive deployment of HeNBs as studied in Subsection 6.2 is that the deployment is performed by a single owner, along with a fixed network infrastructure, that allows cooperation between the femto-cells.

We assume that no fine planning of the deployment can be done. Also, the load of each cell, and the number of active UEs in each cell varies randomly and much more quickly than for larger cells deployment. As a result, any optimization of the radio access network based on a statistical distribution of the users will not provide the best performance.

Thus, the main challenges for these scenarios are to design innovations enabling self organization at the core and radio access networks. We focus here on interference avoidance, and will describe the achievements of three innovations allowing radio resource management with centralized or decentralized implementations.

### 6.3.2 Centralized power control for femto campus

In this section we will introduce and discuss centralized techniques for power setting in the general context of heterogeneous networks. More precisely we will focus on campus deployment of femto base stations by detailing DL power control techniques for two scenarios:

**CSG campus:** Where the base stations of the campus are closed subscriber group base stations, i.e. the UEs in one femto base station coverage are not allowed to connect to the neighboring base station.

**Open campus:** Where the base stations of the campus are open base station, i.e. the UEs of the base station's coverage are allowed to connect to the neighboring base stations.

The techniques we are considering are centralized since it is assumed that a central entity collects information from the base stations of the campus and sets their transmit powers in order to optimize a performance criterion.

#### 6.3.2.1 Optimization problem formulation

Let us denote the cell edge SINR in the cell  $i$  as

$$SINR(i) = \frac{\alpha_{i,i} P_i}{\sum_{j \neq i} \alpha_{i,j} P_j + N_0}$$

Where:

$P_i$  is the transmit power of the node  $i$

$\alpha_{i,i}$  is the cell edge path gain of the node  $i$

$\alpha_{i,j}$  is the path gain of the interfering node  $j$  over the node  $i$  evaluated in the cell edge of the node  $i$ .

$N_0$  is the additive noise power

The centralized optimization problem studied in this section is based on the maximization of the generalized mean of the cell edge SINRs as:

$$\begin{aligned} & \text{maximize } S_{\beta}(SINR(1), SINR(2), \dots, SINR(n)) \\ & \text{subject to } P_i \leq P_0 \end{aligned}$$

The generalized mean function  $S_{\beta}(SINR(1), SINR(2), \dots, SINR(n))$  of the cell edge SINR, with parameter  $\beta$  is defined as

$$S_{\beta}(SINR(1), SINR(2), \dots, SINR(n)) = \left( \sum_{j=1}^n SINR(j)^{\beta} \right)^{\frac{1}{\beta}}$$

The Lagrangian of the optimization problem is given as

$$L(P_1, \dots, P_n, \underline{\lambda}) = S_{\beta}(SINR(1), SINR(2), \dots, SINR(n)) - \sum_{i=1}^n \lambda_i (P_i - P_0)$$

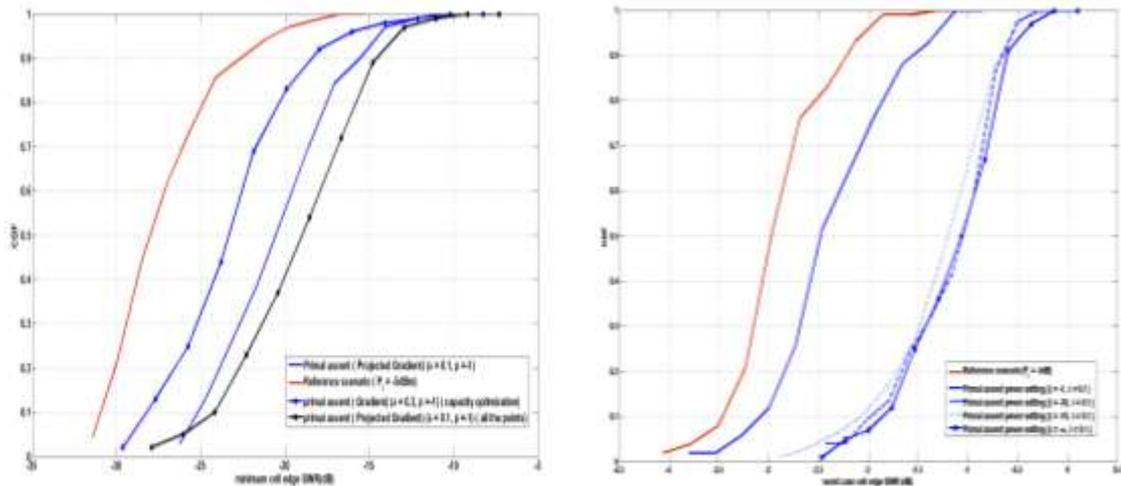
The optimization techniques used for solving the centralized optimization problem described above are based on Gradient ascent. The basic idea behind Gradient ascent is to search iteratively for the transmit powers that optimizes the generalized mean of the SINR's, in the direction of the gradient of the Lagrangian. Since the SINRs are non-convex in the powers we have used the dual parameters  $\lambda_i$  and pre-conditioned conjugate gradient techniques, based on the Polak-Ribiere direction calculation and adaptive step (details can be found in the Appendix A6-1).

We have also considered in this study power setting techniques based on the maximization of standard fairness metrics among the cell edge spectrum efficiencies of the open and CSG campus. The fairness metrics considered are the Rag-Jain index and the generalized proportional fair metric (details can be found in the Appendix A6-1). We have used for the power setting technique similar gradient ascent as described previously. The common simulation assumptions used for the evaluation of both scenarios are recalled in the Table 6.2.

**Table 6.2: Common simulation assumptions**

Parameter	Value
Carrier	2GHz
Antenna Gain	5dBi
Body loss	6dB
Powers	Minimum value : -5dBm Maximum value: 20dBm
Path gain model	Motley Keenan : $PL(\text{dB}) = a + b \log_{10}(r)$ a = 127 b=30 r: radius in Kms
Walls attenuation	10dB
Noise density	-174dBm/Hz
Noise factor	9dB
Shadowing	No Shadowing

The basic results of the power setting algorithm are given in the following figures.



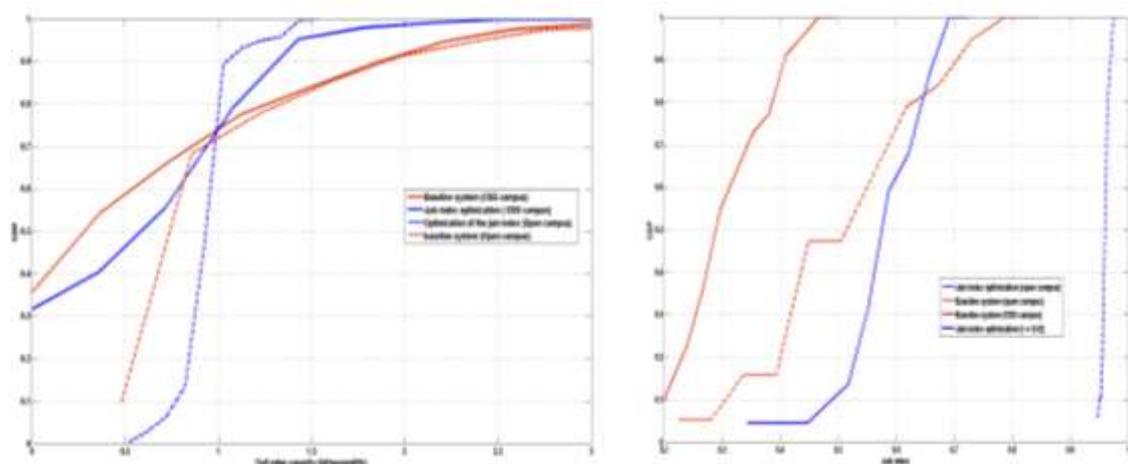
worst case cell edge SINR (CSG campus)

worst case cell edge SINR (Open campus)

**Figure 6.4: WC cell edge SINR performance (CSG campus (a) and open campus (b))**

It is shown that the power setting technique improves the worst case cell edge performance for CSG campus (case (a)) and for open campus (case (b)) when compared to the equal power base line system (red plots in the figures). The improvement shown in the open campus scenario is more significant since the SINR range allows for the open campus spectrum efficiency improvement.

In the Figure 6.5 we have plotted the cell edge spectrum efficiencies for the open and CSG campus scenarios and the corresponding jain index cumulative distribution function.



cell edge spectrum efficiency for CSG and open campus

Jain index CDF for CSG and open campus

**Figure 6.5: Cell edge spectrum efficiencies and Jain index for CSG and open campus**

It is shown in the figure that the cell edge spectrum efficiency is around 0.3 bit/second/Hz for the CSG campus when the reference performance is around 0.28bit/second/Hz. The improvement for open campus is around 0.25 bit/second/Hz for a nominal performance around 0.7bit/second/Hz. Both improvements are evaluated for 50% of the topologies of the simulations. Regarding the Jain index, the overall improvement is 0.3 for a nominal performance of 0.28 for CSG campus and 0.46 for a nominal performance of 0.51.

Focusing on the best performance, i.e. open campus scenario, we report the Jain index evaluated by taken into account all the sampled topologies in Table 6.3. In Table 6.3 we have evaluated the Jain fairness index for both cell edge SINRs and cell edge capacities for various parameters of the generalized mean.

**Table 6.3: Jain index evaluation for the open campus scenario**

$\beta$	Cell edge SINRs		Cell edge capacities	
	Reference system Jain index	Power setting system Jain index	Reference system Jain index	Power setting system Jain index
-1	0.4596	0.6667	0.8144	0.8834
-10	0.4378	0.5449	0.7926	0.8621
-70	0.4589	0.5035	0.8013	0.8439

As a conclusion, we have studied power setting techniques based on the maximization of three fairness metrics for CSG and open campus deployment scenarios. The fairness metrics considered are: The generalized mean of the cell edge SINRs, the Jain index of the cell edge SINRs and the generalized proportional fairness metric, i.e.  $\alpha$ -fair metrics described in the Appendix A.6-1. It is seen that the Jain index based power setting improves the overall cell edge SINRs of the CSG and open campus base stations. Generally, the improvement of the cell edge spectrum efficiency is higher for the open campus and when 400 random positions of the campus are considered.

### 6.3.3 Coordinated scheduling for heterogeneous deployment

In [ARTD13], we have made an analysis to clarify the potential gain brought by coordinated scheduling and CoMP transmissions in a femto campus deployment. We have considered varying cell loads, i.e., not all the resources are being used at each cell. We have considered a campus deployment where all femto cell are in square rooms placed in a grid fashion, separated by a wall with an attenuation of 10dB. Thus, the interference from the neighboring nodes is in general limited and occurs in positions where the interferers are located right behind the wall.

We have considered interference avoidance coordinated scheduling based on the free resource of the neighboring cell. In other word, one cell reserves resource in advance for a subset of selected UEs, and communicates the restriction to the neighboring cell that does not schedule any user on said resource. With this approach, there is no gain for the full load case, which is not realistic for femto cell networks, and the throughput of the worst UEs is increasing when the cell load of the highest interferer decreases.

For the sake of comparison, we have considered a simple joint transmission scheme where the same signal as the serving femtocell is sent from the main interferer as well. We have compared two strategies where the price to pay for the resource needed for CoMP is shared between all the users of the cell, which shows a global degradation of the cell performance. Thus, we then restrict the usage of CoMP on the free resource of the cooperating cell, and show that the gain

with respect to cooperative scheduling only is low. This conclusion is mainly due to the fact that the wall between the two cells limits the SINR boost usually observed at the cell edge by using CoMP. Thus, interference nulling is sufficient in that case, and a practical implementation of distributed schedulers by using a game theoretic approach is detailed in 6.3.4.

### 6.3.4 Resource allocation in slow fading interfering channels with partial knowledge of the channels

In this section, we show how the innovation based on game theory and detailed in [ARTD13] plays an important role in the femto campus scenario. The femto-cells are sharing a common pool of frequency resources, and the deployment is characterized by three key features:

It consists of a large number of nodes with intelligence that take autonomous decisions. This is a fundamental aspect to support self organization and network scalability.

Channel qualities, requests of service, and traffics afferent to individual nodes fluctuate randomly and more quickly than in a macro deployment, due to the low number of active UEs per cell. From the perspective of a global exploitation of the available bandwidth, this property offers the possibility to opportunistically benefit from channel, user, and traffic diversity.

In contrast to CoMP systems based on an intense data sharing on both the data and control plane, it was shown in Subsection 6.3.3 that user plane exchange only provides a limited gain for non fully-loaded campus deployments, and that most of the gain can be obtained with scheduling cooperation. Thus, we consider no sharing on the data plane and limited exchange on the control plane.

The objective of this contribution is the design of resource allocation algorithms capable to take benefit from channel, user, and traffic diversity to the extent enabled by the available control information and implementable locally on each node thanks to the decision capabilities of each femto-node.

On one hand, cross-layer design approaches benefit at the best from different kinds of diversity since they jointly optimize the exploitation of statistical variability. As an example, joint rate and power allocation, scheduling and admission control is a cross-layer design that can simultaneously optimize the system with respect to channel quality fluctuations and traffic randomness.

On the other hand, game theory is a mathematical framework for multiple decisions making to determine the strategies/policy/actions of multiple independent decision-makers (players) in a context where the action of each decision maker affects the system in an intricate way and the behaviour of the full system depends on the policy adopted by each of the decision makers. In game theory, multiple decisions are obtained as the equilibrium point for a system of decision makers where each decision maker aims at maximizing (minimizing) its own utility cost and its utility cost depends on the action taken by all the other players. In contrast to the theory for a single decision maker where the decision is a maximizer (minimizer) of a utility cost function and optimizes the system, the solution offered by game theory to multiple decision making is an equilibrium point. This point does not necessarily coincide with an optimum and may even be inefficient. However, it guarantees the stability of the system in the case where the decisions are taken independently by the decision makers and does not require the existence of a centralized authority to enforce optimum strategies. It is worth noting that the utility functions of players do not need to target conflicting objectives. Thus, the difference between an optimization and a game equilibrium is not the difference between a cooperative and a competitive approach but rather the difference between a system of dumb nodes whose individual instances are irrelevant in determining the working point of the system and a system of rational nodes that interact together to find an equilibrium. While a careful choice of the utility function can avoid inefficient equilibrium points, the philosophy of letting a system working at an equilibrium point is intrinsically characterized by self-organizing and self-healing properties that strongly facilitate issues of deployment, configuration, and maintenance of the global system. In addition to the mentioned benefits, often the search for an equilibrium point can be easier than the search of an optimum point. Therefore, a game theoretic framework is suitable to model network scenarios

where individual decisions impact the performance of every node but in which centralized operations want to be minimized.

With this objective in mind of designing control functions that exploit different kinds of diversity being locally implementable by femto cells and macro eNBs with autonomous decision capabilities, it seems very promising to apply game theory to develop cross-layer design algorithms.

Cross-layer design approaches have been studied in single-cell systems, i.e. for networks where the frequency bands are orthogonally allocated to different entities and not shared, both in uplink and downlink. In such settings, cross-layer design resulted in large performance gains. A straightforward extension of these available results to femto campus would require a complete sharing of data among macro and femto cells on the data plane and additional data exchange on the control plane including instantaneous CSI at the transmitters and instantaneous occupation of the queues. Due to the restrictions on the data and control information exchanges, the existing cross-layer algorithms cannot be directly adapted. Our proposed approach will solve this issue.

To define cross-layer algorithms for a system with distributed intelligence among nodes and limited amount of CSI we resort to a branch of game theory which deals with multiple decision making in contexts where each player has only partial knowledge of the system, more specifically, of the impact that the actions of other decision makers have on its utility and on the full system behaviour. Such a branch is dubbed Bayesian game theory. Bayesian game theory offers the most suitable game theoretical framework to design cross-layer algorithms based only on locally available measurements. In the innovation developed in [ARTD13], we show that a game theoretic approach for computing the transmission power level, the transmission rate, and the user selection (in practice, the packet acceptance/rejection which can be understood as a user selection in a multi-user system) show the best performance for systems satisfying the following conditions

The transmitter has CSI knowledge, which means that the channel has slow time fluctuations, which is generally the case for indoor deployments

The transmitter also has a CSI knowledge of the interferer link, which usually requires control exchange or to take benefit from the channel reciprocity in a TDD system

The nodes has the knowledge of traffic statistics of the other nodes, which can be obtained via long term cooperation (CP\_COOP(S))

Thus, our innovation is particularly suited for the deployment of the femto campus in a TDD band, which is one envisioned possibility to use TDD channels that are usually not preferred for macro cellular networks in Europe. Furthermore, this innovation does not put stringent requirements on the architecture since only a low amount of data with no latency constraints has to be exchanged between the nodes.

## 6.4 eNB/Pico/Relay Heterogeneous deployments

### 6.4.1 Problem statement

Mobile broadband traffic is increasing rapidly, driven by many factors such as the improved performance of mobile networks and the availability of new devices, like smartphones and tablets, and new mobile applications introducing new ways using mobile devices and services. There is clearly a demand for high data-rate services in the mobile environment. Rapidly increasing traffic volume demands and higher QoS expectations can be met in different ways, for example by an improved or densified macro network. Finding sites for macro base stations is however in many cases difficult. Finding sites for smaller base stations can be easier. Deployment of such smaller base stations results in a heterogeneous network.

The deployment of low power nodes improved the user's performances in the macro network since the low power nodes manage to significantly offload the macro layer. This relies on that the low power nodes; such as picos, femtos, and relays; are deployed so that they cover hotspots or coverage holes. To enhance the performance of these heterogeneous deployment, advanced techniques are described which are needed to manage and control interference and deliver the full benefits of such networks. Radio resource management is required to guarantee QoS, the adaptive inter-cell interference coordination is a solution to provide smart resource allocation amongst interfering cells in a heterogeneous network.

An interference alignment technique has been proposed above in order to mitigate cell interference between pico and macro cells.

### 6.4.2 A Practical Iterative Algorithm for Joint Signal and Interference Alignment in Heterogeneous Networks

#### 6.4.2.1 Overview of Interference Alignment

In this section, we consider the case of femto and pico access point (AP) deployed in the coverage area of a macro AP. Each AP serves one user and interferes with all the other users served by the other APs regardless of their type (macro/pico/femto) as illustrated in Figure 6.6. Each AP and each user are considered to have  $M_T$  and  $M_R$  transmit and receive antennas, respectively. Finally, the following assumptions are made about the level of cooperation between the APs:

No useful data exchange is allowed between the different APs;

A limited level of control information exchange is authorized between the APs due to the fact that we can not afford a high speed backhaul.

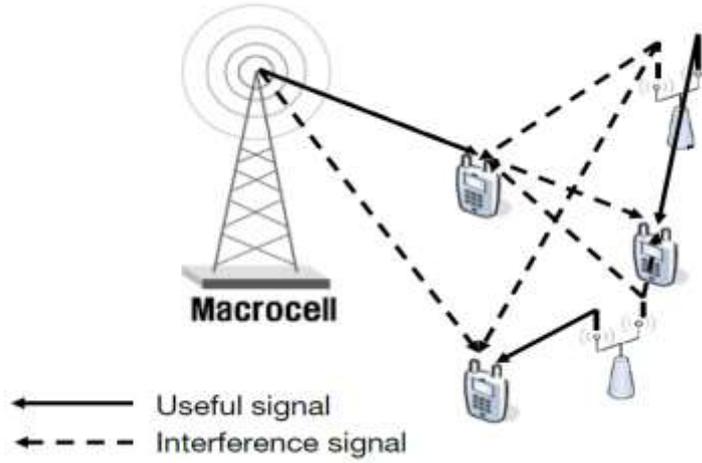


Figure 6.6: an example of the studied scenario

Such a system can be described by an interference channel (IC) [C78] [C75]. In fact, IC models the situation where the transmitter communicates with its intended receiver while generating interference to the other users in the network.

Recently, an introduced promising solution to combat interference for the IC is the interference alignment (IA) technique which has proven to achieve the sum-capacity of the time varying or frequency selective  $K$ -user IC [CJ08]. IA exploits the time/frequency or spatial dimensions to confine the interfering signals in a reduced subspace leaving the remaining space for the desired signal. IA pre-coding design presented in [CJ08],[TJ10],[GMK09] exploits time varying aspect or frequency selectivity of the channel to reach the channel capacity. For constant MIMO interference channels, the provided dimension is limited since IA schemes can exploit the spatial dimension only. Such constraint makes the gain from IA depend on the number of transmit, receive antennas, and the number of active links.

In a  $K$ -user constant MIMO IC, the transmitted signal from AP  $k$  to its intended receiver  $k$  is given by

$$\mathbf{x}_k = \mathbf{V}_k \mathbf{s}_k,$$

where  $\mathbf{V}_k \in \mathbb{C}^{M_T \times d_k}$  is a unitary pre-coding matrix applied to the streams vector intended to user  $k$ ,  $d_k$  is the number of streams transmitted on link  $k$  and  $\mathbf{s}_k \in \mathbb{C}^{d_k \times 1}$  is the vector containing the data symbols intended to user  $k$  and verifying  $E[\mathbf{s}_k \mathbf{s}_k^H] = \frac{P}{d_k}$ , with  $P$  is the total transmit power available at any AP.

The received signal  $\mathbf{y}_k \in \mathbb{C}^{M_R \times 1}$  by user  $k$  is given by

$$\mathbf{y}_k = \mathbf{H}_{kk} \mathbf{V}_k \mathbf{s}_k + \sum_{k' \neq k} \mathbf{H}_{kk'} \mathbf{v}_{k'} \mathbf{s}_{k'} + \mathbf{n}_k,$$

where  $\mathbf{H}_{kk'} \in \mathbb{C}^{M_R \times M_T}$  is the channel matrix between user  $k$  and AP  $k'$  and  $\mathbf{n}_k \in \mathbb{C}^{M_R \times 1}$  is an additive complex white Gaussian noise vector with covariance  $E[\mathbf{n}_k \mathbf{n}_k^H] = \sigma^2 \mathbf{I}$ .

At the reception, each user  $k$  decorrelates the received signal by applying the decorrelation matrix  $\mathbf{U}_k \in \mathbb{C}^{d_k \times M_R}$ , which leads to the following relation

$$\mathbf{U}_k \mathbf{y}_k = \mathbf{U}_k \mathbf{H}_{kk} \mathbf{V}_k \mathbf{s}_k + \mathbf{U}_k \sum_{k' \neq k} \mathbf{H}_{kk'} \mathbf{v}_{k'} \mathbf{s}_{k'} + \mathbf{U}_k \mathbf{n}_k,$$

The IA consists in designing the matrices  $\mathbf{U}_k$  and  $\mathbf{V}_k$  so that they verify the following two conditions

$$\mathbf{U}_k \mathbf{H}_{kk'} \mathbf{V}_{k'} = \mathbf{0} \quad \forall k \neq k'$$

$$\text{rank}(\mathbf{U}_k \mathbf{H}_{kk} \mathbf{V}_k) = d_k$$

Existing IA algorithms for constant MIMO IC either rely on full channel state information (CSI) exchange between the transmit nodes [TJ10], [YS10], [TGR09] or proceed in a distributed approach based on local CSI knowledge only [YTJK08],[ PH09],[ PH11],[ KF10]. Since in our deployment scenario, there is no guarantee to have a high speed backhaul between APs allowing exchange of full CSI, we focus on distributed algorithms.

#### 6.4.2.2 Goal of the Proposed Algorithm

Previous works on design of distributed algorithms for IA sought either perfect alignment of interfering signals which is not always the best strategy for some system SNR regions or tried to reach a tradeoff between useful signal strengthening and IA depending on system signal to noise ratio (SNR) but used parameters that do not have practical significations. Such observations motivated us to design a distributed algorithm that overcomes the limits of the perfect IA algorithm from one side and from another side that uses design parameters that have practical significations compared to the other existing algorithms.

In this work, we propose a distributed algorithm for the constant MIMO IC. The algorithm computes the pre-coding and decorrelation matrices that aim at ensuring a target signal to interference plus noise ratio (SINR)  $\gamma_T$  for all the streams of all the links in the network. In other words, if we define  $\Gamma_k^l$  as the SINR of the  $l^{\text{th}}$  stream of link  $k$ , the goal of our work is to have  $\Gamma_k^l \geq \gamma_T \forall l, k$ .

We show that depending on the choice of the target SINR, the algorithm performs either IA at all the receivers or tries to reach a tradeoff between useful signal strengthening and IA.

#### 6.4.2.3 Simulation Results

Figure 6.7 represents the sum-capacity evolution of the proposed algorithm and the minimum leakage algorithm presented in [YTJK08]. For the proposed algorithm, different target SINRs were chosen to run the iterative process. Figure 6.7 shows that for low values of target SINRs, the proposed algorithm outperforms the minimum leakage algorithm in low SNR region. This can be explained by the fact that this algorithm sacrifices the alignment in favor of the strengthening of the useful signal. Such behavior can be confirmed from Figure 6.8 which represents the total chordal distance of the interference channel of each algorithm. In fact, the chordal distance is a metric used to measure the distance between vector spaces. When two subspaces are aligned, the chordal distance will be zero. Otherwise, it will be different from zero.

Figure 6.7 shows that at high SNR region, high target SINR values should be used to run the proposed algorithm. When doing so, the behavior of this algorithm will be similar to the minimum leakage algorithm seeing that in these SNR regions, the interference alignment is the most important task from sum-capacity point of view.

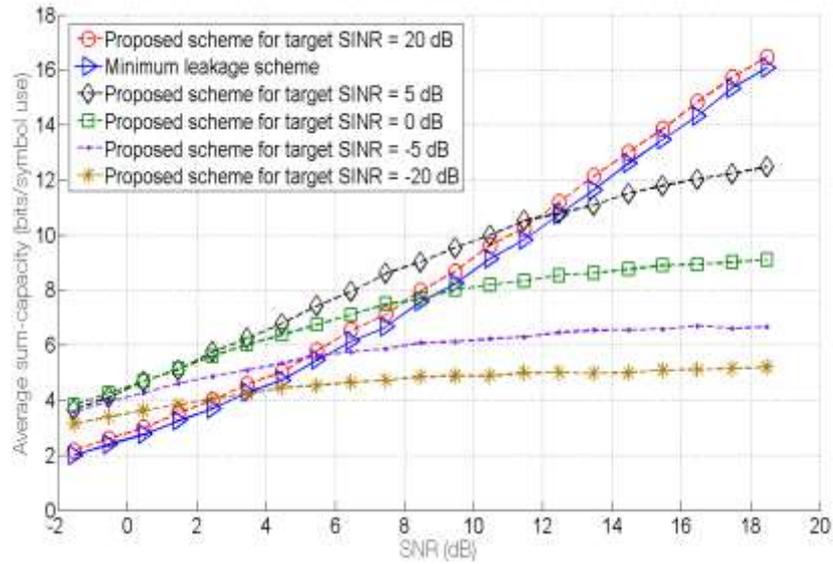


Figure 6.7: Average sum-capacity of the proposed algorithm compared to minimum leakage algorithm

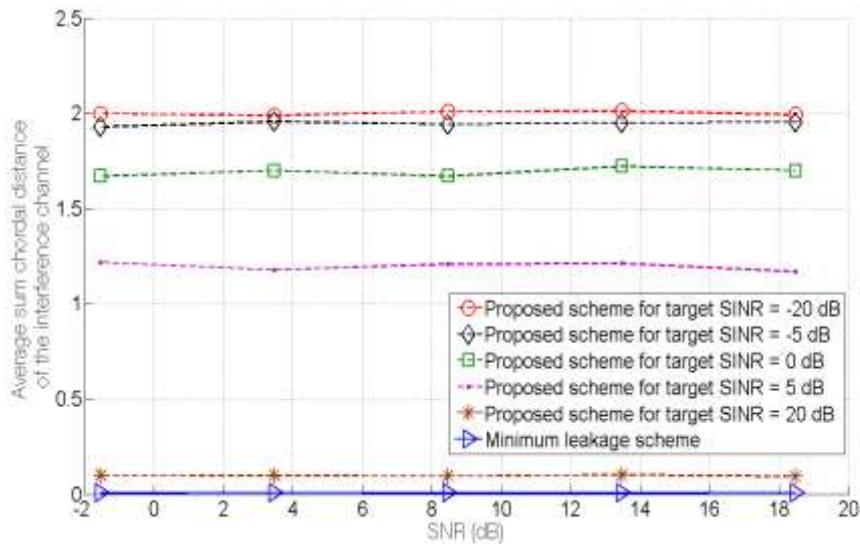


Figure 6.8: Average sum-chordal distance of the interference subspaces of the proposed algorithm compared to the minimum leakage algorithm

## 6.5 Conclusions

In this section, we have addressed the first steps towards heterogeneous networks deployments, mainly focusing on interference avoidance schemes that enable a benefit from the huge potential offloading gain taken from smaller cells. Indeed, a tradeoff exists between increasing the network capacity by multiplying the number of cells, and the interference level they generate. Several scenarios have been investigated targeting short or mid term deployments constrained by the LTE-A architecture.

First, we have considered the co-channel deployment of HeNBs as one of the most promising offloading capability, but also as a potential threat for the eNB network. Thus, we have proposed power control strategies for the protection of the downlink eNB network, and have shown that without any cooperation between the nodes, a good tradeoff between the cell edge throughput of the HeNB and eNB networks can be obtained, which can be applied for short term deployments without any modification of the architecture. Also, we have shown that a minimal amount of cooperation is needed for the sake of protection of the uplink of the eNB network, and have proposed architecture enablers for targeting short to mid-term deployments. Of course, higher degrees of cooperation will allow addressing specific interference situations, and we proposed a light coordinated beamforming and scheduling scheme between HeNBs and eNBs in order to target eNB-UEs lying inside a CSG-HeNB coverage area.

Second, we have targeted the deployment of small cooperating cells in a campus fashion, i.e., non-planned and relying on auto-organizing techniques. We have shown that, because of the wall attenuation between the nodes, and due to the non-full load hypothesis of the cells, most gain is provided by coordinated scheduling and that coordinated joint processing is not necessary. Also, we have investigated distributed scheduling techniques based on game theory, involving some traffic statistics exchange between the nodes. Also, when a centralized implementation is possible, we have shown how the transmit power of the nodes in the system can be optimized so as to maximize the Jain index of the users capacity. These innovations target mid term deployments.

Finally, we have investigated the use of interference alignment techniques for mitigating the interference between an eNB and a hotspot (or pico station).

## 7 Conclusions

This last document deliverable for the WP1 workpackage within the ARTIST4G project has focused on presenting performance assessments of the most promising techniques that were studied within WP1 related to interference avoidance. The results are based on evolved techniques that were identified and classified in deliverable D1.1 and investigated in deliverable D1.2 and D1.3, as well as novel alternative techniques that were introduced and assessed in this document.

Based on the insights from these performance results we provided a synthetic perspective over the most promising solutions for interference avoidance. We argued that these solutions can reach a satisfactory trade-off in terms of performance benefits vs complexity of implementation. Some of these techniques were also identified as complementary techniques towards an integrated interference avoidance concept.

The proposed solutions were grouped into three main sections, Section 4 “Advanced Beamforming and Multi-cell Coordination”, Section 5 “Advanced Joint Transmission Schemes for Multi cell cooperation” and Section 6 “Advanced interference avoidance schemes for small cells deployments”, for which the synthetic perspectives were summarized, with the details for modelling, mathematical derivations and simulation results presented in the Appendix chapters.

We started in Section 3 by outlining the three level optimization framework considering cooperative cellular networks in order to avoid interference. This framework is based on identified degrees of freedom, which we classified into environment, traffic and user property related parameters.

In Section 4, innovations which require either no or only a limited exchange of control plane information and no user plane exchange between eNBs were considered. We considered schemes, which involve different mechanisms and system parameters, and which have different levels of complexity and also different requirements on the architecture and deployment costs: “Advanced Beamforming” schemes that apply specific adaptation of the beam patterns, “Coordinated Beamforming” schemes, which optimize precoding of the horizontal antenna pattern, and “Coordinated scheduling”, which is shown to be able to reduce interference without antenna pattern adaptation. Finally, a distributed scheduling algorithm was combined with vertical beam pattern adaptation. This algorithm requires adjacent cell interference measurements and exchange of scheduling constraints among eNBs. The scheduling constraints exchange can be envisioned for multi-vendor solutions, as an extension in future releases of the 3GPP LTE of the signalling exchanged over the X2 interface for e-ICIC.

With these innovations of relatively low complexity compared to JP CoMP, already reasonable gains in specific application scenarios have been shown and could be partially verified with field measurements. Some of these innovations are transparent to 3GPP standards (even Release 8 and 9) or require only minor enhancements. Therefore they will be applicable at an early stage. Others will require extensions of the standard, especially those requiring adjacent cell interference measurements. Additionally it can be pointed out that the schemes relying on multi-cell coordinated beamforming require multiple antenna elements or directional antennas at each eNB. Thus, although the innovations can not provide the same performance gains as promised by the JP CoMP schemes addressed in section 5, they can be considered as entry schemes feasible with moderate network and system enhancements, even at a stage where the required advanced standard features and RAN capabilities needed for JP CoMP are not fully available.

In Section 5, more complex interference avoidance solutions were presented. These techniques rely on the assumption that user plane data can be shared by cooperating eNBs located within a cooperation cluster, thanks to a suitable backhaul architecture. Under this assumption, so-called Joint Precoding techniques across the cluster eNBs can be applied, which mimic the precoding methods used in multi-user MIMO systems and, in principle, intra-cluster interference can be

fully eliminated. Section 5 addresses some of the key challenges to such approaches such as the design of clusters via suitable user and cell grouping algorithms, the management of inter-cluster interference via power control and robust beamforming methods, and finally the design of latency-robust feedback schemes exploiting channel predictions.

The starting point in Section 5 was the maximization of the system level gains under ideal conditions, i.e. especially assuming ideal channel state information (CSI). Some first steps were taken to understand limitations and requirements regarding channel estimation as well as for CSI feedback. Without suitable enhancements for channel estimation, feedback, precoding, scheduler and inter-cluster interference mitigation a significant part of the system level gains under ideal conditions are easily lost, but for some well justified enhancements of these entities, such as channel prediction, robust precoding in combination with fine tuning of the CoMP scheduler and inter-cluster interference coordination by user centric antenna downtilt or fractional frequency reuse techniques, a significant part of the gains is achievable even under real world conditions. The resulting Interference Mitigation Framework – Advanced relies on 1) Inter-cluster interference control 2) User grouping 3) Robust precoding, 4) Mode-switching between different transmission modes 5) State of the art channel prediction to overcome CSI outdateding due to backhaul latencies and user mobility, and 6) an enhanced feedback scheme that should be combined by as yet unspecified compression schemes to obtain a reasonable uplink feedback overhead. A performance with a spectral efficiency of about 5 bit/s/Hz/cell – equivalently to about 70% gain over a 4x2 MU-MIMO system - seems to be possible. A spectral efficiency of around 6 bit/s/Hz/cell seems to be in reach by quite realistic means, although a price will have to be paid by using a significant uplink feedback overhead and downlink reference signal overhead.

In Section 6, interference avoidance was investigated for the specific case of small cells and heterogeneous networks. There, practical constraints related to ease of implementation, low complexity, and distributed optimization to avoid heavy exchange of information between macro and femto cells were emphasized. In order to satisfy such constraints, techniques were proposed making use of clever power control protocols, together with coordinated scheduling and beamforming. We investigated co-channel deployment of HeNBs as one of the most promising offloading solutions. We showed that without any cooperation between the nodes, a good tradeoff between the cell edge throughput of the HeNB and eNB networks can be obtained, which can be applied for short term deployments without any modification of the architecture. Also, we showed that a minimal amount of cooperation is needed for the sake of protection of the uplink of the eNB network. We also targeted the deployment of small cooperating cells in a campus fashion, i.e., non-planned and relying on auto-organizing techniques. We showed that, because of the wall attenuation between the nodes, and due to the non-full load hypothesis of the cells, most gain is provided by coordinated scheduling and that coordinated joint processing is not necessary. Also, we have investigated distributed scheduling techniques based on game theory, involving some traffic statistics exchange between the nodes. When a centralized implementation is possible, we showed how the transmit power of the nodes in the system can be optimized so as to maximize the Jain index of the users capacity. These innovations could target mid term deployments. Finally, we investigated the use of interference alignment techniques for mitigating the interference between an eNB and a hotspot (or pico station).

In the Appendix chapters we provide additional details for modelling, mathematical derivations and simulation results of all the considered schemes.

## Appendixes

This Appendix contains detailed results of the final investigations within ARTIST4G Work Package 1 (WP1). The presented results are categorized at the top level chapters according to the classes of innovations that were introduced and used in the previous WP1 deliverables [ARTD11], [ARTD12] and [ARTD13].

Some of the presented results are based on refined ideas and investigations with a basis in the deliverables [ARTD11], [ARTD12] and/or [ARTD13]. In those cases the reader is advised to follow such references to get the full understanding of the underlying algorithms and assumptions. For results based on studies that are introduced in this deliverable, a complete description is provided.

### A1. Single-cell MU-MIMO scheme

#### A1-1 Transmit and receive filter design with limited signalling information

##### System Model

In this contribution MU-MIMO downlink system is investigated. It is considered that  $M$  parallel data streams are aimed to be transmitted to  $K$  UEs, where each stream  $m$  is assigned to one UE  $k$ , if  $m \in M_k$ . The parallel data streams, included in vector  $\mathbf{c}$  are independently channel coded and modulated with a certain transmission rate  $R$ . The resulting data symbol vector  $\mathbf{d}$  is assumed to be Gaussian distributed and linearly pre-equalized with the precoding matrix  $\mathbf{B} = [\mathbf{B}_1, \dots, \mathbf{B}_K]$ , where  $\mathbf{B}_k$  is the partition of  $\mathbf{B}$  related to UE  $k$ . The precoded data symbols are mapped to  $N$  eNB antennas and transmitted over the frequency flat channel  $\mathbf{H} = [\mathbf{H}_1^H, \dots, \mathbf{H}_K^H]^H$ , which represents the spatial coupling between eNB and UE antennas. The channel is assumed to be static for the transmission of a data block with  $N_b$  symbols while subsequent channel blocks are uncorrelated. Each UE is equipped with  $N_k$  receive antennas, where  $M = \sum_k N_k$ . The signals received at UE  $k$  are corrupted by additive Gaussian noise  $\mathbf{n}_k$  and multiplied with the receive filter  $\mathbf{U}_k$ . After equalization the data signals are detected. Depending on the detection strategy the average achievable user rate is either

$$R_k = \mathbb{E}_{\mathbf{H}} \left\{ \sum_{l=1}^{U_k} \log_2 \left( 1 + \sigma_s^2[l, k] (\sigma_n^2 + \sigma_t^2[l, k])^{-1} \right) \right\} \quad (1)$$

for linear detection with

$$\sigma_s^2[l, k] = \left| [\mathbf{U}_k \mathbf{G}_k]_{l,l} \right|^2$$

and

$$\sigma_t^2[l, k] = \sum_{j=1, j \neq l}^{U_k} \left| [\mathbf{U}_k \mathbf{G}_{k,k}]_{l,j} \right|^2 + \sum_{j=1, j \neq k}^K \sum_{i=1}^{U_k} \left| [\mathbf{U}_k \mathbf{G}_{k,j}]_{l,i} \right|^2$$

For readability the precoded channel of user  $k$  is denoted as  $\mathbf{G}_k = \mathbf{H}_k \mathbf{B}$  which is separated into the matrices  $\mathbf{G}_{k,l} = \mathbf{H}_{k,l} \mathbf{B}$ . The useful part of this matrix is for  $l = k$  while interference is obtained with  $l \neq k$ . Considering rate optimal receivers the achievable rate can be upper bounded with

$$R_k = \mathbb{E}_{\mathbf{H}} \left\{ \log_2 \left( \det \left( \mathbf{I} + \Phi_S[k] \left( \sigma_n^2 \mathbf{I} + \Phi_I[k] \right)^{-1} \right) \right) \right\} \quad (2)$$

where  $\Phi_S[k] = \mathbf{G}_{k,k} \mathbf{G}_{k,k}^H$  and  $\Phi_I[k] = \sum_{l \neq k} \mathbf{G}_{k,l} \mathbf{G}_{k,l}^H$ . Note, that the latter rate expression inherently assume MMSE equalization. Both expressions are based on the assumption of perfect knowledge of the underlying precoded channel at the receiver.

The basic concept of the MU-MIMO downlink transmission is shown in Figure A.1.

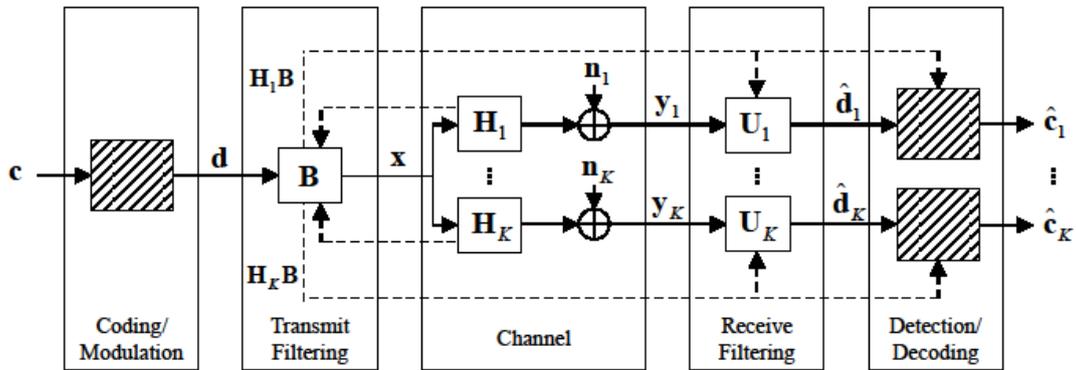


Figure A.1: Block diagram of the considered single cell MU-MIMO downlink system with linear transmit and receive filters.

### Linear Filtering Design

The spatial filter design is based on sum MSE minimization []. Transmit and receive filters are calculated in an alternating fashion as described in [SSV08]. In the first step of iteration the optimal receive filters result in

$$\mathbf{U}_k = \mathbf{G}_{k,k}^H \left( \mathbf{G}_k \mathbf{G}_k^H + \sigma_n^2 \mathbf{I} \right)^{-1}$$

with fixed transmit filters. In the second step the transmit filters are recalculated by fixing the receive filters

$$\mathbf{B} = \nu \mathbf{H}^H \mathbf{U}^H \left( \mathbf{U} \mathbf{H} \mathbf{H}^H \mathbf{U}^H + \frac{\sigma_n^2}{\rho} \mathbf{I} \right)^{-1}$$

with scaling variable  $\nu$  is chosen in order to achieve the sum power constrain  $\text{tr}\{\mathbf{B} \mathbf{B}^H\} \leq \rho$ . Convergence of the alternating filter calculation is shown in [], while the alternating approach does not necessarily achieve global optimality.

### Signaling strategies

In this work it is assumed that the BS has perfect knowledge of the instantaneous channel state, which can be motivated by the application of time division duplex and negligible channel estimation errors. However, at UE side the precoded channel is not known a-priori and

additional information need to be signaled to the users. The knowledge about  $\mathbf{G}_k$  at UE  $k$  is denoted with  $\hat{\mathbf{G}}_k$ .

Furthermore, it is assumed that user position and SNR are constant over a large number of transmission blocks and the second moments of  $\mathbf{G}$  are known at transmitter and receiver side. Additional signaling leads to further information about the precoded channel at the receiver. However, the knowledge of  $\mathbf{G}$  is still imperfect, while its accuracy depends on the amount of signaling.

The rate with imperfect channel knowledge at the receiver can be calculated by constituting the actual channel as random variable  $\mathbf{G}_k = \hat{\mathbf{G}}_k + \mathbf{E}_k$  with estimate  $\hat{\mathbf{G}}_k$  and an unknown error  $\mathbf{E}_k = [\mathbf{e}_{1,k}, \dots, \mathbf{e}_{U,k}]$ . The achievable rate reduces since channel uncertainty can be interpreted as additional interference, while the signal power decreases to the power of the known channel. In order to compare different signaling strategies the net rate

$$R_N = R \left( 1 - \frac{N_S}{N_B} \right)$$

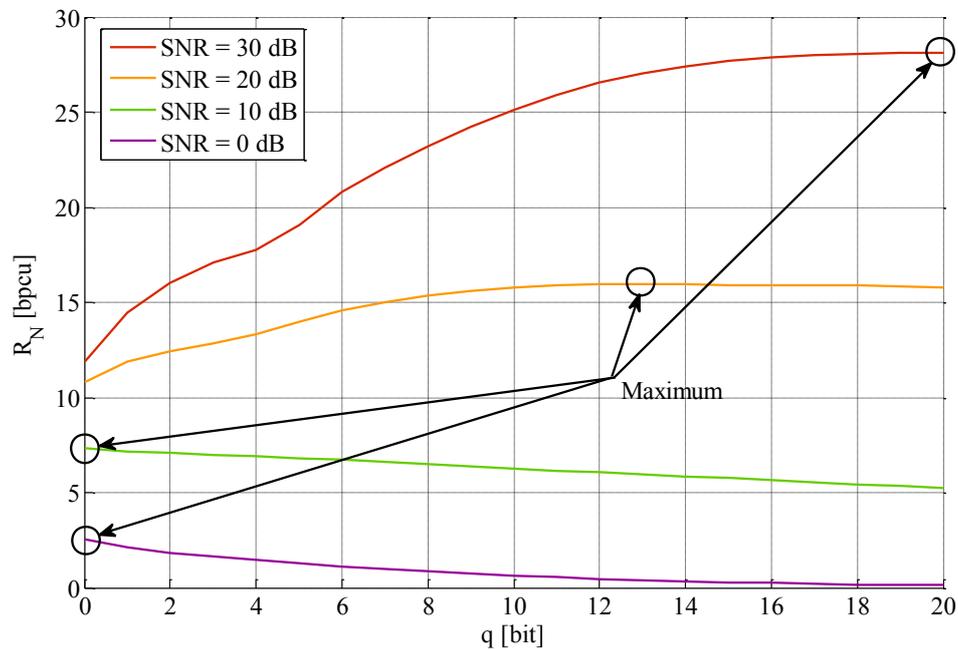
is introduced as a metric reflecting the achieved performance under consideration of the required signaling. In this context  $N_S$  out of  $N_B$  resource elements cannot be used for data transmission but for transmitting overhead information. Using digital feedback signaling information is assumed to be transmitted with zero-forcing (ZF). In this case the precoded channel is always an identity matrix and inherently known at the UE. Therefore, signaling information can always be decoded correctly without any additional knowledge.

#### Receive Filter Forwarding

Since the receive filters are inherently calculated at the eNB side an intuitive approach is the forwarding of the receive filters to the respective UEs. The filters are quantized with  $q$  bits applying vector quantization  $Q(\cdot, q)$ . The filter used at the UEs is  $\mathbf{U}_k = Q(\tilde{\mathbf{U}}_k, q)$ , where  $\tilde{\mathbf{U}}_k$  is the receive filter calculated at the eNB. However, only statistical knowledge of the precoded channel is available at the UE for detection. Since the signaled receive filters are optimized for full knowledge of  $\mathbf{G}$  the resulting performance decrease with the number of signaling bits, i.e. calculating the receive filters at the BS and forward them to the UEs is not meaningful at all.

#### Precoded Channel Forwarding

The previous strategy many suffers from the absence of knowledge about the precoded channel, which is needed for detection. In order to avoid this problem an alternative approach is the forwarding of the precoded channel. The main advantage of that method is that the precoded channel can be used for computation of the receive filter as well as for detection. The performance of the downlink transmission increases with the number of quantization bits. However, the net rate has a maximum at a certain amount of quantization. As can be seen in Figure A.2 signaling is reasonable only in the high SNR regime.



**Figure A.2: Net rate over the number of quantization bits used for signaling the precoded channel to the UEs**

### Precoded Pilots

A quite common strategy for signaling the precoded channel is to use pilot signals which are known at transmitter and receiver and filter them with the precoding matrix. The signals received at the UEs are corrupted by Gaussian noise. However, with increasing SNR more accurate information of the precoded channel is required since the sensitivity against channel uncertainty increases. Additionally, a small amount of signaling can decrease the performance compared to no signaling. This effect appears since the receive filters are calculated based on noisy information which is not known at the transmitter. Hence, outage can occur.

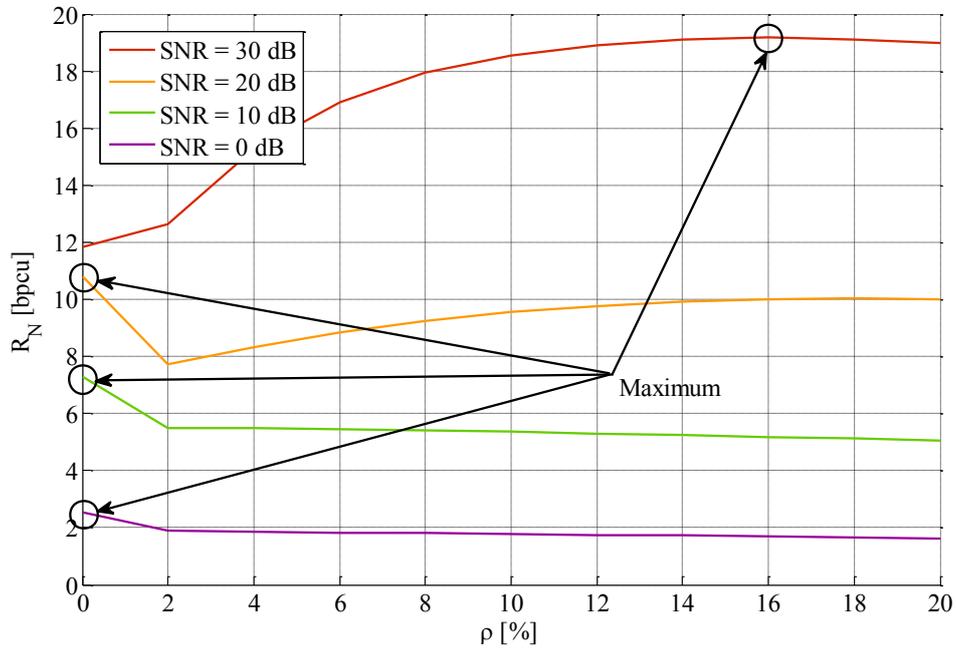


Figure A.3: Net rate over the pilot density for different SNR considering precoded pilots for signaling the precoded channel to the UEs.

### Conclusion

Applying always the optimal amount of signaling for each strategy leads to the figure in Section 4.2.1.1. However, this plot reflects a block size of  $N_B = 25$ . With increasing block size signaling becomes less relevant and the curves of precoded pilots and precoded channel forwarding converge to the upper bound where the precoded channel is perfectly known at the receivers.

## A2. Multi-cell MU-MIMO schemes

### A2-1 Robust linear precoding with per-base-station power constraints

[This Appendix contains algorithms used in Subsection 5.2.3.]

This contribution presents a specific precoding algorithm for multi-cell transmission. A comparison of several precoding schemes can be found in Section 5.2.3. Performance gains in cooperative cellular networks can be achieved by applying MU-MIMO techniques, such as linear precoding. In the downlink, transmission from multiple eNBs implies that the transmit power is restricted per eNB. In contrast single cell MU-MIMO downlink transmission restricts the overall transmit power. Optimization of the precoding filter with sum power constraint is typically easier to solve compared to per-base-station power constraints. However, optimization for precoding under per-base-station power constraints is addressed in [SSV+08].

Linear precoding filters are generally computed based on channel state information (CSI), where the precoder needs to be updated per channel state. In this contribution it is assumed, that CSI is obtained from the UEs and forwarded to a central unit (CU) where the precoding filter is calculated. The CU also obtains data of all jointly served UEs and can directly pre-equalize the data with the precoding filter. Afterwards, the precoded data is fed back to the respective eNBs.

In practical systems CSI is typically impaired by several sources of deterioration, like noisy channel observations, limited feedback rate in FDD systems, limited backhaul rate as well as outdated CSI, due to delays for processing and forwarding CSI from the UEs to the CU. CSI impairments with a focus on latency aspects are discussed in [FOF12].

However, if the CU is aware of the accuracy of the CSI, such information can be incorporated into the precoding filter design, which is commonly denoted as robust precoding [JKG+02]. Recent investigations came up with robust linear precoding solutions in the context of cooperative multi-cell downlink transmission [FF12]. However, the proposed multi-cell solutions are basically coming along with high computational complexity. An intuitive alternative is the application of simpler precoding schemes with sum power constraints with an appropriate scaling in order to ensure per-base-station power constraints. However, a scaled sum power constrained precoding filter exploits full transmit power only at one eNB, while the other eNBs exceed the maximum power level. Cooperatively served users located at the cell edge benefit from coherent precoding in order to avoid inter-user interference even if the transmit power is not exploited at all eNBs. With higher user separation, non-coherent transmission with full power at all eNBs can outperform coherent precoding.

In this work the problem of per base station power constraints in cooperative scenarios is analyzed in more detail. In this regard, two existing MSE based algorithms are compared. The scaled Wiener filter (SWF) is basically optimized according to a sum power constraint, which results in an analytical solution. However the resulting precoding matrix is scaled in order to fulfill the per-base-station power constraints. The generalized MMSE (GMMSE) precoder inherently considers per-base-station power constraints in the optimization. However, the performance advantage comes with significant complexity enlargements, since the solution is obtained by solving a second order cone program. Additionally, a third algorithm is proposed, which performs close the GMMSE but with almost the same complexity as SWF.

#### Precoding Algorithms

The presented algorithms are based on CSI  $\hat{\mathbf{H}}$  which is possibly imperfectly obtained at the CU. At the CU the precoding is performed considering that the transmit power at all  $M$  eNBs is restricted to  $\rho$ .

*Scaled Wiener Filter (SWF):*

By applying the SWF approach the precoding matrix results in:

$$\mathbf{B} = \nu [\tilde{\mathbf{B}}_1, \dots, \tilde{\mathbf{B}}_M] = \nu \hat{\mathbf{H}}^H \left( \hat{\mathbf{H}} \hat{\mathbf{H}}^H + N_U \frac{\sigma_n^2}{\rho} \mathbf{I} \right)^{-1},$$

where  $\sigma_n^2$  and  $N_U$  is the receiver noise and the overall number of receive antennas at all jointly served UEs, respectively. Additionally,  $\tilde{\mathbf{B}}_m$  is the part of the precoding matrix which is applied to eNB  $m$ . The scaling with  $\nu$  is chosen in order to fulfill the per-BS-power constraints:

$$\nu = \min_{\forall m} \rho / \text{tr} [\tilde{\mathbf{B}}_m \tilde{\mathbf{B}}_m^H].$$

With the max function the matrix  $\tilde{\mathbf{B}} = [\tilde{\mathbf{B}}_1, \dots, \tilde{\mathbf{B}}_M]$  is coherently scaled according to that part  $\tilde{\mathbf{B}}_m$  which possesses the largest transmit power  $\text{tr} [\tilde{\mathbf{B}}_m \tilde{\mathbf{B}}_m^H]$ . However, at the remaining eNBs the maximum transmit power is not necessarily achieved.

*Generalized MMSE Filtering (GMMSE):*

In contrast to the SWF approach this algorithm inherently minimizes the MSE by taking into account the per-base-station power constraints. The resulting precoding matrix  $\mathbf{B}$  can be obtained by solving a second order cone program. More details to the algorithm can be found in [ARTD12] as well as in [FF12]. The GMMSE algorithm always outperforms SWF, while the complexity is significantly larger.

*Optimal Scaled Wiener Filter (OSWF)*

Since it can be seen (e.g., in section 5.3.1) that in some situations it is preferable that each eNB transmits to its assigned UEs with full power instead of jointly precoding the user data under imperfect CSI. Such situations can be found when cooperatively served users are located close to their eNBs. Based on that observation, it seems beneficial to do SWF precoding but allowing that each part  $\tilde{\mathbf{B}}_m$  of the precoding matrix can be scaled individually. Based on that, the precoding matrix can be formulated as:

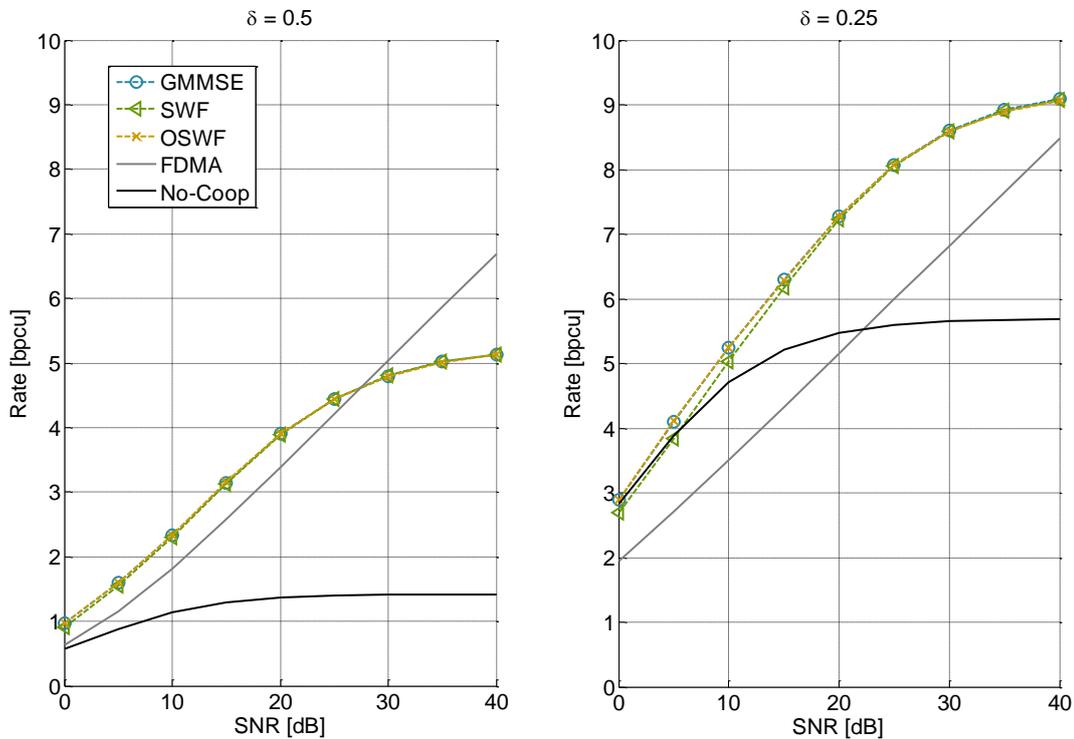
$$\mathbf{B} = \Lambda [\tilde{\mathbf{B}}_1, \dots, \tilde{\mathbf{B}}_M],$$

with diagonal matrix  $\Lambda = [\Lambda_1, \dots, \Lambda_M]$ ,  $\Lambda_m = \nu_m \mathbf{I}$  and the per eNB scaling  $\nu \leq \nu_m \leq \rho / \text{tr} [\tilde{\mathbf{B}}_m \tilde{\mathbf{B}}_m^H]$ . The scaling  $\nu_m$  can basically vary between the scaling of the SWF algorithm  $\nu$  and the value  $\rho / \text{tr} [\tilde{\mathbf{B}}_m \tilde{\mathbf{B}}_m^H]$ , which would achieve the maximum allowed transmit power at eNB  $m$ . This solution reflects a trade-off between consistent precoding and exploiting the maximum transmit power. In this work, scaling matrix  $\Lambda$  is optimized in order to maximize the sum user rate.

For all algorithms robust solutions do exist, which incorporate statistical knowledge of the CSI impairments (see [ARTD12]). This work is focused on the robust algorithms.

## Simulation Results

In this section simulation results are shown, where two eNBs jointly serve two UEs, which are located in between the two eNBs. The parameter  $\delta$  denotes the distance of the UE to its master eNB (the eNB the UE is assigned to) relative to the inter side distance.



**Figure A.4: Rate over the SNR for the different precoding approaches. The left plot refers to the scenario, where both UEs are at the cell edge, the right plot shows the performance where both UEs are located in their cell centers.**

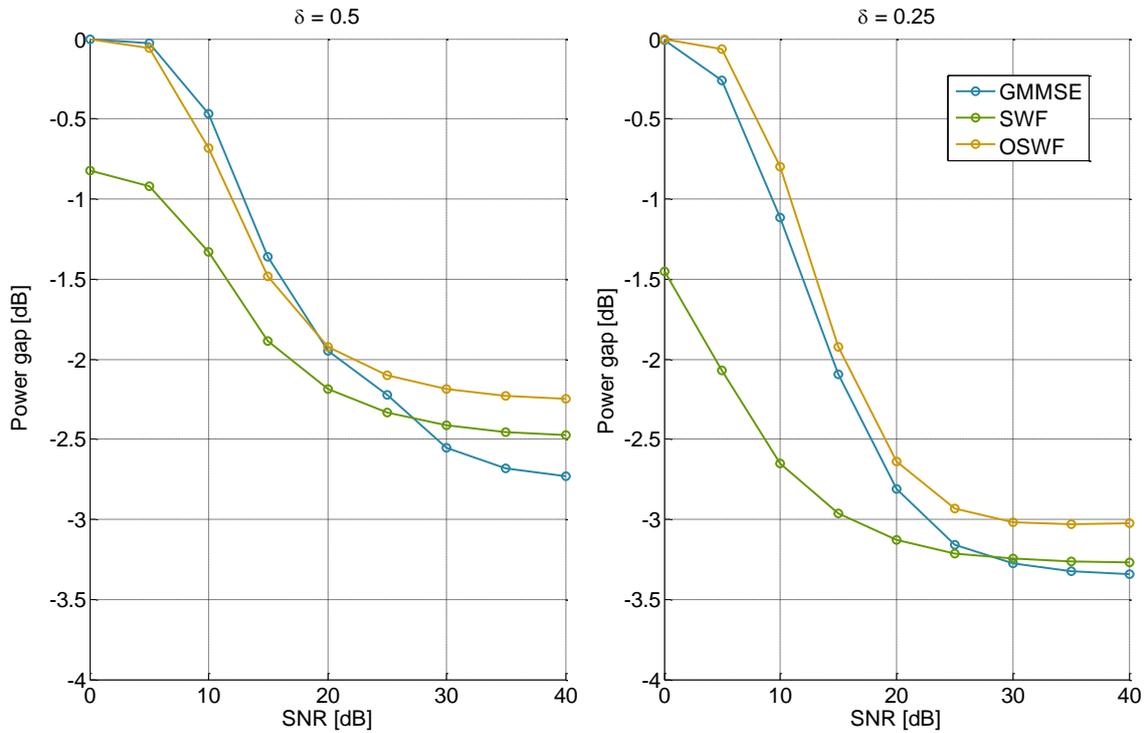
Figure A.4 shows the rate performance of different transmission schemes over the SNR a user would observe at the cell edge. The FDMA refers to the case where the available resources are equally shared between the users in order to avoid inter-cell interference. With No-Coop each eNB transmits to its assigned UE using all available resources without taking care of inter-cell interference. In other words, FDMA refers to reuse 2 (reuse 3 in a 3-D setup), while No-Coop refers to a reuse 1 setup. In contrast to FDMA, No-Coop saturates with increasing SNR, due to the interference limitation. With increasing user separation, inter-cell interference is reduced and No-Coop shows performance advantages compared to FDMA. However, at higher SNRs FDMA always outperforms No-Coop due to the interference limitation.

A similar behavior as for No-Coop can be observed when joint precoding is applied under limited rate for CSI feedback. In this case, the CSI accuracy does not increase with the SNR and inter-cell interference cannot be reduced appropriately by the applied precoding algorithm.

Comparing the different precoding approaches, it can be observed that the rate gap between SWF and GMMSE filter is small but increases with higher user separation, especially for SNRs smaller than 20 dB. This gap can be closed by the proposed OSWF algorithm, which individually scales the parts of the precoding matrix referring to an eNB. However, the OSWF algorithm possesses almost the same complexity than the SWF. In Figure A.5 the actual transmit power of the weaker eNB is plotted over the SNR. In this case the weaker eNB refers to that one which transmits with lowest power.

From Figure A.5 it can be observed that for SNR = 0 dB both eNBs transmit almost with the maximum power when GMMSE or OSWF is applied for precoding, while SWF shows a power gap of up to 1.5 dB. However, with increasing SNR the transmitted power for applying GMMSE and OSWF converges to that of SWF. Hence, performance advantages of GMMSE compared

to SWF can be mainly explained with the power gap of the weaker eNB. Since that gap can be compensated with OSWF, similar performance can be achieved but with significant lower complexity.



**Figure A.5: Shows the gap between the actual power and the maximum possible transmit power. The plots refer to the eNB with the smaller actual transmit power.**

## A2-2 An integrated design for downlink Joint Transmission CoMP

[This Appendix contains algorithms and results used in Subsection 5.2.3, and uses material from Appendix A4-2]

We here propose and evaluate a candidate “Joint Transmission (JT) CoMP baseline”, that is aligned to the general framework of Subsection 5.2: The scheme could be realistically implemented, it takes all important side conditions and constraints of the problem into account, while it strives for good (but not globally optimal) performance under computational constraints.

### A2.2.1 Outline

The core of the design is the robust linear precoder (RLP) that has been developed within the project and also has been presented in [ASA12]. It is here combined with a simple yet powerful user grouping. Its performance is in Section A2.2.6 below evaluated on channel sounding measurements. The performance is there compared to that of single cell transmission and also to Zero Forcing (ZF) joint transmission CoMP. We focus on downlinks of FDD systems, which represent the most challenging case for coherent joint transmission CoMP. The RLP scheme (without user grouping) is also evaluated and compared to other precoders in a simple reproducible simulation environment in Subsection 5.2.3.

We begin by listing the six main challenges for joint transmission downlink CoMP below, and outline the design choices that are made here to meet these challenges:

1. **Clustering.** Global joint optimization is unrealistic. But the use of CoMP only intra-site or within small clusters gives limited gains due to inter-cluster interference.

We here assume the **clustering scheme outlined in Subsection 5.2.1**, that gives good control of inter-cluster interference by semi-static frequency allocation over sets of overlapping clusters of relatively large size. The joint transmission is assumed to be designed and performed independently within these clusters of transmitters.

2. **Backhaul.** Constraints on backhaul bandwidth and delay may limit the application and performance of JT CoMP schemes, see Subsection 5.3.1. The design of precoders should therefore be able to take varying qualities or availabilities of backhaul links into account. Our solution can be extended to **enforce zeros in the precoding matrix** to represent forbidden backhaul links and it minimizes the resulting performance loss.

3. **Multi-channel prediction.** Due to transmission control loop delays, the precoders (beamformers) used in joint transmission should be based on predicted channels, see Subsection 5.3.2. Several downlink channels then have to be predicted for each UE. Furthermore, channels from different sites will have widely differing gains and SNRs.

We here have at least two problems:

- First, obtaining good enough channel predictions for at least pedestrian velocities and realistic prediction horizons.
- Furthermore, the simultaneous prediction of many downlink channels, without requiring excessive downlink reference signal overhead.

We here utilize the **Kalman-based methods** that are presented and evaluated in Appendix A4-2 below. They provide accuracy statistics of direct use for robust precoder design. As discussed in A4-2, the predictions can, and should, primarily be based on orthogonal reference signals. This enables the simultaneous prediction of channels with

very different powers. The assumed use of a clustering scheme that suppresses out-of-cluster interference at the reference signals improves the prediction accuracy.

4. **Feedback of channel estimates/predictions:** The precoding design for a cooperation cluster is here assumed to take place in a central unit (CU) placed in the fixed network. Channel quality prediction reports need to be transmitted to the CU, see Subsection 5.3.3 and Appendices A4-4 and A4-5. Two aspects of the here presented design are of importance in this respect:

- We take the statistics of compression errors that may be generated by these transmissions into account, when optimizing the joint linear precoders.
- When channel predictors are located at the UEs, a two-step strategy can be used for the channel prediction feedback. This strategy is described in the concluding Section A2.2.8 and is also outlined in Subsection 5.3.3. It radically reduces the required amount of feedback information needed for joint transmission downlink CoMP in FDD systems.

5. **Precoder performance and robustness:** The performance of zero-forcing-like linear precoders can be sensitive to CSI uncertainty.

The here utilized **Robust Linear Precoding (RLP) scheme** is designed based on Imperfect CSI due to prediction errors, feedback compression and other causes. It optimizes the performance averaged (marginalized) over all these uncertainties. For channel predictions, the optimal linear prediction as well as the required statistics of the prediction errors can be provided by the Kalman channel estimators.

6. **Complexity of scheduling and user grouping.** Most proposed solutions for the multi-user multi-cell downlink OFDMA optimization problem have infeasible complexity.

We here propose a solution, **RLP with cellular user grouping**, that immediately provides a reasonably good user grouping, and that utilizes the cellular scheduling that is already present in the network. This principle provides an avenue for introducing downlink JT CoMP as an add-on to LTE-Advanced systems. We may reuse present (possibly coordinated) cellular scheduling and beamforming algorithms and methods. The JT-CoMP design by RLP is then performed separately for each resource block, but jointly over the whole cluster.

We assume that a total of  $M$  single antenna UEs are to be served in a temporal subframe by  $N$  transmitters that belong to a fixed cooperation cluster. Such clusters may be designed as proposed in Subsection 5.2.1. See also [MZ11]. The transmitters may represent antennas, antenna ports, fixed beams or beams formed by single cell multi-user MIMO beamformers. In evaluations, we use measured channels from  $N=3$  single antenna base stations. Performance will mainly be evaluated based on the maximum sum-rate criterion under a per-antenna power constraint. The evaluations are based on channel sounding measurements from an urban environment that have previously been used in Subsection 4.4.1 of [ARTD12].

### A2.2.2 Channel model

Out of the users within a cluster, different subsets of users will in general be served within different frequency resource blocks of the downlink channel.

Let  $M_k \subseteq \{1, \dots, M\}$  be a group of  $M_k \leq N$  UEs. They are to be served jointly on the  $k$ th frequency resource block of an OFDM downlink by the  $N$  transmitters belonging to the cluster.

One or several channels to each of these UEs then need to be predicted, and the precoder design will be based on these predictions. As discussed in Section 5.3.2, predictors for downlink channels can either be located in the UEs or in the fixed network.

### Channels, predicted channels and prediction errors

If the predictors for the downlinks are located at UE's, then each UE in  $M_k$  predicts a complex-valued row  $h_i = [h_{i1}, \dots, h_{iN}]$  of the channel matrix  $H$ , where  $h_{ij}$  denotes the complex channel between transmitter  $j$  and UE  $i$  within a resource block of the OFDM downlink. Very low-powered links are typically not predicted but are represented by zero elements in  $h_i$  to reduce the prediction and reporting overhead. The predictions are transmitted on an uplink control channel to the strongest base station of each UE, denoted the *Master base station*.<sup>28</sup> This feedback is performed on the short term fading time scale and needs to have low latency. The prediction error statistics is also estimated and reported, but this can be performed on a much slower time-scale, related to that of the shadow fading. The corresponding feedback reporting overhead is therefore much less than that of the channel predictions themselves.

If predictors are located at the network side, then each UE sends (compressed) measurement reports over an uplink control channel. The calculation of channel predictions will be based on these reports.

The channel predictions and the prediction error statistics are forwarded to the coordinating unit (CU) of the cooperation cluster via control signalling [ARTD43].

The CU is thus provided with the (possibly compressed) predicted row vectors  $\hat{h}_i$  of the channel matrix and the statistics of their uncertainties  $E_{\Delta}[\Delta h_i \Delta h_i^*]$ , for  $i \in M_k$ . Here,  $(\cdot)^*$  is the Hermitian transpose. The average  $E_{\Delta}[\cdot]$  denotes an expectation over the CSI errors,  $\Delta h_i = h_i - \hat{h}_i$ . These covariance matrices represent the second order moments of the *total* CSI errors, including prediction errors, errors due to channel variability within resource blocks and feedback compression errors.

The  $M_k \times N$  channel matrix  $H$  will comprise of two parts (cf. Figure 5.9 in Subsection 5.2.3),

$$H = \hat{H} + \Delta H, \quad (\text{A2.2.1})$$

an estimate  $\hat{H}$  and an error matrix  $\Delta H$  for which the second order statistics has been made available at the CU. The estimate  $\hat{H}$  is assumed to be the mean of  $H$ , given all information available at the CU, so  $\Delta H$  has zero mean.

### Including pre-existing cellular beamformers into the design

In case of multi-antenna cellular base stations, the CoMP design will represent an add-on to any pre-existing cellular MU-MIMO transmit schemes. The channel matrix in (A2.2.1) then represents the channels including such beamformers. Thus, let  $H_0$  represent the physical channel from all transmit antennas to all receive antennas and let  $W$  be a block-diagonal  $N \times M_k$  matrix. The blocks of  $W$  contain the up to  $M_k$  beamformers that might be used by all the sites within the cluster. (In the concept description of Section 5.2.1, fixed sets of beamformers are assumed.) The channel matrix on which the CoMP design operates is then given by  $H = H_0 W$ .

<sup>28</sup> An alternative could be to use uplink CoMP for this feedback channel to improve its capacity and thus to reduce the control overhead over the uplink. However, the demand for a very low latency could make this variant infeasible.

### Linear precoding under per-transmitter power constraints

Let  $s$  be the column vector of symbols intended for the  $M_k$  UEs. Denote the column vector of signals from the  $N$  transmitters by  $u$ . The vector  $y$  of received signals at the  $M_k$  UE's, in an OFDM symbol and on a subcarrier within the resource block  $k$  of a subframe, will then be

$$y = Hu + n = HB_s + n = \frac{1}{c} H\tilde{B}s + n. \quad (\text{A2.2.2})$$

Above, the  $M_k \times N$  matrix  $B = (1/c)\tilde{B}$  represents the complex-valued linear precoding matrix which distributes the  $M_k$  symbols over the  $N$  transmitters. The column vector  $n$  is the sum of out-of-cluster interference and additional noise at the UEs. It is assumed to have known variance and zero mean and is henceforth denoted noise. We shall assume a per-transmitter power constraint, that the scalar scaling  $1/c$  is applied equally to all transmitters to ensure,

$$E |u_j|^2 \leq P_{\max,j}, \quad (\text{A2.2.3})$$

for  $j=1, \dots, N$ . Here  $u_j$  is the transmit signal at the  $j$ 'th transmitter and  $P_{\max,j}$  is the corresponding maximum transmit power, on average over the utilized symbol constellation. We here furthermore, for complexity reasons, assume the power constraint to be known and to be fixed for all time-frequency resource blocks.<sup>29</sup>

#### A2.2.3 Precoding algorithm: The Robust Linear Precoder (RLP)

The precoding scheme outlined below is the one denoted RLP-ACFF in Subsection 5.2.3. Our strategy for robust linear precoding designs transmit filters, leaving the detection and decoding to be adjusted by the receivers.<sup>30</sup>

#### The diagonal target system and scaling of the desired received signal magnitudes

We introduce a target signal vector (cf. Figure 5.9 in Subsection 5.2.3)

$$z = Ds = \frac{1}{c} \tilde{D}s. \quad (\text{A2.2.4})$$

The square target matrix  $D$  is diagonal in the here mainly considered case of single-antenna terminals. Each element of the column vector  $z$  is then a scaled version of the corresponding element in the symbol vector  $s$ .<sup>31</sup> A diagonal structure of  $D$  represents the ideal of full intracluster interference suppression.

<sup>29</sup> Redistribution of the allowed transmit powers between resource blocks, as described in Appendix A2-4 below can be added, but such an additional water-filling optimization is not assumed to be used in the evaluations below. It would significantly increase the complexity of the total scheme and lead to rather small performance gain: See results for cases with one receive antenna per UE in Appendix A2-4 below.

<sup>30</sup> This is based on the assumption that it is problematic to let a coordinating unit within the fixed network design also the receiver filters, based only on imperfect channel predictions obtained from outdated downlink measurements. Since there will be demodulation reference signals (precoded pilots) embedded in the downlink transmissions, the terminals will be able to adjust receiver filters much more accurately based on these current signals and they can also improve the channel and message estimate by iterative channel estimation and decoding of the payload data. See also Subsection 4.2.1.1.

<sup>31</sup> The design is readily generalized to a receiver with multiple antennas and MIMO links. For a MIMO link, the cellular transmit beamformer is assumed known, and it is already part of the matrix factor  $W$  introduced above. The corresponding desired channels to the receiver antennas are likewise assumed known, and are included in  $D$ , which then becomes block-diagonal.

Note that the UEs will likely have differing distances to various transmitters and thus to have different attainable channel gains after CoMP precoding under transmit power constraints. In the problem formulation, the diagonal elements  $d_{ii}$  should reflect this by being set to realistic values.

In particular, we should not ask for the power of the received signal to be equal to that of the transmitted signal ( $D=I$ ), nor should we ask that the power of the signals at all UEs should be equal ( $D=rI$ , where  $r$  is a constant) unless all channels have equal attenuation.

In the design of zero forcing linear precoders, convex optimization of transmit powers is sometimes included. This corresponds to an optimizing adjustment of the diagonal elements of  $D$  in our setting above.

In the evaluations below, we propose a much simpler method for selecting the target received powers. We just set the diagonal elements of the unscaled target matrix  $\tilde{D}$  in (A2.2.4) to the amplitudes of the channels between each UE and their (predicted) strongest transmitter:

$$\tilde{D} = \text{diag}(\max_j |\hat{h}_{ij}|), \quad (\text{A2.2.5})$$

for diagonal elements  $i=1, \dots, M_k$ . The scalar  $c$  is then adjusted after the precoder design to fulfill the power constraint (A2.2.3.).

Another reasonable alternative choice of  $\tilde{D}$  is the one used in [PBG11], in which the UEs are allocated equal fractions of the total sum of transmit powers.

Our evaluations have shown that the use of (A2.2.5) instead of  $\tilde{D} = I$  results in a significantly improved sum-rate, when designing precoding matrices using RLP and ZF. The attainable sum-rate is close to that obtained (with *much* higher computational effort) by optimizing all diagonal elements of  $D$  with respect to sum-rate. For a ZF precoding design, the choice of [PBG11] is marginally superior to the choice in (A2.2.5) and the opposite is true for a RLP design when the errors in  $\Delta H$  are large.

### The basic Robust Linear Precoder criterion and design

The RLP is designed to minimize a scalar quadratic criterion. It is an averaged weighted sum of the powers of the transmit signals  $u$ , and the powers of the difference  $\bar{\varepsilon} = Hu - z$  between the received signal excluding noise, and the target signal vector:

$$J = E_{\Delta} (E \|V\bar{\varepsilon}\|^2) + E \|Su\|^2. \quad (\text{A2.2.6})$$

Here,  $V$  of dimension  $M_k \times M_k$  and  $S$  of dimension  $N \times N$  are penalty matrices while  $E[\cdot]$  denotes averaging with respect to the statistics of the message symbols. It can be shown ([Brä11],[ASA12]) that for any positive definite  $S$  and for a channel error matrix  $\Delta H$  in (A2.2.1) with zero mean and known second order statistics, the criterion  $J$  is uniquely minimized by the linear precoder matrix

$$\tilde{B} = (\hat{H}^* V^* V \hat{H} + S^* S + E_{\Delta} [\Delta H^* V^* V \Delta H])^{-1} \hat{H}^* V^* V \tilde{D}. \quad (\text{A2.2.7})$$

For square invertible and perfectly known systems ( $M_k = N$ ,  $\Delta H=0$ ,  $H$  invertible), with equal user weighting and no transmit power penalty ( $V=I$ ,  $S=0$ ), the RLP precoder by (A2.2.7) reduces to a zero-forcing precoder  $\tilde{B} = H^{-1} \tilde{D}$ , or  $B = H^{-1} D$ .

The resulting precoder matrix is thus explicitly designed to take the second order statistics of the channel uncertainty into account by averaging the criterion over the model uncertainty. Since the channel model and the precoder is linear and the criterion is quadratic, only second order

moments of the pdfs of model uncertainties affect the design. This design constitutes a special case of a general methodology for the averaged robust design of linear-quadratic optimal MIMO feedforward regulators, filters, predictors, smoothers and equalizers, which was developed in [SA93] and [ÖAS95].

The term  $E_{\Delta}[\Delta H^* V^* V \Delta H]$  in (A2.2.7) should represent the second order moments of all relevant types of uncertainties. If Kalman predictors are used, as assumed here, a covariance matrix  $(CP(t + \mathcal{G} | t)C^*)_i$ , see (A4.2.5), is provided for the significant downlink channels to each UE  $i=1, \dots, M_k$ . From these, the total covariance matrix  $E_{\Delta}[\Delta H^* V^* V \Delta H]$  can be calculated as outlined in [ASA12]. Variance terms are also added for channels that are not predicted.

After obtaining the precoder matrix  $\tilde{B}$ , the scalar scaling factor  $c$  is adjusted so that all transmitters fulfil the power constraints. This scales the criterion  $J$ , but does not affect the minimizing linear precoder  $\tilde{B}$  for (A2.2.6). Then,  $B = (1/c)\tilde{B}$  is used as joint linear precoder.

The resulting received signals can by (A2.2.1), (A2.2.2) and (A2.2.4) be expressed as

$$y = (\hat{H} + \Delta H)Bs + n = Ds + (\hat{H}B - D)s + \Delta HBs + n = \frac{1}{c}\tilde{D}s + \frac{1}{c}(\hat{H}\tilde{B} - \tilde{D})s + \Delta H\frac{1}{c}\tilde{B}s + n. \quad (\text{A2.2.8})$$

In the last expression, the first term is the target (A2.2.4). The second term is the error  $\bar{\varepsilon} = Hu - z$  that would be obtained if the true channel coincided with the predicted channel. The third term describes the influence of the model error. The RLP is designed to find an appropriate balance between the second and the third error-related right-hand terms in (A2.2.8).

### The power normalization loss problem

Note that the received signal described by the first three right-hand terms of (A2.2.8) is affected by the scaling  $(1/c)$  that is included in  $B$  and  $D$ , while the noise and out-of-cluster interference  $n$  remains unaffected. A scaling parameter  $c$  with large magnitude may sometimes be required. As a consequence of the rescaling, the transmit power from the nearest (strongest) base stations is then reduced, while transmitters far from all users are allowed to transmit at full power, little of which reaches the users. The received signal in (A2.2.8) can then have a low SNR and the UEs will therefore have low throughput. This phenomenon is known as “power normalization loss”. As will become evident below, it is a severe problem for zero-forcing precoders.

A good choice of  $\tilde{D}$  in (A2.2.4) is a first step to reduce the problem, but it is not sufficient. Furthermore, the model error term in (A2.2.7) will regularize the inverse and leads to much reduced power normalization losses in RLP designs, but also this will not completely solve the problem.

### RLP used as a tool to optimize more general criteria

To further control the power normalization loss, we observe that the first part of the criterion  $J$  may be too narrowly defined. It takes only the intra-cluster interference into account. The noise  $n$  is ignored. Generally, a good precoder design should not only aim to minimize this interference power as in (A2.2.6), but rather a more general objective function that may be expressed by

$$f(E_{\Delta}E(P_{S,i}), E_{\Delta}E(P_{I,i}), E_{\Delta}E(P_{N,i}), i = 1, \dots, M_k). \quad (\text{A2.2.9})$$

Here,  $P_{S,i}$ ,  $P_{I,i}$  and  $P_{N,i}$  are the signal, the intra-cluster interference and the noise plus out-of-cluster interference powers at the  $i$ 'th UE, respectively.

With power constraints, the noise affects the objective via nonlinear relations like (A2.2.3), so closed-form solutions are mostly hard to obtain. Global optimization of the precoder matrix under transmit power constraints could e.g. be performed with the Particle Swarm optimization strategy described in Appendix A2-5. Interestingly, our original criterion  $J$  by (A2.2.6) can also be used as a low-complexity tool for the iterative optimization of (A2.2.9). The penalty matrices  $S$  and  $V$  are then used as design parameters to balance the signal powers, interference powers and noise powers of (A2.2.8) with respect to (A2.2.9).

As an optimization criterion we here maximize the estimated sum-rate within resource blocks

$$\sum_{i=1}^{M_k} \log_2 \left( 1 + \frac{E_{\Delta} [E[P_{S,i}]]}{E_{\Delta} [E[P_{I,i}]] + E_{\Delta} [E[P_{N,i}]]} \right). \quad (\text{A2.2.10})$$

In [ASA12] we show that a *one-dimensional search* with respect to a single diagonal element of the penalty matrix  $S$  in the auxiliary criterion  $J$  improves the sum-rate significantly. Furthermore, for  $M_k = N = 3$  we have observed that a higher dimensional search with respect to several diagonal elements of  $S$  would not cause further noticeable improvements of the cdf of the sum-rate. To limit the complexity, we therefore here focus on one-dimensional search strategies. At each iteration, the averaged powers in (A2.2.9) are calculated as described in [ASA12].

#### Control of the power normalization loss by tuning transmit power penalties using a one-dimensional search.

One strategy, suggested in [ASA12] and used here, is to first calculate the precoder in (A2.2.2) by (A2.2.7) with  $V=I$  and  $S=0$ . Then, we penalize the transmitter that has thereby been given the highest power. If this transmitter is denoted  $n_{max}$  we set

$$S = \text{diag}(\rho_{jj}), \quad \text{where} \quad \begin{cases} \rho_{jj} = \rho & j = n_{max} \\ 0 & \text{else} \end{cases}. \quad (\text{A2.2.11})$$

The optimal balance with respect to (A2.2.9) is searched for in the restricted region  $0 \leq \rho \leq \rho_{max}$  where  $\rho_{max}$  is the lowest real-valued  $\rho$  that will cause a different transmitter to be given the highest transmit power. This solution reduces the difference in the transmit powers from the transmitters and thereby also reduces the required scaling constant  $c$ . Hence the SNR is increased at the cost of a lower intra-cluster SIR, to improve the sum rate (A2.2.10).

#### A2.2.4 Alternatives: Zero Forcing (ZF) and Single Cell transmission

We compare our results with the **zero forcing (ZF)** solution, where

$$B = \hat{H}(\hat{H}\hat{H}^*)^{-1}D. \quad (\text{A2.2.12})$$

The diagonal target scaling matrix  $D = (1/c)\tilde{D}$  and the scalar  $c$  are selected as in (A2.2.5).<sup>32</sup>

Moreover, we also compare our result to that of **single cell transmission**, where each UE is served by its Master base station using full power. Intracell interference is then suppressed: In case of single-antenna eNB's, UEs that have the same Master base station are allocated to orthogonal resources, see Figure A.6. For multi-antenna base stations, beamforming multi-user MIMO transmission can be applied per cell. In all cases, *inter-cell* interference within the cooperation cluster remains uncontrolled.

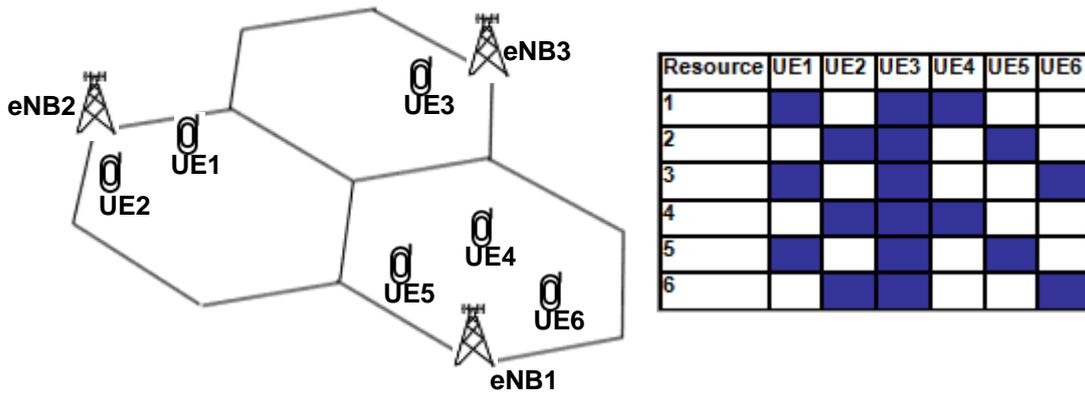


Figure A.6: An example of an UE setup in a scenario with 3 single-antenna base stations (here corresponding to eNBs) at different sites. The UEs will then be served on the resources indicated by blue fields in the right-hand figure. UEs with the *same* Master eNB are served on *different* time-frequency resources in single-cell transmission. The same principle is used by JT CoMP with cellular user grouping.

### A2.2.5 User grouping and scheduling: RLP with cellular user grouping

Within a cooperation cluster, the transmission to a potentially large number of users needs to be organized in the OFDM downlink with respect to time, frequency allocation and spatial resources (beams). Out of  $M$  users,  $M_k \leq N$  users are selected to be served jointly in a time-frequency physical resource block  $k$ . This selection is denoted *user grouping*. The  $M$  users are furthermore to be allocated to the  $K$  available physical resource blocks within a temporal subframe. This step is denoted (frequency) *scheduling*. Our aim will be to propose a strategy for joint scheduling and user grouping that is computationally simple, that forms good CoMP groups of users, that preserves single-cell multiuser scheduling gains and that works well also in highly loaded CoMP scenarios, where  $M_k = N$ , or close to  $N$ .

As a baseline comparison case, we use full load and random user grouping. The user groups are then randomly chosen. In our evaluation,  $M_k=N=3$  users are allocated in each frequency resource block.

This strategy will frequently generate user groups where at least one site has a weak channel to all UEs, so that e.g. the remaining  $< N$  transmitters try simultaneously to serve  $N$  users. Coherent joint transmission will in such cases result either in a poor solution for all UEs or a solution where one or more UEs are not served.

<sup>32</sup> For invertible channel matrices, the zero forcing precoder is  $B = \hat{H}^{-1}D$ . When  $M_k < N$ , the interference cancellation problem is underdetermined, i.e. it has infinitely many solutions, out of which (A2.2.12) picks a pseudo-inverse. The RLP solution (A2.2.7) does in such cases rely on the second and third terms of its inverse,  $S^*S + E_\Delta[\Delta H^*V^*V\Delta H]$ , to assure a full rank and thus a unique solution.

## Use of cellular scheduling, leading to cellular user grouping

To avoid such problems, we introduce a UE grouping scheme based on cellular scheduling: Frequency allocation and allocation of spatial layers is first performed separately per cell, just as in single-cell transmission. UEs with the *same* Master base stations (cells) are thus served on *different* (orthogonal or close to orthogonal) resources. However, in contrast to single cell transmission, all base stations may potentially jointly serve all UEs within the cluster that are allocated to a physical resource block. For example, in Figure A.6, UE1, UE3 and UE4 will be served jointly by coherent JT on resource 1, UE2, UE3 and UE5 will be served jointly by coherent JT on resource 2 and so on. The precoders for joint transmission are designed separately per resource block, for the whole cluster of  $N$  transmitters.

This “*cellular*” *user grouping* scheme is introduced also within the overall JT CoMP framework in Subsection 5.2.2. It provides several advantages.

The different users that are allocated to a resource block will then mostly have *differing* (instantaneously) strongest base stations. The resulting channel matrices therefore tend to be well-conditioned, with reduced spread of singular values as compared to random user grouping. The resulting joint precoder design problem for these user groups becomes easier to solve. Therefore, a lower power normalization loss is generated when the precoder is scaled to fulfil per-antenna power constraints.

We can perform the frequency allocations for users *per cell*, as in present systems. Design of the joint processing is thereafter performed per resource block for the whole cluster, as described above. No additional optimization or calculation is needed, to obtain a joint processing multi-user multi-cell downlink OFDMA and SDMA solution.

We may thus use already existing LTE cellular scheduling and beamforming algorithms. Alternatively, novel coordinated scheduling and (vertical) beamforming mechanisms for interference avoidance (see e.g. Section 4) could be used, and thus be supported by the here proposed joint transmission scheme to further enhance performance.

The multiuser diversity gains that are generated by today’s scheduling algorithms are not reduced by the CoMP user grouping, since the scheduling may remain unaltered. They are instead enhanced, by providing the scheduled users with an improved SINR. This also provides better conditions for the use of intra-cell MIMO transmission.<sup>33</sup>

The scheme improves upon cellular transmission when possible, by suppressing intra-cluster interference. It falls back to cellular transmission automatically in cases when all other channels except those from the Master base station are either too weak or too uncertain to be of use. Thus, when properly designed, the scheme should in no situation be worse than cellular transmission, with respect to the criterion we chose to optimize.

The proposed user grouping scheme will not improve performance in all situations. In the special case when the system is fully loaded and the UEs are unevenly distributed (e.g. 70% of all UEs in 3 cells have the same Master base station), the user grouping scheme could lower system performance since the denser UEs are given a lower individual bandwidth than they would otherwise have obtained by CoMP without this user grouping. This indicates that the basic suggested scheme can be improved further.

## Scheduling

Below, we will also investigate how two simple cellular scheduling algorithms applied at single-antenna base stations affect the performance. First, we use a *Round Robin* (RR) frequency allocation, in which each of  $m$  UEs that have the same Master base station are allocated  $1/m$  of the total bandwidth as in Figure A.6. This *regular frequency pattern* does not take the frequency

---

<sup>33</sup> The link adaptation parameters used by cellular schedulers and adaptive modulation and coding should be set aggressively, to harness the SINR improvement obtained by the added joint CoMP design.

selectivity of each UEs channel into account. Second, we use the *Score-based* (SB) opportunistic scheduler suggested by [Bon04]. This scheduler uses the variations of the channels over frequency. It is related to Proportional Fair scheduling, but is less sensitive to variations of the distributions of the SINR statistics of different users.

For joint processing with random user groups we apply RR scheduling in the sense that  $\text{Int}[M/M_k]$  user groups are first created, and held fixed. All users in a group are then given the same fraction of frequency resources (equally spread over the bandwidth), with no consideration taken to their channels in these resources.

## A2.2.6 Measurement-based performance evaluation

### Measurements and simulation assumptions

The channel matrices in (A2.2.1) are given by channel sounding measurements from an urban environment.<sup>34</sup> The measurements over a 20 MHz bandwidth have been collected by a vehicle. Noisy signals received simultaneously from the  $N=3$  transmitters at the UE positions are generated based on these measured channels, see Appendix A4-2 for details.

In our evaluation we use the upsampled channels and their corresponding Kalman predictions that are explained and evaluated in Appendix A4-2. The transmission latency and corresponding prediction horizon that will mostly be used is 10 ms<sup>35</sup>. The Reference Signals (RS) from the three different transmitters are placed on orthogonal time-frequency resources.

To simulate sets of UEs, 9 positions along the measurement route are randomly choose as the position of  $M=9$  single antenna UEs. For the simulations we assume that the unknown channel term  $\Delta H$  in (A2.2.1) is entirely caused by the Kalman prediction errors.

Each position, or UE, provides one row of the channel matrix  $H$  and one row of the predicted channel matrix  $\hat{H}$  in (A2.2.1). Each UE  $i$  also provides a covariance matrix  $(CP(t+g|t)C^*)_i$  by (A4.2.5), from which the total covariance matrix  $E_{\Delta}[\Delta H^* V \Delta H]$  in (A2.2.7) can be calculated, see [ASA12] for details. As RLP design, we use the iterative optimization of the sum rate (A2.2.10), by tuning a scalar transmit power penalty element (A2.2.11).

As the UEs then move at a maximum speed of 5 km/h along the measurement route for 128 ms, the different schemes are evaluated in terms of Shannon sum-rate, based on the true (unknown) channel gains. We model the noise of (A2.2.2) as i.i.d complex Gaussian variables with zero mean and covariance matrix  $E[n^* n] = \sigma^2 I$  where  $\sigma^2$  are the noise floor levels used for the Reference Signals (RSs) in the predictions of Appendix A4-2. (Please see the noise floors of the first figure of Appendix A4-2.)

On average over these three noise levels, the distribution of the signal-to-out-of-cluster interference ratio is close to that generated by other project partners by system-level simulations in a clustered 3-site environment with interference floor shaping as illustrated by Figure 5.6. Comparisons of performances that are averaged over our three noise floors should therefore indicate the relative performances of different schemes in an interference environment generated by the suggested clustering scheme.

The Shannon rate and the SIR are calculated for each scheme and are averaged over the 128 ms intervals. The evaluation includes a total of 1000 sets of 9 UEs. Scheduling and user grouping is performed over 144 subcarriers à 15 kHz, separated by 45 kHz. For a given precoder and (true) channel, the Shannon rate is calculated separately for all utilized subcarriers.

<sup>34</sup> The data set used by Uppsala University in these investigations was collected by Ericsson Research in December 2008 in the Stockholm suburb Kista [MSK+09], [LMF10]. See also Section 4.4.1 in [ARTD12].

<sup>35</sup> At the Referens Signal (RS) pattern RS18 explained in Appendix A4-2.

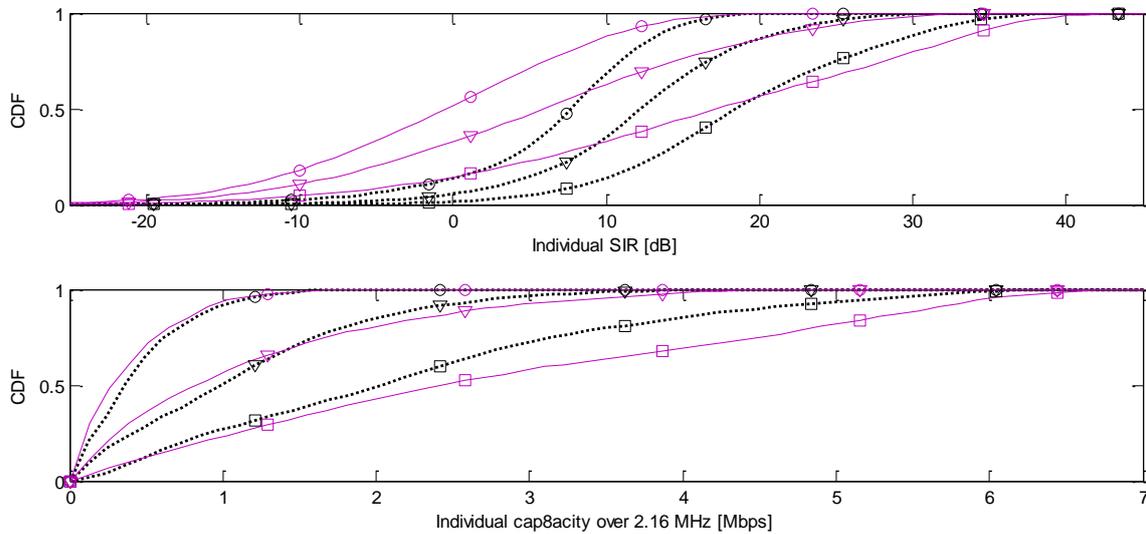
### Results: The impact of robustification in JT CoMP

The average sum-Shannon rates for the different UE grouping schemes and precoding schemes, as well as the 5 % percentiles are presented in Table A.1. Random UE grouping is applied in the first line of the table. In that setting, RLP significantly outperforms ZF, with the largest gains obtained in the 5% percentile<sup>36</sup>, where RLP has a 200 % gain compared to ZF. To find out why this is so, we can study Figure A.7 which shows the CDFs of the SIR and the capacity for UEs with different channel quality. We see that the ZF solution attempts to provide a more “fair” solution by improving the SIR for the UEs with poor channel qualities, compare the dotted to the solid lines. However, these users are often noise limited. Suppressing their interference does not yield any gains: The SIR improves for the worst users, but the SINR is not significantly improved. The RLP strategy instead improves the SINR and thus the rate more for users with medium to high quality channels, which in turn improves the sum-rate. This is seen from the rate CDFs.

**Table A.1: Average and 5 % percentile of the *sum*-Shannon rates for precoding schemes with random and cellular UE grouping with 9 UEs. Use of 10 ms prediction, 5km/h users at 2.66 GHz. Values are given in bps/Hz/cell.**

	RLP, $S=S_p$		ZF		Single cell	
	Ave	5% perc.	Ave	5% perc.	Ave	5% perc.
Random grouping with RR	5.5	2.4	4.7	0.79	-	-
Cellular grouping with RR	7.7	4.0	7.6	3.5	4.9	2.3
Cellular grouping with SB	8.5	5.1	8.5	4.8	5.5	3.5

<sup>36</sup> Note that this is the 5 % percentile of the *sum-rate*. For fairness we must also consider the 5 % percentile of the individual rates.



**Figure A.7: CDFs for SIR (top) and rates for UEs for ZF (black dotted lines) and RLP (solid purple lines) when random user grouping is applied. UEs are sorted into groups of the SNR they would have if all eNBs transmitted non-coherently to that UE only with intervals of 0-10 dB (circles), 10-20 dB (triangles), and 20-30 dB (squares).**

### Results: The impact of user grouping and scheduling

With random user grouping, JT CoMP using the RLP scheme and RR scheduling does not achieve a large gain compared to single-cell transmission with RR. From Table A.1, we clearly see that the switch to a good user grouping strategy (cellular user grouping) changes the situation completely. It generates a large gain for both of the JT precoding algorithms. In fact, it is the single most important aspect for achieving JT CoMP gains, scheduling and robustification only come second and third respectively. This is an important aspect to take into account to obtain fair comparisons between JT CoMP and single cell transmission in system level evaluations.

A striking result is that when cellular UE grouping is used instead of random user grouping, the average gain obtained by RLP as compared to ZF is non significant. When cellular UE grouping is applied, the channel matrix for the users to be served jointly in general becomes diagonal dominant, creating a much easier channel inversion problem for ZF. However, there is still a gain for RLP with respect to ZF in the 5% percentiles of 14 % and 6.3 % for RR and SB scheduling respectively on average.<sup>37</sup> We therefore recommend using a robust scheme also with good UE grouping, at least at higher noise floors.

One of the main advantages of cellular user grouping is that it does not destroy the scheduling gains that can be achieved from multi user diversity. In our investigation, the gains when going from RR to SB scheduling is highest in the 5% percentiles, where we have a 28 % gain for the RLP scheme.

The gains due to user grouping and scheduling are naturally depended of user diversity. A higher number of users allows for a higher gain, which can be seen by comparing the values of Table A.1 with those of Table A.2 in which only three users were considered in the system. We see that the first line of both table are similar (some small variations due to a slightly different user setup occur). However the average “user grouping gain” for RLP (comparing random user

<sup>37</sup> This gain is highest at the highest noise floor.

grouping with RR to cellular user grouping with RR) decreases from 40% to 8.8% as the user number decreases to three. The SB scheduling gain (when cellular user grouping is applied) decreases from 10 % for nine users to 1.6 % for for three users.

**Table A.2: Sum-rates for different precoding schemes with random and cellular UE grouping with 3 UEs within the cooperation cluster. Prediction horizon 10 ms, 5km/h users at 2.66 GHz. Values are given in bps/Hz/cell.**

	RLP, $S=S_p$		ZF		Single cell	
	Ave	5% perc.	Ave	5% perc.	Ave	5% perc.
Random grouping using all frequencies	5.7	2.4	4.8	0.86	-	-
Cellular grouping with RR	6.2	2.8	6.2	2.7	4.7	2.1
Cellular grouping with SB	6.3	3.0	6.4	2.9	4.8	2.7

### Results: Fairness

The above percentiles are given in terms of the sum-rate and do not take user fairness into account. In Table A.3 the 5% percentile and the average of the individual channel capacities of the UEs are given. Here, the fact that we have chosen to let RLP use the sum-rate as an objective function is shown to be less beneficial for the weakest users in the system. When cellular user grouping with RR scheduling is used, ZF outperforms RLP by 32% with respect to the 5% percentile results. Using SB scheduling helps the situation somehow and reduces this figure to 20 %. However, to better take fairness into account, a different objective function must be used in the optimization.

Both JT schemes outperform the single cell solution.

**Table A.3: Individual payload capacity (no overhead included) on 144 subcarriers à 15 kHz = 2.16 MHz. Single-antenna transmitters and receivers are used. M=9 UEs, 10 ms prediction, 5km/h users at 2.66 GHz.**

	RLP, $S=S_p$		ZF		Single cell	
	Ave	5% perc.	Ave	5% perc.	Ave	5% perc.
Random grouping with RR	4.0 Mbit/s	0.20 Mbit/s	3.4 Mbit/s	0.27 Mbit/s	-	-
Cellular grouping with RR	5.5 Mbit/s	0.91 Mbit/s	5.4 Mbit/s	1.2 Mbit/s	3.5 Mbit/s	0.63 Mbit/s
Cellular grouping with SB	6.1 Mbit/s	1.5 Mbit/s	6.1 Mbit/s	1.8 Mbit/s	4.0 Mbit/s	0.96 Mbit/s

Table A.3 also shows that use of cellular user grouping for JT CoMP schemes does not only improve the sum-rate significantly, but equally importantly, increases the 5 % percentiles significantly.

## Results: The impact of noise and out-of-cluster interference levels

All results above represent averages of three constant noise floors used in the simulation environment described in Appendix A4-2: -110 dBm, -120 dBm and -130 dBm. Figure A.8 shows the sum-rate for the individual noise floors in the case of random user grouping and also for cellular user grouping with SB scheduling.

A high noise floor affects the rate performance directly; by the noise term in the rate expression, and also indirectly; due to higher prediction errors in the CSI. Both of these aspects tend to benefit single cell transmission in comparison to JT CoMP. With high noise and out-of-cluster interference, the intracluster interference becomes relatively less important. Therefore there is less need for interference management. The capacity of single-cell transmission is furthermore not affected by the prediction errors in our single-antenna setting.<sup>38</sup> By comparing the performance at the lowest noise floor with that of the highest noise floor, we see that these two aspects cause a large difference in the gain we can achieve by switching to JT CoMP instead of single cell transmission. This supports what was illustrated by Figure 5.4 in Chapter 5.2.1, namely that the largest CoMP gains can be achieved when the noise floor is properly suppressed.

To better understand the performance limitations of ZF when random user grouping is applied, Figure A.8 also shows the performance that ZF would have if provided with perfect CSI. We see that in a high-noise setting, ZF would not perform better when provided with perfect CSI. Hence, it is in fact the problem of badly conditioned matrices, leading to a large scaling factor  $c$  that causes the large performance loss, and not the prediction error. The RLP is here benefitted both from the fact that it has an extra term in the inverse part of (A2.2.7) and by the iterations performed. This agrees with ZF becoming a much better scheme when cellular user grouping is applied: matrices then become well conditioned and inversion does not lead to a need for large scaling constants.

When the noise floor is low, as in the lowest figure, and cellular user grouping is used, we see that the zero-forcing solution actually outperforms the RLP slightly. This is due to a small error in the estimation of the covariance term in (A2.2.7). The term used is the one given directly from the Kalman predictors. However, the actual channel error is lower since the predicted channels are then smoothed over all frequencies.

---

<sup>38</sup> When power control at the transmitter is not used, the theoretical Shannon (ergodic) rate for fading SISO links is not affected by channel state information at the transmitter, or errors in that information [Gol05]. The situation for real performance of adaptive transmission is of course different, but in this investigation, we do not take into account the effects of prediction errors on the transmission rates obtainable by realistic rate adaptation, coding and HARQ schemes.

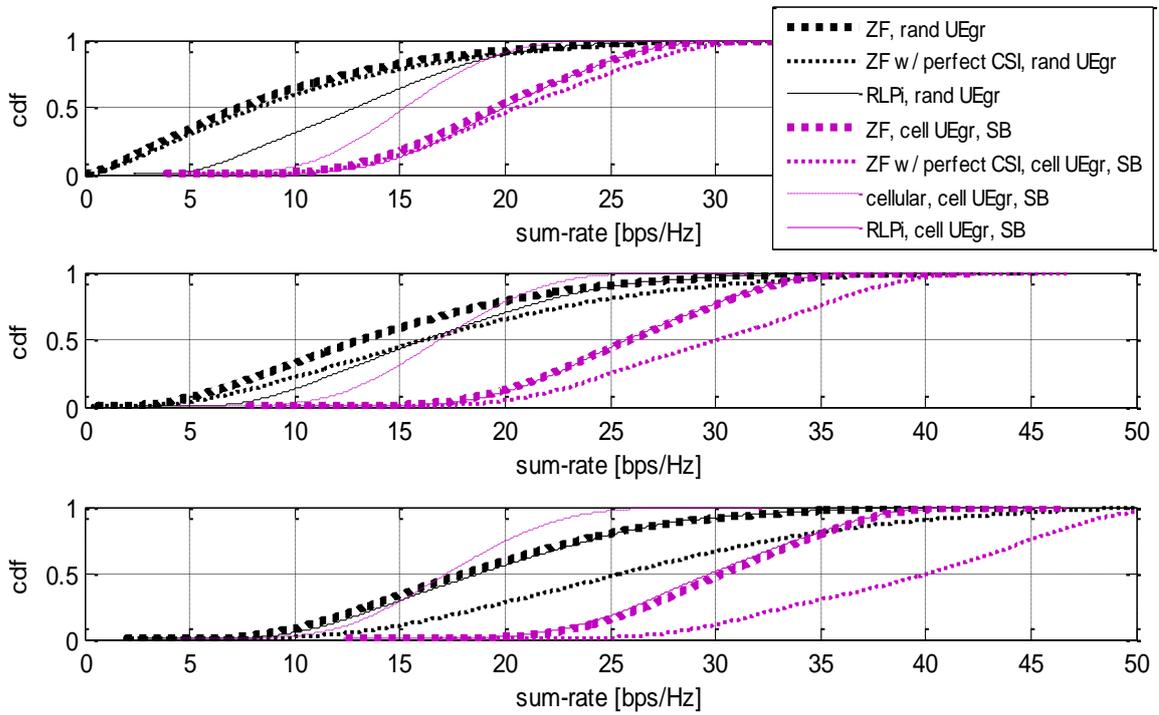


Figure A.8: CDFs of sum-rates for different noise floors are (from the top) -110 dBm, -120 dBm, -130 dBm. Results for random user grouping (black) and cellular user grouping with SB scheduling (purple).

### Results: Impact of a longer transmission latency and prediction horizon

Table A.4 shows what happens when we increase the channel error by using the Kalman prediction at a prediction horizon  $\mathcal{G}=18$  (23 ms) from appendix A.4-2. Comparing these results with those in Table A.1 we see that both JT CoMP schemes loose some performance, but even at this higher latency, single cell transmission is significantly outperformed by JT CoMP.

Table A.4: Average and 5 % percentile of the *sum-rates* using 23 ms prediction horizon ( $\mathcal{G}=18$ ), for precoding schemes with random and cellular UE grouping with 9 UEs, for 5km/h users at 2.66 GHz. Values are given in bps/Hz/cell.

	RLP, $S=S_p$		ZF		Single cell	
	Ave	5% perc.	Ave	5% perc.	Ave	5% perc.
Random grouping with RR	5.0	2.2	4.0	0.67	-	-
Cellular grouping with RR	6.8	3.6	6.7	3.1	4.9	2.3
Cellular grouping with SB	7.6	4.7	7.6	4.5	5.5	3.5

### A2.2.7 Conclusions from the performance evaluation

We have investigated a setting with joint downlink transmission from three sites with single antennas, using macro-cellular channels from channel sounding measurements in an urban environment. We have assumed significant channel uncertainty, due to transmission latencies of 10-23 ms in 2.66 GHz downlinks with 5 km/h user mobility. In this setting, coherent JT CoMP provides significant gains over single-cell transmission with respect to the average sum Shannon (ergodic) rate. Kalman-based channel prediction over the required 10-23 ms horizons provides adequate channel accuracy for useful coherent joint linear precoding designs.

The obtained gains depend to a striking extent on the utilized user grouping scheme. With random user grouping, JT CoMP would outperform the single-cell average sum rate (with Round Robin scheduling) by only 12%. With cellular user grouping, the gains grow to 54% when using Score-based scheduling and to 57% when using Round Robin scheduling.

We have seen that for coherent JT CoMP, **a good user grouping scheme is a first order effect** to achieve gains. Hence, system level simulations that do not provide good user groups for the JT CoMP schemes will not provide a fair comparison when comparing to single cell transmission.

We have shown that the user groups that are obtained when each cell schedules its own users independently of the other cells of the cooperation area will provide a good basis for JT CoMP. This simple method of user grouping has two additional strengths:

It does not destroy the scheduling gains provided by multi user diversity.  
Already implemented scheduling algorithms can be applied.

Moreover, we have seen that cellular user grouping increases the 5 % percentile of the individual users' performance in terms of capacity. For the RLP scheme, cellular user grouping resulted in an impressive gain of 360 % in capacity for the weakest users when RR scheduling was applied.

**A second order effect for obtaining JT CoMP gains is to choose a good scheduling algorithm.** We have here seen this by comparing RR scheduling with SB scheduling. The gains from using SB scheduling were especially high in the 5 % percentiles. For 9 users, it gives a system level (sum-Shannon rate) performance gain of 28 % and a user capacity performance gain of 65 % when compared to RR scheduling.

**The use a precoder that takes CSI uncertainty into account is a useful safety net.**

With cellular user grouping, the performance difference between RLP and ZF is minor in cases when everything goes well. However, we have shown that the RLP scheme provides good sum-rate and outperforms single cell transmission, even when random user grouping is applied. This is not the case when using zero-forcing.

The system might fail to provide a good CoMP group in all fully loaded scenarios. The channel eigenvalue spread may still be quite substantial. This has been observed when MU-MIMO is used within cells, with aggressive overbooking strategies, see Subsection 5.2.2.3. The robust linear precoder provides safety against catastrophic failures of the precoder design in such situations.

In order to achieve increased fairness on a single user level, we recommend that an objective function different from the sum-rate (A2.2.10) is used. We also recommend that a fair scheduler such as SB scheduling is used. We have shown that this improves the 5 % percentiles of the individual user capacities significantly, and also increases average capacity.

### A2.2.8 Outline of the transmission scheme for FDD downlinks

The downlink joint transmission scheme relies on mechanisms that work on three separate time-scales. We summarize the scheme here and also describe which parts work on what time-scale in the spirit of the distinctions introduced in Section 3. A strategy for reducing the feedback reporting overhead is also outlined and motivated.

#### ***Semi-static resource allocation.***

Overlapping sets of cooperation clusters are formed statically or semi-statically. The  $N$  transmitters within the cluster are logically organized into  $N_c$  cells, each served by a base station with one or several antenna elements.

Interference control between clusters that use the same time-frequency resources is also performed statically or semi-statically by power control, fractional frequency reuse and/or (frequency-dependent) antenna downtilt.

#### ***Slow (0.5 seconds - 2 seconds) timescale control processes:***

- *Association:* Terminals are (re)allocated to one or several cooperation clusters. They are to use the time-frequency resources that have been allocated to the cluster for downlink transmission.
- *Mode switching:* Depending on channel quality statistics, terminal velocities and types of data traffic, UEs are to use either coherent joint downlink transmission, non-coherent joint transmission or cellular transmission, possibly combined with coordinated scheduling. Out of the total user pool,  $M$  users are to participate in joint coherent downlink transmission.
- *Cell association/handover:* Each terminal is also (re)allocated to one specific Master eNB within the cluster that is responsible for scheduling as well as downlink and uplink control signalling. It is the eNB associated with the cell with the highest downlink power averaged over the slow timescale.
- *User-plane coordination:* Queues for the user-plane traffic of all data streams for each of the  $M$  UEs that are to participate in joint transmission are maintained within the cluster. Queues are also set up/terminated for UEs that are reallocated to/leave the cluster.<sup>39</sup>
- *Channel estimation:* The fading statistics may change on this timescale, so models of the fading statistics used for channel prediction need to be re-estimated, see Appendix A4-2.
- *Reporting of prediction accuracy.* The channel predictors report the accuracy parameters (such as covariance matrix elements) to the central unit that calculates joint precoding matrices.

---

<sup>39</sup> Logically, we here assume coordinated per-stream queues implemented at each eNB within the cooperation cluster. This mechanism could in practice be implemented as per flow class queues, which handle the RLC layer segments of several streams. It could also be implemented as one set of centralized queues per cooperation cluster, if latencies in the communication to all  $N$  antenna elements can be kept low. Likewise, the medium access control (MAC) layer and the physical layer (PHY) processing are also assumed to be implemented separately in the eNB/cells within the cluster. Also for MAC and PHY, centralized implementations are possible. The cooperation cluster would then consist of  $N$  remote radio heads, fibre-optically connected to the central unit.

The situation after these allocations is that  $M$  users are to share FDD downlinks within a cluster, and are to use coherent joint downlink transmission that use some of the  $N$  transmitters. This allocation will remain stable until the next update of the slow-scale association process.

We now describe the control sequence that results in  $M_s$  out of these  $M$  users receiving downlink transmissions in one specific subframe of duration 1 ms. (In most of the numerical examples above,  $M_s = M = 9$ .) We here assume channel predictors to be located at the UE. At the start of the sequence (subframe 0), the subframe of interest lies  $L_1$  subframes, or milliseconds, into the future. where  $L_1$  could be 6-20 ms.

### **Fast (1 ms) subframe timescale processes:**

1. *Channel prediction* is performed over the  $L_1$  ms horizon, for the  $N_p$  strongest downlink channels to users within the cluster. This is done by and for all UEs that might need downlink transmissions within that subframe. As discussed in Section 5.3.2, CSI-RS are used for this purpose, using orthogonal RS locations for RS from different cells of the cluster.
2. *Uplink reporting of coarse CSI* for scheduling, link adaptation and intra-cell beamforming and MIMO precoder selection. This CSI feedback may use the same format and uplink control channels as are used today for cellular transmission. It could also use enhanced formats with more detailed reporting of channel variations over frequency and/or phase information. Anyway, its quality is improved compared to current state-of-art since it is based on channel predictions generated at the UEs, rather than on outdated instantaneous measurements.
3. *Cellular scheduling and cellular beamforming* (both possibly coordinated with other cells in the cluster) is then performed per subframe by all the eNBs within the cooperation cluster, for all frequency resources and spatial resources within each cell. Link adaptation parameters are set aggressively, taking into account that CoMP will be used to suppress interference. Thus, in total  $M_s$  out of the  $M$  terminals are scheduled to receive downlink transmission within the cluster in one subframe of the OFDM downlink. As a consequence of this cellular scheduling, a CoMP user grouping is generated: In each frequency resource block  $k=1, \dots, K$ , transmissions to  $M_k$  out of the  $M_s$  UEs will have to share non-orthogonal frequency-spatial resources within the cluster. These groups of  $M_k$  terminals form our resource-block-specific CoMP groups.
4. *Downlink control signalling and request for detailed CoMP CSI:* The  $M_s$  scheduled UEs are notified of their decided downlink transmission, and the corresponding transmission parameters, in the resource block that now still lies  $L_2$  ms into the future. A request for detailed CoMP CSI accompanies this transmission.
5. *Updated channel predictions:* These  $M_s$  scheduled UEs, and only these, update their channel predictions, now using the (much shorter) prediction horizon  $L_2$  ms, where  $L_2 < L_1$ . This needs to be done only for the frequency resource blocks that will actually be scheduled for that UE.<sup>40</sup> The prediction horizon  $L_2$  must take into account the time that will be needed to distribute the joint precoder beamformers to all participating sites in the cluster.
6. *Uplink reporting of fine-grained CSI from relevant UEs:* The channel predictions for horizon  $L_2$  ms are reported over (expanded) uplink control feedback channels, likely using random access channels. Only information from the  $M_s$  scheduled UEs, and only for the relevant scheduled resource blocks, is fed back, appropriately quantized.

---

<sup>40</sup> Predictions, according to eq. (A4.2.4) in Appendix A4-2 only need to be calculated for the relevant resource blocks. However, the state estimators (A4.2.3) of the Kalman filters must be continuously updated for all frequency resource blocks that might be potentially used by the scheduler. The state estimation recursions cannot be quickly “cold-started”. They need some initial transient time after initialization.

7. *Design of coherent JP CoMP precoders* is then performed for all resource blocks in the subframe, using the RLP algorithm described in Section A2.2.3.
8. *Distribution of the sets of JP CoMP beamformers* to all transmitters in the cooperation cluster.
9. *User-plane data has in the mean time been distributed* to all queues within the cluster so that all queues are synchronized and prepared in time for transmission. The MAC and PHY cellular processing has also been performed. Only the CoMP beamformers are added as a last step.
10. *Data transmission is then performed* jointly over the  $N$  transmitters. The CoMP processing can also be applied to the downlink control signalling within the subframe.

In a variant where the channel predictors are located at the network side, steps 5 and 6 would be eliminated. Instead, during step 1, measurements at RS positions would have to be reported over the uplink by *all*  $M$  UEs that are potentially to be scheduled, for *all* resource blocks that may potentially be used. Based on these reports, predictors at the network side then predict the channels, using the longer prediction horizon  $L_1$  ms. (This horizon could be made somewhat shorter, since the second round of feedback is here eliminated). Such a variant would lead to both much higher feedback load and more uncertain predictions due to a longer prediction horizon.

In the above scheme, it has been assumed that Kalman predictors are used. Although this provides the best performance, it is not necessary. Any channel prediction scheme could be used, even simple extrapolation of outdated estimates, as long as the predictions are also accompanied by an indicator of their accuracy, which can be used in robust precoders. CSI in the absence of such quality indicators would represent information of very dubious value.

### A2-3 Joint scheduling and power control with non-coherent transmission

[This Appendix contains algorithms and results used in Subsection 5.2.3]

#### System Model

Consider the downlink of a CoMP cluster, in which  $N$  BSs are connected via backhaul links to a CU.  $M$  users are grouped together using a particular resource slot, e.g., a time slot or a subchannel. The  $N$  BSs are assumed to have the same maximum per-BS power constraint  $P_{max}$  and to share the same resource slot. The BSs and the users are assumed to have one transmit antenna and one receive antenna, respectively.

Let  $\mathbf{S} = [s_{nm}]$  denote a user selection indicator matrix of size  $N \times M$ , where  $s_{nm}$  is interpreted as

$$s_{nm} = \begin{cases} 1, & \text{if BS } n \text{ transmits data to user } m \\ 0, & \text{otherwise} \end{cases}$$

It is assumed that a BS can transmit data to at most one user in any given resource slot. Then, at most one single element in each row of  $\mathbf{S}$  is non-zero. Hence, the  $N$  BSs within a cluster are grouped into several subclusters, forming a group of user-specific cooperative BS sets (CBS). Each user has a CBS, which is formed by the BSs of the cluster that provide data transmission service to it. Hence, for each user, its CBS may include zero, one, or multiple BSs. Denote  $\text{CBS}_m$  as the CBS of user  $m$ , that is

$$\text{CBS}_m = \{n \mid s_{nm} = 1, n \in \{1, \dots, N\}\}.$$

The data symbol of user  $m$  is transmitted non-coherently from the BSs in  $\text{CBS}_m$  without phase adjustment. Therefore, inter-user interference can not be mitigated by cancellation.

Let  $\mathbf{x} = [x_1, \dots, x_N]$  denote the signal vector transmitted from all  $N$  BSs, with  $P_n = x_n^H x_n \leq P_{max}$  for all  $n$  ( $n = 1 \dots N$ ). The received signal at user  $m$  ( $m = 1 \dots M$ ) can then be expressed as

$$y_m = \sum_{i \in \text{CBS}_m} h_{mi} x_i + \sum_{j \notin \text{CBS}_m, j \in \{1, \dots, N\}} h_{mj} x_j + n_m,$$

where  $h_{mi}$  denotes the channel vector between user  $m$  and BS  $i$ . Above,  $n_m$  is the sum of the thermal noise and the uncoordinated out-of-cluster interference, modeled as independent complex additive Gaussian noise with zero mean and covariance  $\sigma^2$ . Then, the SINR of user  $m$  is given as

$$\gamma_m = \frac{\left\| \sum_{i \in \text{CBS}_m} h_{mi} \sqrt{P_i} \right\|_2^2}{\left\| \sum_{j \notin \text{CBS}_m, j \in \{1, \dots, N\}} h_{mj} \sqrt{P_j} \right\|_2^2 + \sigma^2},$$

Hence, the sum rate of the CoMP cluster can be calculated by

$$C = \sum_{m=1}^M \log_2(1 + \gamma_m).$$

#### Joint Scheduling and Power Control

With the objective of maximizing the sum rate of the cluster, the CU needs to jointly determine the user selection indicator matrix  $\mathbf{S}$  and the power allocation vector  $\mathbf{P} = [P_1, \dots, P_N]$  based on

the CSI information available at the CU. The optimization problem under per-BS power constraints can be formulated as

$$\begin{aligned} \max_{\mathbf{S}, \mathbf{P}} \quad & \sum_{m=1}^M \log_2(1 + \gamma_m) \\ \text{s.t.} \quad & 1) 0 \leq P_n \leq P_{\max}, \forall n, \\ & 2) s_{nm} \in \{0, 1\}, \forall n, \forall m, \\ & 3) \sum_{m=1}^M s_{nm} \leq 1, \forall n. \end{aligned}$$

Note that the formation of the CBS for each user is included in the optimization problem. The constraint 3) guarantees that a BS transmits to at most one user. This optimization problem is non-convex. Based on [LSB+11], a suboptimal binary power control (BPC) is considered for power allocation, i.e.,  $P_n = 0$  or  $P_n = P_{\max}$ , for  $\forall n$ . Then, the relaxed problem becomes an exhaustive binary search. The CU searches all the possible values of the user selection indicator matrix  $\mathbf{S}$  and all feasible boundary point sets for binary power control. The chosen matrix  $\mathbf{S}^*$  and transmit power vector  $\mathbf{P}^*$  will be the ones that achieve the highest  $C$ . Here, the non-coherent joint transmission with BPC is named as NCJT-BPC.

### Simulation Results

The performance of the NCJT-BPC scheme in terms of sum rate is studied. For the sake of comparison we also include the sum rate achieved when using coherent zero-forcing joint transmission scheme with optimal power allocation [BH06], named as ZF-OPA. Single cell transmission scheme without BS coordination, denoted as SC, is used as baseline.

As shown in Figure A.9, we consider the downlink of a CoMP cluster with  $N=3$  neighboring sectors and  $M=3$  users. Assume that sector  $n$  is the serving sector of user  $n$ , with  $n=\{1,2,3\}$ . Each user is moving from the cluster center to the sector center of its serving sector along the dashed line in Figure A.9. The cluster radius is  $R=500\text{m}$ . The path loss model is  $PL(d)=128.1+37.6\log_{10}(d)$  dB, with  $d$  given in km [3GPP36814]. Long-term shadowing is log-normally distributed with zero mean and standard deviation 8 dB. The system SNR is set to 18 dB, which is defined as the received SNR at the boundary of the cell, assuming full power transmission  $P_{\max}$  from the BS, accounting only for pathloss gain and ignoring shadowing. The carrier frequency is 2 GHz.

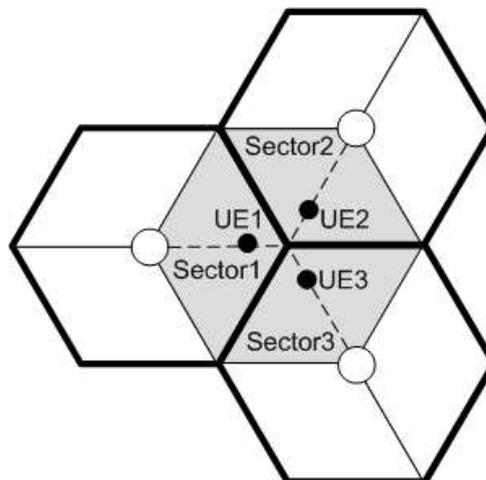


Figure A.9: A CoMP cluster of 3 neighboring sectors (the shadowed area)

### Sum rate performance with perfect CSI

Let  $d$  be the distance between a user and the center of its serving sector as shown in Figure A.10. Assume that perfect CSI is available at the CU. In Figure A.10, the average sum rate of each transmission scheme is plotted versus the normalized distance ( $d/R$ ). Compared with the SC scheme, the CoMP transmission schemes (ZF-OPA, NCJT-BPC) provide a significant average sum rate gain, especially for the users located at the cluster center areas or cell-edge areas (the users with large values of  $d$ ).

Note that ZF-OPA achieves superior performance compared to the NCJT-BPC scheme. That is because, with perfect CSI at the CU, zero-forcing precoding performed in ZF-OPA can completely remove the interference for all users within the CoMP cluster. In the NCJT-BPC scheme, a single user receives data symbols from a subset of BSs (i.e., its CBS), hence, the BSs outside its CBS would still introduce interference. In addition, the data symbols of each user are transmitted without phase adjustment, which would also result in performance degradation.

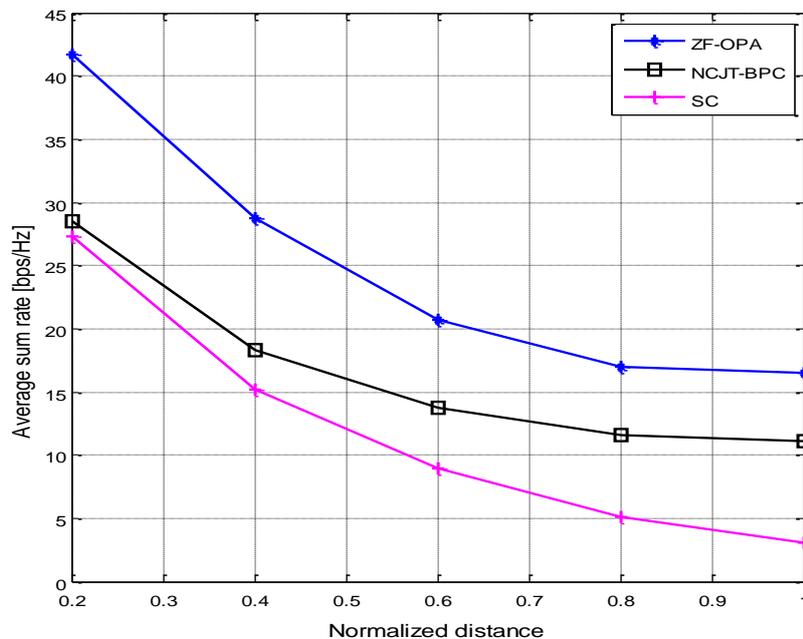


Figure A.10: Average sum rate vs. different normalized distance,  $d/R$ . Assume perfect CSI is available at the CU side.

### Sum rate performance with predicted CSI

Due to practical issues (e.g., synchronization constraints, feedback and backhaul constraints, user mobility), only imperfect CSI is available at the CU, which affects the performance of CoMP transmission schemes. In this subsection, the performance of the considered CoMP transmission schemes (ZF-OPA, NCJT-BPC) is evaluated with predicted CSI. The effects of total feedback and backhaul latency, and the user mobility are considered. Channel prediction accuracy is obtained according to [Aro11].

Figure A.11 shows the effect of total feedback and backhaul latency,  $\Delta t$ , on the performance of the considered CoMP transmission schemes. The user speeds are set to 5 km/h. The average sum rate of each scheme is plotted versus normalized distance ( $d/R$ ) for  $\Delta t = 10.2, 20.4,$  and  $30.6$  ms, respectively. Compared with the achieved performance under perfect CSI (see Figure

A.10), the average sum rate of the sector center users when  $\Delta t=30.6$  ms is decreased approximately by 30.2% and 1.2% for ZF-OPA and NCJT-BPC, respectively. For the sector-edge users with  $\Delta t =30.6$  ms, the average sum rate of ZF-OPA dramatically decreases to 53.4%; while the performance loss due to imperfect CSI for NCJT-BPC is 10.8%. Hence, the NCJT-BPC scheme is more robust with respect to the effect of latency. Note that in a realistic CoMP system, backhaul links can be implemented via high-latency X2 interfaces [3GPP-R1112339]. Considering the feedback latency and the data sharing among coordinated BSs, the total latency may be greater than 30.6 ms. Therefore, for high-latency backhaul links, NCJT-BPC is a better choice for CoMP transmission design.

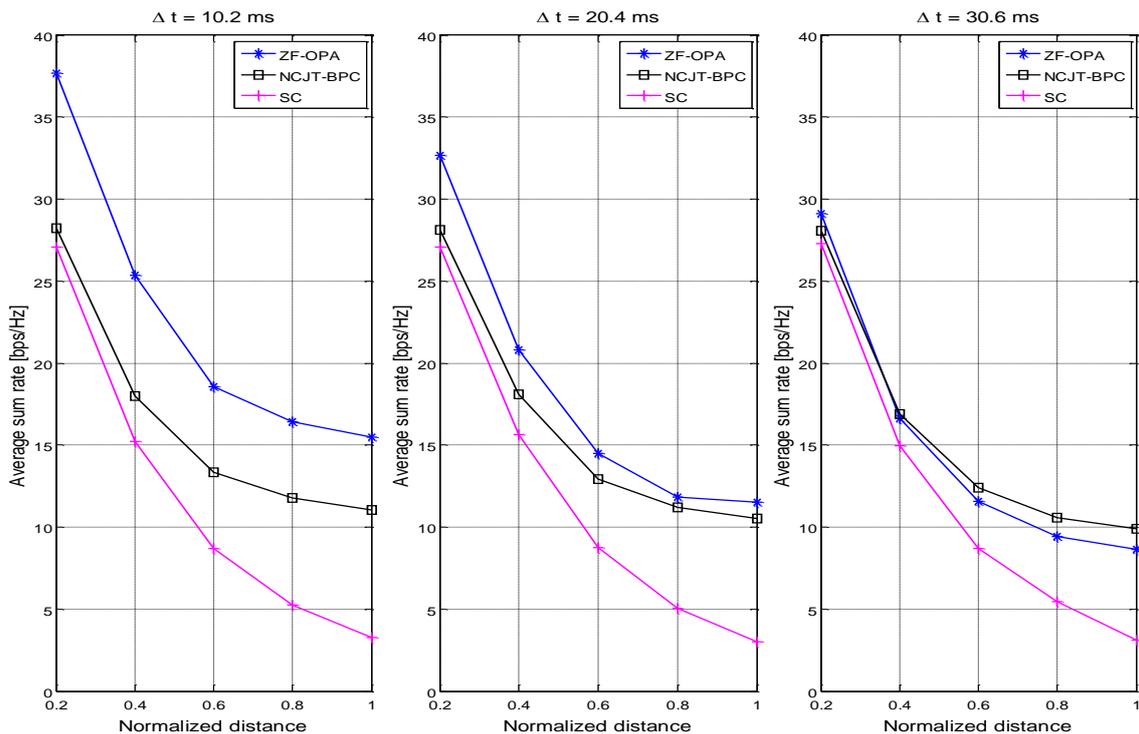


Figure A.11: Average sum rate vs. different normalized distance,  $d/R$ .  $v=5$  km/h.  $\Delta t =10.2, 20.4, 30.6$  ms.

## A2-4 Waterfilling schemes for Zero-Forcing coordinated transmission

[This Appendix contains algorithms and results used in Subsection 5.2.3]

### System Model

The system model corresponding to one of the Advanced Joint Transmission Schemes for Multi cell cooperation proposed in Section 5 is described below. It assumes a coordinated transmission downlink cellular scenario based on OFDM, where  $M$  cooperating eNB serve  $N$  users or UE. Each base station has  $t$  transmit antennas and each user has  $r$  receive antennas. In the following, the analysis will be applied to eNB-user pairs, therefore the case  $M = N$  will be considered. We consider a CoMP system with OFDM where the whole channel (linear time-invariant channel with frequency selective fading and additive Gaussian noise) is known to all eNB. As long as the length of the cyclic prefix is chosen longer than the longest impulse response, the channel seen by each user can be decomposed into  $N_{OFDM}$  independent flat subchannels with frequency response  $\mathbf{H}_k^p$  for the user  $k$  ( $k = 1 \dots N$ ) and the subchannel  $p$  ( $p = 1 \dots N_{OFDM}$ ). So, the received signal on the  $p$ -th subcarrier is as follows

$$\mathbf{y}^p = \mathbf{H}^p \mathbf{x}^p + \mathbf{n}^p$$

where  $\mathbf{y}^p$  is the received  $Nr \times 1$  signal vector on the  $p$ -th subcarrier,  $\mathbf{x}^p$  is the  $Mt \times 1$  signal vector transmitted from all the eNB on the  $p$ -th subcarrier,  $\mathbf{n}^p$  is the  $Nr \times 1$  i.i.d complex Gaussian noise vector on the  $p$ -th subcarrier, with variance  $\sigma^2$ , and  $\mathbf{H}^p = [\mathbf{H}_1^{pT} \ \mathbf{H}_2^{pT} \ \dots \ \mathbf{H}_N^{pT}]^T$  is the  $Nr \times Mt$  matrix channel on the subcarrier  $p$  ( $p = 1 \dots N_{OFDM}$ ) whose matrix coefficients represent the fading from each transmit antenna in the eNB to each receive antenna at the user side. For this scenario we define  $\mathbf{x}^p$  as follows

$$\mathbf{x}^p = \sum_{i=1}^r b_{1i}^p \mathbf{w}_{1i}^p + \sum_{i=1}^r b_{2i}^p \mathbf{w}_{2i}^p + \dots + \sum_{i=1}^r b_{Ni}^p \mathbf{w}_{Ni}^p = \mathbf{W}^p \mathbf{b}^p$$

where  $b_{ki}^p$  represents the  $i$ -th symbol for user  $k$  transmitted with power  $P_{ki}^p$  on the  $p$ -th subcarrier, and  $\mathbf{w}_{ki}^p = [w_{ki}^{p,1}, \dots, w_{ki}^{p,(m-1)t+j}, \dots, w_{ki}^{p,Mt}]^T$  are the precoding vectors being  $w_{ki}^{p,(m-1)t+j}$  the weight of  $j$ -th transmit antenna ( $j = 1 \dots t$ ) of the  $m$ -th base station for the  $i$ -th symbol of the user  $k$  transmitted on the  $p$ -th subcarrier. The precoding matrix  $\mathbf{W}^p = [\mathbf{w}_{11}^p, \dots, \mathbf{w}_{1r}^p, \dots, \mathbf{w}_{k1}^p, \dots, \mathbf{w}_{kr}^p, \dots, \mathbf{w}_{N1}^p, \dots, \mathbf{w}_{Nr}^p]$  will be obtained under a Zero-Forcing criteria defined as in [KFO6], to guarantee that

$$\mathbf{H}_k^p [\mathbf{w}_{q1}^p, \mathbf{w}_{q2}^p, \dots, \mathbf{w}_{qr}^p] = \begin{cases} \mathbf{0} & k \neq q \\ \mathbf{U}_k^p \mathbf{S}_k^p & k = q \end{cases}$$

$$\|\mathbf{w}_{ki}^p\|^2 = 1, \quad k = 1 \dots N; i = 1 \dots r; p = 1 \dots N_{OFDM}$$

where  $\mathbf{U}_k^p$  is a unitary matrix and  $\mathbf{S}_k^p = \text{diag} \left\{ (\lambda_{k1}^p)^{1/2}, (\lambda_{k2}^p)^{1/2}, \dots, (\lambda_{kr}^p)^{1/2} \right\}$  is a diagonal matrix that contains the square roots of the nonzero eigenvalues of the matrix  $\mathbf{Q}_k^p \mathbf{Q}_k^{p\dagger}$  being  $\mathbf{Q}_k^p$  the

part of the channel matrix  $\mathbf{H}_k^p$  orthogonal to the subspace spanned by other users' channels  $\mathbf{H}_q^p$  ( $q \neq k$ ).

Therefore, the received signal on  $p$ -th subcarrier can be expressed as

$$\mathbf{y}^p = \begin{bmatrix} \mathbf{U}_1^p \mathbf{S}_1^p & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2^p \mathbf{S}_2^p & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{U}_N^p \mathbf{S}_N^p \end{bmatrix} \mathbf{b}^p + \mathbf{n}^p$$

Each user may independently rotate the received signal and decouple the different streams so the signal obtained by  $k$ -th user on  $p$ -th subcarrier can be expressed as

$$\tilde{\mathbf{y}}^p = \mathbf{U}_k^{p\dagger} \mathbf{U}_k^p \mathbf{S}_k^p \mathbf{b}_k^p + \tilde{\mathbf{n}}_k^p = \begin{bmatrix} (\lambda_{k1}^p)^{1/2} \mathbf{b}_{k1}^p \\ \vdots \\ (\lambda_{kr}^p)^{1/2} \mathbf{b}_{kr}^p \end{bmatrix} + \tilde{\mathbf{n}}_k^p$$

where the noise  $\tilde{\mathbf{n}}_k^p$  remains white with the same covariance because of the unitary transformation. It can be observed that the overall system is then a set of parallel noninterfering channels. Hence, the achievable rates per user can be modelled as follows

$$R_k = \frac{I}{N_{OFDM}} \sum_{p=1}^{N_{OFDM}} \sum_{i=1}^r \log_2 \left( 1 + \frac{\lambda_{ki}^p P_{ki}^p}{\sigma^2} \right)$$

The aim is to maximize a weighted sum of the rates  $R_k$  for the set of users, that requires solving the following optimization problem in terms of the power  $P_{ki}^p$  allocated to the  $i$ -th ( $i = 1 \dots r$ ) stream of user  $k$

$$\max \left\{ \frac{I}{N_{OFDM}} \sum_{k=1}^N \alpha_k \sum_{p=1}^{N_{OFDM}} \sum_{i=1}^r \log_2 \left( 1 + \frac{\lambda_{ki}^p P_{ki}^p}{\sigma^2} \right) \right\}$$

subject to a constraint on the maximum available power for transmission from each base station  $m$   $P_{max}$

$$P_{BSm} = \underbrace{\sum_{j=1}^t \sum_{p=1}^{N_{OFDM}} \sum_{k=1}^N \sum_{i=1}^r P_{ki}^p \left| w_{ki}^{p,((m-1)t+j)} \right|^2}_{j \text{ transmit antenna power}} \leq P_{max}, \quad \forall m = 1 \dots M$$

where the values  $\alpha_k$  can be seen as indicating the priority of the users. It is a convex problem that can be solved by standard convex optimization techniques [BOY04]. The application of the Lagrange multiplier technique allows solving this problem (see [GSC09] for details) leading to the general solution given by:

$$\begin{cases} P_{ki}^p = \sigma^2 \left[ \frac{\alpha_k}{\ln(2)L_{ki}^p} - \frac{1}{\lambda_{ki}^p} \right] \\ L_{ki}^p = - \sum_{m=1}^M \sum_{j=1}^t \mu_m \left| w_{ki}^{p,((m-1)t+j)} \right|^2 \\ \sum_{j=1}^t \sum_{p=1}^{N_{OFDM}} \sum_{k=1}^N \sum_{i=1}^r P_{ki}^p \left| w_{ki}^{p,((m-1)t+j)} \right|^2 = P_{max} \end{cases}$$

where  $k = 1 \dots N$ ,  $i = 1 \dots r$ ,  $p = 1 \dots N_{OFDM}$  and  $\boldsymbol{\mu}$  is the vector of the Lagrange multipliers. This solution resembles the well-known waterfilling distribution. However, here the waterlevel is different for each symbol  $i$  to be transmitted to each user  $k$  on each subcarrier  $p$ . Even though the values of the waterlevels can be found again by convex optimization techniques, we still have a similar computational complexity that we would like to reduce.

### Modified waterfilling

By considering the most stringent of the constraints on the maximum available transmission power from each base station, we can reduce the problem to an “equivalent” base station  $m_0$  whose weights are calculated as follows

$$\Omega_{ki}^p = \max_{m=1 \dots M} \left( \sum_{j=1}^t \left| w_{ki}^{p,((m-1)t+j)} \right|^2 \right)$$

Application of the Lagrange multiplier technique gives the new function whose solution is given by

$$P_{ki}^p = \left[ K \frac{\alpha_k}{\Omega_{ki}^p} - \frac{\sigma^2}{\lambda_{ki}^p} \right]^+ \quad \text{with} \quad K = \frac{-\sigma^2}{\ln(2)\mu}$$

where  $[ ]^+$  denotes the maximum between zero and the argument. This corresponds again to a waterfilling distribution with variable waterlevel. However, for given user priorities  $\alpha_k$  and channel realization determining  $\lambda_{ki}^p$  and  $\Omega_{ki}^p$ , the problem reduces to finding a constant  $K$  that can be solved with the same algorithms that solve standard waterfilling (see for example [C07]).

### Waterfilling

We can further simplify the previous solution considering the fact that in a practical realization the values of  $\Omega_{ki}^p$  are close to each other for all  $k$ ,  $i$  and  $p$ . Then we can simplify the modified waterfilling solution to give

$$P_{ki}^p = \left[ K \alpha_k - \frac{\sigma^2}{\lambda_{ki}^p} \right]^+$$

which corresponds to a waterfilling distribution with the waterlevel modified only by the user priorities. In particular for equal priorities  $\alpha_k = 1/N$  it corresponds to a standard waterfilling.

## Numerical Results

Next, we compare the performance of the power allocation schemes in the cooperative processing for interference cancellation in terms of achievable rates of the proposed waterfilling (WF), modified waterfilling (MWF) and the optimum solution found by convex optimization (CVX). For the sake of comparison we also include the rates achieved when using a uniform power distribution (UP).

The channel models studied are a simple frequency-selective channel with an exponential power-delay profile (PDP) and the SCM Micro Urban NLOS channel specified by 3GPP [3GPP25996] for evaluating MIMO system performance in 4G networks, both with  $N_{path} = 6$  paths.

The proposed algorithms are studied in different eNB-UE deployment scenarios:

- Simple scenario with  $M = N = 2$  hexagonal cells.
- More realistic scenario with  $M = N = 16$  hexagonal cells arranged to form a torus [GSC09].

In the first case ( $M = N = 2$ ), different configurations are defined in order to study the influence of the user position in the cell on the algorithms performance. These configurations are as follows:

Configuration 1: both users are placed close enough to the eNB so that the dominant received signal from the base stations is the one that is paired to that particular user.

Configuration 2: one of the users is placed near its eNB and the other one is placed in between the two cell boundaries.

Configuration 3: both users are placed in between the two cell boundaries.

Configuration 4: users are randomly placed in the cell with a uniform distribution.

In the second case ( $M = N = 16$ ), only a uniform distribution, where users are randomly placed in the cell, is considered.

The Figure A.12 shows the mean achievable rates for a two users scenario ( $M = N = 2$ ) where the users, randomly placed in the cell, are assigned different priority. Each of the achievable rates pairs correspond to a pair of values of  $\alpha_1$  and  $\alpha_2$  that  $\alpha_1 + \alpha_2 = 1$ . This figure shows that the achievable rates obtained with WF and MWF are very close to the optimal solution CVX and in most cases far from the performance of the UP.

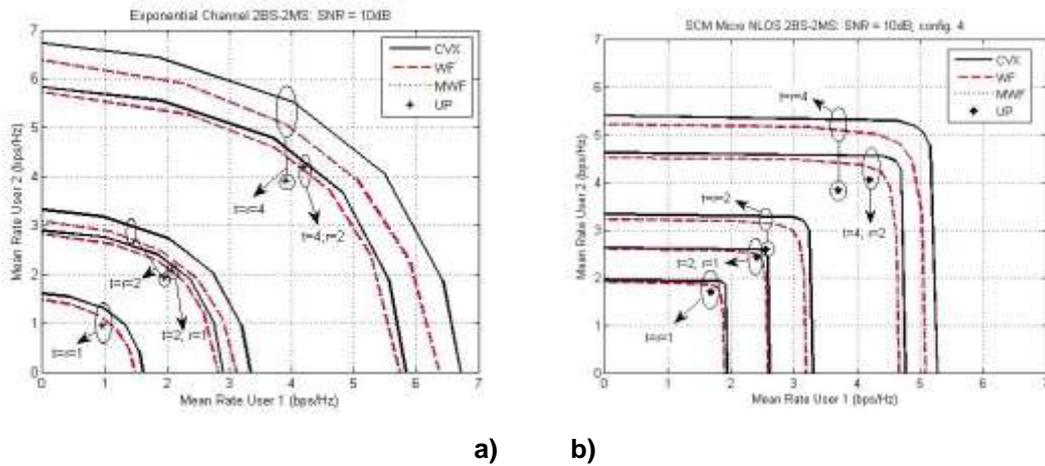


Figure A.12: Mean achievable rates with  $M = N = 2$  and several values of  $t$  and  $r$ : a) Exponential Channel; b) SCM Micro Urban NLOS.

The comparative performance of the users in different positions is given in Figure A.13. It should be noted that due to the optimization criteria, where the sum rate of all users is maximized, the power allocation schemes assign most of the available power in each eNB to serve the user in the best situation and this leads to the user in the boundary to get very low rates. In this figure the effect of the values of the channel eigenvalues in the power allocation schemes is observed.

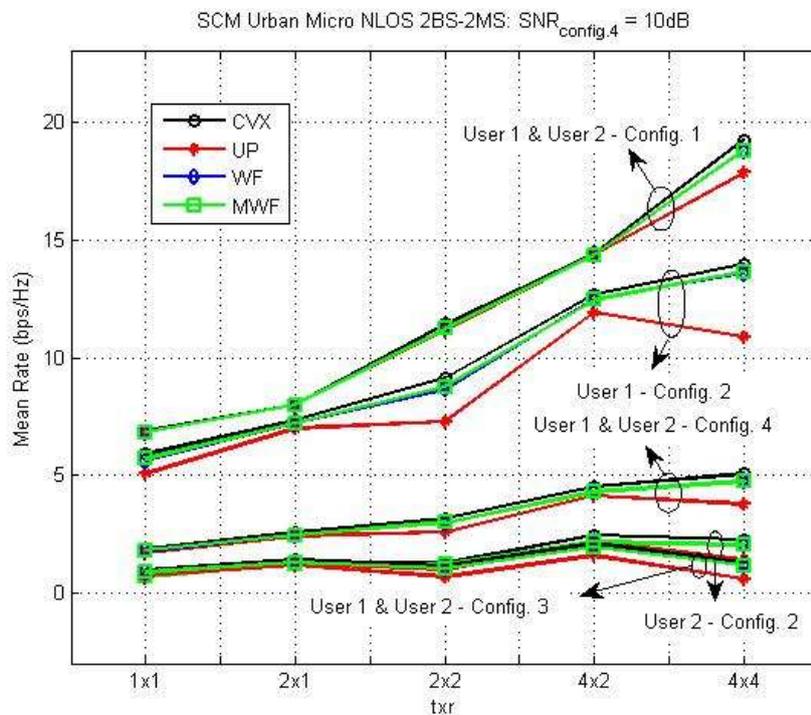
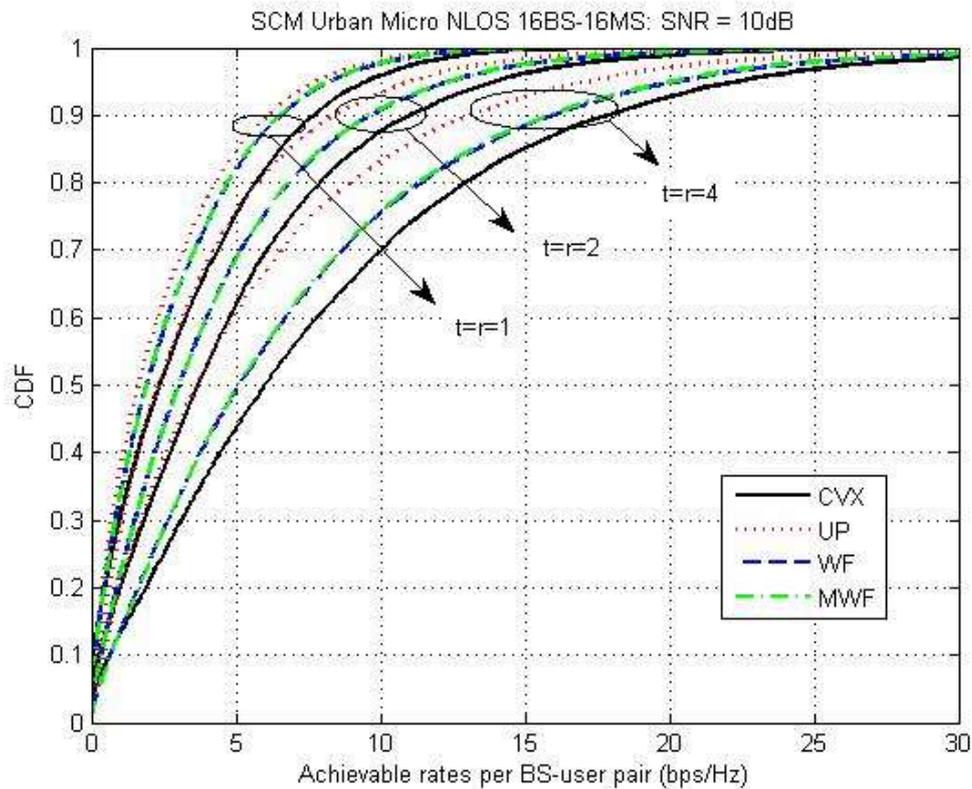


Figure A.13: Mean achievable rates in a SCM Micro Urban NLOS channel with  $M = N = 2$  and  $SNR = 10dB$ . Different configurations and values of  $t$  and  $r$  are considered.

The Figure A.14 shows the CDF of the achievable rates in a more realistic scenario cellular system defined by  $M = N = 16$  hexagonal cells arranged to form a torus. This particular shape avoids the boundary effect that causes cells at the border of the cellular deployment to receive less interference. The users are randomly deployed and an SNR= 10 dB is guaranteed for 90% of the users. This figure shows that the proposed algorithm (WF, MWF) performs close to the optimal (CVX) and outperforms a uniform power allocation scheme.



**Figure A.14: Achievable rates in a SCM Micro Urban NLOS channel with  $M = N = 16$  and SNR = 10dB. Different values of  $t$  and  $r$  are considered.**

## A2-5 Dynamic Partial Joint Processing

[This Appendix contains algorithms and results used in Subsection 5.2.3]

This appendix presents an algorithm for local optimization of precoder matrices, Particle Swarm Optimization, which can be used to adjust the linear precoder matrices used in JT-CoMP with respect to various criteria, under various constraints. In Section 5.2.3, the method is used for optimizing a robust sum throughput criterion under per transmit antenna power constraints, and is compared to other linear precoding schemes.

In this section, the focus is mainly on improving the sum-rate of the cell-edge users who are prone to intercell interference, under backhaul constraints. Joint Processing (JP) is one of the downlink CoMP techniques that can overcome this interference. To obtain the benefits of CoMP, the User Equipments (UEs) need to feed back the Channel State Information (CSI) to its serving Base Station (BS). In our system model, we consider a Central Coordination Node (CCN), where the BSs forward the CSI to be aggregated at the CCN for precoding. Figure A.15 illustrates this process. To reduce the overhead for realizing CoMP, BSs are formed into clusters. Clustering is classified into user-centric clustering and network centric clustering, depending on where the clustering decisions are formed based on the availability of CSI. When the clustering decision is taken by the UE it is called user-centric clustering.

Formation of clusters reduces the overhead for the UEs to estimate and feedback the CSI. This load can be reduced further if UEs could choose the BSs to feedback. In this regard, Partial Joint Processing (PJP) involves selective feedback of CSI from the UEs, where dynamic overlapping subclusters are formed. The BSs could instruct the UEs to feedback the CSI of only those BSs that fall within a threshold, relative to the strongest BS. The overlapping clusters form an active set depending on the UEs that have fed back the CSI. This is illustrated in Algorithm 1. Transporting the CSI coefficients from the UE to the BS is defined to as the *feedback load* while transporting the precoding weights from CCN to the cooperating BSs is defined the *backhaul load*. Even though the backhaul comprises of transporting the CSI from the BSs to the CCN and the user data, we restrict ourselves to the above definitions as we show later that the feedback load due to CSI coefficients is equivalent to the backhaul load due to precoding weights.

It should be noted that a very similar approach based on relative thresholding is proposed in [PBG+08]. The main difference is that we have considered the best link could be from any BS not necessarily the serving BS. For example, when the user is at the cell edge, there is a high probability that the serving BS is not the strongest. And in step 5 of Algorithm 1, the difference is  $\|h(m, k')\|^2 \geq \varepsilon \|h(m, k)\|^2$ , where the  $m$ th user feeds back to the serving  $k$ th BS such that  $k \neq k'$  and  $\varepsilon$  is in the range  $[0,1]$ . Nevertheless, [BSX+10] and [PBG+08] aim to reduce the feedback load under different assumptions.

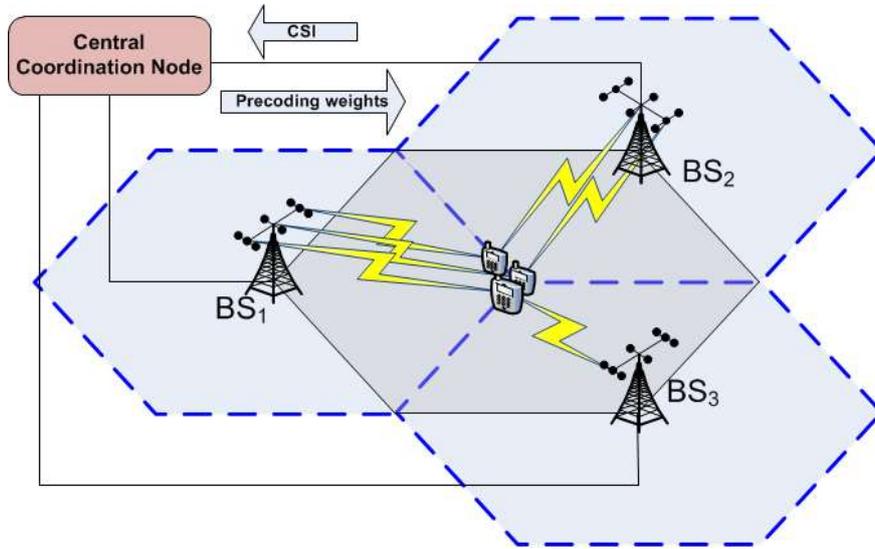


Figure A.15: BSs involved in PJP in serving the UEs at the cluster centre.

**Algorithm 1: Active set relative thresholding for limited feedback based on [BSX+10]**

1:	Choose: <i>threshold</i> = 10 dB
2:	<b>for each UE do</b>
3:	Measure the channel gain from all BSs
4:	<i>bestLink</i> = max{channel strength from all BSs}
5:	<b>if</b> ( <i>bestLink</i> – <i>otherLink</i> ) <= <i>threshold</i>
6:	UE feeds back the CSI of <i>otherLink</i>
7:	CCN marks this link as <i>active</i>
8:	<b>else</b>
9:	<i>Feedback load reduction:</i>
10:	UE does not feed back the <i>otherLink</i>
11:	CCN marks this link as <i>inactive</i>
12:	<b>end if</b>
13:	UE feeds back the <i>bestLink</i>
14:	CCN marks this link as <i>active</i>
15:	<b>end for</b>

PJP reduces the amount of CSI feedback. But due to this, the CSI that is not reported at the CCN are modelled as zeros. This puts additional burden on the precoder design. To cope with the sparsity of the aggregated channel matrix at the CCN, we investigated joint partial zero-forcing (ZF) beamformer (BF), as the basic ZF BF performs poorly under sparsity, i.e., the right inverse could fail,  $\tilde{\mathbf{H}}^H (\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H)^{-1}$ , where  $\tilde{\mathbf{H}}$  is the sparse aggregated channel matrix and  $(\cdot)^H$  is the conjugate transpose operation. The partial ZF BF is designed and evaluated as shown in [LBS+10]. This is summarized in [ARTD12].

The active sets formed in the PJP approach can be applied to every Resource Block (RB) in a frequency selective channel, where the channel gains could vary in different subcarriers. This approach is called Frequency Adaptive (FA) thresholding but when it is averaged for all the RBs then it is called Non Adaptive (NA) frequency thresholding. FA and NA were demonstrated with a WINNER+ channel model in [ARTD13].

Inverting a sparse aggregated channel matrix is presented as a classic right inverse in [PBG11] with a PHY layer approach and a MAC layer approach. The PHY layer approach breaks down when there is no scheduling or block-diagonalization of the aggregated channel matrix. Moreover, the ZF solution gives rise to unnecessary backhauling, i.e., the precoding weight is assigned to BSs where the CSI feedback was not performed by the UE. This also overwhelms the network with unnecessary user data routing based on the network model considered in Figure A.15. In [LBS12a], Particle Swarm Optimization (PSO) is proposed to overcome the difficulties of inverting a sparse aggregated channel matrix without any constraints on scheduling. With the objective of minimizing interference, the PSO proposed in [LBS12a] performs better than [PBG11], given the cases when the closed-form solution exists.

Consider Table A.5 where the UEs feedback the CSI that is aggregated at the CCN as  $\tilde{\mathbf{H}}$ . With a ZF approach, a BF weight is generated for UE3 from BS1 even though UE3 did not feedback the CSI for the link with BS1. And also, UE1 has fed back the CSI for BS1 and BS2, but there is no BF weight generated for BS2. This results in wasting uplink feedback resources when the BF weight is not generated. The user data that is routed via the CCN would also take the same path as that of the precoding weight and thereby causing unnecessary backhauling due to precoding weights. The PSO approach presented in this work takes care of this too, as illustrated in Table A.5.

**Table A.5: An illustration of the precoding matrices generated from limited CSI**

Agg. Channel Matrix at CCN (Zeros ~ feedback savings)				ZF Precoding				PSO based Precoding (Zeros ~ backhaul savings)			
$\tilde{\mathbf{H}}$	BS1	BS2	BS3	$\tilde{\mathbf{W}}$	UE1	UE2	UE3	$\tilde{\mathbf{W}}$	UE1	UE2	UE3
UE1	h11	h12	0	BS1	w11	w12	w13	BS1	w11	0	0
UE2	0	h22	h23	BS2	0	w22	w23	BS2	w21	w22	0
UE3	0	0	h33	BS3	0	0	w33	BS3	0	w32	w33

### Particle Swarm Optimization (PSO)

Scientists, who were simulating the movement of a flock of birds or a shoal of fish trying to find food or escape from a predators, discovered that the birds/fishes were performing optimization. In this work, we regard the birds as particles and finding food to their little ones as the objective function. This is illustrated in Figure A.16 as the birds bringing food their young ones (chicks). The chicks are the result of evaluating the objective function. In our work we consider the objective of *Weighted Interference Minization* where the interference faced by the users is minimized while simultaneously improving the weak Signal to Interference plus Noise Ratio (SINR) user, keeping in mind the fairness for users.

PSO can be seen as a tool used for forming the precoding weights with the constraint of obtaining the zeros in the precoding matrix where needed. In other words, the precoding weights in the backhaul are equivalent to the CSI coefficients. It should be noted that in section 5.2.3, the very same PSO proposed here has been used when there is no constraint on the CSI

feedback, i.e., the zeros in the precoding do not exist. The proposed algorithm automatically takes care of these constraints.

The basic PSO is implemented as shown in Algorithm 2, where the precoding weight is stochastically initialized. Then, the real and imaginary parts of the precoding weight are mapped to the particle, i.e. every particle carries a stochastically initialized precoding matrix. The particles fly through the search space evaluating the objective function. Note that the Figure A.16 also captures the precoding weights in Table A.5. The particles ascertain their velocity and position based on their self belief called the cognitive component ( $c_1$ ) and the belief in other particles through the social component ( $c_2$ ). Each particle also remembers the best precoding coefficients to determine its velocity and position based on this belief.

The algorithmic complexity of the PSO is similar to that of ZF but the complexity of the PSO also grows with the problem size. Hence, it linearly depends on the number of iterations required for the convergence of the algorithm. Recall, that the complexity of a ZF approach is that of the pseudoinverse. The PSO does not invert the aggregated channel matrix as such but instead iteratively always finds an equilibrium solution. Refer [LBS12b] for more details.

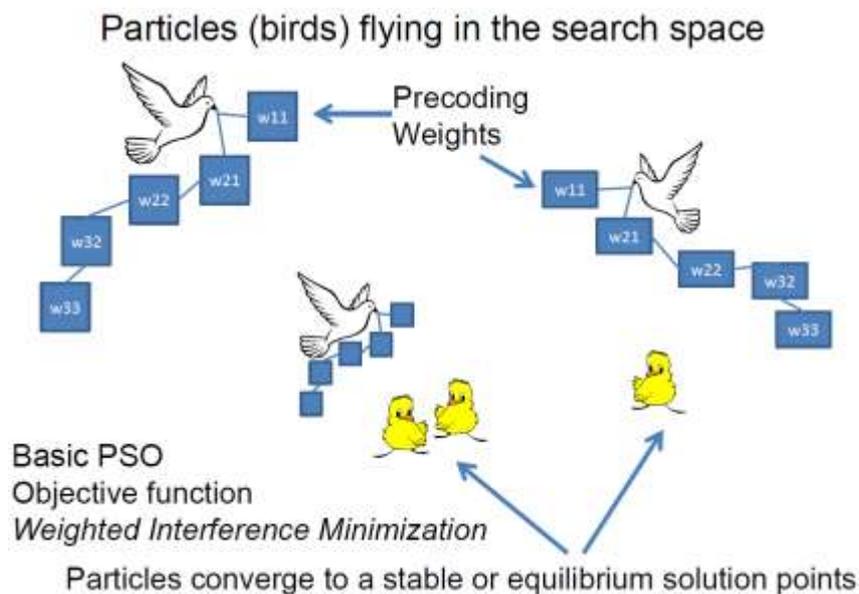


Figure A.16: PSO illustration

### Algorithm 2: Pseudocode for PSO implementation

- 1: *Initialization:*
- 2: Determine the number of non-zero coefficients  $n$  needed in the BF matrix,  $\tilde{\mathbf{W}}$
- 3: Map the BF to the particle:
- 4:  $\mathbf{X}(i, j) \leftarrow \Re\{\tilde{\mathbf{W}}(l, m)\}, l \in \{1, \dots, KN_T\}, m \in \{1, \dots, M\}$
- 5:  $\mathbf{X}(i, j+1) \leftarrow \Im\{\tilde{\mathbf{W}}(l, m)\}$
- 6: Stochastically initialize particles with BF coefficients:
- 7:  $x_{\max} = 1/\max|\hat{\mathbf{H}}(i, j)|$
- 8:  $x_{\min} = -x_{\max}$
- 9: Position:  $\mathbf{X}(i, j) = x_{\min} + r \cdot (x_{\max} - x_{\min})$
- 10: Velocity:  $\mathbf{V}(i, j) = \frac{1}{\Delta t} \left( -\frac{(x_{\max} - x_{\min})}{2} + s \cdot (x_{\max} - x_{\min}) \right)$
- 11: **while** Termination Criterion **do**
- 12:   **for** the  $i$ th particle in the swarm **do**
- 13:     Demap the variables in a particle to form the BF matrix
- 14:      $\tilde{\mathbf{W}}(l, m) \leftarrow \{\mathbf{X}(i, j)\} + i \cdot \{\mathbf{X}(i, j+1)\}$
- 15:     Evaluate the objective function  $f(\mathbf{X}(i, :))$
- 16:     Store:
- 17:     **if**  $f(\mathbf{X}(i, :)) < f^{pb}(\mathbf{X}(i, :))$  **then**
- 18:       Particles' Best:  $\mathbf{X}^{pb}(i, :) \leftarrow \mathbf{X}(i, :)$
- 19:     **end if**
- 20:     **if**  $f(\mathbf{X}(i, :)) < f^{sb}(\mathbf{X}(i, :))$  **then**
- 21:       Swarm's Best:  $\mathbf{x}^{sb} \leftarrow \mathbf{X}(i, :)$
- 22:        $\tilde{\mathbf{W}}^{sb}(l, m) \leftarrow \{\mathbf{x}^{sb}(j)\} + i \cdot \{\mathbf{x}^{sb}(j+1)\}$
- 23:     **end if**
- 24:   **end for**
- 25:   **for** Each particle in the swarm with BF coefficients **do**
- 26:     Update:
- 27:     Velocity:  $\mathbf{V}(i, j) \leftarrow w \cdot \mathbf{V}(i, j) + c_1 \cdot p \cdot \left( \frac{\mathbf{X}^{pb}(i, j) - \mathbf{X}(i, j)}{\Delta t} \right) + c_2 \cdot q \cdot \frac{\mathbf{x}^{sb}(j) - \mathbf{X}(i, j)}{\Delta t}$
- 28:     Restrict velocity:  $|\mathbf{V}(i, j)| < v_{\max}$
- 29:     Position:  $\mathbf{X}(i, j) \leftarrow \mathbf{X}(i, j) + \mathbf{V}(i, j) \cdot \Delta t$
- 30:   **end for**
- 31:    $w \leftarrow w \cdot \beta$
- 32: **end while**
- 33: **return** BF Weight Matrix,  $\tilde{\mathbf{W}}^{sb}$

**Table A.6: Simulation assumptions**

Simulation Parameters	Value
Number of BSs/UEs, $K/M$	3/6
Number of antennas at BS/UE, $N_T/1$	3/1
Shadow fading	$N(0, 8 \text{ dB})$
Pahloss model	$128.1 + 37.6 \log_{10}(d)$
Rayleigh fast fading	$CN(0,1)$
BS antenna gain	9 dBi
Max. BS Tx Power with cell-edge SNR = 15 dB	17.8 dBm
Noise bandwidth	1 MHz
Noise figure	0 dB
Active set threshold for LFB	10 dB
PSO Parameters	Value
Number of particles	30
Number of variables	Number of $\Re$ and $\Im$ parts of the BF coefficients
$x_{\max} = -x_{\min}$	$1/\max\{ \hat{H}(m,l) \}$
Time step length, dt	1
Max. velocity	$(x_{\max} - x_{\min})/dt$
Cognitive factor, c1	2
Social factor, c2	2
Inertia weight	1.4 $\rightarrow$ 0.4
Constant decay factor	0.99
Max. number of iterations	500

Consider 3 BSs with 3 antennas each serving 6 single antenna UEs. These UEs are uniformly dropped around an ellipse at the cluster center as shown in Figure A.15. The simulation parameters and the PSO settings are summarized in Table A.6.

Assuming the CSI coefficients are fed back every 1 ms scheduling interval for a group of subcarriers and OFDM symbols, the feedback rate can be defined as

$$f_r = \frac{(\text{Average number of coefficients}) \cdot (\text{Number of bits})}{\text{Scheduling interval}}$$

where 16 bits are assumed to be required for every complex coefficient for CSI feedback. Similarly the backhaul rate is defined as the number of precoding weights being scheduled for the interpolated frequency-time resource. In that sense, the feedback load is equivalent to the backhaul load as they are applicable for the same resource. Figure A.17 captures this notion for different active set thresholds. The limited feedback (LFB) gives rise to the number of CSI coefficients that are equivalent to the number of precoding weights with limited backhaul (LBH) achieved with the PSO. While it can be observed that the ZF performs poorly with the increase in the backhaul rate due to non-zero precoding weights being generated for the BSs not in the active set. Providing a leverage of 20 dB for the ZF case where the absolute value of the precoding weight is less than 20 dB, it can be treated as a null precoding weight. But even with this leverage the ZF approach still consumes significant backhaul resources.

Figure A.18 shows the CDF of the sum rate when serving 6 UEs, with the objective of the PSO being weighted interference minimization. The sum rate is evaluated as shown in the x-axis label. In the LFB and LBH scenario, the PSO outperforms the ZF by more than 4 bps/Hz/cell. Other combinations of limited/full FB (FFB) and limited/full backhauling (FBH) are also shown in the Figure A.18. It is interesting to note that the PSO based precoder outperforms the ZF when there is FFB. For more interesting discussions on the other curves relevant to other state of art scenarios is captured in [LBS12b]. Figure A.19 shows the CDF of the SINR of any user for the various precoders with the objective of the PSO being sum rate maximization. Figure A.19 captures the notion that sum rate maximization does not benefit the poor SINR users. Hence, sum rate maximization alone is biased towards weak SINR users.

The PSO being used as a tool for precoder design can achieve backhaul load reduction equivalent to the feedback load reduction. As the user data is routed at the CCN based on the non-zero precoding weights, the PSO based precoder also achieves a reduction in the need for the user data to be available at all the BSs. Thereby using the backhaul resources more efficiently compared to the ZF approach.

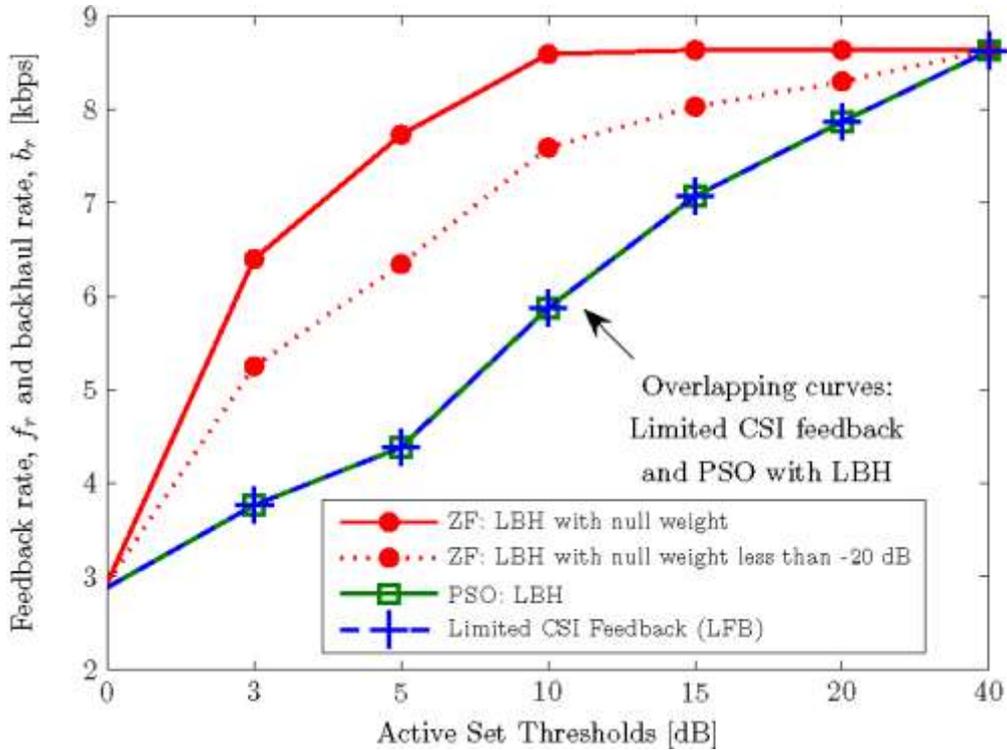


Figure A.17: Feedback rate being equivalent to the backhaul rate

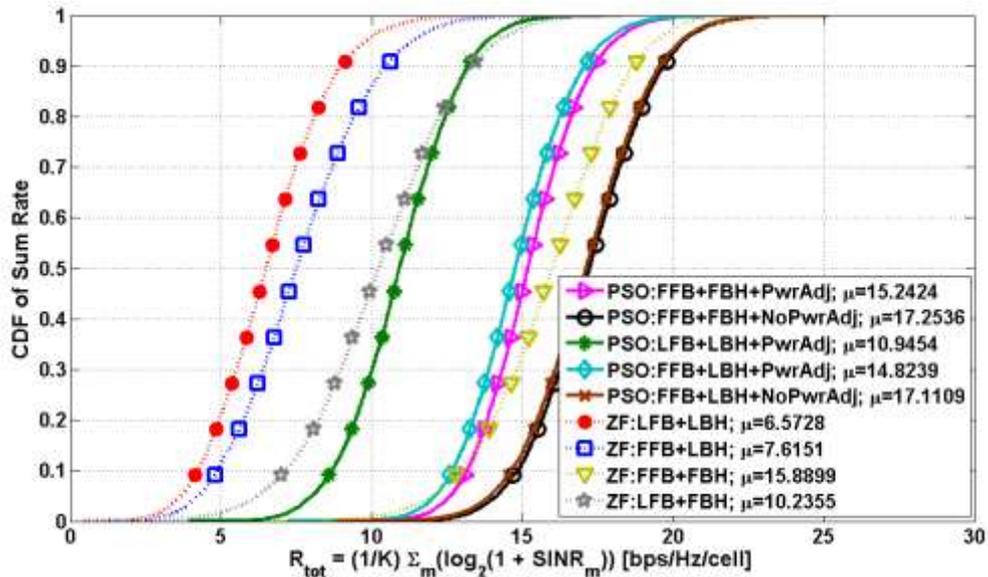


Figure A.18: CDF of sum rate with the PSO objective of Weighted Interference Minimization

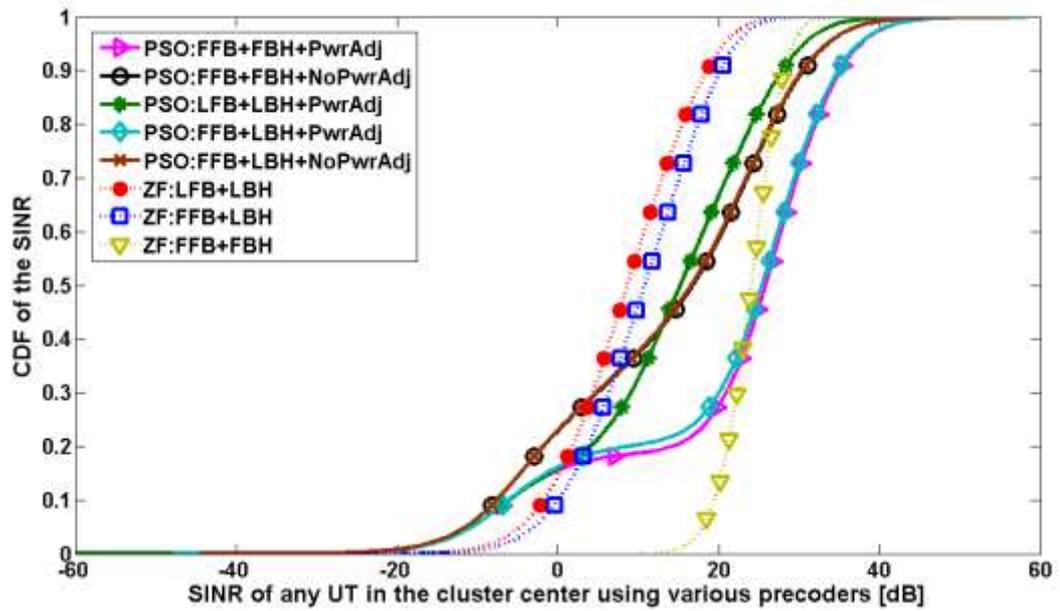


Figure A.19: CDF of sum rate with the PSO objective of sum rate maximization

## A2-6 Resource allocation for OFDMA Joint Processing CoMP

[This Appendix contains additional algorithms and results relevant for the JP-CoMP framework as discussed in Subsection 5.1.1.]

The resource allocation problem with joint processing CoMP has been widely addressed for the systems with a flat-fading channel. Considering zero-forcing or block diagonalization joint processing, it has been shown in [BH06] that the resource allocation problem, with the objective of maximizing the weighted sum rate under per-BS or per-antenna power constraints, is a convex problem, if the total number of transmit antennas is more than or equal to the total number of receive antennas. In a realistic system, the total number of users in the system could be very large such that the above condition does not hold. Therefore, different joint user selection schemes are proposed to determine the selected user group [LSB+11], [ZCA+09]. Orthogonal frequency division multiple access (OFDMA) has been adopted as the downlink access technology for LTE-based systems. With a frequency-selective fading channel, OFDMA adds another dimension in the design of the resource allocation scheme. Thus, the resource allocation schemes proposed based on the assumption of a flat-fading channel may not be suitable for the multiple subchannels case.

We here focus on the downlink resource allocation for OFDMA joint processing CoMP. Two iterative resource allocation algorithms are designed to jointly optimizing the scheduled users and the transmit power across multiple subchannels and multiple cells [LXB+12]. The objective is to maximize the weighted sum rate under per-BS power constraints.

### System model

Consider the downlink of a cluster with  $N$  single-antenna BSs and  $M$  single-antenna users in a multi-cell OFDMA system. The system bandwidth  $B$  is divided into  $K$  subchannels. Assume that the BSs share both data and perfect CSI of all  $M$  users. Joint transmission is provided from all  $N$  BSs to a subset of the  $M$  users  $S(k) \in \{1, \dots, M\}$  using the same subchannel  $k$ . Let  $\mathbf{b}^k \in \mathbb{C}^{|S(k)|}$  denote the complex data symbol vector of the users in set  $S(k)$  on subchannel  $k$ , where  $|S(k)|$  denotes the cardinality of the set  $S(k)$ . The beamforming matrix,  $\mathbf{W}^k \in \mathbb{C}^{N \times |S(k)|}$ , is used to map the user data symbols to the transmit signals, with  $w_{nm}^k$  denoting the beamforming weight for user  $m$  from BS  $n$  on subchannel  $k$ . Assume  $|S(k)| \leq N$ , zero-forcing beamforming is adopted here as the joint process CoMP technique, i.e., the precoding matrix is obtained as the pseudo-inverse of the channel matrix. Thus, the data rate of the scheduled user  $m \in S(k)$  on subchannel  $k$  can be expressed by

$$r_m^k = B/K \times \log_2(1 + p_m^k / \sigma_m^2),$$

where  $B/K$  is the bandwidth of each subchannel.  $p_m^k = E(|b_m^k|^2)$  is the symbol power allocated to user  $m$  from all the  $N$  BSs on subchannel  $k$ .  $\sigma_m^2$  is the sum power of the thermal noise and the uncoordinated out-cluster interference. The transmit power of BS  $n$  on subchannel  $k$  is given by

$$P_n^k = \sum_{m \in S(k)} |w_{nm}^k|^2 p_m^k,$$

Assume each BS has a maximum transmit power constraint as  $P_{\max}$ . Then, we have

$$\sum_{k=1}^K P_n^k \leq P_{\max}, \forall n = \{1, \dots, N\}.$$

In order to maximize the weighted sum rate of the cluster under per-BS power constraint, the coordinated  $N$  BSs need to jointly determine the set of selected users for each subchannel, and the symbol power allocated to each selected user, i.e.,  $S(k)$  and  $p_m^k$  for all  $k$  and  $m$ . Let  $\alpha_m$  denote the weight assigned to user  $m$ . The optimization problem can be formulated as

$$\begin{aligned} \max \quad & \sum_{k=1}^K \sum_{m \in S(k)} \alpha_m r_m^k \\ \text{s.t.} \quad & 1) \sum_{k=1}^K \sum_{m \in S(k)} |w_{nm}^k|^2 p_m^k \leq P_{\max}, \forall n, \\ & 2) p_m^k \geq 0, \forall m, \forall k, \\ & 3) |S(k)| \leq N, \forall k, \\ & 4) S(k) \in \{1, \dots, M\}, \forall k. \end{aligned}$$

In general, the optimization problem is non-convex, since it needs to find the optimal scheduled user set for each subchannel. Specifically, let  $\mathcal{S}$  denote the set of all feasible user sets that satisfy the constraints 3) and 4) per subchannel. Then, we have  $|\mathcal{S}| = \sum_{i=1}^N \frac{M!}{(M-i)!} = O(M^N)$ ,

which is the number of user sets that needs to be considered for each subchannel. With  $K$  subchannels in total, we need to roughly search over  $O(M^{NK})$  user set combinations. Within each user set combination, the optimal power allocation needs to be jointly designed for these scheduled user sets across the  $K$  subchannels. For a CoMP cluster with a large number of BSs and a large number of subchannels, the computational complexity for this exhaustive search is unacceptably high. In the following, we will discuss several sub-optimal algorithms that require lower complexity.

### Identical user set resource allocation

In this approach, resource allocation is designed by restricting an identical user set to be scheduled for all subchannels. Hence, the complexity for selecting the user sets for all subchannels can be reduced to  $O(M^N)$ . For each feasible user set, the power allocation problem becomes a convex problem, since the objective function becomes concave and the remaining constraints 1) and 2) are linear. Hence, the optimal power allocation for each candidate user set subject to per-BS power constraints can be solved via standard convex optimization techniques [TCJ08]. Then, the user set that achieves the highest sum rate will be chosen for joint data transmission.

In order to further reduce the complexity for power assignment design, per-BS power constraints can be relaxed to per-subchannel power constraints [BSX+10]. Thus, the power allocation problem is decoupled into  $K$  independent per-subchannel subproblems.

We refer these two identical user set resource allocation algorithms as IUS&PBSPC and IUS&PSCPC, respectively.

### Iterative joint resource allocation

In this case, the user scheduling and power allocation is jointly optimized over all subchannels. Based on general duality theory [YL06] for non-convex optimization problems in multi-subchannel systems, dual decomposition can be used to solve the resource allocation optimization problem. The dual objective function of the primal problem can be expressed by

$$g(\boldsymbol{\lambda}) = \sum_{k=1}^K g_k(\boldsymbol{\lambda}) + \sum_{n=1}^N \lambda_n P_{\max},$$

where  $g_k(\lambda) = \max_{S(k) \in \mathcal{S}} \{ \sum_{m \in S(k)} \alpha_m r_m^k - \sum_{n=1}^N \sum_{m \in S(k)} \lambda_n |w_{nm}^k|^2 p_m^k \}$ . With a fixed vector of dual variables,  $\lambda$ , the dual objective function is decomposed into  $K$  independent convex subproblems. Hence, the complexity for finding the optimal user sets over all subchannels is reduced to  $O(M^N K)$   $O(M^N \cdot K)$ . For a selected user set  $S(k)$  on subchannel  $k$ , the value of  $p_m^k$  that maximizes  $g_k(\lambda)$  can be derived by  $p_m^k = [\mu_m - \sigma_m^2]^+$ , where  $\mu_m = \alpha_m B / (K \ln 2 \cdot \sum_{n=1}^N \lambda_n |w_{nm}^k|^2)$ . The optimal  $\lambda^*$  that minimizes  $g(\lambda)$  can be found via the subgradient-based dual update method [YL06].

For each iteration, solving per-subchannel optimization problem still requires a complexity of  $O(M^N)$ . A greedy algorithm can be used for finding the optimal user sets for each subproblem, with a complexity in the order of  $O(MN)$ . We refer these two joint resource allocation algorithms as JRA and LC-JRA, respectively.

### Simulation Parameters

The performance in terms of weighted sum rate of the four considered schemes (IUS&PBSPC, IUS&PSCPC, JRA and LC-JRA) is evaluated via Monte-Carlo simulation. We consider the downlink of an OFDMA cellular system, where a cluster of two coordinated BSs simultaneously transmits on multiple subchannels. The users are randomly dropped in the cell-edge area with equal weights, i.e.,  $\alpha_m = 1$  for all  $m$ . Cell-edge SNR, denoted as  $\gamma$ , is defined to be the received SNR at the boundary of the cell, assuming full power transmission  $P_{max}$  from the BS, accounting for only pathloss gain  $PL(R)$  and ignoring shadowing and Raleigh fading. The sum rate is averaged over 1000 independent locations of the users. The main simulation parameters are listed in Table A.7.

The average sum rate performance of the different algorithms is plotted versus different number of users for  $K = 10, 10, 18$   $K = 10 \gamma = 2, 10, 18$  dB respectively. As shown in Figure A.20, the performance of LC-JRA via greedy user selection is very close to the performance of JRA. Compared with the identical user set resource allocation schemes (IUS&PBSPC and IUS&PSCPC), JRA and LC-JRA provide a large gain by jointly optimizing the user selection across multiple subchannels, e.g., the sum rate gain provided are roughly 45% and 20% when cell-edge SNR are 2 and 18 dB, respectively. In addition, it is shown that the performance gain increases as the number of users increases due to multi-user diversity. Hence, we can conclude that joint optimization of user scheduling across multiple subchannels is crucial for the design of OFDMA joint processing CoMP.

Table A.7: Simulation parameters

Simulation Parameter	Value
Transmission bandwidth	$B=1$ MHz
Number of subchannels	$K=10$
Carrier frequency	2 GHz
Radius of cell	$R=500$ m
Minimal distance from UE to eNB	$d_{min}=35$ m
Path Loss Model	$PL(d)=128.1+37.6\log_{10}(d)$ dB, $d$ in km.
Shadowing variance	8 dB
Fast Fading Model	Rayleigh Fading
Noise power	$\sigma_m^2 = -105$ dBmW
Cell-edge SNR	$\gamma = 2, 10, 18$ dB
Number of UEs per cell	$M=2$ to 10
Number of cooperating BSs	$N=2$

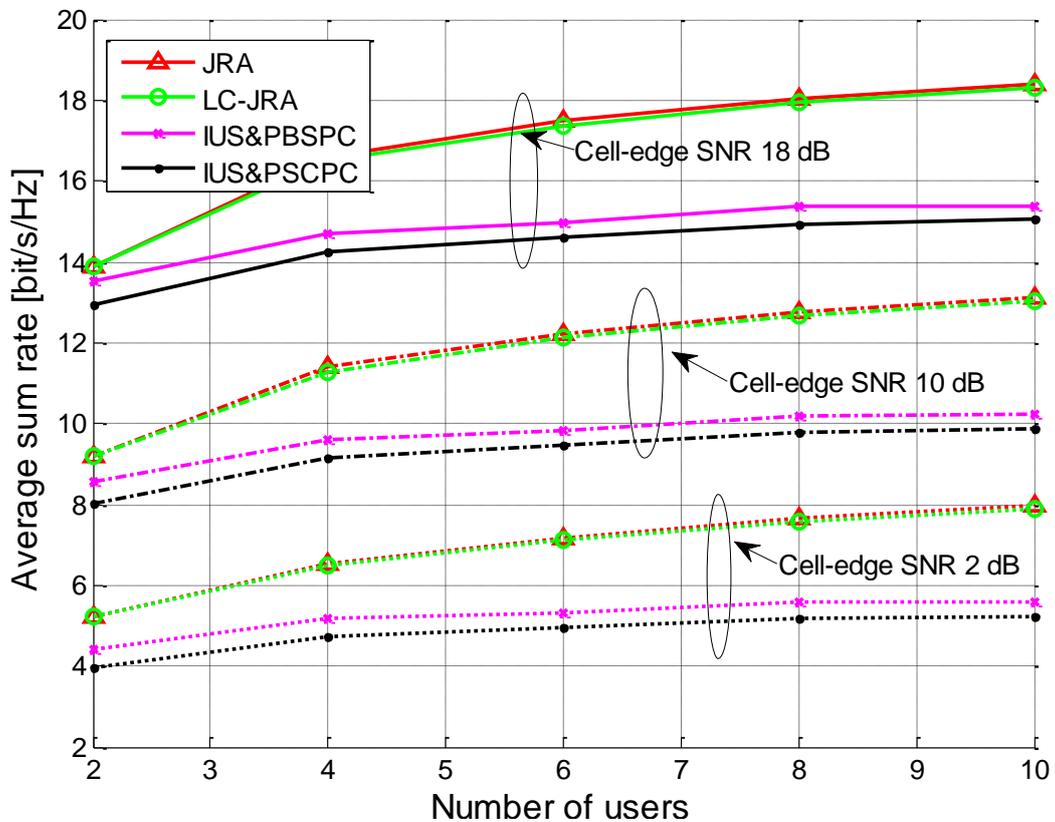


Figure A.20: Average sum rate vs. different number of users in the system

## A2-7 Robust Precoding with Distributed Channel State Information

[This Appendix contains a detailed presentation of results outlined in Subsection 5.3.1]

In Section 5.3.1, we study the joint precoding across two non colocated eNBs, sharing the knowledge of the data symbols to be transmitted, to two UEs, each equipped with one antenna. We consider a distributed CSI configuration where each eNB has its own local estimate of the multi-user channel and no communication is possible between the eNBs. In this distributed CSI configuration, we introduce a concept of distributed MIMO precoding. We focus on the high SNR regime where the two eNBs aim at jointly reducing the level of interference and we show that the conventional Zero Forcing precoder does not perform well when the CSI quality is non-uniform across the eNBs. In contrast, we develop a new precoding scheme which manages to recover the maximal pre-log factor which would be obtained without the CSI inconsistencies. Beyond the established pre-log optimality, simulations show that the proposed precoding schemes achieve better performances at intermediate SNR than known linear precoders.

We study a joint downlink transmission from two eNBs to two UEs using linear precoding and single user decoding. For ease of exposition, the eNBs and the UEs are equipped with only one antenna, such that the received signal is written as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1^H \\ \mathbf{h}_2^H \end{bmatrix} \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

where  $y_i$  is the signal received at the  $i$ -th UE,  $\mathbf{h}_i^H$  of size 2 by 1 is the channel from the eNBs to the  $i$ -th UE. We also denote by  $\tilde{\mathbf{h}}_i$  its normalized version.  $\eta_i$  is the noise at the  $i$ -th UE and is distributed as i.i.d. complex circularly symmetric Gaussian noise. The channel is block fading and its entries are distributed as i.i.d. complex circularly symmetric Gaussian with unit variance to model a Rayleigh fading channel. The vector  $\mathbf{t}_i$  designates the beamforming vector used to transmit symbol  $s_i$  to the  $i$ th UE.

We consider a per-user power constraint  $\|\mathbf{t}_i\|^2 = P/2$ . Note that normalizing the individual columns does not alter the ability to zero-force the interference so that it does not affect the pre-log factor.

We also define formally the pre-log factor at user  $i$  as

$$\text{MG}_i = \lim_{P \rightarrow \infty} \frac{R_i(P)}{\log_2(P)}$$

so that the total pre-log factor is  $\text{MG} = \text{MG}_1 + \text{MG}_2$ . We study the long-term average throughput over the fading distribution and also over the realizations of the random codebooks  $W_i^{(j)}$  used for the random vector quantization of the channel  $\mathbf{h}_i$  at eNB  $j$  (see [KG2011b] for more details). Denoting by  $\bar{i}$  the complementary index of  $i$ , the throughput for UE  $i$  reads as

$$R_i(P) = \mathbb{E}_{\mathbf{h}, \mathbf{w}} \left[ \log_2 \left( 1 + \frac{|\mathbf{h}_i^H \mathbf{t}_i|^2}{1 + |\mathbf{h}_i^H \mathbf{t}_{\bar{i}}|^2} \right) \right]$$

To achieve the maximal pre-log we aim at removing all the interference, i.e. at having

$$I_1(\mathbf{t}_2) = |\mathbf{h}_1^H \mathbf{t}_2|^2 = 0, \text{ and } I_2(\mathbf{t}_1) = |\mathbf{h}_2^H \mathbf{t}_1|^2 = 0.$$

### Distributed Channel State Information

We consider that the eNBs have only access to an imperfect channel estimate resulting from a digital quantization with a finite number of bits. The *distributed* CSI is defined here in the sense that each eNB has a different estimate of the normalized channel  $\tilde{\mathbf{h}}_i$  from all eNBs to UE  $i$ . Moreover, the estimates for  $\tilde{\mathbf{h}}_1$  and  $\tilde{\mathbf{h}}_2$  are also a priori of statistically different qualities. We denote by  $\tilde{\mathbf{h}}_i^{(j)}$  the estimate of the normalized channel vector  $\tilde{\mathbf{h}}_i$  acquired at eNB  $j$ . Furthermore, the number of quantizing bits for  $\tilde{\mathbf{h}}_i^{(j)}$  is given by  $B_i^{(j)}$ .

For the transmission of multiuser single eNB it is shown in [J06] that the number of quantization bits should scale indefinitely with the SNR in order to achieve a positive pre-log factor with ZF. It also holds in a distributed CSI configuration so that we focus on the *scaling in the logarithm of the SNR* of the number of quantization bits

$$\alpha_i^{(j)} = \lim_{P \rightarrow \infty} \frac{B_i^{(j)}}{\log_2(P)}.$$

Since  $\alpha_i^{(j)}=1, \forall i, j \in \{1,2\}$  is shown later in Theorem 1 to be sufficient to achieve the maximal MG, we will consider for the ease of exposition that  $\alpha_i^{(j)} \in [0,1]$  even when not stated explicitly.

### Distributed Precoding

In the distributed CSI setting, each eNB has a different estimate of the channel, which it uses to compute the precoding matrix. We denote the precoder computed at eNB  $j$  as

$$\mathbf{T}^{(j)} = [\mathbf{t}_1^{(j)} \quad \mathbf{t}_2^{(j)}] = \begin{bmatrix} T_{11}^{(j)} & T_{12}^{(j)} \\ T_{21}^{(j)} & T_{22}^{(j)} \end{bmatrix}.$$

**Key remark:** Note that although a given eNB  $j$  may compute the whole precoding matrix  $\mathbf{T}^{(j)}$ , only the  $j$ -th row will be used in practice since the other row corresponds to the coefficients being implemented at the other eNB. Practically, it means that

$$\mathbf{T} = \begin{bmatrix} T_{11}^{(1)} & T_{12}^{(1)} \\ T_{21}^{(2)} & T_{22}^{(2)} \end{bmatrix}.$$

### Main Results

The conventional ZF precoder applied distributively consists in transmitting symbol  $i$  with the beamformer  $\mathbf{t}_i^{\text{ZF}} = [t_{1i}^{\text{ZF}(1)}, t_{2i}^{\text{ZF}(2)}]^T$ , with its elements defined as

$$\mathbf{t}_i^{\text{ZF}(j)} = \begin{bmatrix} t_{1i}^{\text{ZF}(j)} \\ t_{2i}^{\text{ZF}(j)} \end{bmatrix} = \sqrt{\frac{P}{2}} \frac{\Pi_{\tilde{\mathbf{h}}_i^{(j)}}^\perp(\tilde{\mathbf{h}}_i^{(j)})}{\|\Pi_{\tilde{\mathbf{h}}_i^{(j)}}^\perp(\tilde{\mathbf{h}}_i^{(j)})\|}}, \quad j \in \{1,2\}.$$

Intuitively, this means that each eNB applies ZF with its own CSI, implicitly assuming that the other eNB is sharing it.

**Theorem 1:** Conventional ZF achieves the following pre-log factor:

$$M_G^{\text{ZF}} = 2 \min_{i,j \in \{1,2\}} \alpha_i^{(j)}.$$

This result is very pessimistic as the pre-log factor corresponds to the one achieved with the *worst* accuracy across the two eNBs and across the two channel estimates without consideration of the other channel estimates.

To improve over this property, we now propose a scheme called **cooperative Zero Forcing (cZF)** which consists in the precoder whose beamformer to transmit symbol  $i$  is given by

$$\mathbf{t}_i^{\text{cZF}} = \sqrt{\frac{P(1+\rho_i^{(2)})}{2 \log_2(P)}} \mathbf{u}_i^{\text{cZF}} = \sqrt{\frac{P}{2 \log_2(P)}} \begin{bmatrix} \frac{1}{\tilde{h}_{i1}^{(2)}} \\ -\frac{\tilde{h}_{i1}^{(2)}}{\tilde{h}_{i2}^{(2)}} \end{bmatrix}$$

where  $\tilde{\mathbf{h}}_i^{(2)\text{H}} = [\tilde{h}_{i1}^{(2)}, \tilde{h}_{i2}^{(2)}]$ ,

$$\rho_i^{(2)} = |\tilde{h}_{i1}^{(2)}|^2 / |\tilde{h}_{i2}^{(2)}|^2,$$

$\|\mathbf{u}_i^{\text{cZF}}\|^2 = 1$ , and we have assumed wlog that  $\alpha_i^{(2)} \geq \alpha_i^{(1)}$ . We don't elaborate on the insight behind the design of that precoder or on the proof of the result and we refer to the full work [KG2011b] for more details. We simply provide the main result.

**Theorem 2:** Cooperative ZF achieves the maximal pre-log factor of

$$M_G^{\text{cZF}} = \max_{j \in \{1,2\}} \alpha_1^{(j)} + \max_{j \in \{1,2\}} \alpha_2^{(j)}.$$

We show in Theorem 2 that cooperative ZF achieves potentially a much better pre-log factor than conventional ZF. However, this comes at the cost of using only a small share of the available power, as can be observed from the definition of the beamformer. This is clearly inefficient and leads to bad performances at finite SNR. To improve the performances, the eNB with the worst accuracy needs to adapt its power consumption to the channel realizations. In the following, we propose two possible solutions.

- Firstly, eNB 1 can use its local CSI to normalize the beamformer which is then given by

$$\mathbf{t}_i^{\text{cZF}} = \sqrt{\frac{P}{2}} \begin{bmatrix} \frac{1}{\sqrt{1+\rho_i^{(1)}}} \\ \frac{\tilde{h}_{i1}^{(2)}}{\sqrt{1+\rho_i^{(2)}} \tilde{h}_{i2}^{(2)}} \end{bmatrix}$$

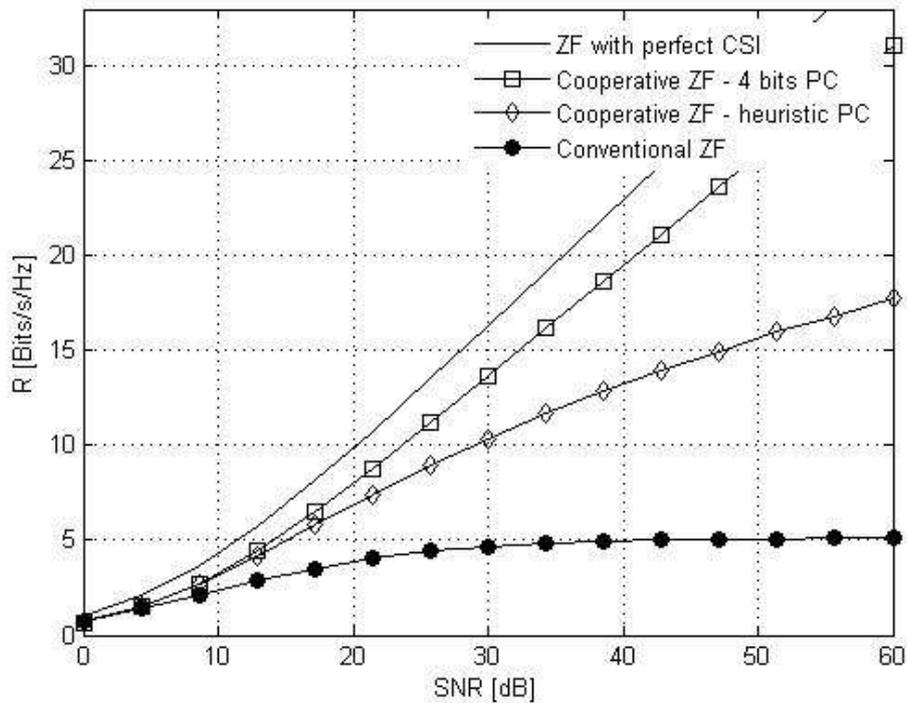
with  $\rho_i^{(j)} = |\tilde{h}_{i1}^{(j)}|^2 / |\tilde{h}_{i2}^{(j)}|^2$ . This beamformer is not pre-log maximizing because the local CSI is used at eNB 1 so that eNB 2 cannot adapt to it to cancel the interference. Thus, the beamformer computed is not orthogonal to  $\tilde{\mathbf{h}}_i^{(2)}$ . Yet, this solution achieves good performance at intermediate SNR.

- Another possibility consists in assuming that eNB 1 receives the scalar  $\rho_i^{(2)}$  (or  $\rho_i$ ) and uses it to control its power. This means that either the UE or eNB 2 needs to feedback this scalar. It requires an additional feedback, but only a few bits are necessary, because it is only used to improve the power efficiency and does not impact the pre-log factor. Thus, the feedback of this scalar does not change the scaling of the CSI in terms of the SNR nor the performances, and appears thus as an interesting practical solution.

## Simulations

In the simulations, we use a Random Vector Quantization scheme and we consider a number of quantizing bits either numerically given or obtained from the CSI scaling as  $q_i^{(j)} = \lfloor \alpha_i^{(j)} \log_2(P) \rfloor$ . We average the performance over 100 codebooks and 1000 channel realizations. Furthermore, we consider the following precoders: ZF with perfect CSI, conventional ZF, and cooperative ZF with heuristic power control and with 4-bits power control.

In Figure A.21, which shows the sum rate over the SNR, we consider the CSI scaling  $[\alpha_1^{(1)}, \alpha_1^{(2)}] = [1, 0.5]$  and  $[\alpha_2^{(1)}, \alpha_2^{(2)}] = [0, 0.7]$ . To emphasize the pre-log (i.e., the slope of the curve in the figure), we let the SNR grow large. As expected theoretically, conventional ZF saturates at high SNR, while cooperative ZF performs close to perfect ZF with a slope only slightly smaller than the optimal one.



**Figure A.21: Sum rate in terms of the SNR with a statistical modeling of the CSI estimation errors using  $[\alpha_1^{(1)}, \alpha_1^{(2)}] = [1, 0.5]$  and  $[\alpha_2^{(1)}, \alpha_2^{(2)}] = [0, 0.7]$ .**

In Figure A.22, we plot the sum rate achieved with the CSI feedback  $[B_1^{(1)}, B_1^{(2)}] = [6, 3]$  and  $[B_2^{(1)}, B_2^{(2)}] = [3, 6]$ . From the theoretical analysis the pre-log factor is equal to zero for all the precoding schemes when a finite (not scaling with the SNR) number of feedback bits is used. This can be observed by the saturation of the sum rate as the SNR grows. Yet, the saturation occurs at higher SNR for cooperative ZF compared to conventional ZF, which leads to an improvement of the sum rate even at intermediate SNR.

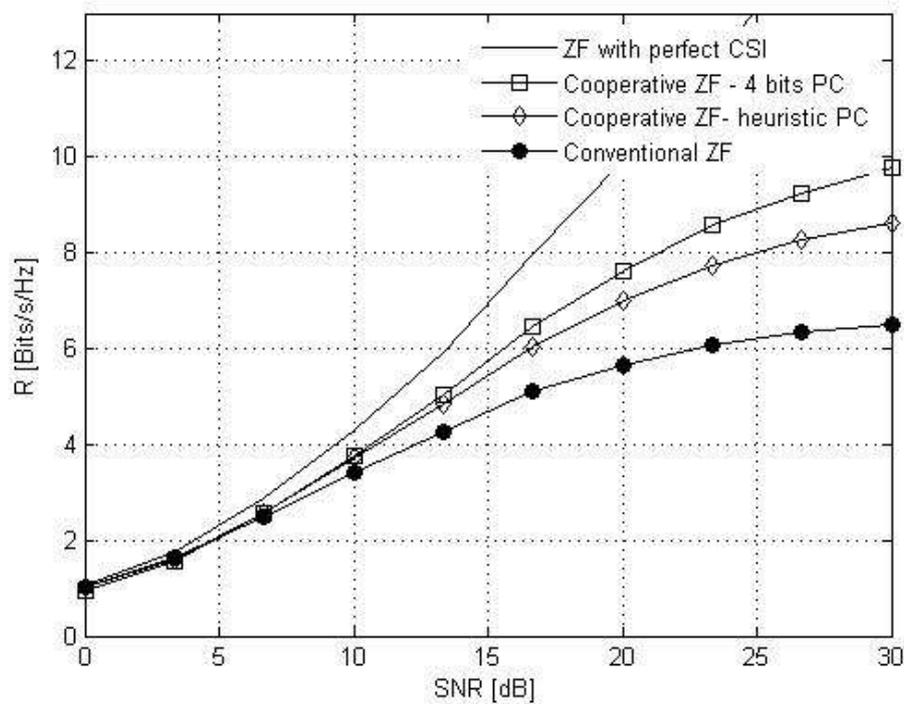


Figure A.22: Sum rate in terms of the SNR using  $[B_1^{(1)}, B_1^{(2)}] = [6, 3]$  and  $[B_2^{(1)}, B_2^{(2)}] = [3, 6]$ .

## Conclusion

In this work, the pre-log factor in a two-cell downlink channel where the eNBs have different estimates of the multi-user channel has been studied. We have shown that conventional Zero Forcing precoding applied without taking into account the differences in CSI quality achieves far from the maximal pre-log factor. We have also derived the value of the maximal pre-log in that distributed CSI configuration and provided a pre-log maximizing precoding scheme. Moreover, we have shown by simulations that the new precoding approach outperforms known linear precoding schemes at intermediate SNR. The extension to multiple cooperating eNBs is done in [KG2011b]. In fact, the impact of the CSI discrepancies between the eNBs becomes even more critical as the number of eNBs increases.

## A2-8 Precoding optimization algorithm for coordinated beamforming

In this section the coordinated beamforming scheme is considered where the beamforming vectors and the downlink transmit powers are jointly designed in each cluster in order to improve the cluster-edge user throughput.

The proposed algorithm exploits the uplink-downlink duality to calculate the transmit powers needed to achieve the SINR targets. We consider the implementation of our algorithm on clustering based solution. The cluster construction could be dynamic or fixed, in this work we assume that the clusters are fixed and are consisted by the same number of cells. We also assume that these clusters are connected by limited capacity backhaul. Only information on the interfered inter-cluster users is allowed to be exchanged between the clusters. The detailed description of this work and performance results are given in the next section:

### System model

We consider the 3GPP scenario1 where each cluster is formed by three cells belonging to the same base station and there are  $N$  cells in the whole network. We consider  $K$  user equipments (UE). Each cell is equipped with  $N_t$  transmit antennas, and  $K$  UEs are selected within each cell to be scheduled. The data symbol available at the  $i$ -th cell intended for the  $j$ -th UE is represented by a complex scalar number  $x_{i,j}$ , precoded by a complex vector  $W_{i,j}$  corresponding to the transmit beamforming vector. The received signal of  $j$ -th UE in the  $i$ -th cell is:

$$y_{i,j} = h_{i,i,j}^H W_{i,j} x_{i,j} + \sum_{l \neq j} h_{i,i,l}^H W_{i,l} x_{i,l} + \sum_{m \neq i} \sum_n h_{m,i,j}^H W_{m,n} x_{m,n} + n_{i,j}$$

where  $h_{i,i,j} \in C^{N_t}$  is the channel between the  $i$ -th cell and the  $j$ -th UE and the  $n_{i,j}$  is the additive white circularly symmetric Gaussian complex noise with variance  $\sigma_{i,j}^2$ .

The optimization problem in the  $n$ -th cluster is stated as following:

$$\min_{\alpha_i, W_{i,j}, \mathbf{h}_{i,c}^l} \sum_{i,j} \alpha_i W_{i,j}^H \left( I + \sum_{c=1}^{N_c} (\mathbf{h}_{i,c}^l)^H \mathbf{h}_{i,c}^l \right) W_{i,j}$$

$$\text{sb SINR}_{i,j} = \frac{|W_{i,j}^H \mathbf{h}_{i,i,j}|^2}{\sum_{k \neq j}^K |W_{i,k}^H \mathbf{h}_{i,i,j}|^2 + \sum_{m \neq i}^N \sum_n^K |W_{m,n}^H \mathbf{h}_{m,i,j}|^2 + \sigma^2} \geq \gamma_{i,j} \quad \forall i, j$$

The term  $\sum_{i,j} \alpha_i \sum_{c=1}^{N_c} (\mathbf{h}_{i,c}^I)^H \mathbf{h}_{i,c}^I$  represents the interference created by the  $i$ th cell

to the other clusters. Instead of only minimizing the transmit power in the considered cell, we minimize at the same time the interference that could be created by this cell to the other UEs in

the other clusters and ensuring the SINR targets  $\gamma_{i,j}$  for the scheduled UEs in the  $i$ th cell. The characterization of the solution is derived using the Lagrangian duality.

The following algorithm aims at finding the optimal precoding vectors and the corresponding power allocation:

### Step 1: Initialization of the Lagrange multipliers

Initialize the UL powers  $\lambda_{i,j}^{(t_0)}$

### Step 2: Update the Lagrange multipliers

According to the following equation find the optimal uplink power allocation  $\lambda_{i,j}$

$$\lambda_{i,j}^{(t+1)} = \frac{1}{\left(1 + \frac{1}{\gamma_{i,j}}\right) \mathbf{h}_{i,i,j}^H \left( \alpha_i I + \alpha_i \sum_{c=1}^{N_c} (\mathbf{h}_{i,c}^I)^H \mathbf{h}_{i,c}^I + \sum_{m,n} \lambda_{m,n}^{(t)} \mathbf{h}_{i,m,n} \mathbf{h}_{i,m,n}^H \right)^{-1} \mathbf{h}_{i,i,j}}$$

Repeat step2 until convergence  $\lambda_{i,j}^{(t)} \approx \lambda_{i,j}^{(t+1)}$

### Step 3: find UL beamforming vectors

It is well known that the optimal receive beamforming vector is given by MMSE receiver:

$$\hat{W}_{i,j} = \left( \alpha_i \sigma^2 I + \alpha_i \sigma^2 \sum_{c=1}^{N_c} (\mathbf{h}_{i,c}^I)^H \mathbf{h}_{i,c}^I + \sum_{m,n} \lambda_{m,n} \sigma^2 \mathbf{h}_{i,m,n} \mathbf{h}_{i,m,n}^H \right)^{-1} \mathbf{h}_{i,i,j}$$

Step1 to step3 necessitates a perfect synchronisation between the coordinated cells, which is implicit in our clustering algorithm. However the algorithm in [DY10] needs a central scheduler for all the cells in the network.

### Step 3: Calculate the DL beamforming vectors

$$W_{i,j} = \sqrt{p_{i,j}} \hat{W}_{i,j}$$

$\sqrt{p_{i,j}}$  is found by replacing the  $\hat{W}_{i,j}$  in the SINR expression satisfying the equality with the constraint  $\gamma_{i,j}$ .

### Simulation parameters

The throughput performance of this algorithm is provided in 4.2.2.1.

All the simulations are run with i.i.d small scale Rayleigh fading channel coefficients. The precoding vectors are computed for each channel realization assuming perfect CSI. The simulations are run for the case of 10 users per cell and 30 per clusters as shown in Table A.8.

**Table A.8: Simulation parameters**

System parameter	Value
Number of eNBs or clusters	21
Nuber of cells	57
UE speed	0-30 km/h
Carrier frequency	2.66 GHz
Distance between eNBs	500 m
Number of transmit antenna per cell $N_t$	4
Number of user per cell	10
SINR targets	15 dB

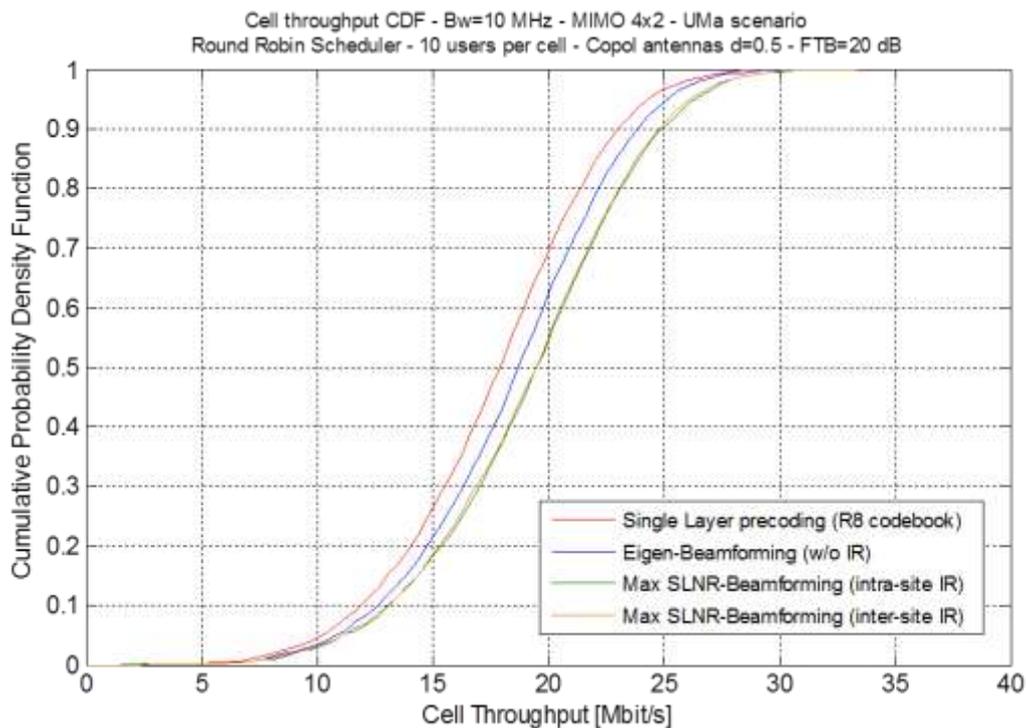
## A2-9 Coordinated beamforming for interference rejection

In this section the performance of the multi cell L1 interference rejection scheme has been evaluated by means of system level computer simulations in order to assess the obtainable enhancements of spectral efficiency related to specific parameterizations of the algorithm, system configuration and specific propagation scenarios. In the simulations  $K=10$  users are dropped in each cell and each user is allocated to one subband formed by  $M=5$  adjacent Physical Resource Blocks (PRBs). The considered system bandwidth is  $B=10$  MHz and thus the number of available PRBs is equal to 50. The base station employs  $n_T$  transmit antennas and each user is equipped with multiple antennas as well. Let  $n_R^{(i)}$  denote the number of receive antennas at the  $i$ -th user. In our simulations  $n_T = 4$  and  $n_R^{(i)} = 2 \quad \forall i$ . A single data layer is transmitted to each UE. In the following Table A.9 the parameterizations and assumptions considered when running the computer system level simulations, have been collected.

**Table A.9: Simulation assumptions.**

Simulation Parameter	Assumption
Coordination scheme	Multi-cell Interference Rejection based on Coordinated Beamforming (CB)
UE CSI reporting	Ideal (channel matrix known for each RB at the transmitter)
Cellular Layout	Hexagonal grid, 19 sites, 3 sectors per site
Simulated Link	Downlink
Deployment scenario	Urban Macrocellular (UMa)
Traffic model	Full buffer
Scheduler	Round Robin (non coordinated among different cells)
Bandwidth	10 MHz
Channel model	Spatial Channel Model (SCM)
Number of antenna elements (BS, UE)	(4, 2)
Antenna separation (BS, UE)	$(\lambda/2, \lambda/2)$
Link to system interface	MIESM
HARQ	Embedded into link level performance curves (Max. 4 transmission)
UE channel Estimation	Embedded into link level performance curves
Number of PDCCH control symbols (L)	2

The cumulative distribution function of the cell throughput obtained by simulation is shown in Figure A.23 for both the intra-site and inter-site application of the interference rejection scheme. In the inter-site configuration the interfering cells can place their null in the antenna radiation diagram also towards users that are served by a different eNode B (i.e. site). The simulated scenario is Urban Macrocellular (UMa) and the antenna setup employed at the eNode B and the UE is formed by co-polarized antennas with half wavelength spacing ( $d=\lambda/2$ ). The front-to-back (FTB) ratio of the eNode B antennas is set to 20 dB. The threshold  $T$  for the interference rejection mechanism is set to 6 dB. For comparison in the Figure is also reported the performance of a LTE non-coordinated network that employs the Release 8 Closed Loop Single Layer precoding (LTE Transmission Mode 6) and the Release 8 Single Layer Beamforming based the eigen-beamforming technique (LTE Transmission Mode 7).



**Figure A.23: Cell throughput distribution (UMa environment)**

The analysis of the results shows that the Intra-site IR scheme provides a gain of about 9% in terms of median cell throughput over Single Layer Precoding and of about 4% over the Eigen-Beamforming. The inter-site IR scheme does not provide significant gains over intra-site IR. This result can be explained as follows: the application of the inter-site IR with the same threshold ( $T=6$  dB) leads to an increase of the users for which the IR is activated (from 10% to about 28%). However, the insertion of nulls in the radiation diagram causes a distortion of the main antenna lobe with a reduction of the throughput performance especially for the users that are in good channel conditions. The Figure A.24 shows the percentage and the spatial distribution of the users that require the activation of the IR scheme in the case of the intra-site application of the technique. In black color are indicated the users that do not request the activation of the IR scheme, while in blue and red color are indicated the users for which the IR is activated by 1 or 2 interfering BS respectively.

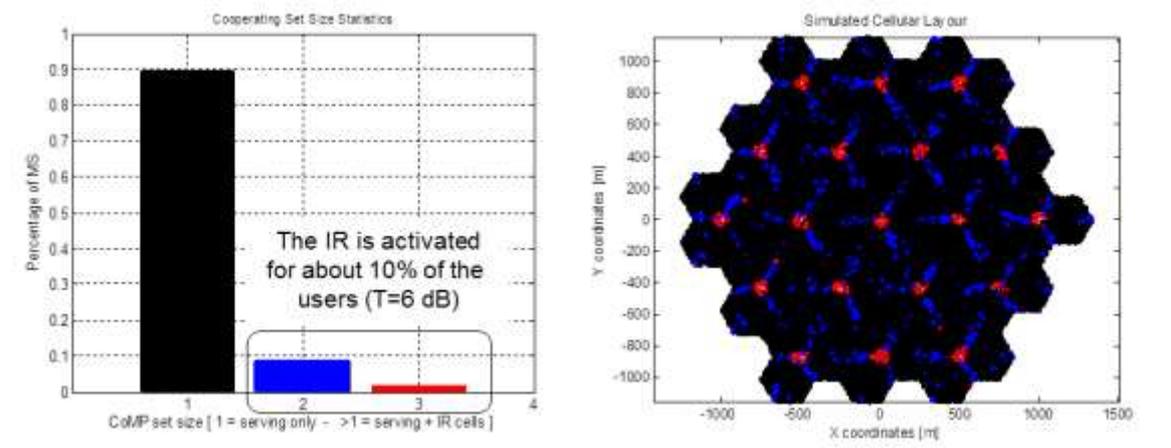


Figure A.24: Percentage (left) and spatial distribution (right) of the users for which the IR is activated (intra-site scheme)

The results shows that the Intra-site IR scheme is activated for about 10% of the users when using a radio quality threshold T of 6 dB. Among these users about 8% is located on the edges among adjacent sectors of the same eNode B (blue color) and for them the IR is activated only by one interfering cell. For the remaining 2% of the users the scheme is activated simultaneously by the other two sectors (different than the serving one) of the same eNode B and these users are mainly located near to the eNode B (red color). The interference for these users can be related to the low antenna Front To Back (FTB). The sensitivity to this parameter is investigated in Figure A.25 where an FTB value of 20 dB (left), 25 dB (center) and 30 dB (right) has been used in the simulations

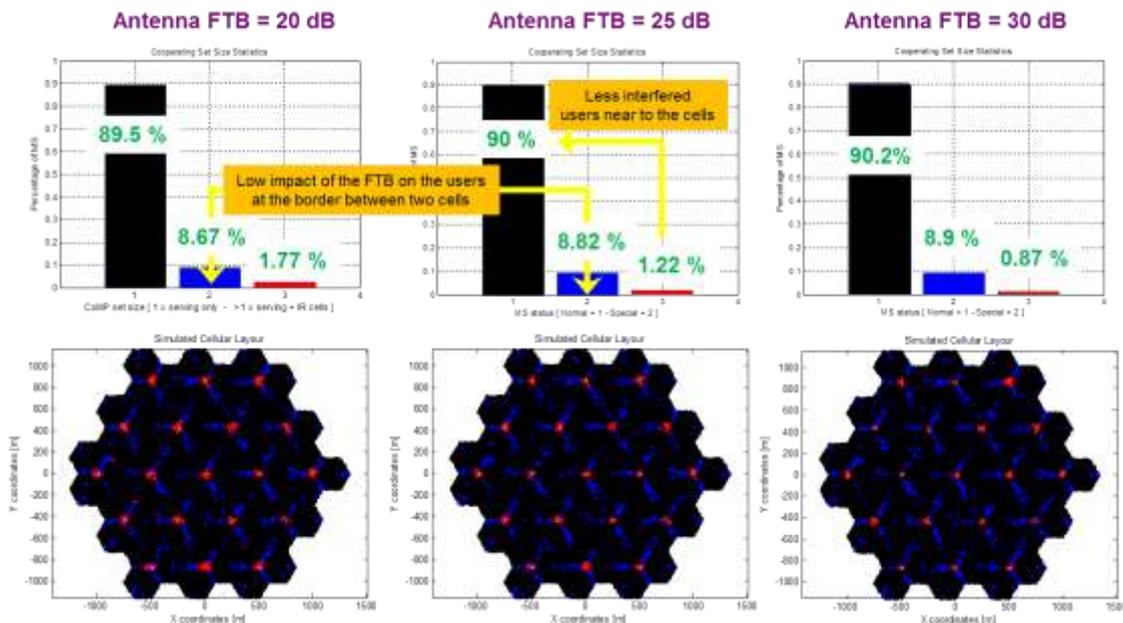


Figure A.25: Sensitivity of the eNode B antennas Front to Back on the percentage and spatial distribution of the users for which the intra-site IR scheme is activated

The results shows that the increment of the FTB has a limited impact on the percentage of interfered users located at the edge of two sectors of the same eNode B while it has a positive effect on the interfered users located near to the eNode B, for which the percentage is nearly halved by increasing the FTB from 20 dB to 30 dB.

The results provided in Figure A.23 are obtained for a fully loaded system with Round Robin (RR) resource allocation non coordinated among the cells of the same site. In this situation the scheduler has not degrees of freedom for mitigating the interference coming from the other two cells of the same site. In Figure A.26 it is analyzed the performance of the proposed Layer 1 IR scheme in conjunction with a simple intra-site coordinated Layer 2 Resource Allocation (RA) algorithm. Basically the coordinated RA works by starting the allocation of the PRBs in the sector of the same site by introducing a suitable offset that allows to obtain a frequency reuse factor of three in low load conditions. The expected effect is that, in low load condition, the intra-site interference is mitigated at L2 while the inter-site interference (for cell edge users) is managed at L1 by the IR scheme.

The Figure A.26 shows the user throughput distribution in case of an average cell load of 20% (i.e. 20% of the PRBs are used on average in each sector). The left part of the Figure shows the case without coordination of the Resource Allocation among the sectors of the same site (i.e. the starting point for the PRB allocation is the same for the three sectors), while on the right the coordinated RA is used by introducing a suitable PRB displacement.

The analysis of the simulation results leads to the following conclusions about the proposed IR scheme. In a fully loaded condition the intra-site IR rejection scheme provides a gain of about 9% for the average cell throughput and a gain of 14% for the cell edge user throughput with respect to the Transmission Mode 6 of LTE R8 (single layer precoding based on codebook), under the ideal CSI assumption. In a fully loaded condition the inter-site IR scheme does not provide significant gains over intra-site IR to justify the related overhead on the backhauling. The simulations also show that a significant part of the users is affected by the interference from the cells of the same site (about 10% by setting a threshold  $T$  of 6 dB and an antenna FTB of 20 dB). The effect of the intra-site interference can be partially mitigated by improving the antenna FTB and implementing intra-site coordinated scheduling algorithms. The challenge in the design of intra-site coordinated scheduling algorithms is to find the best trade-off between orthogonalization of the users under different cells of the same site and the unavoidable reduction of the frequency diversity. The simulations show that with a simple static Fractional Frequency Reuse (FFR) intra-site resource allocation algorithm and in a fractional load condition (load 20%) it is possible to achieve a cell edge user throughput gain of about 41% by using the inter-site L1 Interference Rejection scheme. It then follows that a cross-layer joint L1 & L2 design is crucial for the maximization of the system performance.

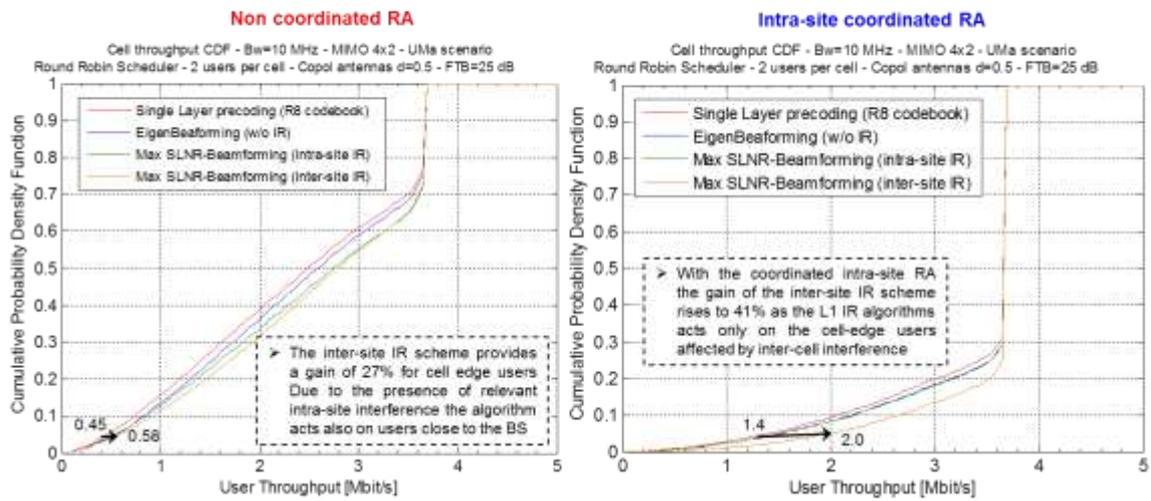


Figure A.26: Combination of the inter-site IR scheme operating at physical layer level with a simple coordinated layer 2 resource allocation algorithm among the sectors of the same site (20% load – FTB = 25 dB)

## A2-10 A Practical Iterative Algorithm for Joint Signal and Interference Alignment in Heterogenous Networks

### Operation of the Proposed Algorithm

The proposed distributed algorithm exploits a duality result from [YXL09] which states that for the MIMO IC, if a set target SINRs is achievable in the downlink (DL) direction (under a sum-power constraint), these target SINRs will be achievable in the reverse direction under the same power constraint. The aim of the design is to ensure that all the streams of all the links can reach a minimum target SINR  $\gamma_T$  already chosen.

Considering the system model, the SINR  $\Gamma_k^l$  of stream  $l$  of user  $k$  is given by

$$\Gamma_k^l = \frac{\frac{P}{d_k} \mathbf{U}_k^{[l*]} \times \mathbf{H}_{kk} \mathbf{V}_k^{[*l]} \mathbf{V}_k^{[*l]H} \mathbf{H}_{kk}^H \mathbf{U}_k^{[l*]H}}{\sigma^2 + I_{kl}^{InterStream} + I_{kl}^{InterLink}},$$

where  $\mathbf{A}^{[*l]}$  and  $\mathbf{A}^{[l*]}$  are used to denote the  $l^{\text{th}}$  column and the  $l^{\text{th}}$  row of matrix  $\mathbf{A}$ , respectively and  $(\cdot)^H$  is the conjugate transpose of a given matrix or vector. Finally, we have

$$I_{kl}^{InterStream} = \frac{P}{d_k} \mathbf{U}_k^{[l*]} \left( \sum_{l' \neq l} \mathbf{H}_{kk} \mathbf{V}_k^{[*l']} \mathbf{V}_k^{[*l']H} \mathbf{H}_{kk}^H \right) \mathbf{U}_k^{[l*]H},$$

$$I_{kl}^{InterLink} = \mathbf{U}_k^{[l*]} \left( \sum_{k' \neq k} \sum_{l'} \frac{P}{d_{k'}} \mathbf{H}_{kk'} \mathbf{V}_{k'}^{[*l']} \mathbf{V}_{k'}^{[*l']H} \mathbf{H}_{kk'}^H \right) \mathbf{U}_k^{[l*]H}.$$

Based on the SINR expression of every stream of a given link, in order to allow each stream of each user reach the target SINR  $\gamma_T$ , we must have

$$\left\| \left( \frac{1}{\gamma_T} + 1 \right) \mathbf{U}_k^{[l*]} \mathbf{H}_{kk} \mathbf{V}_k^{[*l]} \right\|_F^2 - \sum_{k'} \sum_{l'} \left\| \mathbf{U}_k^{[l*]} \mathbf{H}_{kk'} \mathbf{V}_{k'}^{[*l']} \right\|_F^2 \geq \frac{d_k \sigma^2}{P}$$

Based on the previous inequality, and for a given target SINR  $\gamma_T$ , at every iteration the de-correlation matrices are computed so that we have

$$\mathbf{U}_k^{[l*]} = \arg \max_{\tilde{\mathbf{U}}_k} \left\| \left( \frac{1}{\gamma_T} + 1 \right) \frac{P}{d_k} \tilde{\mathbf{U}}_k^{[l*]} \mathbf{H}_{kk} \mathbf{V}_k^{[*l]} \right\|_F^2 - \sum_{k'} \sum_{l'} \left\| \frac{P}{d_{k'}} \tilde{\mathbf{U}}_k^{[l*]} \mathbf{H}_{kk'} \mathbf{V}_{k'}^{[*l']} \right\|_F^2$$

Given fixed values of  $\mathbf{U}_k$ , at each iteration, the matrices  $\mathbf{V}_k$  must verify

$$\mathbf{V}_k^{[*l]} = \arg \max_{\tilde{\mathbf{V}}_k} \left\| \left( \frac{1}{\gamma_T} + 1 \right) \frac{P}{d_k} \tilde{\mathbf{V}}_k^{[*l]} \mathbf{H}_{kk}^H \mathbf{U}_k^{[l*]H} \right\|_F^2 - \sum_{k'} \sum_{l'} \left\| \frac{P}{d_{k'}} \tilde{\mathbf{V}}_k^{[*l]} \mathbf{H}_{k'k}^H \mathbf{U}_{k'}^{[l*]H} \right\|_F^2$$

Based on the previous analysis, the algorithm operates as shown below in Algorithm 1.

### Algorithm 1: Joint Signal and Interference Alignment

- 1: Start with arbitrarily pre-coding vectors  $\mathbf{V}_1, \dots, \mathbf{V}_K$  such as  $\mathbf{V}_k^H \mathbf{V}_k = \mathbf{I} \quad \forall k$
- 2: For each stream  $l$  of receiver  $k$ , compute the following matrix

$$\mathbf{Q}_{kl} = \left( \frac{1}{\gamma_T} + 1 \right) \mathbf{H}_{kk} \mathbf{V}_k^{[*l]} \mathbf{V}_k^{[*l]H} \mathbf{H}_{kk}^H - \sum_{k'} \sum_{l'} \mathbf{H}_{kk'} \mathbf{V}_{k'}^{[*l']} \mathbf{V}_{k'}^{[*l']H} \mathbf{H}_{kk'}^H$$

- 3: For each stream  $l$  of receiver  $k$ , compute the de-correlation vector  $\mathbf{U}_k^{[*l]}$  which corresponds to the  $l^{\text{th}}$  row of matrix  $\mathbf{U}_k$  and is given by

$$\mathbf{U}_k^{[*l]H} = \mathbf{v}_{\max}^1(\mathbf{Q}_{kl}),$$

with  $\mathbf{v}_{\max}^1(\cdot)$  is the eigenvector corresponding to the maximum eigen value of a given matrix

- 4: Reverse the communication direction and set  $\tilde{\mathbf{V}}_k = \mathbf{U}_k^H$
- 5: For each stream  $l$  of new receiver  $k$ , compute the following matrix

$$\tilde{\mathbf{Q}}_{kl} = \left( \frac{1}{\gamma_T} + 1 \right) \mathbf{H}_{kk}^H \tilde{\mathbf{V}}_k^{[*l]} \tilde{\mathbf{V}}_k^{[*l]H} \mathbf{H}_{kk} - \sum_{k'} \sum_{l'} \mathbf{H}_{k'k}^H \tilde{\mathbf{V}}_{k'}^{[*l']} \tilde{\mathbf{V}}_{k'}^{[*l']H} \mathbf{H}_{k'k}$$

- 6: For each stream  $l$  of new receiver  $k$ , compute the de-correlation vector  $\tilde{\mathbf{U}}_k^{[*l]}$  which corresponds to the  $l^{\text{th}}$  row of matrix  $\tilde{\mathbf{U}}_k$  and is given by

$$\tilde{\mathbf{U}}_k^{[*l]H} = \mathbf{v}_{\max}^1(\tilde{\mathbf{Q}}_{kl})$$

- 7: Reverse the communication direction and set  $\mathbf{V}_k = \tilde{\mathbf{U}}_k^H$
- 8: Repeat the process until the convergence of the algorithm

### Simulation Assumptions

All the simulations are run with i.i.d small scale Rayleigh fading channel coefficients for every link. The pre-coding and de-correlation matrices are computed for each channel realization assuming perfect CSI. The simulations are run for the case of the 3-user  $2 \times 2$  MIMO IC. Therefore, each AP has 2 antennas and is paired with a user having 2 receive antennas leading to a total of three  $2 \times 2$  MIMO links in the network.

### A3. Advanced 3D Beamforming

#### A3-1 UE-specific horizontal and vertical beamsteering

The basic behaviour, the capabilities and potential benefits of the UE-specific horizontal and vertical beamsteering have been investigated based on the downlink of a 3GPP LTE system. The main parameters are summarized in Table A.10.

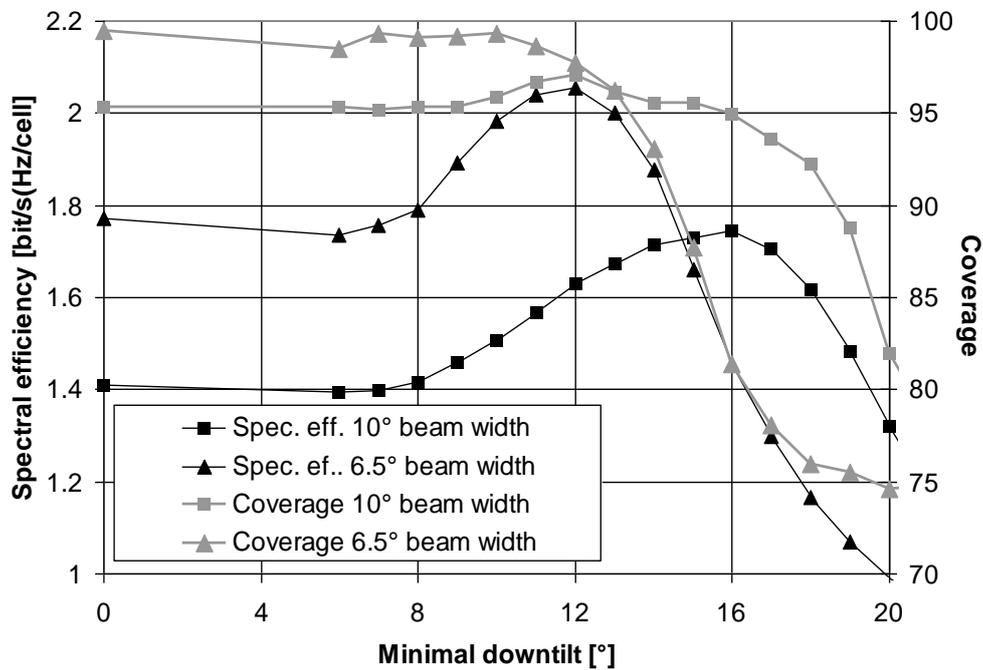
**Table A.10: Main system parameters**

Parameter	Value
Number of sites and cells	- 19 sites with three 120° cells each - Single cell
Inter site distance	500 m
Antenna height	Base station: 32 m; Terminal: 1.5 m
Base station antenna pattern	Shape according to 3GPP TR 36.814 [11] Horizontal HPBW 70° Vertical HPBW 4.7°, 6.5° and 10°
Propagation model	Line-of-sight
Terminals per cell	10 randomly distributed
Frequency resources	10 MHz bandwidth; Re-use 1
Scheduling algorithm	Round robin

As realization options for vertical beam steering the case of two fixed downtilts (DT) has been compared to the case of exact main lobe steering to the terminal (also called synonymously “direct steering” in the following), with and without limitation of the smallest DT, and fixed DT as a baseline.

#### **Multi-cell scenario**

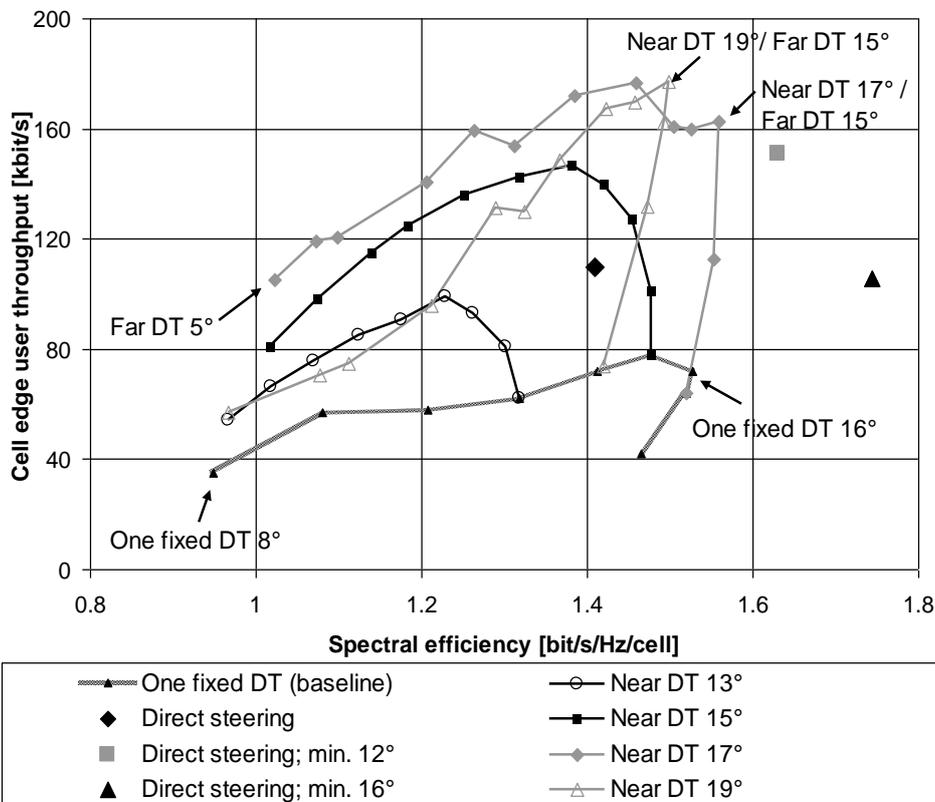
The most significant parameters influencing the overall performance are the vertical half-power beamwidth (HPBW) of the antenna, the inter-site distance (ISD) and the eNB antenna height. The eNB antenna height and ISD impact directly results from the geometry, since it defines the available range of vertical angles between the eNB and the UE. More detailed investigations can be found in [ARTD12]. The example presented here focuses on a classical macro scenario as described in Table A.10. The impact of the vertical HPBW is shown in the Figure A.27 below. Spectral efficiency and coverage were optimized by varying the minimal possible DT that can be applied in the “direct steering” operation mode.



**Figure A.27: Spectral efficiency and coverage for direct main lobe steering depending on the HPBW and the lowest possible DT**

Both coverage and spectral efficiency can be improved with smaller HPBW, but the spectral efficiency is also more sensitive to the optimum minimum DT selection. Obviously, a smaller HPBW allows steering the beam closer to cell edge (smaller minimal DT) without causing too much interference in the adjacent cell, a setting, which generally also improves the coverage. The outcome of this first study is that a small HPBW is preferable, assuming that the location of the UE can be estimated with sufficient accuracy.

The cell edge user throughput versus the spectral efficiency has been investigated for 10° HPBW and different realization options. The curves shown in Figure A.28 represent different combinations of near DT and far DT of the two-fixed-DT realization option. For comparison, the points related to exact main lobe steering with different DT limitation angles are shown, and also the curve for one fixed DT (near DT = far DT).



**Figure A.28: Cell edge user throughput over spectral efficiency for different operation modes and configurations**

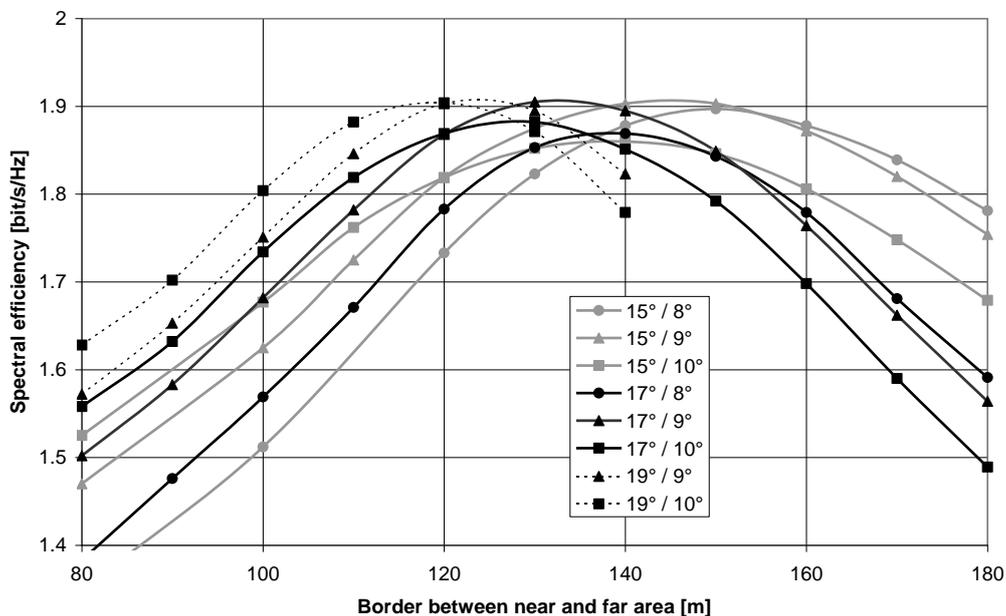
Whereas the cell edge throughput for one fixed DT remains low, even if spectral efficiency increases, it is obvious that with an appropriate combination of near and far DT significant improvements are possible. The combinations of 19° near DT / 15° far DT and 17° near DT / 15° far DT give almost similar performance as the case of exact main lobe steering with a minimum DT limitation to 12°. Exact main lobe steering with 16° DT limitation further increases spectral efficiency, but only at lower cell edge throughput, since this limitation to too steep DTs leads to coverage holes at the cell edge. On the other side, too flat DTs, e.g. the case of exact steering without limitation, creates too high inter-cell interference when serving cell edge users. So the spectral efficiency as well as cell edge throughput remains far behind optimum. The important message of this investigation is that both spectral efficiency and cell edge UE throughput can be improved simultaneously if the system parameters are selected properly.

### Single cell scenario:

Operators wish to keep their site locations even if they extend their existing hardware infrastructure with an additional communication system, e.g. LTE, which utilizes other frequency bands. The cellular network optimized for one carrier frequency (in our example 2.1 GHz) may be subject to performance degradation due to insufficient coverage if a higher carrier frequency (in our example 2.6 GHz) is applied. It has been investigated whether dynamic vertical beam steering is capable to compensate the additional path loss induced by the higher frequency, which is approximately 2 dB. For this, we have used the simulation set-up as described in Table A.10, but with only one single cell, i.e. without interference. The baseline system operates at 2.1 GHz with one optimized but fixed DT of 10°. The transmit power at the eNB was selected such that it was barely sufficient to cover the entire cell, i.e. at cell edge the minimal possible SNR for a successful transmission was perceived. This value is 10 dBm.

This baseline case has been compared with a system operating at 2.6 GHz, but having the ability to apply optimized near- and far DT. Cell size (ISD = 500m), vertical HPBW (10°) and transmit power remain unchanged. Our approach with the near- and far DT provides the additional option to adapt the ratio of the underlying near- and far area in the cell, i.e. the border between both regions can vary. This means that three parameters have to be optimized: near- and far DT as well as the distance of the border between near and far area from the eNB.

The main outcome of this optimization is that the achievable maximal spectral efficiency does not depend on the distance on the border between both areas. This fact can be clearly seen in Figure A.29. However, for each border a different combination of near- and far DT provides this optimal value for the spectral efficiency. Overall, the combination near DT = 17° (serving 23% of the cell) and far DT = 9° (serving 77% of the cell) has been found out to be the most appropriate setting.



**Figure A.29: Spectral efficiency depending on the applied two fixed DTs and the distance of the border between near- and far area from the base station**

Finally, the performance of both systems (baseline case at 2.1 GHz and improved system operating with two fixed DTs at 2.6 GHz) with individually optimized parameters has been compared. The key results are summarized in Table A.11. A more detailed evaluation can be found in [SH11].

**Table A.11: Performance comparison of two different radio systems operating at different carrier frequencies with individually optimized parameters**

Performance metric	Baseline system 2.1 GHz	Vertical beam steering 2.6 GHz
Spectral efficiency	1.86 bit/s/Hz	1.91 bit/s/Hz
Cell edge user throughput	437 kbit/s	399 kbit/s

In spite of the increased path loss, the spectral efficiency can be slightly improved with vertical beam steering. The reason is the significantly higher receive power in the near area due to the steep near DT 17°. The much more important fact is that the cell edge user throughput can be almost conserved. This answers the initial question of this study: Vertical beam steering is capable to compensate a 2 dB path loss increase. This result implies that existing site locations can be reused without a costly new radio network planning and without acquisition of additional site locations.

#### **Verification of vertical antenna pattern in real propagation environment provided by WP6**

For first tests in real LOS/NLOS propagation environment a special antenna array with 4 single columns has been used. Each column can be tilted separately. Each single antenna element is fed with signals with dedicated and distinguishable pilots from an eNB emulator. With this arrangement we could simultaneously measure up to four data streams with a dedicated radio network analyzer during drive tests. The main outcome is summarized in Figure A.30. For small Rx SNR values, related to far distant locations from the eNB, the best performance is achieved with low DT angles (red and green curve), whereas for large Rx SNR values, related to locations close to the eNB, best performance is achieved with larger DT angles (yellow and blue curve). This indicates that the basic behaviour assumed in our system simulations can be observed in practical propagation conditions. However, the parameters of the measurement location differ from the simulation assumptions, in particular the measurements were done in a single-cell scenario without interference. Therefore a direct comparison is not possible at this stage. More detailed work of WP6 on this topic is described in [ARTD61, ARTD62, HSK+12].

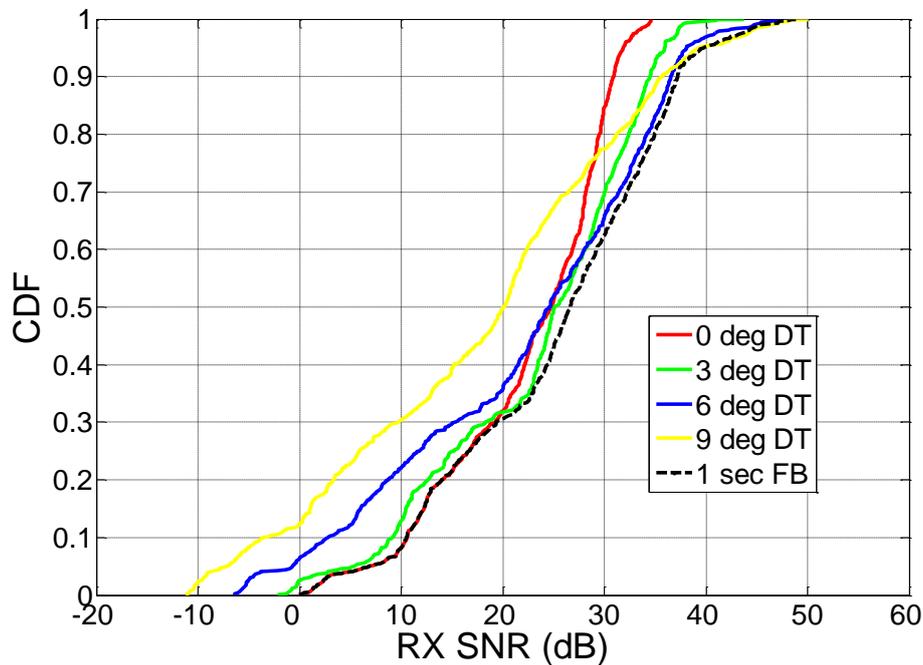
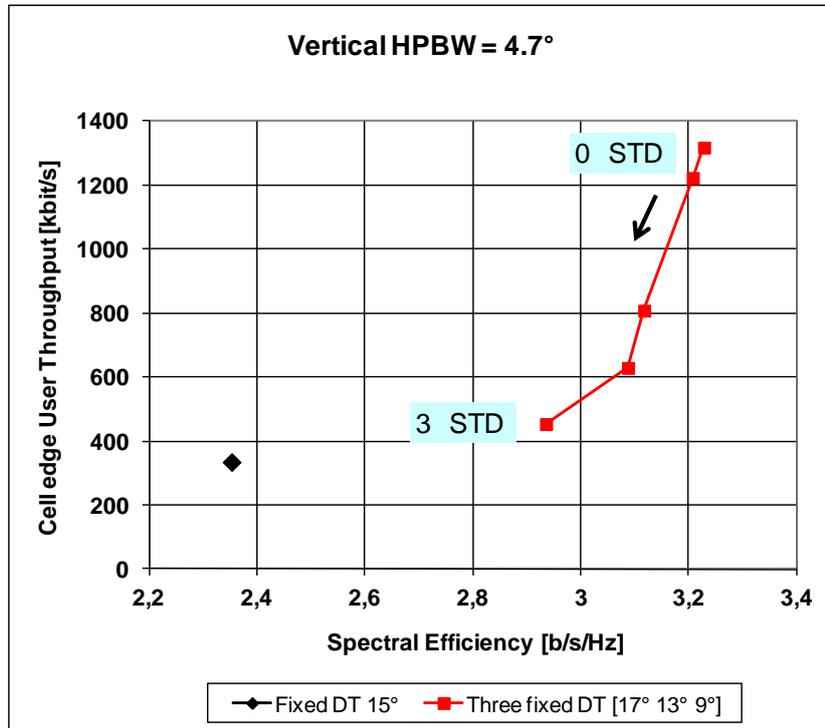


Figure A.30: CDF of Rx SNR measured during drive tests in real propagation environment

#### Impact of UE location errors:

Whereas the simulation studies described earlier in this section have been carried out with the assumption of ideal knowledge of the UE location at the eNB, this paragraph investigates the impact of UE location estimation errors. Our simulation model calculates the estimate of the elevation angle for each UE as the sum of the actual elevation angle and a normally distributed random number with zero mean and arbitrary but fixed variance. The selection of the DT to be applied for a specific UE at the eNB is now based on this estimate and not anymore on the actual angle. Consequently, it can happen that a UE located in the near area is accidentally served with the far DT and vice versa.

The impact of the UE location error has been investigated for the case of three fixed DTs ( $17^\circ$  for the near area,  $13^\circ$  for the central area, and  $9^\circ$  for the far area). The vertical HPBW has been set to  $4.7^\circ$ . Figure A.31 shows the achieved cell edge user throughput over the spectral efficiency for increasing standard deviation of the UE location error from  $0^\circ$  to  $3^\circ$ . The performance for one fixed DT is shown as baseline.



**Figure A.31: Cell edge user throughput over spectral efficiency depending on the standard deviation of the UE location error**

Obviously, spectral efficiency is fairly robust against location estimation errors. The reason is the significant dependence of the overall spectral efficiency on the achieved high data rates in the near sector. As the near sector covers by far the biggest range of possible elevation angles ( $17^\circ < \varphi < 90^\circ$ ), the number of near UEs served with a wrong beam is rather small. On the other hand, cell edge user throughput drops relatively fast down to the performance level of one fixed DT. This metric strongly depends on the data rates of UEs located in the far area, which covers a much smaller range ( $6^\circ < \varphi < 9^\circ$ ), and according to this the probability for being served with an erroneous beam is big.  $6^\circ$  is the known elevation angle of the cell border, i.e. angles below  $6^\circ$  cannot occur.

The conclusion is that the estimation of the elevation angle of the UEs at the eNB must be quite accurate to achieve the theoretically possible gains in cell edge user throughput, in particular if the vertical HPBW is small.

### A3-2 Distributed scheduling for beam coordination

#### Implicit Coordination:

This approach for horizontal and vertical beam coordination works without explicit exchange of scheduling information between eNBs. The basic idea is a mapping of UE locations to radio resources, e.g. frequency sub-bands, as described in 4.2.3.2. Implicit coordination can be combined with UE specific horizontal and vertical beam steering. Both DT selection and implicit coordination on their own can reduce the interference. But also the superposition of both effects is possible and promises additional gains. This has been investigated in a system simulation study. The main system parameters are summarized in Table A.12.

**Table A.12: Main system parameters**

Parameter	Value
Number of sites and cells	19 sites with 3 cells each
Inter site distance	500 m
Antenna height	eNB: 32 m; Terminal: 1.5 m
eNB antenna pattern	4-element uniform linear array, half-wavelength spacing (horizontal) vertical HPBW = 10°
Propagation model	Line-of-Sight (LOS)
Terminals per cell	12
Frequency resources	10 MHz bandwidth; Re-use 1
Scheduling algorithm	Round Robin (RR)

Figure A.32 shows the key performance metrics spectral efficiency and cell edge user throughput. The colors correspond to the different implementation options of 3D beamforming. The squares represent the different implementations of vertical beam steering without any coordination, the triangles correspond to the vertical sorting, the circles to horizontal sorting. So each point in the diagram represents a certain combination of UE specific beam steering schemes and implicit coordination schemes.

A first general observation is that gains achieved from UE specific beam steering (indicated as “3D-BF gain” in Figure A.32) and implicit coordination are additive. Another important outcome is that horizontal sorting is superior to vertical sorting. The reason is the reduced capability for beam separation in the vertical plane. Although the vertical HPBW is normally significantly smaller than the horizontal HPBW, the beams overlap due to the small range of applicable tilt values. With the parameters in Table A.12, 95% of the area is served with DTs between 6° and 20° [KHS+12].

The highest value for *spectral efficiency* is achieved with horizontal sorting combined with exact steering of the beam (labeled as direct steering in Figure A.32), where the minimum DT that can be applied is 15°. This operation mode yields 12% coordination gain on top of 4% 3D-BF gain for spectral efficiency and 42% coordination gain for cell edge throughput. Three fixed DTs is the better solution if the *cell edge user throughput* shall be maximized (33% coordination gain over 31% 3D-BF gain). Obviously, both key performance metrics can be improved simultaneously. The reason for this behavior is that exact beam steering maximizes the desired user signal, which consequently boosts the spectral efficiency for UEs close to the eNB, whereas

the “three fixed DT” case is primarily capable to reduce interference, which is beneficial for the throughput of cell edge users.

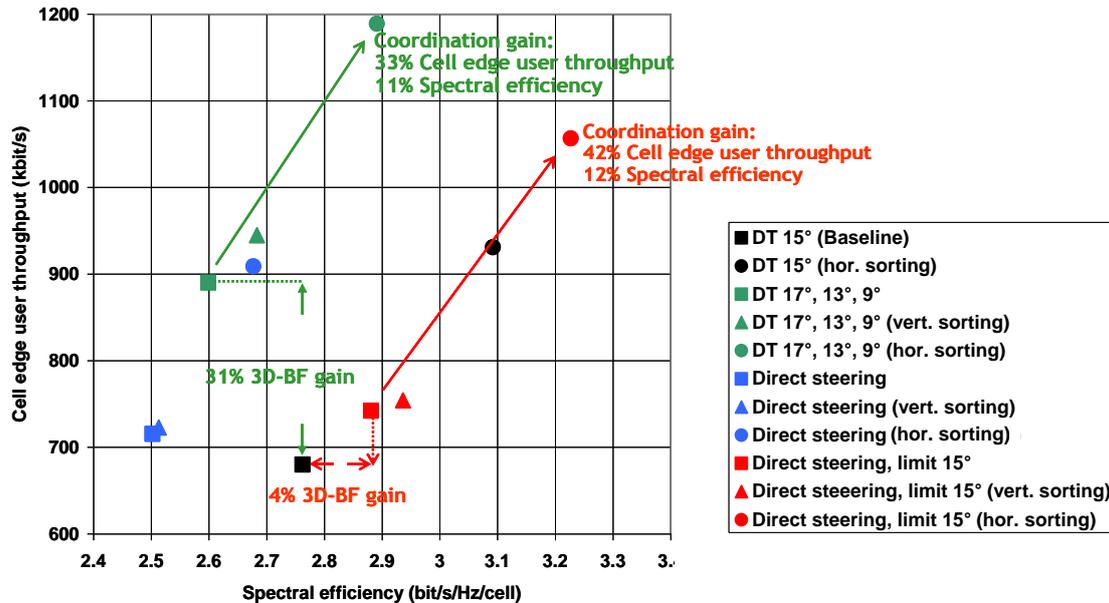


Figure A.32: Cell edge user throughput over spectral efficiency for different combinations of UE specific beam steering schemes and implicit coordination schemes

### Distributed Horizontal and Vertical Beam Coordination:

The approach based on horizontal and vertical beam coordination uses a distributed algorithm working on overlapping coordination areas around each cell represented by groups of seven cells, as described in section 4.2.3.2, Figure 4.10. The goal was to define a decentralized L1/L2 coordination scheme working with exchange of scheduling information, but calculating the individual scheduling decisions individually per cell. For the simulation analysis in the following section the availability of specific UE feedback on adjacent cell interference is assumed. Each UE reports the worst case interfering PMI of each adjacent cell to its own eNB, which exchanges this information with the adjacent eNBs. Ideally this information is available for each possible DT and in combination with the expected SIR degradation. This functionality is not feasible with the current LTE-A Rel-10 standard [3GPP36211] and would require an extension. The proportional fair schedulers of the eNBs use this information to put further constraints on the scheduling decisions to avoid resource assignments with PMIs leading to high interference in adjacent cells. Scheduler parameters are the number of PMI constraints to be considered (0 constraints = no coordination, 1 ... 3 constraints = 1 ... 3 worst case PMIs taken into account for coordination) and a fairness parameter  $\alpha$  for weighting the estimated data rate based on CQI reporting. So each eNB schedules the UE with the *highest score, for which none of the PMI constraints is violated*. The score is calculated from the average rate  $r_{av_t}$ , the expected instantaneous rate  $r_{inst_t}$  and the weighting parameter  $\alpha$  according to

$$score_t = r_{inst_t} / (r_{av_t})^\alpha$$

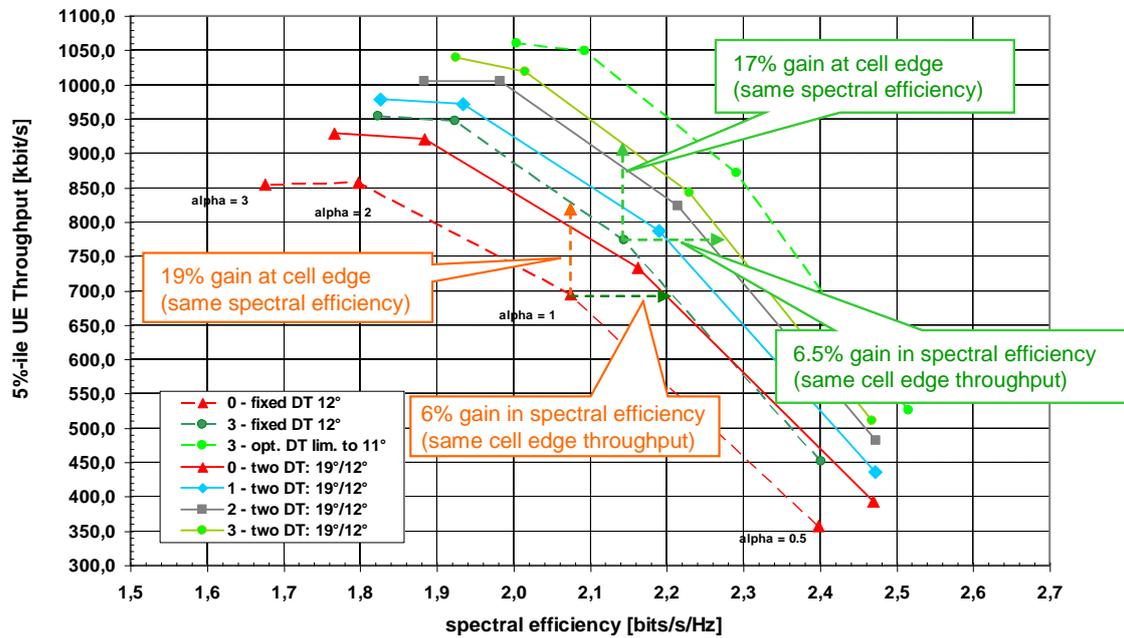
System simulations for different parameter settings have been conducted. The main simulation parameters are summarized in Table A.13.

**Table A.13: Main system parameters**

Parameter	Value
Number of sites and cells	7 sites with 3 cells each
Inter site distance	500 m
Antenna height	eNB: 32 m; Terminal: 1.5 m
eNB antenna pattern	- 4-element uniform linear array, half-wavelength spacing (horizontal) vertical HPBW = 6.5°
Propagation model	3GPP SCME case 1 (non light-of-sight, NLOS)
Terminals per cell	In average 15, randomly distributed
Frequency resources	10 MHz bandwidth; Re-use 1
Scheduling algorithm	Proportional Fair (PF)

For discussion of the results, in Figure A.33 the cell edge user throughput versus the spectral efficiency is shown. The case of two fixed DTs and the case of exact main lobe steering to the UE are compared to the baseline case of one fixed DT of 12°. For ideal main lobe steering the selected DT is  $DT = \max(\theta, \theta_{\text{limit}})$ , with  $\theta$  = exact DT angle to UE and  $\theta_{\text{limit}}$  = DT limitation, with  $\theta_{\text{limit}} = 11^\circ$  in the following simulation.

Each curve represents the performance of a specific scenario in dependency of the values of the parameter  $\alpha$ , where  $\alpha = 1$  corresponds to exact proportional fair scheduling, whereas larger or smaller values shift the fairness towards low rate or high rate UEs, respectively. The case of fixed DT has been simulated without and with horizontal beam coordination. The case of exact main lobe steering has been simulated with 3 constraints. The corresponding curves are the dashed lines (red, dark green and light green). The cases of two fixed DTs of 12° and 19° have been simulated with 0, 1, 2 and 3 constraints taken into account for scheduling (solid lines). Performance gains can be quantified as the difference between curves while keeping either cell edge throughput or spectral efficiency constant. This means that also the parameter  $\alpha$  is modified slightly between the compared curves. In this sense, the performance gain of coordination with 3 constraints and fixed DT (dark green dashed line in Figure A.33), and the gain of adaptive DT without coordination (solid red line) is almost the same. If in addition coordination is applied (additional solid lines), further gains are seen. This leads to the conclusion that coordination gain and gain through vertical beam pattern adaptation are independent effects that can be combined. This leads almost to an addition of the gains, even with the simplified implementation option of using only two different fixed DTs.



**Figure A.33: Exemplary simulation results for distributed horizontal and vertical beam coordination**

With the described simulation environment also simulations with other parameter settings, as well as with single-cell MU-MIMO operation have been conducted. An overview on the achievable gains of the different implementation options is given in the following Table A.14. The gains are given in percent increase of spectral efficiency or cell edge throughput relative to the fixed DT case. The gains are defined as in the Figure A.33, i.e. keeping cell edge throughput or spectral efficiency constant while measuring the gain of spectral efficiency or cell edge throughput, respectively. The number of constraints for coordination and the limitation of the DT angle for exact steering are chosen to give optimum performance. For MU-MIMO the number of constraints for optimum gain is lower than for single user operation. 3D beamforming without coordination shows gains comparable to single user operation, while coordination on top shows only moderate gains in cell edge throughput and small gains in spectral efficiency.

**Table A.14: Relative gains of 3D beamforming compared to fixed DT**

Tx modes	2 fixed DT		3 fixed DT		exact beam steering with DT limitation	
	spectral efficiency	cell edge throughput	spectral efficiency	cell edge throughput	spectral efficiency	cell edge throughput
Single User 0 constraints	7,0	19,0	8,7	21,4	10,1	25,0
Single User 3 constraints	12,0	39,0	15,0	47,0	16,0	52,0
MU-MIMO 0 constraints	5,3	13,2	8,0	18,0	9,5	22,5
MU-MIMO 2 constraints	6,2	20,5	9,0	29,0	10,6	30,9

## A4. Enablers: channel estimation & feedback design

### A4-1 Centralized/decentralized joint transmission with limited signalling information

*[This Appendix contains results that are partly summarized in Subsection 5.3.1]*

In a cellular downlink system with cooperating eNBs joint transmission (JT) has the potential to achieve high performance gains compared to non-cooperating systems, especially for cell edge users.

JT can be realized by means of linear precoding, including multiple transmission nodes, while the transmit power of each node is restricted separately.

This work focusses on frequency division duplex (FDD) systems, where the users downlink channel cannot be derived from observations of its uplink channel, but is obtained by CSI feedback. It is assumed that a user estimates the channel from all eNBs from which it is cooperatively served. Although a user is cooperatively served by multiple eNBs, there is only one master eNB the user is assigned to. This eNB is typically the one with the highest receive signal strength. The user's CSI feedback is only decoded by its master eNB. Hence, without any information exchange between nodes, an eNB has only CSI of the users that are assigned to it. However, precoding requires that knowledge of all user channels as well as all the users' data is available at least at one point.

Precoding with multiple transmission nodes is typically distinguished between centralized joint transmission (C-JT) and distributed joint transmission (D-JT). In the former case channel state information (CSI) of all jointly served users is forwarded from the respective master eNBs to a central unit (CU). Based on CSI of the whole channel the precoding matrix can be calculated. For C-JT data of all users need to be available either at the CU or at each of the cooperating eNBs. In the former case the user data is precoded at the CU while only a part of the precoded data vector is transmitted to the respective eNBs. If the user data is available at each eNB the precoding matrix need to be forwarded. Since the precoding matrix only depends on the channel state, it needs to be updated only per transmission block, which refers to a block of symbols transmitted over a constant channel. Applying D-JT the precoding matrix is computed at each eNB separately, where CSI as well as data of all users need to be available at each eNB. Assuming an ideal backhaul without latency and any capacity restrictions, both approaches result in the same overall precoding matrix.

In contrast to theory, practical systems are affected by several limitations, which lead to impaired CSI. Such limitations include channel estimation errors due to noisy channel observations, limited rate for feedback and backhaul forwarding. Additionally, feedback, backhaul and signal processing introduce delays which impair CSI as soon as the channel varies in time.

For exchanging CSI between eNBs or between a CU and eNBs, typically the core network is used. Unfortunately, such connections can include multiple routers, which lead to high backhaul latency. Since in practice JT mainly suffers from backhaul delays, alternative low latency connections are of high interest for network operators. However, the deployment of direct fibre optical connections is typically quite costly. Options for reducing backhaul latencies are prioritization or directive wireless links. On the contrary, these alternatives are typically more restricted in terms of capacity. In this contribution, the trade-off between backhaul latency and rate is analysed.

## System Model

The mathematical structure of the underlying model is described in detail in [FOF12]. In general a block static channel is assumed, where the channel states  $h[t]$  and  $h[t + \Delta]$  of two blocks with time distance  $\Delta$  have covariance  $\mathbb{E}\{h[t]h^*[t \pm \Delta]\} = c[\Delta]$  according to the jakes model. Since downlink data transmission is precoded based on the current states of all user channels, the complete CSI matrix needs to be available for precoding. System delays lead to outdated CSI, which can be partially compensated by channel prediction. For this purpose previous channel observations as well as the channels time correlation is used.

## Rate-Latency Trade-off

The channel uncertainty  $\varepsilon$  (MSE between actual channel and known channel) which reflects the degree of CSI impairment for a single link is shown in Figure A.34 dependent on the ratio of delay  $\Delta$  and coherence time  $T_c$ . From this illustration it can be observed, how limitations in rate and latency are coupled. From the left plot it can be seen, that increasing the number of quantization bits from 4 to 6 per complex channel coefficient by a required channel uncertainty of less than 0.1,  $\Delta/T_c$  can be doubled (from 0.25 to 0.5). Hence, either the channel coherence time can be halved or the allowed delay can be doubled. Applying that relation to a cooperative scenario, where a certain CSI quality is required for certain links a network operator can optimize the backhaul in order to cooperatively serve users with a certain channel fluctuation.

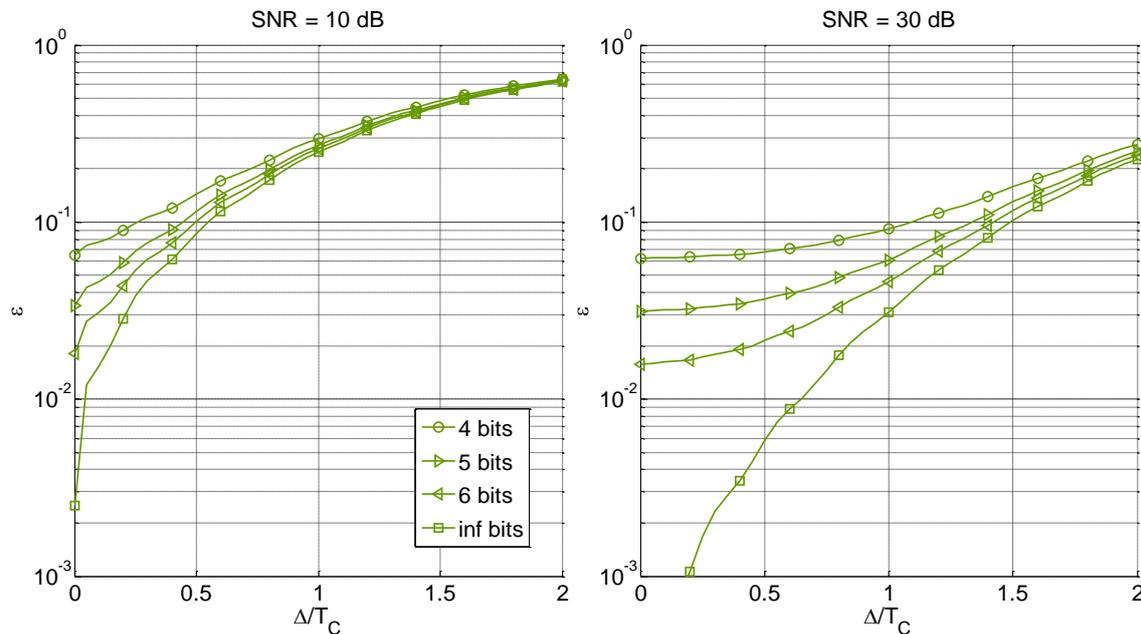


Figure A.34: Illustration of the channel uncertainty of a single link, where the SNR corresponds to the pilot signal-to-noise-ratio.

## Point of Channel Prediction

As mentioned in the previous sections the accuracy of outdated CSI can be increased by channel prediction. Considering D-JT, where each eNB computes the precoding matrix on its own, CSI of all users needs to be available at each eNBs. For this purpose, a user first feeds back CSI to its master eNB. Afterwards, the master eNB forwards the obtained CSI to all other cooperating eNBs by using the backhaul. As soon as the backhaul is affected by latency, the CSI based on a certain channel observation is available at different points in time at the

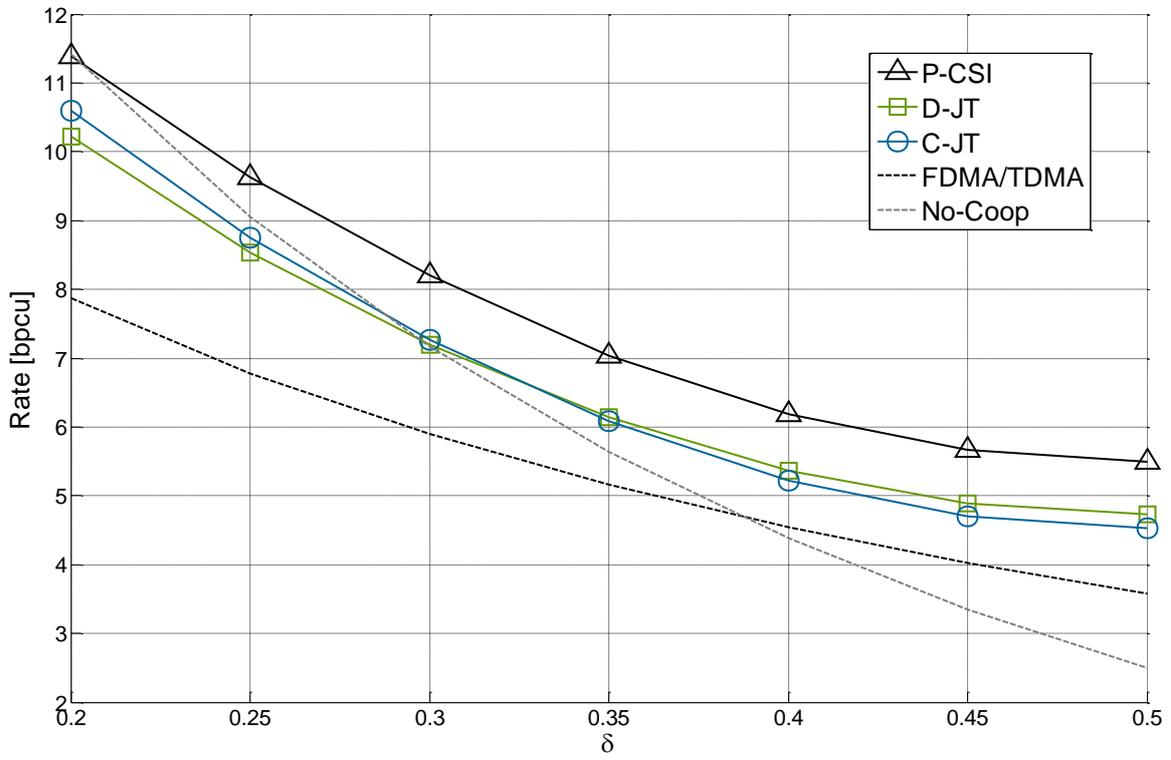
respective eNBs, while the master eNB always has the newest CSI version of its users. In a distributed setup an eNB applies the newest available CSI version of each link for calculating the precoding matrix. In this context the question arises of where to predict the channel. Prediction at the UE can be computed based on uncompressed channel observations. However, the CSI needs to be predicted for one specific point in the future, while it arrives at different eNBs at different points in time. The feedback of several CSI versions would increase the feedback overhead. In contrast, channel prediction at the eNB would bring the flexibility to predict for the point in time which is required based on the so far occurred delay. The drawback of such option is that the information used for channel prediction is already quantized. A third option which might come in mind is the joint prediction at UE and eNB side. In [FOF12] it is shown, that prediction before quantization lead to the same impairment as prediction after quantization. Also joint prediction at UE and eNB result in the same performance. From those results it is shown that channel prediction for D-JT should be computed at the eNB side. Performance gains compared to prediction at the UE are shown in [FOF12].

### Centralized- vs. Distributed Precoding

The previous section focused on distributed precoding. The advantage of that approach compared to C-JT is, that the master eNB can use more accurate CSI of its assigned users. However, since each eNB has different CSI versions, also the resulting precoding matrices the eNB calculates based on its CSI are different. The combination of the respective parts of the unequal matrices to an overall precoding matrix results in inconsistencies. This trade-off is analyzed in this section.

The basic assumption for comparing centralized and distributed precoding is that the backhaul delay between collaborating eNBs is equal for each eNB-eNB connection and is denoted as  $\Delta_B$ . The delay from each eNB to the CU and back to each eNB is also  $\Delta_B$ . The delay for feedback and any kind of signal processing is  $\Delta_F$ . With that assumptions the overall delay for CSI which is transmitted either to the CU or to another eNB which is not the master eNB of the respective user results in  $\ddot{\Delta} = \Delta_F + \Delta_B$ . CSI which is directly applied at the master eNB is delayed with  $\dot{\Delta} = \Delta_F$ . Hence, for C-JT all links are affected by the delay  $\ddot{\Delta}$ . In contrast, for D-JT only CSI of other-cell UEs has delay  $\ddot{\Delta}$ , while UEs assigned to an eNB have delay  $\dot{\Delta}$ . For logical point of view C-JT does not necessarily refer to the case where a CU is place at the core network. Equivalently, C-JT reflects the case where each eNB computes its CSI on its own like for D-JT, but applies older CSI versions of its assigned users so that each CSI is affected by delay  $\ddot{\Delta}$ .

In Figure A.35 simulation results for the mean user rate is plotted over the relative distance  $\delta$  between a UE and its master eNB. For the simulations, two 2-antenna-eNBs jointly transmit to two 2-antenna-UEs. For  $\delta = 0.5$  both UEs are located in the center between the two eNBs, i.e. at the cell edge, while  $\delta = 0$  refers to the case, where both users are located directly at its respective master eNBs. In Figure A.35 it is shown, that D-JT outperforms C-JT only until a certain user separation. If users are closer to its master eNB C-JT shows performance advantages. This behavior is not obvious, since D-JT has better CSI of its own users, while CSI accuracy of other cell users is equal to that of C-JT. However, inconsistencies in the precoding matrix are responsible for the performance degradation. Additionally, the black straight line refers to precoding with perfect CSI (P-CSI). FDMA/TDMA reflects non-coherent reuse 2 transmission (reuse 3 for 3-D setups) while non-coherent reuse 1 is labeled with No-Coop.



**Figure A.35: Mean achievable user rate over the relative distance between eNB and its assigned UE.**

## A4-2 Kalman prediction of multi-site MIMO channels for CoMP

*[This Appendix contains algorithms and results used in Subsections 5.2.3, 5.3.2 and A2-2.]*

For coherent Coordinated Multipoint (CoMP) joint transmission (JT) to multiple users, high quality Channel State Information (CSI) at the transmitter side is crucial. Outdated CSI can cause significant performance loss in CoMP precoder design and also in link adaptation, see e.g. [SFS+05], [SSO+07].

The downlink transmission control loop in LTE and LTE-advanced requires the channel to be known at least 5 ms ahead in time. CoMP processing will add to the delays, and thus to the required channel prediction horizon. Use of outdated present channel estimates are then of questionable value, except for stationary users. At pedestrian and low vehicular velocities, a large performance improvement can be obtained by using predictors: The time-frequency channel is then extrapolated based on its statistical properties, using past and present channel measurements. A prediction horizon  $L$  in time corresponds to a prediction of  $Lv/\lambda = Lv f_c/c$  carrier wavelengths ahead in space, where  $v$  is the relative velocity,  $\lambda$  is the carrier wavelength at frequency  $f_c$ , and  $c$  is the speed of light.

We here focus on OFDM channel prediction that would enable CoMP to work reliably for terminals with high pedestrian velocities at 2.66 GHz or for vehicular velocities at 700 MHz. We mainly discuss FDD downlinks, but the method is relevant also for TDD and for uplinks. For FDD systems, predictions of downlinks must be based on downlink reference signals.

We here present and use a Kalman state-space channel prediction algorithm, making extensive use of recent results from [Aro11] and [ASA12]. Kalman predictors provide two advantages:

First, they provide optimal linear predictions of multiple channels that are based on measurements over an unlimited past time window, and a selectable frequency window. Second, at no extra cost they also provide the second order statistics of the prediction accuracy. This statistical information can then be included into a robust precoder design, resulting in less intracluster interference (see Subsection 5.2.3 and Appendix A2-2).

The Kalman predictor is based on a linear state space model. This model is here based on Auto-regressive (AR) models for each estimated channel component. The state space model represents the second order statistics of the variability over time and frequency of all OFDM channels of interest. A vector of state estimates is updated every time a Reference Signal (RS) is received, using channel measurements at the RS time-frequency locations. From the state estimate, the predicted CSI can then be calculated for a given prediction horizon. As discussed in Subsection 5.3.2, the prediction algorithms for downlinks can be located either in the UEs or in the fixed network.

The prediction performance is below evaluated using channel sounding measurements from an urban environment. This represents a continuation of the limited preliminary investigation that was presented in Subsection 4.4.1 of [ARTD12]. The set of measurements and the prediction results are also used in Appendix A2-2 for the evaluation of robust precoders for joint transmission. See also [ASA12].

### A4.2.1 Model of the measurements and of the fading channel

We assume that  $N_c$  base stations (cells), each with  $N_t$  antennas, transmit known time-frequency reference signals (RS) to a terminal over  $\omega$  RS-bearing subcarriers, for which a single Kalman predictor is designed. For each receiver antenna, we chose to predict downlink channels from  $N_p \leq N_c N_t$  transmitters.

### Measurement model

The received signals  $y \in \mathbb{C}^{\omega \times 1}$  in the frequency domain at one receiver can then be represented by

$$y(t) = \Phi(t)h(t) + n(t), \quad (\text{A4.2.1})$$

where  $\Phi(t) \in \mathbb{C}^{\omega \times \omega \times N_p}$  is a known matrix containing the RS and zeros. The column vector  $h \in \mathbb{C}^{\omega \times N_p \times 1}$  contains the complex frequency-domain channel gains from the relevant transmit antennas, at each of the  $\omega$  subcarriers. Here, the discrete time index  $t$  counts time increments between reference signals.<sup>41</sup>

The noise vector  $n \in \mathbb{C}^{\omega \times 1}$  represents noise plus interference from transmitters outside of the assumed cooperation cluster, at the RS positions. We assume that RS time-frequency positions will be reused by clusters that belong to the same frequency set/cover shift. If the RS pattern is reused *within* the cluster, then the noise term will also include interference from the transmitters within the cluster. In addition, if Kalman predictors are placed on the fixed network side, the measurements  $y(t)$  have been reported by feedback from terminals. In that case, the assumed statistics of  $n(t)$  should also include any contributions from compression errors and feedback errors. In all cases, we will assume that the covariance matrix  $R_n = E(nn^*)$  is reasonably well adjusted to the actual sample correlations<sup>42</sup>.

If terminals have several antennas, the design of lowest computational complexity is to use separate Kalman predictors for the channels to each receiver antenna. If the temporal fading of channels to different receiver antennas is correlated, then higher performance can be obtained by a single Kalman predictor with extended state space, which uses a measurement equation (A4.2.1) that encompasses all antenna signals. However, this increases the computational complexity significantly.

### Downlink reference signals

We can choose the RS patterns from the different transmitters that appear in  $\Phi(t)$  to be fully orthogonal, i.e. placed on orthogonal time/frequency resources. Alternatively, the RS can be selected quasi-orthogonal (also denoted overlapping, or code-orthogonal). In the latter case the RS from different transmitters are placed simultaneously on the same resources, but vectors of  $\omega$  RS symbols from different transmitters are selected to be orthogonal over  $\omega \geq N_p$  RS-bearing subcarriers.<sup>43</sup> If the subcarriers were flat fading,  $N_p$  channels could then be perfectly estimated from the measurements  $y(t)$  in a noise-free scenario. Frequency-selective fading will however degrade the performance. This effect of different types of reference signals is evaluated further in Subsection A4.2.3 below.

---

<sup>41</sup> We here describe modelling of frequency-domain channel coefficients. Kalman prediction of OFDM channels can also be set up in the time domain [Aro11]. The significant contributions to each channel impulse response are then predicted and collected in  $h(t)$ . The measurement matrix  $\Phi(t)$  then contains time-domain signals obtained by subsampled IFFT from the frequency domain measurements at RS locations. (These locations should then be regularly spaced over frequency). Alternatively, the vector  $h(t)$  could represent the singular values of a SVD of the channel [ESB+98], [HE06]. We have so far found no performance advantage for time domain prediction or for using the SVD for prediction, so we here concentrate on direct frequency-domain model prediction.

<sup>42</sup> The prediction performance is not overly sensitive to the norm of  $R_n$ .

<sup>43</sup> We might also choose a combination of the two methods to allocate RS, e.g. to use orthogonal RSs for the  $N_c$  base stations but to use quasi-orthogonal RS for each set of  $N_t$  transmitters at the same base station.

## Channel model

The predictors are based on a linear dynamic model which is based on the autoregressive (AR) state-space models for all the fading subcarriers in all  $N_p$  downlink channels that are to be predicted. It is represented by a state realization vector  $x(t)$  and matrices  $A, B, C$  that specify a set of coupled linear first-order difference equations driven by noise:

$$\begin{aligned} x(t+1) &= Ax(t) + Be(t), \\ h(t) &= Cx(t). \end{aligned} \quad (\text{A4.2.2})$$

Here  $e(t)$  is the zero mean process noise with known covariance matrix  $R_e = E(ee^*)$ . The dynamics of each element of  $h(t)$  is modelled by  $n_a$  elements of the state vector  $x(t) \in \mathbb{C}^{n_a \times N_p \times 1}$ , where  $n_a$  is the AR model order. This model is preferably realized as a diagonal state-space model, with  $A$  being a diagonal matrix and  $B, C$  being block-diagonal.

The matrices  $R_e, A, B$  and  $C$  represent the statistics of the channel variations over time and over frequency. These properties vary slowly over time, due to the shadow fading, which causes propagation paths to change in strength, appear and disappear. For pedestrian velocities, this occurs on a time scale on the order of a few seconds. The model parameters can be accurately estimated from past channel estimates (not predictions) on regular intervals of around 0.5-2 s for pedestrian velocities. For details on how to set up the state-space models and how to estimate the corresponding model parameters, please see [ASA12] and [Aro11]. When varying the model complexity, we have found the prediction performance on measured channels to peak at AR model orders of  $n_a = 4-6$ . Higher orders lead to overfitting and also to higher computational complexity, see Section 8.3.2 in [Aro11]. We here use  $n_a=4$  in the evaluations.

### A4.2.2 Kalman channel prediction

Based on the above measurement model equation (A4.2.1) and channel model equation (A4.2.2), the state realization vector at a given discrete time  $t$  can be estimated, using measurements up to that time instant.

#### State estimation

The MMSE-optimal linear state estimator for (A4.2.1) and (A4.2.2) is the Kalman filter:

$$\begin{aligned} \hat{x}(t|t-1) &= A\hat{x}(t-1|t-1), \\ P(t|t-1) &= AP(t-1|t-1)A^* + BR_eB^*, \\ K_f &= P(t|t-1)C^*\Phi(t)^*(R_n + \Phi(t)CP(t|t-1)C^*\Phi(t)^*)^{-1}, \\ \hat{x}(t|t) &= \hat{x}(t|t-1) + K_f(y(t) - \Phi(t)C\hat{x}(t|t-1)), \\ P(t|t) &= (I - K_f\Phi(t)C)P(t|t-1). \end{aligned} \quad (\text{A4.2.3})$$

Here,  $\hat{x}(t_1|t_2)$  and  $P(t_1|t_2)$  are the estimate of the channel state realization vector and the corresponding covariance matrix of the state estimation error  $E(\hat{x}(t_1|t_2) - x(t_1))(\hat{x}(t_1|t_2) - x(t_1))^*$  at time  $t_1$ , given measurements up to time  $t_2$ .<sup>44</sup> The state

<sup>44</sup> For known models (A4.2.1) and (A4.2.2) and Gaussian noises  $n(t)$  and  $e(t)$ , the Kalman filter is the MSE-optimal estimator in the class of all (linear and nonlinear) estimators. Without the Gaussian assumption, the Kalman filter is still the MSE-optimal *linear* estimator [AM79].

estimate represents the conditional mean of the pdf of the state vector:  $\hat{x}(t_1 | t_2) = E(x(t_1) | y(t_2), y(t_2 - 1), \dots)$ .

To initialize the filter  $\hat{x}(0 | 0)$  and  $P(0 | 0)$  are needed. The filter will converge relatively fast (in a few samples) for SNR values and autoregressive models relevant for OFDM downlink estimation [Aro11]. Therefore,  $\hat{x}(0 | 0)$  can be chosen as an all zero vector when a new terminal enters the system. A discussion on how to choose  $P(0 | 0)$  can be found in Chapter 3 of [Aro11].

### Channel prediction and the prediction error covariance matrix

The prediction  $\mathcal{G}$  steps ahead of the complex channel vector at a time  $t + \mathcal{G}$  can now be obtained by iterating the model (A4.2.2) forward in time, starting from the Kalman filter estimate  $\hat{x}(t | t)$  of the state vector. Utilizing that  $\hat{x}(t + 1 | t) = A\hat{x}(t | t)$  and since  $E(e(t)) = 0$ , we obtain the closed-form expression

$$\hat{h}(t + \mathcal{G} | t) = E(h(t + \mathcal{G} | t)) = C\hat{x}(t + \mathcal{G} | t) = CA^{\mathcal{G}}\hat{x}(t | t). \quad (\text{A4.2.4})$$

The covariance matrix of the prediction error vector  $\Delta h(t + \mathcal{G} | t) = h(t + \mathcal{G}) - \hat{h}(t + \mathcal{G} | t)$  is

$$E(\Delta h(t + \mathcal{G} | t)\Delta h(t + \mathcal{G} | t)^*) = CP(t + \mathcal{G} | t)C^*, \quad (\text{A4.2.5})$$

where the matrix  $P(t + \mathcal{G} | t)$  is calculated iteratively from  $P(t | t)$  using (A4.2.2), which gives

$$P(t + k | t) = AP(t + k - 1 | t)A^* + BR_e B^*. \quad (\text{A4.2.6})$$

The prediction error variance for an element of  $\hat{h}(t + \mathcal{G} | t)$  is given by the corresponding diagonal element of (A4.2.5). Note that choosing  $\mathcal{G} = 0$  provides the filter estimate of the channel, and the corresponding estimation error covariance matrix.

Prediction error variances will increase with an increasing prediction horizon  $\mathcal{G}$ . They also increase with a decreasing SNR of the measurements. Finally, the prediction accuracy depends markedly on the properties of the fading over time. Fading processes with flat Doppler spectra are harder to predict than fading processes with large single peaks. The predictability of the theoretical Jakes Doppler spectrum (with peaks at  $\pm$  the maximal Doppler frequency) is intermediate between these cases [Ekm02],[Aro11], see Figure 5.16.

The prediction performance is below evaluated in terms of the Normalized Mean Square prediction Error (NMSE). This performance indicator is calculated separately for each RS-bearing subcarrier  $k$  and channel  $j$  by averaging the corresponding element of  $h(t)$ , with index  $l$ :

$$NMSE_{k;j} = \frac{\sum_{t_0}^T |\Delta h_l(t)|^2}{\sum_{t_0}^T |h_l(t)|^2}, \quad (\text{A4.2.7})$$

where  $T - t_0$  is the number of samples in an averaging interval and  $t_0$  is the first sample of that interval. An NMSE of  $-x$  dB implies that if the channel is used for coherent JT, then interference due to that channel can on average be suppressed by at most  $x$  dB by cancellation.

### A4.2.3 Performance evaluation

#### Measurements and simulation assumptions

The prediction performances in this section are based on channel sounding measurements from three omnidirectional single antenna transmitters located at different sites in an urban area.<sup>45</sup>

The measurements over a 20 MHz bandwidth are collected by a vehicle driving at 0-30 km/h. Measurements are obtained for all subcarriers, with RSs from different transmitters being placed on orthogonal subcarriers. The system settings for the measurements are presented in Table A.15 and the powers of the three received signals along the measurement route is shown in Figure A.36. Additional details on the channel soundings were described in Section 4.4.1 of [ARTD12]. The measurements of frequency-selective complex channel gains are of high quality: All transmitted signals were used as reference signals and no interference was present. They will for our purposes be considered to be the true channels.

**Table A.15: System parameters for channel measurements that are here used in simulations**

System parameter	Value
Number of base stations	3
Type of base stations	Omnidirectional single antenna
UE speed	0-30 km/h
Carrier frequency	2.66 GHz
Distance between base stations	360-600 m
RS spacing in time	5.33 ms
Subcarrier bandwidth	15 kHz
RS pattern	Orthogonal, Quasi-orthogonal

In our evaluations, we assume an RS pattern with reference signals on every subcarrier and every  $L$ 'th OFDM symbol (with a symbol time of 71 ms). The fading should correspond to fast pedestrian velocities at 2.66 GHz and the subcarrier bandwidth should be 15 kHz, as in LTE. We generate fading channels with these properties, by up-sampling the channel data set 25 times in time using the Fast Fourier Transform (FFT).

This up-sampling in time generates channels that will be used at the RS-bearing OFDM symbols. Up-sampling of the measured channels by a factor 25 is e.g. appropriate for two fast pedestrian scenarios:

- A system with UEs at speeds up to 10 km/h and an RS time spacing of 0.64 ms (corresponding to 9 OFDM-symbols at 15 kHz subcarrier spacing), denoted **RS9**.<sup>46</sup>
- A system with UEs at speeds up to 5 km/h and an RS time spacing of 1.3 ms (corresponding to 18 OFDM-symbols at 15 kHz subcarrier spacing), denoted **RS18**.

<sup>45</sup> The data set used by Uppsala University in these investigations was collected by Ericsson Research in December 2008 in the Stockholm suburb Kista [MSK+09], [LMF10].

<sup>46</sup> Each 5.3 ms interval in measured time then corresponds to  $3 \times 5.3 = 15.9$  ms at the new max 10 km/h velocity. It represents 225 OFDM symbols, out of which  $225/9 = 25$  contain reference symbols. In the RS18 case, the 5.3 ms interval corresponds to 31.8 ms, or 450 OFDM symbols, out of which  $450/18=25$  contain reference symbols.

To investigate the effect of using quasi-orthogonal pilot patterns, we also compare two RS patterns:

- **Orthogonal RS.** The three single-antenna transmitters send RS on every third subcarrier of the RS-bearing OFDM symbol.
- **Quasi-orthogonal RS.** All three transmitters send RS on every subcarrier of the RS-bearing OFDM symbol, with equal total transmit power as compared to the orthogonal case. The RS patterns are here orthogonal over 4 adjacent subcarriers and are time-invariant. To generate this situation, the original channel data has been up-sampled 3 times in frequency.<sup>47</sup>

To generate the measurement signal  $y(t)$ , the up-sampled channels are used in equation (A4.2.1), where the measurement matrix  $\Phi$  is formed by using the appropriate reference signals, and noise is added. The measurement noise  $n(t)$  is modelled as i.i.d. Gaussian variables with covariance matrix  $R_n = \sigma^2 I$ . In evaluations, we use three levels for  $\sigma^2$ , shown (dotted) in Figure A.36.

These noise levels are intended to span a reasonable range of noise and out-of-cluster interference powers that could be expected in a clustered JT CoMP system of the type described in Subsection 5.2. The lowest of these three levels would correspond to either efficient interference floor shaping or a low load in interfering clusters. The highest of the three levels could approximately correspond to a clustered system without additional interference floor shaping that also has high load.<sup>48</sup>

Predictions are performed for 432 subcarriers over the full 7 min. measurement route, corresponding to 21 min. and 42 min. at the assumed velocities of the RS9 and RS18 cases, respectively.

At every 800 RS in time (0.5 s in case RS9 and 1.0 s in case RS18), fourth order AR models are re-estimated. The frequency correlation required to form the matrix  $R_e$  for equation (A4.2.1) is estimated by averaging over all utilized subcarriers. Parallel Kalman predictors each estimate and predict the channels from all base stations over 4 adjacent RS-bearing subcarriers. Their prediction outputs are then Wiener-smoothed in the frequency domain to counteract the rather small frequency window  $\omega=4$  per Kalman filter. The measured performances exclude the initial transient phases of the Kalman algorithms, but these transients are fast and would change the total averages insignificantly.

---

<sup>47</sup> For the orthogonal RS pattern, the original measurements already use a pattern with RSs from the three transmitters placed on every third subcarriers. Therefore no resampling in frequency is needed for that case.

<sup>48</sup> The use of constant interference/noise levels is an approximation. In a real environment, the interference level would vary with position within the cluster and also be markedly affected by shadow fading. For example, the locations encountered between time 10.5 min and 13 min in Figure A.36 are in a courtyard shadowed by buildings. The out-of-cluster interference would likely be reduced by shadowing in such positions as well.

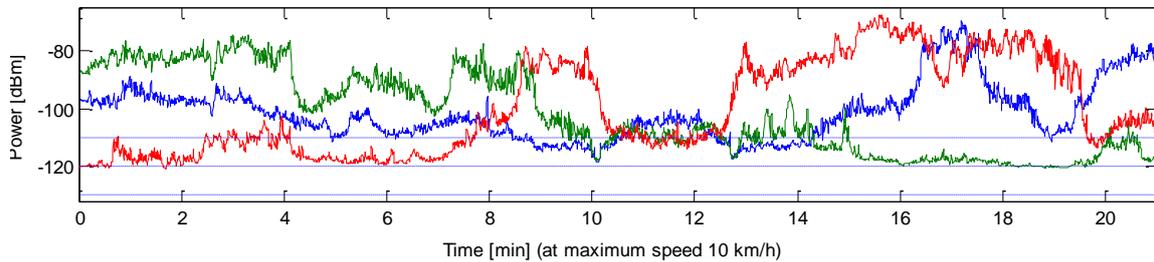


Figure A.36: Signal strengths of received signals from the three omnidirectional single antenna transmitters (full lines) and the three power levels of noise added in the simulations (dotted lines). Transmitters 1, 2 and 3 are represented by blue, green and red respectively.

Table A.16: Average NMSE performance of Kalman prediction and for using outdated CSI for  $\mathcal{G} = 0$  (Kalman *filter* estimates of channels) and for prediction horizons  $\mathcal{G} = 4, 8, 12$  and  $18$  (corresponding to 5 ms, 10 ms, 15 ms and 23 ms for case RS18).

Prediction horizon	RS pattern	Noise level (dBm)	Outdated CSI: Average NMSE for all base stations	Predicted channels:	
				Average NMSE for all base stations	Average NMSE for the weakest base station
$\mathcal{G} = 0$	Orthogonal	-110	-17.8 dB	-17.8 dB	-7.1 dB
		-120	-23.9 dB	-23.9 dB	-12.7 dB
	-130	-30.9 dB	-30.9 dB	-20.0 dB	
	Quasi-orthogonal	-110	-16.8 dB	-16.8 dB	-5.6 dB
-120		-19.6 dB	-19.6 dB	-6.9 dB	
$\mathcal{G} = 4$ (5 ms)	Orthogonal	-110	-10.5 dB	-12.8 dB	-5.9 dB
		-120	-12.5 dB	-15.3 dB	-9.4 dB
	-130	-14.0 dB	-17.6 dB	-13.3 dB	
	Quasi-orthogonal	-110	-10.1 dB	-12.4 dB	-5.2 dB
-120		-11.2 dB	-13.9 dB	-6.6 dB	
$\mathcal{G} = 8$ (10 ms)	Orthogonal	-110	-6.9 dB	-11.0 dB	-4.8 dB
		-120	-7.9 dB	-12.9 dB	-7.4 dB
	-130	-8.6 dB	-14.8 dB	-10.3 dB	
	Quasi-orthogonal	-110	-6.6 dB	-10.8 dB	-4.4 dB
-120		-7.3 dB	-12.2 dB	-5.9 dB	
$\mathcal{G} = 12$ (15 ms)	Orthogonal	-110	-4.4 dB	-9.6 dB	-4.0 dB
		-120	-5.0 dB	-11.2 dB	-5.9 dB
		-130	-5.3 dB	-12.8 dB	-8.2 dB
$\mathcal{G} = 18$ (23 ms)	Orthogonal	-110	-1.8 dB	-7.9 dB	-3.0 dB
		-120	-2.1 dB	-9.2 dB	-4.1 dB
		-130	-2.2 dB	-10.3 dB	-5.4 dB

### Summary of results over the whole test route: Average prediction NMSE versus the prediction horizon, the RS pattern and the noise level

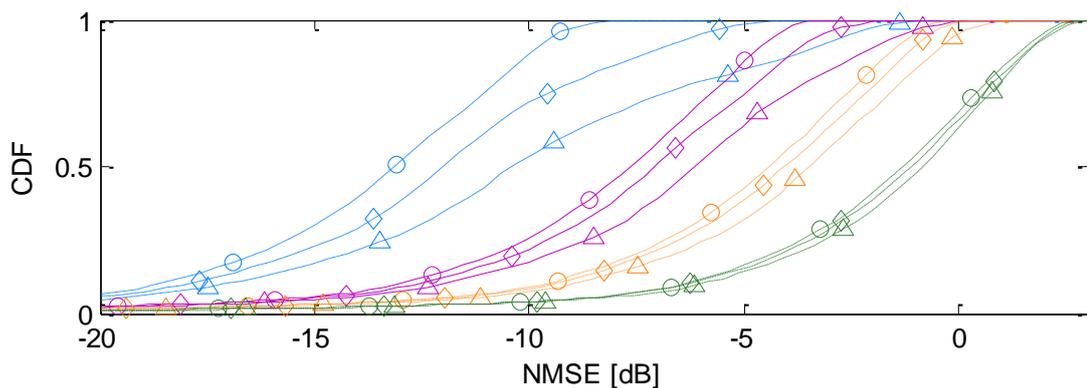
Prediction horizons of  $\vartheta = 0, 4, 8, 12$  and 18 RS symbols are evaluated. These horizons correspond to prediction horizons of 0, 2.5, 5.0, 7.5, 11.5 ms for the RS9 case and 0, 5.0, 10, 15, 23 ms for RS18 respectively. The horizon  $\vartheta=0$  is a basic comparison case of *filter* estimates.

The resulting prediction NMSE by equation (A4.2.7) is calculated for the strongest (with highest channel gain), the second strongest and weakest signals from the three base stations, averaged over the estimation time interval, for every subcarrier. Total averages of these results are summarized in Table A.16. The prediction performance is also compared to the NMSE experienced when simply using the outdated Kalman filter estimate, denoted as outdated CSI.

#### Prediction versus using outdated CSI

Clearly the CSI is significantly improved by using predicted channels as opposed to extrapolating outdated channels, especially in situation with a low noise floor. Specifically we see that for the longer prediction horizons, corresponding to 10-23 ms at velocities up to 5 km/h ( $\vartheta=8, 12, 18$ ), the outdated CSI will on average have an NMSE worse than -8 dB. This accuracy was in [SFS+05] found to be required to achieve most of the multi-user scheduling gains that are obtainable with perfect channel estimates. It is here used as a loose limit of “acceptable NMSE”.

For the prediction horizon  $\vartheta = 4$ , corresponding to 5 ms at velocities up to 5 km/h, the average accuracy when using outdated CSI would be above -8 dB. However, studying Figure A.37, which shows the CDF of the NMSE for the outdated CSI for orthogonal RS patterns, it is clear that we need a very low noise floor to ensure good prediction performance for most channels. From this figure we also see that as the prediction horizon increases a lower noise floor, though causing a better instant estimate, only results in small NMSE gains. Already for  $\vartheta=8$ , only 46 % of the predictions have an NMSE below -8 dB even at the lowest noise floor.

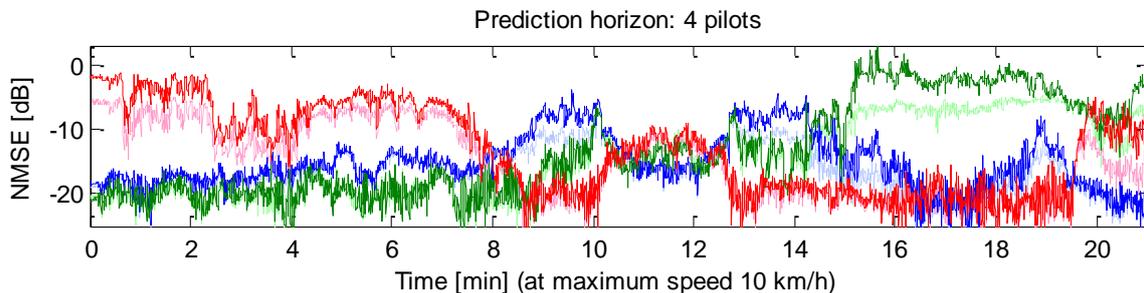


**Figure A.37: CDF of NMSE when using outdated CSI for  $\vartheta = 4$  (blue dashed lines),  $\vartheta = 8$  (purple solid lines),  $\vartheta = 12$  (orange dashed lines) and  $\vartheta = 18$  (green dashed lines) using an orthogonal RS pattern. Results for noise floors of -130 dBm (circles), -120 dBm (diamonds) and -110 dBm (upwards triangles). The prediction quality increases towards the left.**

### Impact of using orthogonal versus quasi-orthogonal reference signals

The RS pattern has a significant impact on how well we can predict the weakest of the three channels. At a noise floor of -120 dBm, the use of an orthogonal RS pattern improves this prediction performance by 5.8, 2.8 and 1.6 dB for the weakest channel for prediction horizons of 0, 4 and 8 RS symbols, respectively. Figure A.38 below shows in more detail how the prediction performance for the three base stations varies over time. The performance gap between orthogonal RSs and quasi-orthogonal RSs is especially high for the weakest channel when there is also a large difference between the received powers from the weakest and the strongest received channel. (Compare e.g. the prediction performance of the weakest base station at time 17 min with that at time 5 min).

The phenomenon has the following explanation. Since the subcarriers are not perfectly flat fading, state estimation errors of one channel will to some extent influence state estimation errors of all other channels. This leakage of estimation errors acts as additional noise which makes the weak signals hard to estimate, and therefore also hard to predict. This is a near-far effect similar to that encountered in CDMA detection in uplinks. In the experiments with the highest noise level, this phenomenon is less obvious: The signal from the weakest base station is then already influenced by noise with high power as compared to its own power (see Figure A.36).<sup>49</sup>

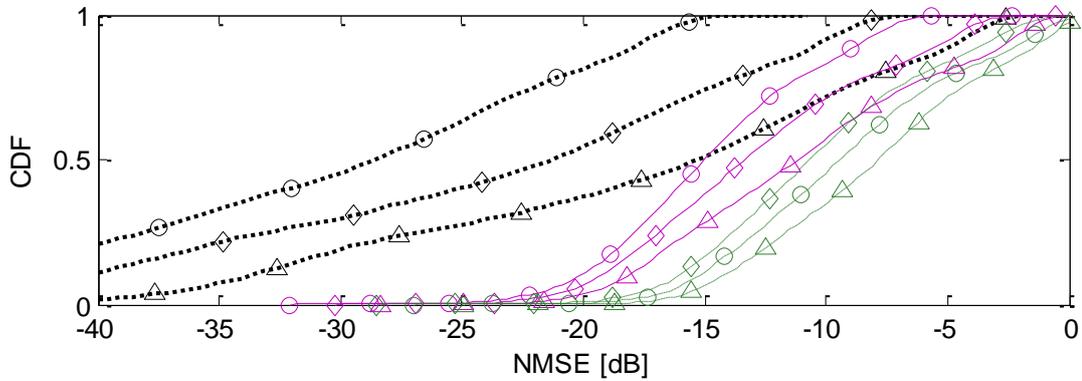


**Figure A.38: NMSE of the predicted channels averaged over all subcarriers for  $\mathcal{G}=4$  and a noise floor of -120 dBm. The colours represent the same base stations as in Figure A.36. Darker shades represent a quasi-orthogonal RS pattern and lighter shades an orthogonal RS pattern.**

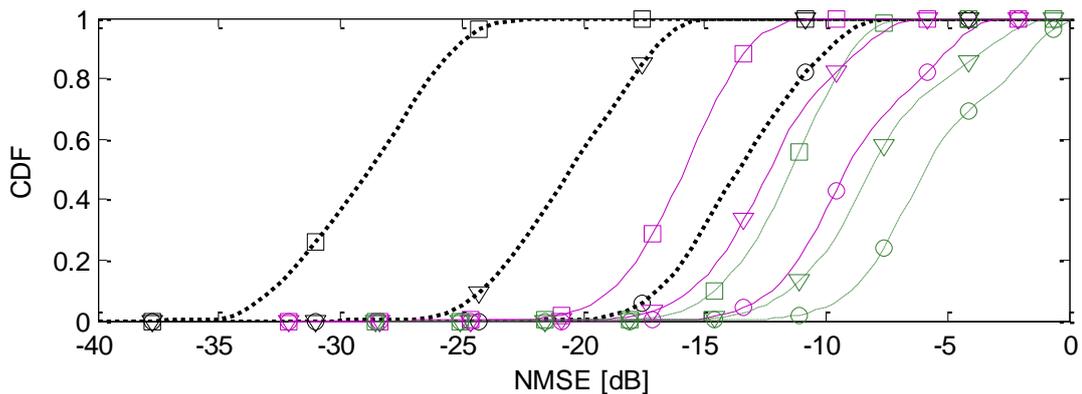
### Kalman prediction performance versus noise and out-of-cluster interference power

As expected, the prediction performance improves as the noise floors decrease, especially for orthogonal RS patterns. In Figure A.39 we see how the performance depends on the noise floors for orthogonal RS patterns. In particular, we see that for the lowest noise floor, 70 % of the  $\mathcal{G}=18$  predictions, 93 % of the  $\mathcal{G}=8$  predictions and 100% of the estimates ( $\mathcal{G}=0$ ) can be obtained with an NMSE below -8 dB. However, for the highest noise floor these numbers decrease to 49%, 69 % and 79 % respectively. This further motivates the need of efficient noise floor suppression/shaping schemes, as those discussed in Subsection 5.2.

<sup>49</sup> In the preliminary results of Section 4.4.1 of [ARTD12], only the channels between 10.5 and 12.5 min were used. Although some tendencies to a performance loss for quasi-orthogonal reference signals (overlapping pilots) was present, this effect was less evident since, in this particular subset of the measurements, base stations have similar channel gains.



**Figure A.39:** Prediction performance for  $\mathcal{G}=0$  (black dotted lines),  $\mathcal{G}=8$ , corresponding to  $0.123\lambda$  by eq. (5.3.1) or  $L=10$  ms at 5 km/h (purple solid lines) and  $\mathcal{G}=18$  corresponding to  $0.28\lambda$  in space or  $L=23$  ms in time (green dashed lines). An orthogonal RS pattern is used. Cdf's for noise floors of -130 dBm (circles), -120 dBm (diamonds) and -110 dBm (upwards triangles).



**Figure A.40:** Cdf's for RS SNR intervals of [0, 10] dB (circles), [10, 20] dB (downwards triangles) and [20, 30] dB (squares). Colours and line-styles indicate prediction horizons as in Figure A.39.

In addition to the noise floor, the prediction horizon and the velocities of the UEs are of great importance for the prediction performance. It is clear from Table A.16 that as  $\mathcal{G}$  increases, the improvements obtained by lower noise floors and by the use of orthogonal pilots diminishes: With an increasing prediction horizon, the main factor limiting performance will increasingly be the statistics of the short term fading process, represented by the Doppler spectrum. This dominating effect of the Doppler spectrum is illustrated by Figure 5.16, which is based on theoretical results in [Aro11]. It is also seen in Figure A.40. Here, the RS SNR has less impact as the prediction horizon increases.

Comparing the performance of  $\mathcal{G}=4$  with that of  $\mathcal{G}=8$  at orthogonal RS, we have gains of 1.8 dB, 2.4 dB and 2.8 dB for the noise floors -110 dBm, -120 dBm and -130 dBm respectively. Hence, prediction accuracy gains caused by introducing a fast backhaul would be largest when the noise floor is low. Again this points to the importance of the inter-cluster interference management.

The comparison between the prediction horizons  $\mathcal{G}=4$  and  $\mathcal{G}=8$  can also be viewed as the difference between UEs moving at 0-5 km/h (RS18) and UEs moving at 0-10km/h (RS9) for the

prediction horizon 5 ms.<sup>50</sup> A higher velocity increases the required prediction horizon in terms of number of wavelengths and thereby also decreases the predictability, yielding a higher NMSE.

#### A4.2.4 Summary, conclusions and consequences for CoMP system design

**Performance.** Use of predicted channels as opposed to outdated CSI is crucial for prediction horizons of 5 ms or more even for slow moving UEs (up to 10 km/h). Kalman predictors can provide sufficient channel prediction accuracy for these “pedestrian” UEs. In our setting, the use of prediction instead of using outdated CSI almost triples the prediction horizon for which acceptable accuracy is obtained on average. The useful range corresponds to prediction horizons (feedback loop latencies) of 15 ms for 5km/h users at 2.66 GHz and to horizons of 5 ms for velocities of 57 km/h at 700 MHz. In Appendix A2-2, large JT CoMP gains are obtained also for 23 ms prediction at 5 km/h at 2.66 GHz.

**Effects of noise and delays.** A well designed suppression of the inter-cluster interference and the use of fast backhaul links that limit the required prediction horizons are both important. If these properties are improved, for example by reducing the basic feedback latency of LTE from 5 ms to perhaps 3 ms, in combination with maximum coordination latencies within (fibre-connected) cooperation clusters of 1 ms, then useful predictions could be achieved also at higher velocities.

**Appropriate reference signals.** Prediction of multiple channels that have very different channel gains benefits from orthogonal RS patterns, so we primarily recommend the use of orthogonal reference signals. This has consequences for the best types of downlink reference signals to be used for CoMP. The CRS defined in LTE Release 8 accommodate only four orthogonal reference signals. This number is too small, since downlinks are likely to have to be predicted from typically 9 base stations in a large cooperation cluster. The CSI-RS defined in LTE release 10 provide adequate orthogonal reference signal locations, but to limit overhead, CSI-RS would be transmitted infrequently, every 5 ms. This channel sampling rate is still adequate for fast pedestrian users at 2 GHz and for vehicular users at 700 MHz.

**Proposed reference signal pattern.** For cooperation clusters comprising 9 cells at 3 sites, each with up to 4 antenna ports/beams, we propose the use of CSI-RS that are repeated at most every 5 ms. Each cell/base station is allocated 4 out of the 40 available time-frequency symbols so 36 out of 40 positions are used. The 4 RS per cell may use either orthogonal signals or may use four quasi-orthogonal sequences, to distinguish up to four antenna ports per base station. The channel predictors for the 9 base stations can then be implemented as separate Kalman predictors. This decreases the computational complexity.

At present, the CSI-RS are envisioned to be transmitted at most every 5.0 ms, which is longer than the 1.3 ms used for evaluation here. The resulting performance would therefore be somewhat lower due to less noise averaging. It would be desirable to allow these reference signals either to be transmitted more frequently or to use RS power boost, when this is motivated from a total performance perspective that includes downlink RS overhead and uplink reporting overhead.

---

<sup>50</sup> In such a comparison, the RS18 case has a built-in disadvantage, since the RS9 case has twice the amounts of RSs per time unit to use for averaging out noise in the channel estimates. This means that the difference in prediction performance for UEs moving at a maximum speed of 5 km/h and UEs moving at a maximum speed of 10 km/h with *equal* RS spacing would be slightly larger than the one presented here.

### A4-3 Advanced channel prediction

*[The methods in this Appendix relate to Subsection 5.3]*

Powerful and accurate channel prediction techniques over 10-20ms would be extremely helpful in the context of JP CoMP as it would avoid CSI outdated due to feedback delay and might significantly reduce the feedback overhead. From simulations and measurements it is known that - more than CSI quantization errors - the mismatch between reported radio channels and those at the time of transmission of precoded data signals might spoil any CoMP gains [ZMS+09]. This is true even more for our advanced interference mitigation framework IMF-A as described in Subsection 5.2.

Channel prediction techniques like 2-dimensional Wiener filter or Kalman [MF11][Aro11] based filtering are well known, exploiting optimally statistical correlations in time and frequency. In [RHG06] a tensor based concept was proven to be helpful in case of MIMO systems by higher order singular value decomposition of the radio channel into the frequency-, time-, Rx-and Tx-antenna domain. This raises the question: 'Is there any room for further improvements?'

We see as potential enhancements - at least theoretically - the so called model based channel prediction (MBCP), which can be seen also as a parametrical compression scheme. It has been inspired by chaos theory and the interesting aspect is that in chaos theory a single start value allows to reconstruct infinite complexity. In case of MBCP the single feedback value is the 3 dimensional location vector of the UE position - instead of CSI values - with respect to an accurate building vector data map (BVDM) of the surrounding of an eNB. As illustrated in Figure A.41 eNB and UE share the same BVDM, either being broadcasted by the eNB or being preloaded at UEs. So the UE estimates the correct location within the BVDM and the eNB can ideally reconstruct all wideband channel components for the reported UE location.

For JP CoMP even more interesting is that all channel components of all cooperating eNBs and antenna elements might be fed back with a single value. For channel prediction additionally a moving vector is fed back allowing to estimate future locations and according channel conditions over quite a long time duration.

In contrast to Wiener or Kalman filters MBCP enables - at least theoretically - to predict birth and death of channel components and therefore promises an improved channel prediction horizon with accordingly higher performance. Figure A.42 plots the evolution of the amplitude of the five strongest multi path components (MPC) over a distance of 50cm for one of our measured CCs. Applying the most simple linear prediction based on previous observations works quite fine over the first 30cm, but then there is a salient point indicated by the black circles leading afterwards to a completely different slope for the linear prediction of two of the MPCs. The most probable reason is the birth or death of some subpaths. Auto covariance based methods relying on previous observations will fail in such occasions completely while MBCP ideally promises full predictability.

The UE estimates its three dimensional position based on GPS and/or channel estimation taking antenna patterns, RF frequency etc. into account. To compare the artificial location with respect to the BVDM and the measured CSI a raytracing simulation has to be done. Raytracing is memory and processor hungry, but in the future Moore's law should help. In addition for a certain area a simplified model can be found, giving the novel technique its name.

This technique raises many issues and in the following sub chapters some are shortly addressed.

### Validity of BVDM over time

In Figure A.43 the stability of typical radio channels is illustrated for a specific real world measurement on our LTE testbed in Munich. The figure contains the channel transfer function of one component at the start of the measurement and after 40 minutes. In between the UE had been moved back and forth over 50cm. The high level of similarity is quite encouraging meaning that a BVDM at least for this location would be valid for a reasonably long time. At the same time and as expectable there is not a 100 percent similarity. This small variations might be due to a small changes of the channel conditions as such, e.g. due to a relocation of some cars or closing/opening of certain windows or doors etc. For that case it is proposed to use a learning algorithm adapting to long term variations of the channel conditions. 'Long term' variations are in the range of minutes, hours or even longer, e.g. effected by newly parking cars or changes in the foliage of trees. In contrast short term variations would be caused e.g. by moving cars.

Another source of small mismatches might be CSI measurement errors e.g. due to frequency synchronization deviations, slight UE dislocations, slightly differently steered antenna elements or any other parasitic effects. As these effects are in a real world system unavoidable the MBPCP has to be combined with a fast and accurate delta reporting feedback, i.e. the UE reports the deviation with respect to the reconstructed CSI for a certain location. As long as these deviations are small the overhead for this delta reporting can be kept small as well, but this issue needs further research.

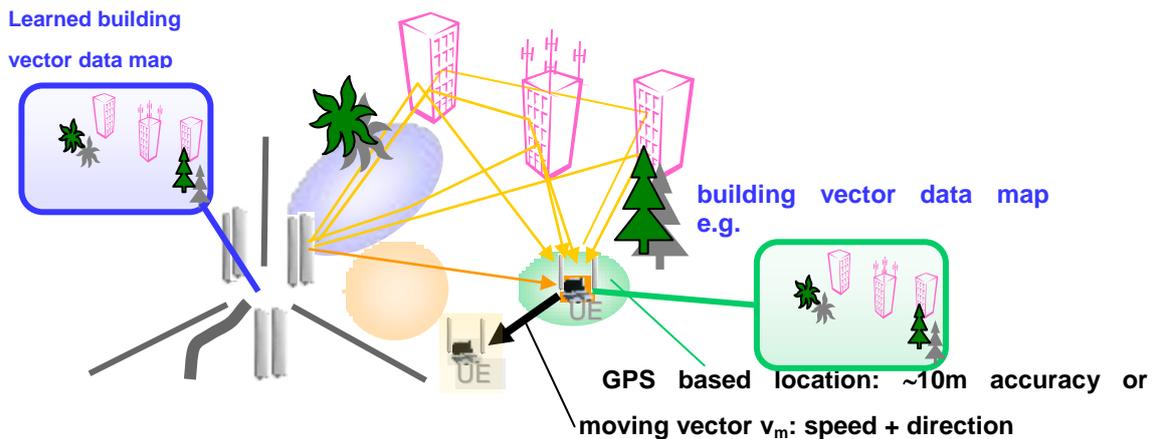


Figure A.41: Illustration of the model based channel prediction concept

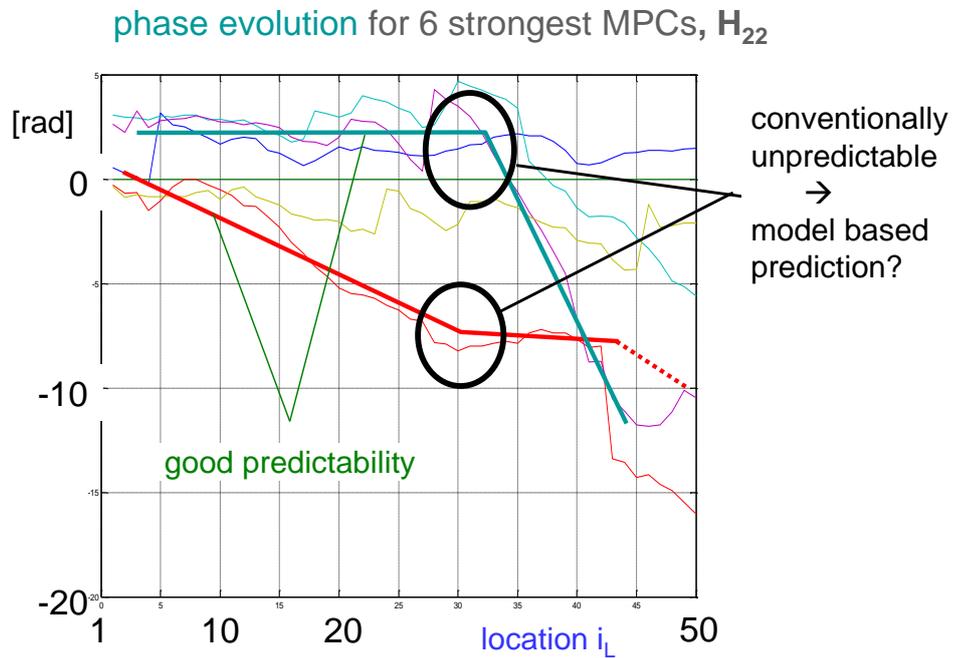


Figure A.42: phase of 5 strongest MPCs measured for moving UE over 50cm distance. At circles slope of phase evolution changes, which is unpredictable with Wiener filtering.

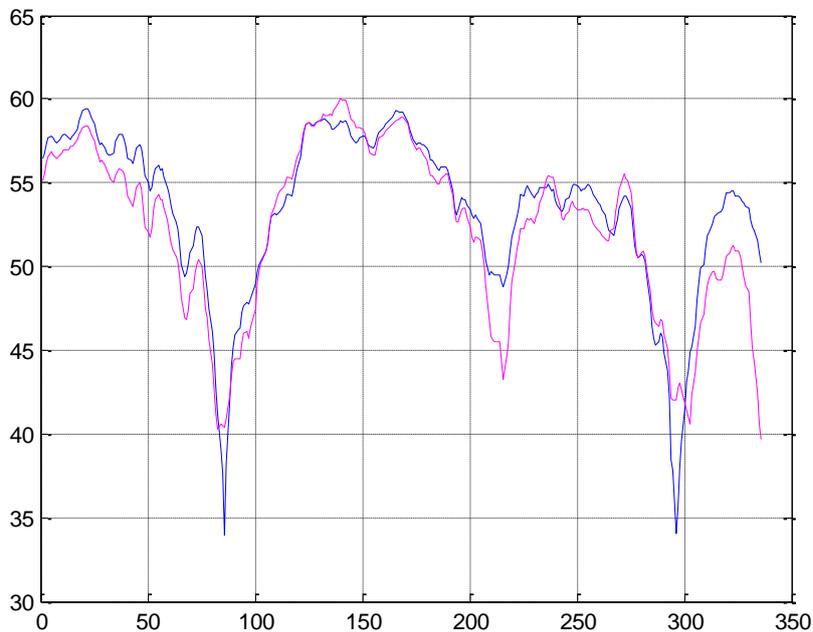


Figure A.43: measured CTF in [dB]; 336 OFDM reference signals with SC spacing of 90 kHz resulting in overall 30MHz bandwidth; carrier frequency =2.6GHz; time instances  $t=0$  (blue) and 40min (magenta).

## Analysis MBCP versus Measurements

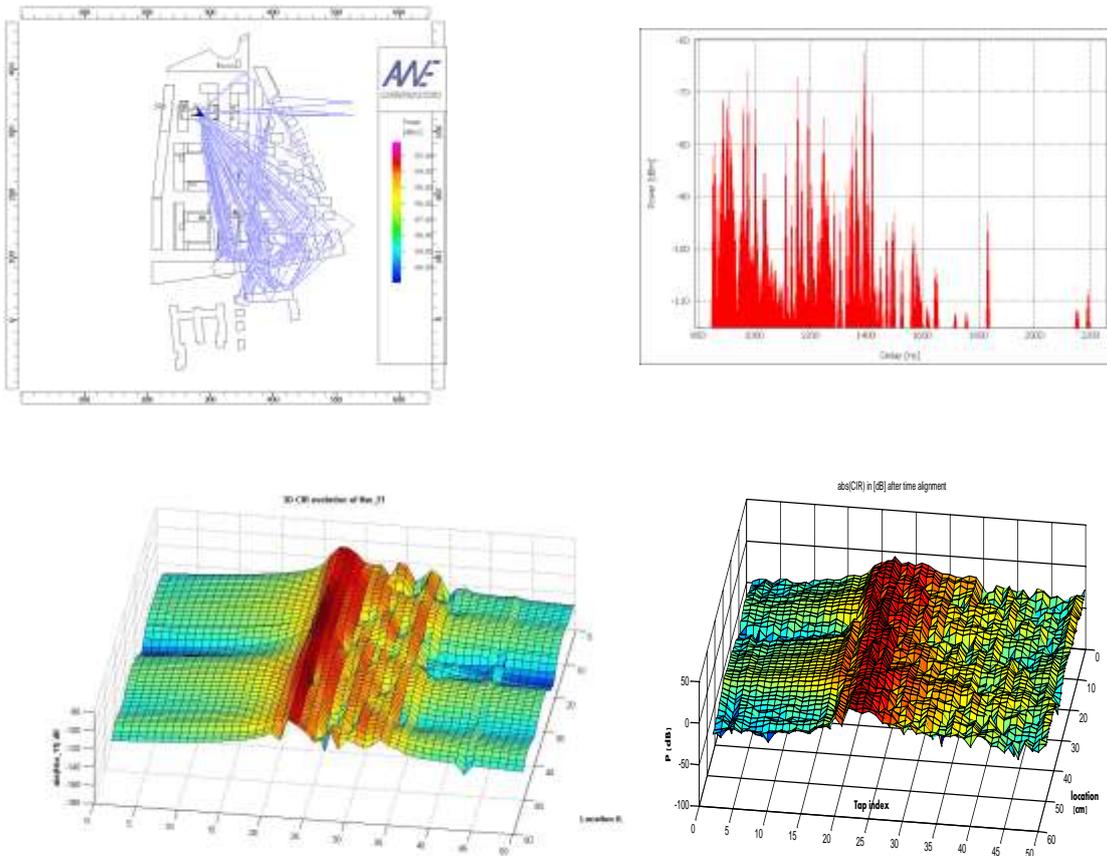
One of the main goals of our MBCP research is to validate the concept against real world measurements and to learn the potential and the limits of the concept. This is a long term research area and only intermediate results are available. Small scale CSI estimation is extremely sensitive to the smallest errors. For that reason it is not expected that MBCP can work as a stand alone solution but should be combined with state of the art techniques like Wiener or Kalman filtering. In addition it is expected that a fast more conventional feedback of small CSI deviations will be extremely helpful. Important is to find out the relation between each of the different above mentioned schemes, i.e. whether MBCP can provide a noticeable and otherwise unachievable part of the overall CSI information. As has been elaborated in [Aro11] it is important to combine the different CSI informations taking into account their second order statistics, i.e. the relative reliability is important. To test the reliability of MBCP CSI information MBCP is compared with realworld measurements in the NSN campus in Munich.

Evaluation of the measurements for the Munich testbed as being introduced in [ARTD12] requires advanced algorithms for extraction of the channel information as such from measurements being impaired by artefacts like time and frequency offsets, noise, etc. In Figure A.44 bottom right the measured and suitably pre-processed channel impulse response (CIR) evolution over a distance of 50cm can be seen.

Enhanced ray tracing simulations indicate that there is at least a very rough similarity between the measured and the raytraced CIR evolution for the same place. Especially the CIR is of similar length and main components show a similar high level characteristic of their evolution. At the same time the measurements still exhibit much more small scale variations, even so it is not fully clear yet how far this is a contribution of the channel or of the measurement procedure. However, even after noise reduction by suppressing of taps outside the guard interval and by suitable averaging over 15 to 25 measurements per location there was mainly a slight smoothing effect visibly, but no significant different channel characteristics.

Therefore currently there is a focus on improvements of the the modeling of the BVDM, e.g. by adding further elements like the measurement van or by more accurate modeling of the antenna patterns.

In parallel measurements have been conducted investigating typical effects like moving persons or cars in the vicinity of a UE with the goal to make their behaviour predictable, but evaluations are still ongoing.



**Figure A.44: raytraing versus measurement results: left top: raytracing of subpaths based on simple BVDM; right top: raytraced CIR; left bottom: raytraced CIR evolution over 50cm; right bottom: CIR evolution over 50cm for measured channel**

### Analysis of relevant Subpaths

As mentioned above there are similarities between MBCP and chaos theory and unfortunately channel estimation seems to be similar sensitive to any model errors as known from the famous Butterfly effect [D03]. For the conventional Wiener or Kalman filtering this leads to a very small useful prediction horizon of few tenth of  $\lambda$ , in case high accuracy is required (see e.g. [Aro11])) and Figure A.45. The reason are the high number of multi path components, which are difficult to estimate individually as they overlap in the measurement domain. This is true for prediction based on measurements as well as based on model based channel estimation. Therefore it is planed to combine in the final concopt conventional measurement with model based prediction, where the combiner should take the expected or estimated uncertainties for each scheme into account.

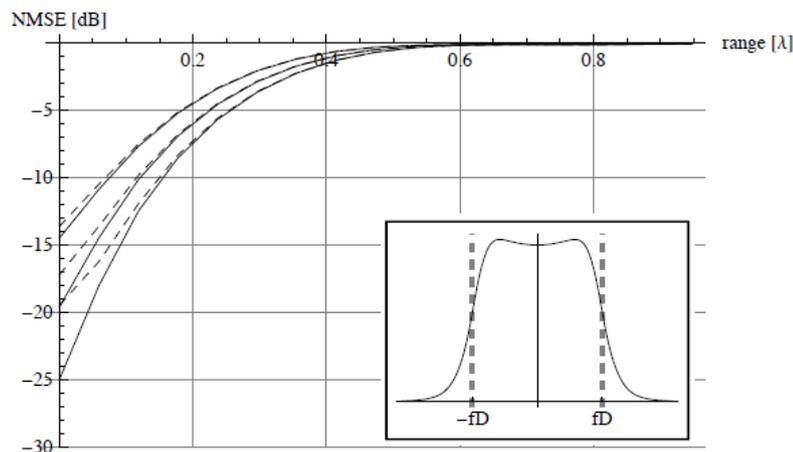
To get a better understanding of the fundamental limits of channel estimation and prediction it is necessary to understand the degree of complexity of a typical radio channel. This complexity is defined by the number of subpaths  $N_{SP}$  per channel component (CC) or more importantly of the mean number of subpaths per tap  $N_m = N_{SP}/N_{Tap}$ , where  $N_{Tap}$  is the number of relevant taps of the sampled CIR.

For large  $N_m$  each tap of the CIR will be the result of the superposition of many subpaths being difficult to estimate. The limited prediction horizon is then just the result of high number of small

estimation errors. Contrary for an ideally single subpath per tap a simple linear prediction for each tap would result in a very large prediction horizon in a stationary scenario. To get a feeling we analyzed the above mentioned raytracing results with respect to relevant subpaths  $N_{SP}$  for different threshold values  $TH_{SP}$ . The reason why we used raytracing results is simply that the measurements do not directly provide any information about the relevant subpaths. Due to the similarities of the length of the measured and raytraced CIR it is expected that the raytracing results hold more or less also in reality.

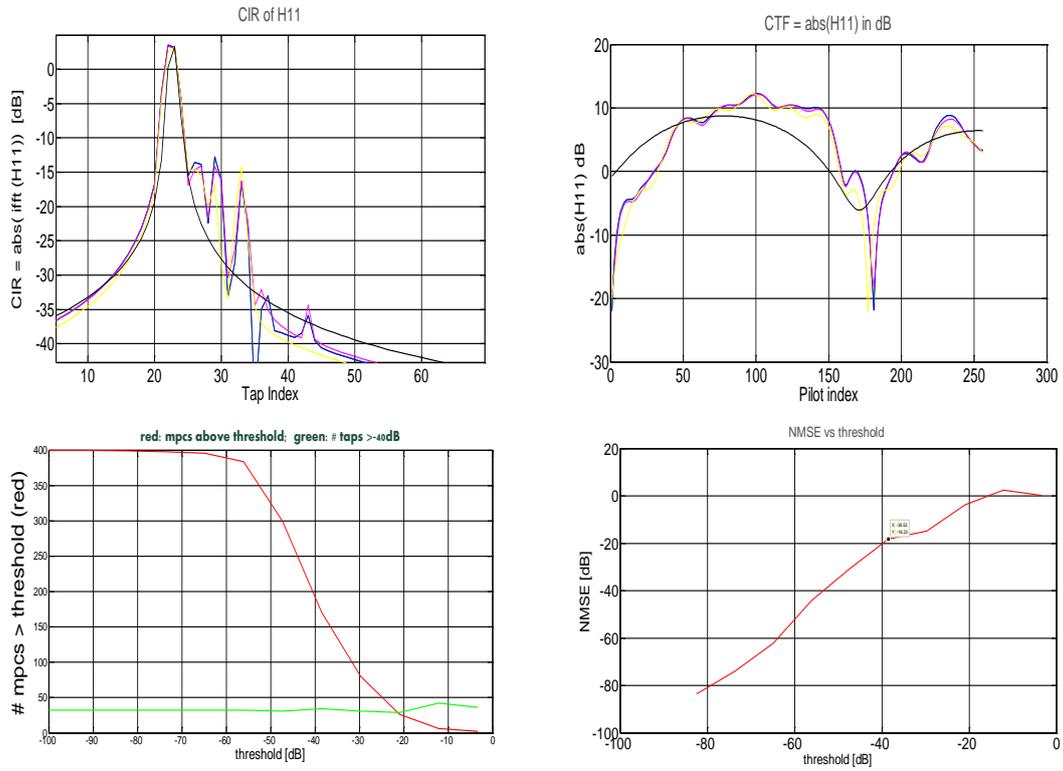
In Figure A.46 some of the results are plotted, e.g. the CIR at the top left and the CTF for decreasing  $TH_{SP}$  at the top right.  $TH_{SP}$  is defined in [dB] relative to the amplitude of the strongest subpath as being provided by the raytracing simulation. The reference is the case when all subpaths – here  $N_{SP} = 400$  – are being used for the calculation of the CIR or CTF. With increasing  $TH_{SP}$  less and less subpaths are being considered. Figure A.46 left bottom shows the number of subpaths above  $TH_{SP}$  and right bottom has the NMSE of the final CIR versus  $TH_{SP}$ . From this it can be concluded that for an NMSE of e.g. -20dB the threshold  $TH_{SP}$  has to be -40dB meaning that about 195 relevant subpaths have to be taken into account. Comparing this with the number of about 35 relevant taps of the CIR - see Figure A.46 top left - there would be around  $N_m = 5.5$  subpaths per tap. This explains the difficulties for a longterm prediction, but gives also hope that there might be means to estimate these limited number of taps in the future more accurately. Especially the combination with MBCP might be fruitful as it hopefully provides some indications on potential subpaths and their evolution over time.

Currently ongoing is the improvement of the BVDM with the goal to get closer to the measurement results and the search for an optimum combining/comparing strategy of measured and modeled CSI information. As a general conclusion it seems that the most interesting part of the research in this area is just starting. For the time being it is unclear at what point in time in the future BVDMs might be available with sufficient accuracy and how much gain it will bring in the end.



Source: Daniel Aronsson

Figure A.45: NMSE over prediction range in  $[\lambda]$  for different channel conditions



**Figure A.46: analysis of relevant subpaths for raytraced channel: left top: main taps of CIR with  $TH_{SP}$  as parameter; right top: same for CTF; left bottom: # of relevant subpaths over  $TH_{SP}$ ; right bottom: achievable NMSE over  $TH_{SP}$**

#### A4-4 Feedback compression

Realistic coordinated multi-point networks (CoMPs) are constrained to limited feedback resources. In this work, the goal is study the effect of CSI quantization and optimal feedback bit allocation on the throughput of the CoMP networks. The results are obtained for different channel models and bit allocation strategies under the block-fading assumption, where the channels remain constant during a codeword transmission period.

##### System model

We consider a communication setup with  $N$  base stations (BSs) and one user. The BSs are assumed to have full cooperation, provide joint transmission to the user and constrained to have the same limited transmission power  $P$ . In this way, the channel can be modeled as

$$Y = \sum_{n=1}^N l_n H_n X_n + Z$$

where  $Y$  denotes the received signal,  $l_n$  is the long-term channel gain between the  $n$ -th BS and the user, consisting of path loss and shadow fading,  $X_n$  is the data transmitted from the  $n$ -th BS and  $H_n$  represents the independent and identically distributed (iid) small-scale fast fading variable between the  $n$ -th BS and the user. Also,  $Z \sim \mathcal{CN}(0,1)$  denotes additive iid complex Gaussian noise plus the interference received from the other base stations/clusters. Moreover, we define  $G_n = |H_n|^2$  which is denoted the  $n$ -th channel gain in the following. Although the results are valid also if the small-scale fast fading channels have different distributions, for simplicity we assume that they experience the same probability density function (pdf). Then, this scenario can be, with no loss of generality, mapped into the case where the channels experience identical pdfs but the BSs have different power constraints  $|X_n|^2 \leq P_n$ , where  $P_n = P |l_n|^2, n=1, \dots, N$ . The gain pdf:s are represented by  $f_{G_n}(g), n=1, \dots, N$ . Finally, the channels are assumed to be known at the receiver, while the BSs are provided with quantized CSI.

##### System throughput with quantized CSI feedback

At the beginning of each fading block, the receiver implements  $N$  deterministic quantization functions

$$C_n(g_n) = i_n \text{ if } g_n \in S_{i_n} = [\tilde{g}_{i_n-1}, \tilde{g}_{i_n}), i_n = 1, \dots, I_n$$

which partitions the gain pdf of the  $n$ -th link,  $n=1, \dots, N$ , into  $I_n$  regions  $S_{i_n}, i_n = 1, \dots, I_n$ . Here,  $\tilde{g}_{i_n}$ 's denote the quantization boundaries where we have  $\tilde{g}_0 = 0, \tilde{g}_{I_n} = \infty$ . Also,  $\pi_{i_n} = \Pr\{g_n \in S_{i_n}\}$  represents the probability that the  $n$ -th channel gain falls into the region  $S_{i_n}$ . Then, the quantization indices are fed back to all BSs. Therefore, assuming unicast to each BS, the total number of feedback bits per block would be

$$B = \sum_{n=1}^N b_n, b_n = \lfloor \log_2 I_n \rfloor + 1$$

where  $\lfloor \cdot \rfloor$  denotes the lower integer value.

Getting the quantization indices  $\{i_n, n=1, \dots, N\}$ , some codeword is created by cooperation of the BSs such that  $Q_{i_1 \dots i_N}$  information nats is sent to the receiver. Therefore, as the length of the codewords is  $L_c$ , the transmission rate would be  $R_{i_1 \dots i_N} = Q_{i_1 \dots i_N} L_c^{-1}$ . The data is successfully decoded by the receiver if the channel gains support the rate, i.e.,  $R_{i_1 \dots i_N} \leq \log(1 + \sum_{n=1}^N g_n P_n)$ , where  $\gamma = \sum_{n=1}^N g_n P_n$  is the instantaneous signal-to-interference-and-noise ratio (SINR) at the receiver. Hence, with some manipulations, the system throughput is obtained as

$$\begin{aligned} \eta &= \sum_{i_1=1}^{I_1} \dots \sum_{i_N=1}^{I_N} \Pr\{g_n \in S_{i_n}, n=1, \dots, N\} r_{i_1 \dots i_N} \\ &= \sum_{i_1=1}^{I_1} \dots \sum_{i_N=1}^{I_N} \left( \prod_{n=1}^N \pi_{i_n} \right) r_{i_1 \dots i_N} = \sum_{i_1=1}^{I_1} \dots \sum_{i_N=1}^{I_N} R_{i_1 \dots i_N} Y_{i_1 \dots i_N} \end{aligned}$$

where

$$\begin{aligned} r_{i_1 \dots i_N} &= R_{i_1 \dots i_N} \beta_{i_1 \dots i_N} \\ \beta_{i_1 \dots i_N} &= \Pr \left\{ R_{i_1 \dots i_N} \leq \log(1 + \sum_{n=1}^N g_n P_n) \mid g_n \in S_{i_n}, n=1, \dots, N \right\} \\ Y_{i_1 \dots i_N} &= \left( \prod_{n=1}^N \pi_{i_n} \right) \beta_{i_1 \dots i_N} = \int_{\tilde{g}_{i_1-1}}^{\tilde{g}_{i_1}} \dots \int_{\tilde{g}_{i_N-1}}^{\tilde{g}_{i_N}} \left( \prod_{n=1}^N f_{G_n}(g_n) \right) \mathbf{1}(g_1, \dots, g_N; R_{i_1 \dots i_N}) dg_1 \dots dg_N \\ \mathbf{1}(g_1, \dots, g_N; R_{i_1 \dots i_N}) &= \begin{cases} 1 & \text{if } \log(1 + \sum_{n=1}^N g_n P_n) \geq R_{i_1 \dots i_N} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

More details can be found in [MLE12].

### Simulation results

Simulation results are presented for Rayleigh-fading and channels with Nakagami-2 gain distributions respectively given by

$$f_G(g) = \lambda e^{-\lambda g}, g \geq 0$$

and

$$f_G(g) = 8\lambda^{-2} g^3 e^{-2\lambda^{-1} g^2}, g \geq 0$$

where we set  $\lambda = 1$ . Notice that, as described in the system model, the effect of user location can be mapped into different transmit power constraints of BSs. Considering three base stations with different mapped transmission powers, Figure A.47 and Figure A.48 respectively demonstrate the system throughput for Rayleigh-fading and Nakagami channels as a function of number of feedback bits. Further, the system throughput with perfect CSI at the BSs is

considered as the system performance upper bound. In each point, the triple  $(b_1, b_2, b_3)$  shows the optimal bit allocation maximizing the system throughput. These values have been determined by exhaustive search among all possible bit allocation strategies. Then, Figure A.49 studies the effect of BSs distances from the user on the system throughput. For this reason, a distance factor  $\alpha$  is considered where, while the first BS power is kept fixed, the other BSs powers change as  $(P_2, P_3) = (10.6 - \alpha, \alpha), \alpha \in [0, 10.6]$ . Simulation results show that:

Considerable performance improvement can be achieved by using joint transmission CoMP with limited number of feedback bits, particularly when the fading severity is not too high (Nakagami distribution, Figure A.48). For instance, for both Nakagami and Rayleigh-fading channels, 95 percent of the system full-CSI throughput is reached by 3 bits per base station feedback. The effect of optimal bit allocation becomes more important when the user experiences different channel qualities from the BSs, i.e., the user is close to the CoMP cluster edge area (Figure A.47 and Figure A.48, for the case of  $P_1 = 1, P_2 = 3, P_3 = 12$ ). However, this effect diminishes when the channel qualities are the same, that is, the user is close to the CoMP cluster center area.

In harmony with practical schemes [ME10], the simulation results show that in almost all cases the optimal bit allocation matches the one obtained by the Zonal sampling approach [ME10] (Figs. 2-4). This is a simple and efficient technique which distributes the bits based on the channel variances (here, base stations mapped transmission powers). Starting from the element with highest power, one bit is given to the component with the highest power and then its power is divided by 4. This procedure continues until all available feedback bits are distributed. Thus, in order to have the highest throughput, the base station with the best channel should get the largest proportion of the feedback bits (Figure A.47 – Figure A.49).

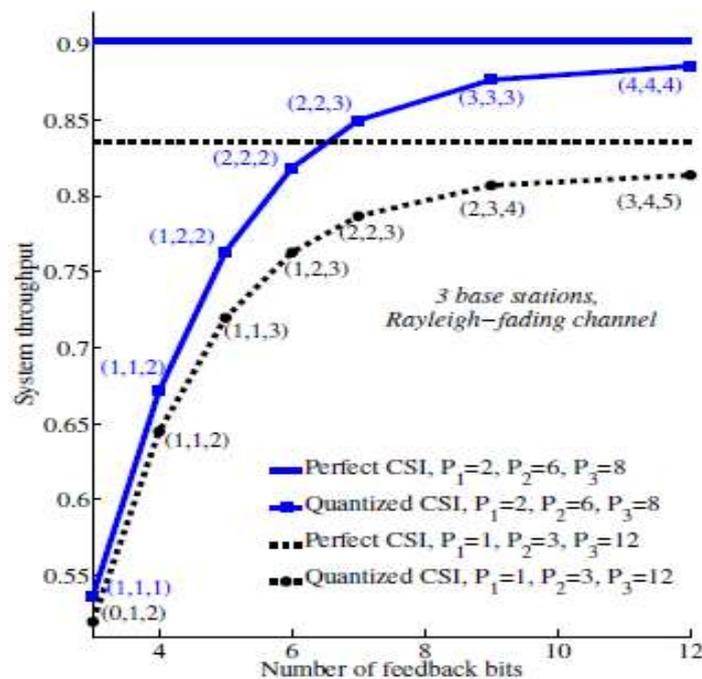


Figure A.47: System throughput vs number of feedback bits, 3 BSs, Rayleigh fading channel.

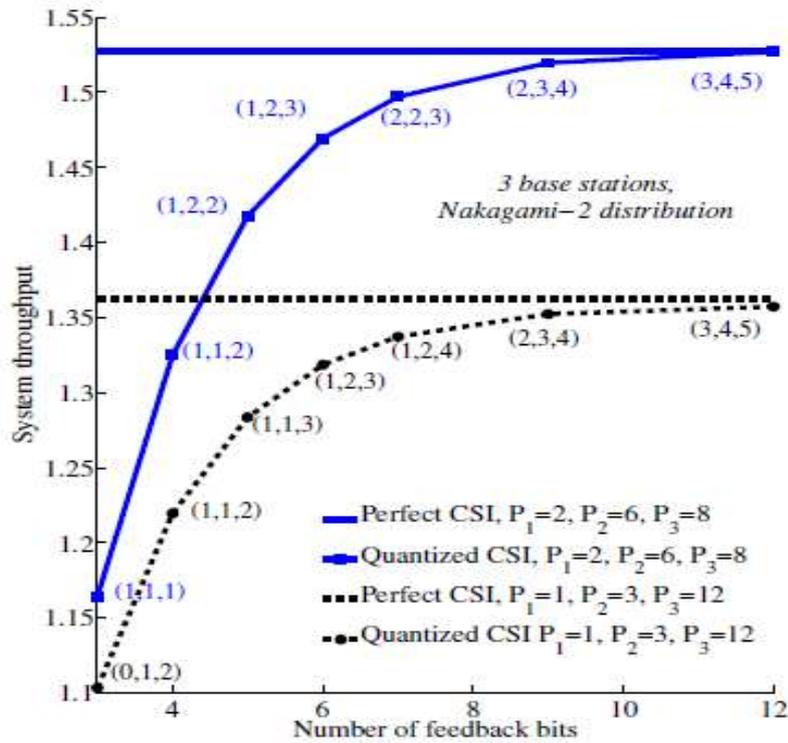


Figure A.48: System throughput vs number of feedback bits, 3 BSs, Nakagami fading channel.

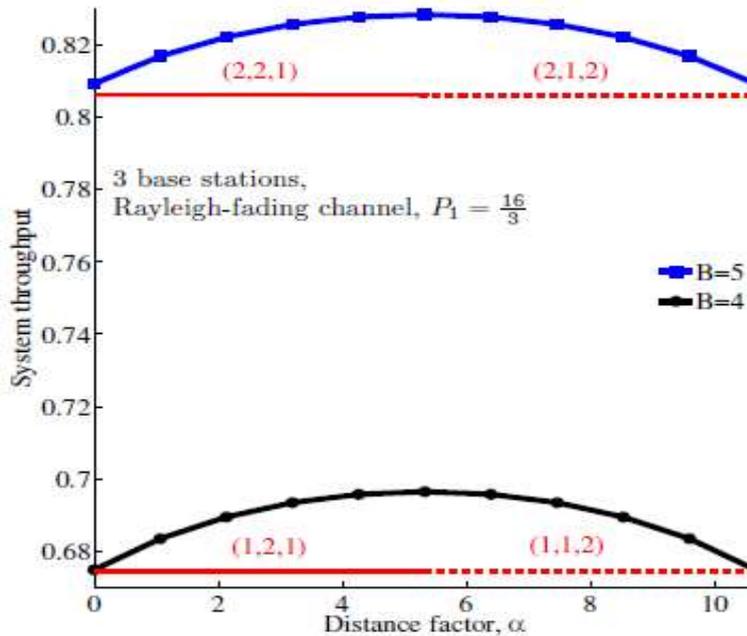


Figure A.49: Throughput vs the distance factor  $\alpha$  with different number of feedback bits  $B$ .

## A4-5 Advanced feedback compression schemes

*[The results in this Appendix are relevant to the discussion in Subsection 5.3]*

Beside the channel prediction as such the important aspect of model based channel prediction is the potentially and ideally extremely reduced overhead for reporting of CSI values. So far it is not clear how much information can be extracted by the proposed scheme in a real world scenario. We are still in the analysis phase of the concept. In the end one will have to combine the model based feedback – i.e. the location information – with some form of efficient delta reporting.

Taking the most advanced interference mitigation concept into account UEs will have to report the most relevant CCs of their cooperation area consisting of 3 sites a' 3cells, i.e. overall 9 cells and in case of 4x2 MIMO for 4 wideband beams per cell. As has been found in [ARTD25] the number of relevant CCs varies per UE between e.g. few to even 10 or 20. Accuracy requirements are extremely high to the above mentioned deliverable with an NMSE in the range of -25 to -30dB per CC. This naturally effects the feedback overhead as well. In chapter 5 a worst and best case feedback overhead has been derived for straight forward reporting versus an ideal MBCP scheme requiring not any further feedback. The final overhead will be somewhere in between, but where exactly is unclear for the time being.

As the main aspects are already reported in chapter 5.3.3 these are not being repeated here once again. Instead in the following a lower bound achievable under ideal conditions for the above described model based channel prediction (MBCP) scheme will be derived.

### Use of Model-based Channel Prediction to reduce feedback load by Parametrized Feedback (MBCP)

The model based channel prediction feedback scheme is used here to estimate a lower bound on the required CSI feedback rate. Theoretically and under ideal conditions, MBCP promises extreme strong compression ratios. MBCP as a stand alone solution would require static channel conditions without any moving objects like cars or tree leaves together with an absolutely perfect model fully reflecting all details of the environment of the UE. Currently research is ongoing to understand the potential and limits of MBCP in real world scenarios, but it is already clear that MBCP will have to be suitably combined with other state of the art channel prediction techniques like the above mentioned Kalman filtering.

The calculation of the minimum feedback rate  $FB_{MBCP}$  is quite different compared to the estimates discussed in Section 5.3.3.. Instead of explicit CSI the UEs feed back their 3-dimensional location vector  $\mathbf{l}_{x,y,z} \in \mathbb{R}^{3 \times 1}$  with respect to a model of the UE environment known by the UE and the eNBs. Additionally the UEs report a moving vector  $\mathbf{v}_m \in \mathbb{R}^{3 \times 1}$  being 3 dimensional as well and being the tangent to the current movement of the UE. This allows the eNBs to predict the locations of the UEs for the near future so that only seldom full location feedback is required. A further 3 dimensional steering or rotation vector  $\boldsymbol{\theta}_{AE} \in \mathbb{R}^{3 \times 1}$  will provide the information of the current boresight direction of the UE antennas.

The location of the UE is assumed to be known with an accuracy of  $\Delta \mathbf{l}_{coarse} = 1\text{m} \times 1\text{m} \times 1\text{m}$ , e.g. from a global navigation satellite system (GNSS) like GPS. For an accurate reconstruction of the wideband CSI an accuracy  $\Delta \mathbf{l}_{fine}$  of about 1mm – i.e. about  $1/100$  of  $\lambda$  - will be needed, ensuring a maximum phase error of  $\leq \pm 1.3^\circ$ . Therefore each element of  $\mathbf{l}_{x,y,z}$  should be reported with  $N_l = \text{round}(\log_2(\Delta \mathbf{l}_{coarse} / \Delta \mathbf{l}_{fine})) = 10\text{bit}$ .

Due to the moving vector  $\mathbf{v}_m$  the feedback rate can be much lower compared to explicit feedback and here we assume a report every 100ms i.e.  $FB_{rate,MBCP} = 10$  reports per second leading to following UE feedback rate:

$$FB_{MBCP}(UE) = 3 * (\mathbf{v}_m + \mathbf{\theta}_{AE} + N_{loc}) * FB_{rate,MBCP} = 900b / s$$

Assuming for example  $K_{cell} = 200$  active UEs the feedback rate per cell would be still very low with:

$$FB_{MBCP}(cell) = FB_{MBCP}(UE) * 200 = 180kb / s.$$

Beside the constant reporting of  $FB_{MBCP}$  there will be typically the need to transfer in a first step the model of a certain environment to the UEs of a cell. In case the model is preloaded from UEs similar as known for navigation maps in car navigation systems this overhead can be avoided, but a broadcast solution has the benefit of being adaptable to changes in the environment. It is expected that it is sufficient to update the model parameters within minutes, which would lead even for very complex models to DL broadcast rates  $BC_{MBCP}$  of few 100 bits/s to few kb/s (area:  $1 \times 1 \text{ km}^2$ ; number of buildings: 100...1000; points per building: 20...100; bits per point: 20; update rate: 300s).

Applying some fancy compression algorithms a further reduction might be possible, but as already noted above MBCP is not seen as stand alone solution. The overhead for  $FB_{MBCP}$  is expected to be the minor part of the overall feedback so that a further compression is probably not so relevant.

## A5. Clustering and user grouping

### A5-1 Clustering and interference floor shaping based on partial CoMP

[The algorithms and results of this section are used in Subsection 5.2.]

Clustering of a network - see chapter 5.2.1 - is an essential part of each CoMP scheme as it turns a potentially interference free system in case of network wide cooperation into an interference limited one due to the inter cluster or inter cooperation area (CA) interference. As for typical macro cellular mobile radio networks this inter cooperation area interference cannot be avoided it is essential to maximize the relation of intra cooperation area signal-  $S_b$  to inter cooperation area interference-  $I_b$  power for all users, being best achieved by user centric defined cooperation areas.

User centric cooperation is a real challenge as it requires that all UEs of a cooperation area have exactly the same set of strongest cells, as JT per definition means common precoding over a given set of cells. For realistic radio channel conditions with strong non line of sight (NLOS) probability, finding of such sets of UEs is extremely seldom, leading to a very low penetration rate (PR) of so called 'happy', i.e. user centric served users. The penetration rate PR is here defined as

$$PR = \frac{\text{number of user centric served UEs}}{\text{all UEs}}$$

and an UE is said to be 'user centric' in case it is served from its e.g. 3 strongest cells. Note other definitions would be possible, but from SL simulations it is known that 3 strongest cells bring already a large part of the potential performance gains [MF11].

Figure A.50 illustrates the main issue of clustering as well as the basic idea of the partial CoMP concept. Visible are several sites  $s_i \in \mathbb{N}^{N_s}$  with three cells  $c_i \in \mathbb{N}^{N_{\text{cell}}}$  each, being the three sectors of the sites.  $N_s$  and  $N_{\text{cell}}$  are the overall number of sites and cells respectively. In case of intra site cooperation (thick arrows) the cooperation areas are defined by the three cells of each site. For line of site (LOS) conditions significant performance gains are achievable as all UEs  $UE_k, k=1..N_{\text{UE}}$ , of a cooperation area  $CA_b \in \mathbb{N}^{N_{\text{ca}}}$  are served from their nearest – and therefore strongest – cells. This is the important general condition for user centric clustering:

$$RSRP(c_{i,b}, UE_k) \geq RSRP(c_i, UE_k), \forall c_i \neq c_{i,b},$$

where the reference signal received signal strength (RSRP) as defined in 3GPP LTE [3GPPTR2] is a measure of the Rx power or equivalently - for fixed Tx power - of the path loss (PL).  $c_{i,b}$  are the cells of cooperation area  $CA_b$  with cooperation area index b while  $c_i$  are all the cells of the network with cell ID i.

Under realistic channel conditions with strong shadowing and many NLOS components equation above will be often violated for intra site cooperation, indicated by the narrow inter site arrows in Figure A.50 requiring a new clustering and a common precoder  $\mathbf{W}$  spanning over several sites s. Optimized clustering is a hot research topic and often known optimization algorithms are adapted accordingly, for example relying on Graph coloring schemes [BSX+10].

*Partial CoMP* has a more structured approach, i.e. in a first step the size of the cooperation areas is expanded from three cells to e.g. three sites a' three cells, i.e. to overall nine cells. From a practical point of view beneficially there are only two inter site backhaul connections and only between adjacent sites, motivating the expansion to nine cells. Nonetheless cooperation areas of size nine are just an example and other criteria might suggest other numbers for the cluster size. Enlarged CAs increase the number of UEs having their e.g. three strongest cells inside the cooperation area - or equivalently the penetration rate PR - significantly. Note, this

definition of PR is a relaxation of equation 1, which requested that all cells  $c_{i,b}$  have to be strongest cells.

Reporting of channel components for nine cells raises a new challenge as it might generate excessive feedback overhead. For that reason as a further means only the e.g. three strongest from the overall nine cells of a cooperation area are reported by the UEs. By the way this *partial* reporting was the motivation to call the scheme *partial CoMP*. Further below a more detailed analysis will reveal that reporting of three channel components (CC) is often too restrictive. Better is to report e.g. all CCs above a certain threshold leading to a somewhat higher reporting effort, but more predictable performance. The maximum number of CCs per CA depends on the number of Tx- and Rx- antennas per eNB and per UE. Assuming a 4x2 MIMO system the cooperation area would have overall 72 CCs.

The main motivation for the threshold based scheme is that unreported low power channel components will generate intra cooperation area precoding errors, even so these errors will be typically relatively small as they are per definition low(er) power channel components. Enlarging cooperation areas increases the penetration rate but leaves an important challenge unsolved. There is still significant inter cooperation area interference – indicated in Figure A.50 by double line arrows - and unfortunately this typically spoils most of the JT performance gains. Note, this inter cooperation area interference will be there even for very large cooperation areas, which are due to geometrical reasons always a large part of the overall area of the cooperation area.

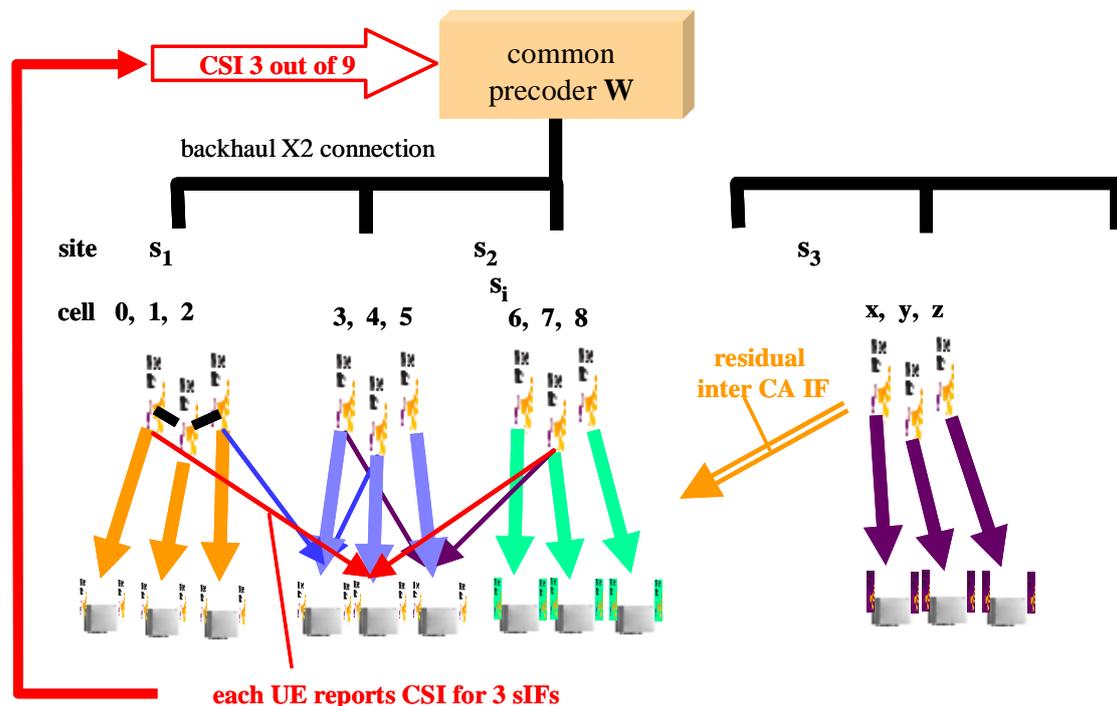


Figure A.50: partial CoMP over 3 sites / 9 cells and UEs reporting their 3 strongest interferers each

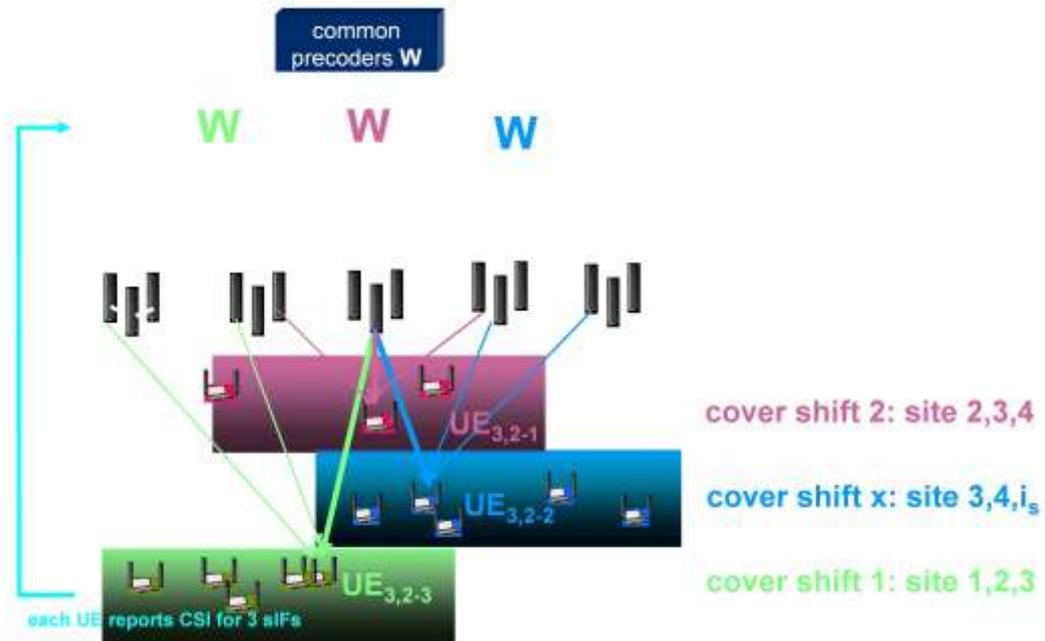


Figure A.51: overlapping cooperation areas or cover shifts

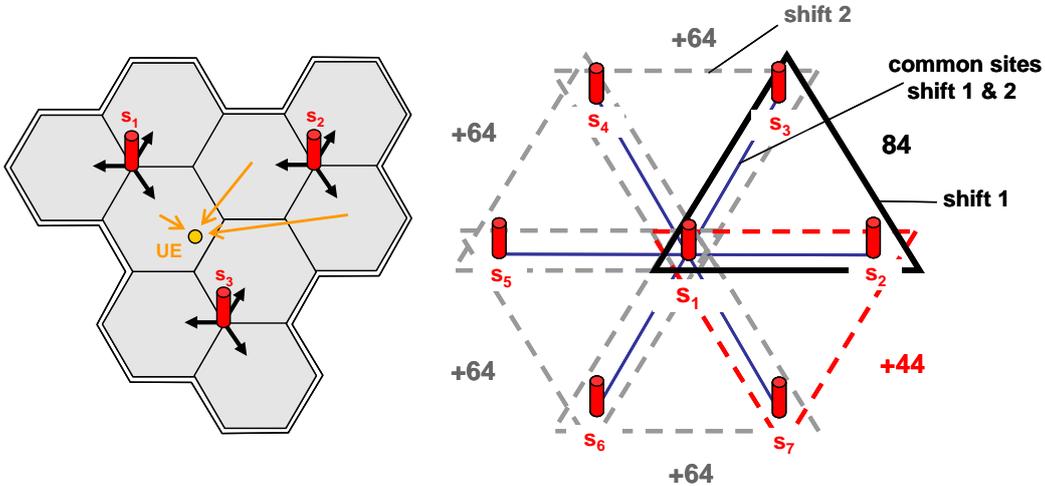


Figure A.52: *left* single CA of 3 sites; *right* 6 cover shifts for  $s_1$ . Each triangle defines the 3 sites for cooperation with the solid triangle defining CS  $CS_1$ .

To deal with inter cooperation area interference we have to open up a new dimension for optimization, i.e. overlapping cooperations areas, which we call cover shifts.

Under the assumption of a relative high number of users we can group UEs into different resources like different subbands and/or subframes, what we call in the following cover shifts  $CS_o$ ,  $o=1\dots6$ . Into each cover shift only cooperation area centric UEs will be scheduled, i.e. those UEs having all there e.g. 3 strongest interferers within one of the cooperation areas of

$CS_0$ . As illustrated in Figure A.51 in case of six<sup>51</sup> cover shifts  $CS_0$  almost all UEs can be served without inter cooperation area interference, at least with respect to the e.g. three strongest interferers. Figure A.51 sketches three different cover shifts/frequency subbands (magenta, light blue and green). For each cover shift only one cooperation area a' three sites is being depicted, but each cover shift will neatly pave the full network by similar but non overlapping cooperation areas.

Each site is involved in all - and we will see later for homogeneous networks it should be six - cover shifts. For illustration see site 3 in Figure A.51. All UEs being served from site 3 are connected to site 3 by at least one bold arrow indicating that site 3 contains at least one of the strongest cells.

The scheduler in site three makes a pre allocation of UEs to cover shifts based on the reported three strongest cells, indicated by the colored lines. All UEs having reported cells from site 2, 3 and 4 will be scheduled into cover shift 2, from site 3, 4,  $c_i$  in cover shift  $CS_0$  and so on. In the end each site serves all cover shifts simultaneously, meaning that it will be a real frequency reuse one system. From the figure it is clear that cover shifts are just differently predefined setups of cooperating sites in different resources like frequency subbands and/or time slots, while eNBs still keep full flexibility to schedule UEs into any cover shift.

Generally UEs  $UE_k$  are scheduled into their best fitting  $CS_0$  based on the reported set of strongest cells  $\{c_{i,r}(1) \dots c_{i,r}(N_r)\}_k = \arg(\max_{c_i} (RSRP(c_i, UE_k)))$ .

$c_{i,r}$  are the indices of the reported cells, e.g. of the  $N_r=3$  strongest cells seen by the UE  $UE_k$ . In a second step a per cover shift scheduler has to find for each physical resource block (PRB) of  $CS_0$  the best fitting groups of simultaneously served UEs, maximizing the multi user (MU) throughput, basically defined by the mutual spatial orthogonality of the UEs.

Figure A.52 helps to understand the much higher penetration rate PR of user centric served UEs for *partial CoMP* compared to conventional setup of cooperation areas. The main issue is that conventionally the goal is to find  $N_r$  UEs simultaneously being served by  $N_R$  identical strongest cells, i.e. in case of  $N_R=3$  there have to be three UEs seeing the same three cells as strongest interferers.

$$N_c(1) = \frac{9!}{6!3!} = 84;$$

$$N_{c,overlap}(n_{c-1}, n_c) = \frac{6!}{3!3!} = 20;$$

$$N_c(n_c) = N_c(1) - N_{c,overlap}(n_{c-1}, n_c) = 64; \quad n_c = 2, \dots, 5;$$

$$N_c(6) = N_c(1) - 2N_{c,overlap}(n_{c-1}, n_c) = 44;$$

$$N_c(1 \dots 6) = \sum_{n_c} N_c(n_c) = 384;$$

For *partial CoMP* beneficially all UEs with three strongest interferers out of nine cells can cooperate. According to the equation above evaluating the binomial coefficient  $n$  over  $k$  for  $n = 9$  and  $k = 3$  strongest cells, this is fulfilled for  $N_c(1)=84$  cell combinations for each cover shift (the 1 indicates one cover shift). Figure A.52 right illustrates the overall servable cell combinations  $N_c$  with respect to site  $s_1$ .  $s_1$  is connected with its adjacent sites  $s_2$  to  $s_7$  forming the different

<sup>51</sup> six cover shifts allow to form all possible CAs of size three sites with all directly adjacent sites of an eNB. It covers all border areas of the basic CA as CA center areas in one of the other cover shifts. All UEs with strongest cells spreaded over a maximum of three sites will be served user centric.

cooperation areas for the different cover shifts  $CS_o$  indicated by according triangles. Let's take the solid line black triangle as basis cooperation area of  $CS_1$  comprising sites  $s_1$ ,  $s_2$  and  $s_3$ .  $CS_1$  covers  $N_c(1)=84$  different combinations of three strongest cells. Overall we have six cover shifts to serve an UE, but between two adjacent cover shifts there are always two common sites. Therefore all cell combinations three out of six for these two sites are common between adjacent cover shifts. To calculate the overall possible cell combinations  $N_c(1...6)$  for all six cover shifts serving an UE user centric with three strongest interferers, one has to subtract these common cell combinations  $N_{c,overlap}(CS_{o-1}, CS_o)$ . The last cover shift has common sites with the left and right cover shift and therefore has to subtract  $N_{c,overlap}$  two times. Finally  $N_c(1...6)$  can be calculated to 384. This leads to a very high penetration rate PR of CoMP UEs (UEs gaining by JP CoMP), as each UE can select the best out of 384 different 3-cell combinations, i.e. there is a high likelihood for finding a suitable cell combination. At the same time the sites of the network are connected only between direct neighbors (blue lines) avoiding very complex network structures. Especially a permanent restructuring of the backbone connections as being required for many cluster optimizations can be avoided.

In combination with strong antenna tilting penetration rates  $P_3$  up to 90% – the subscript 3 indicating number of strongest interferers - have been found [MZ11]. This might be further improved by an interference rejection combining (IRC) receiver at the UE, allowing to cancel one more – potentially far off – strongest interferer with according performance gains of about 3dB higher SINR [TJ08].

In reality there will be load variations for different cover shifts, leading potentially to under utilization of some of the resources. Fortunately some of the UEs can be scheduled into different cover shifts without performance loss, i.e. those UEs being served by one single site (intra site cooperation) or only 2 sites. The according relative probabilities  $P_c(1s)$ ,  $P_c(2s)$ ,  $P_c(3s)$  for clusters covering 1, 2 or 3 sites can be found in the equation further below.  $P_c(1s)$  is for the most valuable intra eNB UEs, i.e. those which can be scheduled into any cover shift. Assuming similar probability for each cell combination  $P_c(1s)$  would be very low and just 3.6%.

Fortunately in real macro networks the collocation of cells at one site in combination with distance dependent path loss leads to the much higher probability  $P_{ch}(1s)$  of about 40% as can be concluded from many system level simulations. Note, the subscript 'ch' of  $P_{ch}$  indicates realistic channel conditions while the subscript 'c' indicates the assumption of a uniform distribution over all potential cell combinations.  $P_{ch}(2s)$  is probably in the range of 30-40% and allows to schedule UEs at least into 2 different cover shifts. If there are still load imbalances then some few UEs will have to be scheduled into second best cover shifts.

$$\begin{aligned}
 P_c(1s) &= k_1/N_c(1) = 3/84 = 3.6\% ; & P_{ch}(1s) &\approx 40\% ; \\
 P_c(2s) &= k_2/N_c(1) = 40/84 = 47\% ; & P_{ch}(2s) &\approx 30 - 40\% ; \\
 P_c(3s) &= k_3/N_c(1) = 41/84 = 49\% ; & P_{ch}(3s) &\approx 20 - 30\% ;
 \end{aligned}$$

Note, instead of frequency subbands any other set of orthogonal resources like time domain or a combination of frequency and time domain is possible.

### **Interference floor shaping**

By using overlapping cooperation areas - as being introduced by the cover shift concept - inter cooperation area interference is significantly reduced. But, this is true only with respect to the e.g. three strongest cells. Homogeneous networks in urban environments exhibit strong path loss variations leading to a high number of received cells with quite similar Rx power. Cell center UEs are less affected due to their strong RX power from their serving cells, but cell edge UEs suffer significantly from the inter cell interference floor. The challenge is the high number of

interferers together with the distribution of interfering cells over large geographical areas. Integrating far off cells into cooperation areas is more or less impossible as it would lead very fast to network wide cooperation.

Figure A.53 illustrates the detrimental effect of a strong inter cell interference floor for an interference limited scenario - i.e in case of a very low noise floor - exemplary for the rate region of two UEs. The small blue rectangular is the rate region for the conventional cellular system. Without interference floor and under the assumption of a very low noise floor the black line indicates the achievable rate region with  $R1_{max}$  and  $R2_{max}$ , ideally being just limited by the maximum modulation and coding schemes (MCS).

Assuming now a strong interference floor indicated by the red lines most of the CoMP gains will be concealed leaving only minor rate gains (see red area). For that reason a simple and robust inter cell interference floor shaping technique has been of great interest for a long time.

Antenna tilting is being used since many years for optimized network planning. In case of LTE Release 10 for SL simulations a considerably strong tilting of  $15^\circ$  has been defined with quite impressive performance gains. However, per cell tilting is a compromise between reducing inter cell interference and maximizing signal power from the serving cell.

In combination with the above mentioned partial CoMP concept we found a much more powerful solution, the so called 'Tortoise' concept as being illustrated in Figure A.54. It uses a cell specific vertical antenna tilting per cooperation area with the goal to minimize outbound interference and to maximize inbound signal strength. For that purpose each cell forms one, two or more wideband (WB) beams, which are semistatically or statically setup and being formed by according WB precoders, similar as known from the grid of beam concept:

Outbound beams are given a strong down tilt of e.g.  $T_{OB}=15^\circ$  and inbound beams use a small vertical tilt of e.g.  $T_{IB}=7^\circ$ . As a further means one might reduce the Tx power  $P_{OB}$  of the outbound beams by e.g.  $P_{OB} = P_{IB} - 3-6$  dB, concentrating the Rx power even further into the center of the cooperation area. As a result we achieve a tortoise like shape of the interference power, i.e. the interference leakage outside the cooperation area decreases significantly within a short distance.

The tortoise concept makes only sense in combination with the above mentioned partial CoMP scheme. For partial CoMP in each cover shift only the cooperation area centric UEs are being served. Therefore it doesn't matter that cooperation area edge UEs might suffer from stronger outbound tilting, as these UEs can be scheduled into another cover shift. In addition the low inbound tilting would normally generate strong inter cell interference, but this will be fully cancelled by the intra cooperation area precoder and even helps to improve the coverage.

The tortoise concept affects the effective channel matrix  $\mathbf{H}_b$  of the UEs  $UE_k$  of the cooperation area  $CA_b$  according to the equation below by the antenna tilting matrix  $\mathbf{A} \in \mathbb{C}^{K \times M}$  and the unitary block diagonal  $\mathbf{V} \in \mathbb{C}^{K \times M}$ . The elements  $a_{km}$  reflect mainly the beam individual downtilt dependent path loss as well as the beam individual Tx power  $P_{IB,OB}$ . Further it includes the tilt depending variation of multi path components. 'o' stands for element wise multiplication. The diagonal blocks of  $\mathbf{V}$  contain the cell specific unitary precoders for the static WB beams.

$$\mathbf{y}_b = (\mathbf{H}_b \circ \mathbf{A}) \mathbf{V} \mathbf{W} \mathbf{x}_b + \mathbf{N};$$

$$\mathbf{W} = \text{pinv}((\mathbf{H}_b \circ \mathbf{A}) \mathbf{V}) = \frac{(\mathbf{H}_b \circ \mathbf{A}) \mathbf{V}}{[(\mathbf{H}_b \circ \mathbf{A}) \mathbf{V}]^H (\mathbf{H}_b \circ \mathbf{A}) \mathbf{V}}$$

Here  $\mathbf{y}_b$  is the receive vector,  $\mathbf{x}_b$  the user data vector for all UEs in cooperation area  $CA_b$ ,  $\mathbf{W}$  being the ZF precoding matrix and  $\mathbf{N}$  the additive white Gaussian noise (AWGN).

The above mentioned tilt values are just examples, having been used for a macro cellular environment with an inter site distance of 500m. They have been chosen based on some rule of

thumbs, but could be justified by a simple ray tracing simulation. Figure A.55 provides the decay of Rx power over distance for different tilt angles of  $7^\circ$ ,  $10^\circ$  and  $15^\circ$ . The goal is to serve all UEs with the highest MCS of 64QAM, requiring approximately a SINR of 20dB. Consequently the inter cooperation area interference within each cover shift should be ideally less than 20 dB as well. Each cover shift serves mainly cooperation area centric UEs and schedules edge UEs into a better fitting cover shift. The distance between two cooperation area centers is just the inter site distance (ISD) of 500m. From geometrical reasoning one can conclude that the Rx-power of outbound beams should be within the inter site distance of 500m and that of inbound beams within 1000m below the power of the serving cell at 250m, the half ISD. As can be concluded from Figure A.55 the combination of  $7^\circ$  and  $15^\circ$  tilting fulfills just these requirements.

Ray tracing simulations as well as real world outdoor measurements verified the basic functionality of the tortoise concept, especially the effect of vertical tilting was quite close to the results in Figure A.55, even so with some deviation in the very close proximity of the Tx-antennas. This effect is not fully understood yet, but does not really affect the overall outcome.

### Performance Analysis

The benefits of the proposed concepts so far can be expressed by making a high level comparison between network centric-, full-, partial CoMP and conventional cellular networks with respect to the achievable signal to interference and noise ratios (SINR).

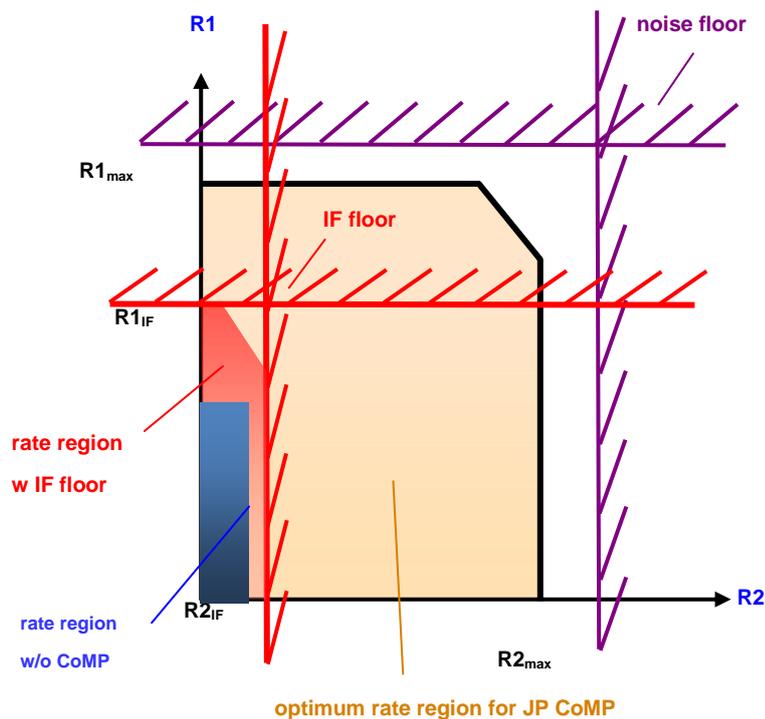


Figure A.53: rate regions with and without IF floor

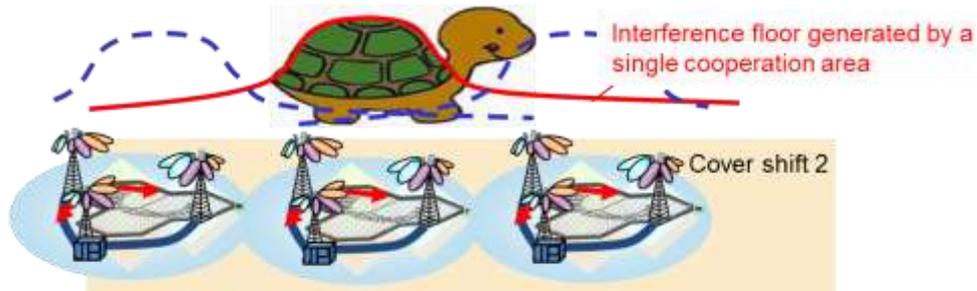


Figure A.54: power and tilt allocation of tortoise concept. Size of beam indicates Tx power.

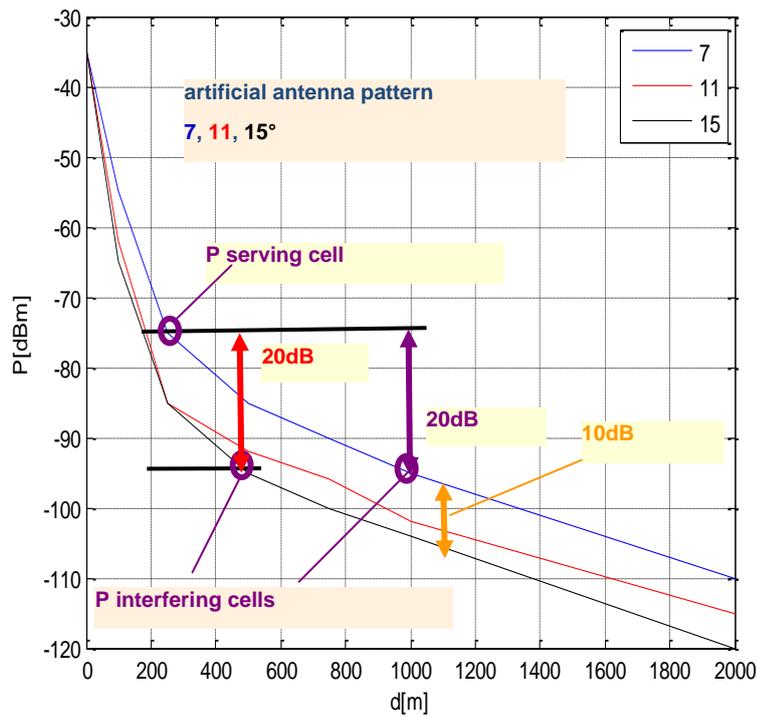


Figure A.55: effect of vertical tilting to Rx power over distance  $d$  from one eNB site with vertical tilt in  $[\text{°}]$  as parameter as outcome of a ray tracing simulation.

In the equations further below the achievable SINRs of the proposed schemes are compared to each other with the conventional single cell transmission ( $\text{SINR}_{\text{SC}}$ ) as reference case,  $\text{SINR}_{\text{intra}}$  for network centric cooperation over the cells or sectors of a single site  $s_i$ ,  $\text{SINR}_{\text{full}}$  for full cooperation over an enlarged cooperation area  $\text{CA}_b$  of e.g. size nine cells and  $\text{SINR}_{\text{partial}}$  for a partial reporting of the channel components of  $\text{CA}_b$ .

The overall network consists of  $N_{\text{cell}}$  cells and  $K=N_{\text{cell}}$  UEs, where each cell  $c_i$  is assumed to serve exactly one UE  $\text{UE}_k$ . Additionally it is assumed that the serving cells are for these UEs the strongest cells (perfect handover) and that we have full buffer traffic. For evaluation of basic clustering effects we use here simple single antenna eNBs and UEs. The resulting overall channel matrix is  $\mathbf{H} \in \mathbb{C}^{N_{\text{cell}} \times N_{\text{cell}}}$  with  $h_{km}$  as channel component from cell  $c_m$  to UE  $\text{UE}_k$ . The precoding weights  $w_{mk}$  of the precoding matrix  $\mathbf{W} \in \mathbb{C}^{K \times M}$  are normalized per UE to avoid increased interference power into other network areas and naturally to fulfill the per eNB power constraint  $P_{\text{Tx}}$ , i.e.

$$\sum_M |w_{mk}|^2 = 1;$$

In case of a cooperative transmission we use a simple zero forcing (ZF) precoder  $\mathbf{W}=\text{pinv}(\mathbf{H}_b)$ , where  $\mathbf{H}_b \in \mathbb{C}^{K \times M}$  is the channel matrix of the cooperation area  $\text{CA}_b$ . Analytical wise ZF is beneficial as cancelled interference terms can be clearly identified, but one should keep in mind that ZF is in many scenarios not the optimum precoder [SH09]. The term  $N$  in the equations below is the additive white Gaussian noise (AWGN).

### ***Single cell transmission***

Sub-equation (a) provides the geometry factor as reference case as being defined in [3GPPT2]. For the geometry factor there exist a lot of SL simulations including various antenna tiltings. It is therefore very useful for the calibration of the SL simulator (see results further below). The precoding weights  $w_{mk}$  reflect a potential link adaption. For SL simulations further below they are set to '1'.

### ***Intra site Cooperation***

Sub-equation (b) provides  $\text{SINR}_{\text{intra}}$  for intra site cooperation being the simplest network centric clustering scheme. The precoder includes all cells of the site and all other sites are acting as interference power.

### ***Full Cooperation***

$\text{SINR}_{\text{full}}$  in sub-equation (c) is for full cooperation over 3 adjacent sites. The term 'full' indicates that all 9 channel components  $h_{km}$  of the cooperation area  $\text{CA}_b$  will be reported. Interestingly sub-equation (c) is true for conventional clustering as well as for the case of overlapping cooperation areas according to the cover shift concept. This is a good verification that the concept is mainly a specific scheduling solution, i.e.  $\text{UE}_k$  is scheduled into the cover shift with the highest number of strongest channel components. Thereby the sum in the nominator is maximized and the sum in the denominator minimized.

### Partial Cooperation

$SINR_{\text{partial}}$  for partial cooperation in sub-equation (d) has basically the same setup as full cooperation with the main difference that the number of reported channel components is limited to the cells  $c_r$ , where the  $r$  stand for reported cells. As only a part of the cells of the cooperation area  $CA_b$  are being reported the sum in the numerator is decreased slightly and more importantly the denominator is increased by the unreported interference terms.

$$SINR_{SC}(UE_k) = \frac{h_{km}w_{mk}}{\sum_{l:l \neq m} h_{kl}w_{lk} + N}; l = 1..N_{\text{cell}}; m = \text{-serving cell}; \quad \text{a)}$$

$$SINR_{\text{intra}}(UE_k) = \frac{\sum_{m \in c_i(s_i)} h_{km}w_{mk}}{\sum_{l \neq m, j \neq k} h_{jl}w_{lj} + N}; l = 1..N_{\text{cell}}; j = 1..K; \quad \text{b)}$$

$m = \text{cells of serving site } s_i;$

$$SINR_{\text{full}}(UE_k) = \frac{\sum_{m \in c_i(CA_b)} h_{km}w_{mk}}{\sum_{l:l \neq m} h_{jl}w_{lj} + N}; l = 1..N_{\text{cell}}; j = 1..K; \quad \text{c)}$$

$m = \text{cells of serving CA } CA_b;$

$$SINR_{\text{partial}}(UE_k) = \frac{\sum_{m \in c_r(CA_b)} h_{km}w_{mk}}{\sum_{l:l \neq m} h_{jl}w_{lj} + N}; l = 1..N_{\text{cell}}; j = 1..K; \quad \text{d)}$$

$m = \text{reported cells of CA } CA_b;$

In [MZ12] SL simulations have been done for the above mentioned CoMP schemes and have been compared with the geometry factor as reference. The cumulative distribution functions (CDF) are given in Figure A.56 (see also chapter 5.2.1 Figure 5.6) and have been achieved for the simulation parameters of an urban macro cellular network according to Table 5.1. They verify the above explained general behavior. Noteworthy is that in our simulation tool all cooperation areas are doing real precoding so that a realistic interference floor is being generated. Evaluated is only the central cooperation area. For 57 cells and enlarged cooperation areas of nine cells the interference floor might be too optimistic as potential far off interferers might be missing. As wrap around in combination with clustered cooperation areas is challenging, we verified the results for a 128 cell scenario and there were only minor deviations to the 57 cell setup.

The geometry factor is significantly outperformed by all schemes. Full CoMP achieves the largest gains, but would need the highest reporting effort. Network- or intra site CoMP looses over the whole SINR range almost 10 dB compared to full CoMP. This is especially bad for the cell edge UEs and indicates that simple intra site CoMP is missing a large part of the potential gains. Partial CoMP looses due to the intra cooperation area interference, but mainly in the higher SINR region. Especially above 20dB this loss is of low relevance for systems with limited MCSs, while the cell edge gains help the whole system. For full CoMP the cover shift concept has been applied to minimize inter cooperation area interference. We did simulations, where the cover shift concept has been switched off. The CDF almost fell back to the network centric curve, indicating the importance of overlapping cooperation areas.

Figure A.56 contains two further thick lines. The red one is that of an ideal CDF given a maximum MCS of 64QAM, i.e. all UEs have an SINR of about 20dB. The magenta curve has been achieved for the tortoise concept for the slightly varying setup, but again for one single UE per cell. We see that the partial CoMP in combination with the tortoise scheme provides for about 70% of the UEs this ideal or even a larger SINR. Interference floor shaping in combination with cooperation areas seems to be an extremely powerful tool.

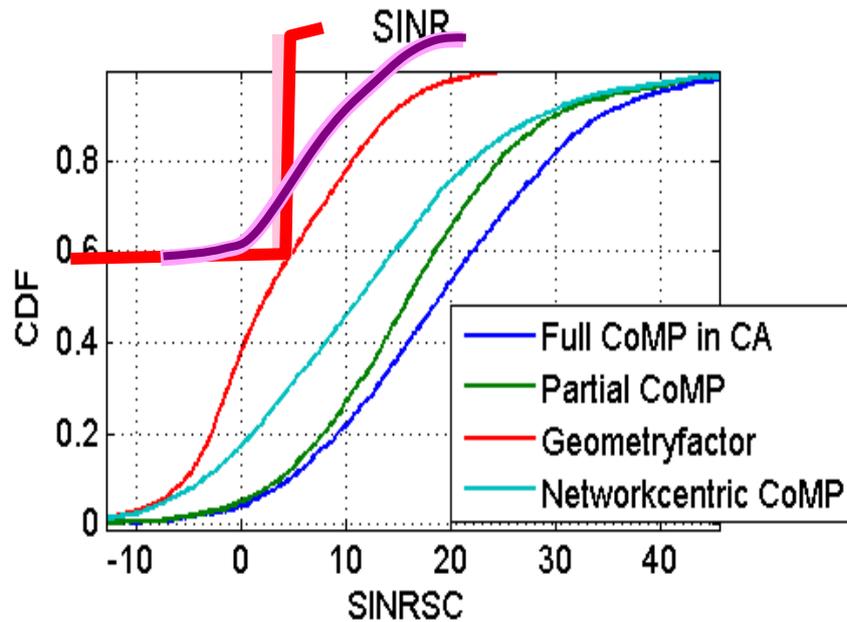


Figure A.56: Optimum CDF for an 'interference free' system (red) and ideally achieved performance with interference floor shaping (magenta). Thin lines are for the other proposed CoMP schemes.

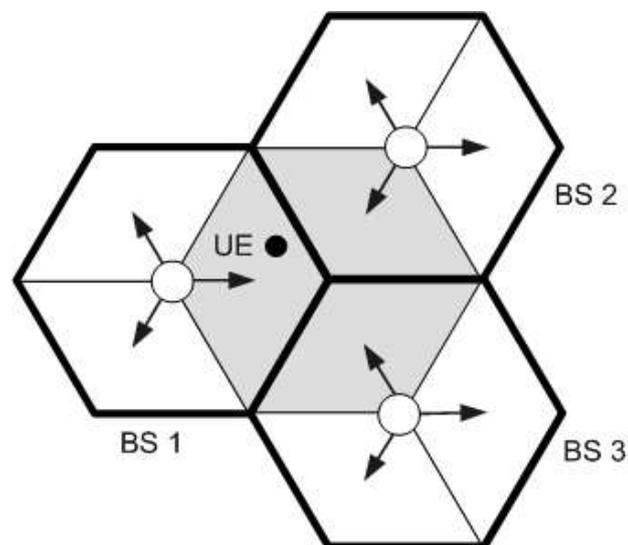
## A5-2 Inter-cluster coordination with fractional frequency reuse

*[The results in this Appendix are relevant to inter-cluster interference control in Subsection 5.2]*

Considering feedback, backhaul and synchronization constraints for a realistic system, the use of joint processing CoMP is usually restricted to a limited number of cells or areas of the system. The system is typically divided into clusters of cells. The cluster formation can be static [ZCA+09], [LXC+10] or dynamic [PA10], [PGH08]. Static clustering specifies a predefined set of clusters that do not change in time, whereas the dynamic clustering forms the clusters based on the varying channel conditions that users experience to different cells. Compared with static clustering, dynamic clustering typically achieves better system performance; however, with higher feedback and backhaul overheads, as well as higher computational complexity. In this section, we consider static clustering. Two fractional frequency reuse schemes are introduced to reduce the inter-cluster interference.

### System Model

Consider the downlink of a multi-cell multi-user OFDMA system with 3-sector base stations (BSs). Each sector has one directional antenna, and it is associated with a directional cell area. The three antennas of each BS are located at the same site. The system is statically divided into disjoint clusters of sectors, where each cluster consists of three neighboring sectors belonging to different BSs (see Figure A.57).



**Figure A.57: An example of a coordinated cluster of 3 neighboring sectors (the shadowed area)**

Due to practical issues, inter-cluster information exchange may not be feasible. Targeting practical scenarios, joint processing CoMP is independently performed in each cluster without inter-cluster information exchange. Assume perfect channel state information (CSI) at the transmitter side, zero-forcing precoding is considered as the joint processing CoMP approach to eliminate the intra-cluster co-channel interference.

Note that a coordinated cluster also introduces inter-cluster interference to the users in the neighbouring clusters, especially to the cluster edge users. Two fractional frequency reuse schemes are introduced here to coordinate the inter-cluster interference [LBS12c].

### User partition process

Users within each cluster are divided into cluster-edge users (CEU) and cluster-center users (CCU), based on the long term channel gain, including passloss and shadow fading. A user partition threshold,  $\Delta I$ , is defined for classifying CEU and CCU.

A user  $k$  in cluster  $c$  estimates and feeds back to its serving sector the long term channel gains from its serving sector and from four candidate neighboring sectors, that is, the two neighboring sectors within cluster  $c$ , and the other two neighboring sectors within the serving BS. For example, the UE in Figure A.57 needs to measure and feedback the long term channel gains from the three sectors belonging to BS1 and from the other two sectors within its coordinated cluster (the shadowed area).

After obtaining these values from user  $k$ , cluster  $c$  finds out the weakest long term channel gain within the cluster (denoted by  $l_k^{in}$  in dB) and the strongest long term channel gain from the two candidate neighboring sectors outside the cluster (denoted by  $l_k^{out}$  in dB). Note that  $l_k^{in}$  reflects the weakest link within the cluster, which is the dominant link that affects the performance gain provided by intra-cluster zero-focusing precoding [ASA12].  $l_k^{out}$  reflects the strongest interference link outside the cluster. Hence, user  $k$  is considered as a CEU if  $l_k^{out} - l_k^{in} \geq \Delta I$ , i.e., the inter-cluster interference would compromise the intra-cluster joint processing gain; otherwise, user  $k$  is regarded as a CCU. One approach to further reduce the feedback would be to obtain  $l_k^{in}$  and  $l_k^{out}$  at the user side. Hence, these values or  $l_k^{out} - l_k^{in}$  could be instead fed back for user partition.

The threshold value,  $\Delta I$ , can be predefined by each cluster or by the network, and it can be a parameter to optimize according to the network design objective.

### Fractional Frequency Reuse Scheme 1 (FFR-1)

Assume  $N$  subchannels are available for the system. The  $N$  subchannels are divided into two orthogonal sets,  $G$  and  $F$ . Subchannels in set  $G$  are assigned for CCU with a frequency reuse factor of one for each cluster. Set  $F$  is further divided into 3 orthogonal subsets, denoted by  $f_1$ ,  $f_2$  and  $f_3$ , which are used for CEUs with a frequency reuse factor of 1/3. This frequency reuse rule, named as FFR-1, is shown in Figure A.58. Note that neighboring clusters assign orthogonal subchannels for CEU, which results in a significant inter-cluster interference mitigation.

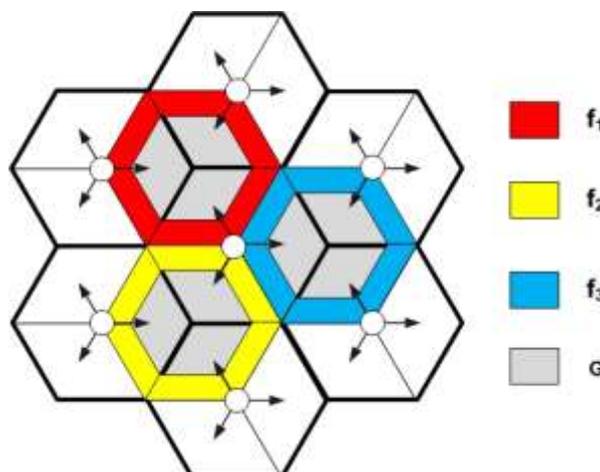


Figure A.58: Frequency reuse rule for FFR-1

## Fractional Frequency Reuse Scheme 2 (FFR-2)

In FFR-1, the reuse factor for CEU is 1/3. Here, we introduce a second fractional frequency reuse scheme, named as FFR-2, where the frequency reuse factor for CEU is 2/3 [LZX+10].

Assume that three neighboring clusters are grouped together and marked as Cluster 1, Cluster 2 and Cluster 3, respectively (see Figure A.59). Given the mark of each cluster, CEU in each cluster are further divided into two types according to their dominant interfering clusters. The dominant interfering cluster for user  $k$  is defined as the one that the neighboring sector with the strongest long term channel gain ( $I_k^{out}$ ) belongs to.

The  $N$  subchannels are divided into two orthogonal sets,  $G$  and  $F$ . Subchannels in set  $G$  are assigned for CCU with a frequency reuse factor of one for each cluster. Set  $F$  is further divided into 3 orthogonal subsets, denoted by  $f_1$ ,  $f_2$  and  $f_3$ . Then,  $f_i$  is assigned for the CEU whose dominant interfering cluster is marked with Cluster  $i$ . This frequency reuse rule, named as FFR-2, is shown in Figure A.59. In FFR-2, two subsets of  $F$  are available for CEU within each cluster. Note that different types of CEU are assigned with different subchannel sets. Thus, the subchannels used for CEU in neighboring clusters are not orthogonal any more. However, the inter-cluster interference coming from its dominant interfering cluster can be eliminated. As illustrated in Figure A.59, the subchannels belonging to subset  $f_2$  can be used for CEU1, when the dominant interference cluster of CEU1 is Cluster 2. Since  $f_2$  is not available in Cluster 2 according to the frequency reuse rule of FFR-2, there will be no inter-cluster interference introduced by Cluster 2 for CEU1.

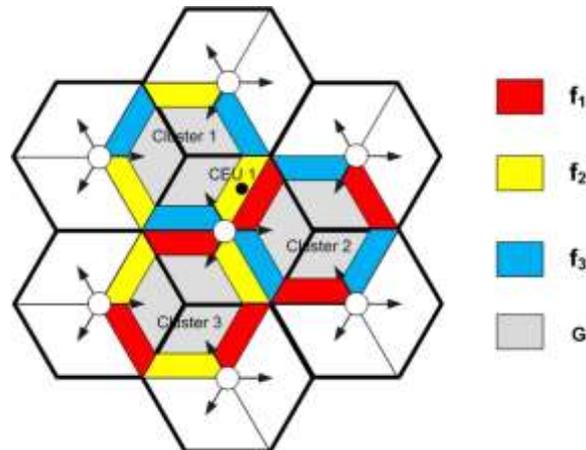


Figure A.59: Frequency reuse rule for FFR-2

## Simulation Parameters

We consider the downlink of an OFDMA cellular system with 19 clusters. A wrap-around technique is adopted to avoid the boundary effect, which causes the clusters in the boundary of the cellular system to receive less interference. The users are uniformly dropped in the cell-edge area. Instead of only considering the interference power from one ring of 6 neighboring clusters [LBS12c], here inter-cluster interference from two rings of 18 neighboring clusters are considered. The main simulation parameters are listed in Table A.17. The following performance metrics defined in 3GPP are considered here [3GPP36942]:

- Throughput cumulative distribution function (CDF) or user average data rate CDF, which is the CDF of the average data rate including all the users in the system.
- Cell-average user data rate ( $R_{ave}$ ), which is the 50% point of the user average data rate CDF.
- Cell-edge user data rate ( $R_{edge}$ ), which is the 5% point of the user average data rate CDF.

**Table A.17: Simulation parameters**

Simulation Parameter	Value
Number of clusters	19
Number of sectors per cluster	3
Number of transmit antennas per sector	1
Number of users per cluster	27
Number of receive antennas per user	1
Maximum transmit power per sector	43 dBm
Bandwidth per subchannel	180 kHz
Number of subchannels	50
Carrier frequency	2 GHz
Radius of cell	500 m
Minimal distance from UE to eNB	35 m
Path Loss Model	$PL(d)=128.1+37.6\log_{10}(d)$ dB, $d$ in km.
Antenna gain pattern	$A(\theta)= -\min[(\theta / 70)^2, 20]$ dB, $\theta = [-180^\circ, 180^\circ]$
Shadowing variance	8 dB
Multipath channel Model	Typical Urban [3GPP36814]

In order to balance the system throughput and the user fairness, a utility function,  $U_k()$  is defined for each user  $k$ . In our simulation, we choose  $R_k(t) / \bar{R}_k(t-1)$  as the utility of user  $k$  in time slot  $t$ , where  $R_k(t)$  denotes the instantaneous data rate for user  $k$  at time slot  $t$ , and  $\bar{R}_k(t-1)$  is the average data rate of user  $k$  at time slot  $t-1$ . Joint processing CoMP is independently performed in each cluster without inter-cluster information exchange. For each cluster, user scheduling is performed to maximize the sum utility of all users within this cluster. The total transmitted power  $P_{\max}$  in each sector is equally pre-allocated to all the available subchannels within the sector. Equal user power allocation is performed to assign transmit power to the scheduled users on each subchannel [LBS+10]. Inter-cluster interference is not considered when performing the user scheduling and power allocation. However, when calculating the achievable data rate of each user, inter-cluster interference from two rings of 18 neighboring clusters is considered.

First, we consider the effects of frequency partition on the performance of FFR-1 and FFR-2. Assume  $|f_1|=|f_2|=|f_3|=|F|/3$ , with  $|F|+|G|=50$ . Note that the subchannels in set  $G$  are used for CCU, while subchannels in set  $F$  are used for CEU. The user partition threshold ( $\Delta$ ) is set to -2dB. The user average data rate CDF curves including all users in the system are plotted for FFR-1 and FFR-2 in Figure A.60 and Figure A.61 respectively. Note that  $R_{\text{edge}}$  is defined based on the user average data rate (the 5% point of the user average data rate CDF including all the users in the system) according to 3GPP, instead of the user category (CEU). Hence,  $R_{\text{edge}}$  is not equivalent to the average data rate of CEU. We can see that there is a tradeoff between the performance of  $R_{\text{ave}}$  and  $R_{\text{edge}}$  when choosing  $|G|$  for FFR-1 and FFR-2.

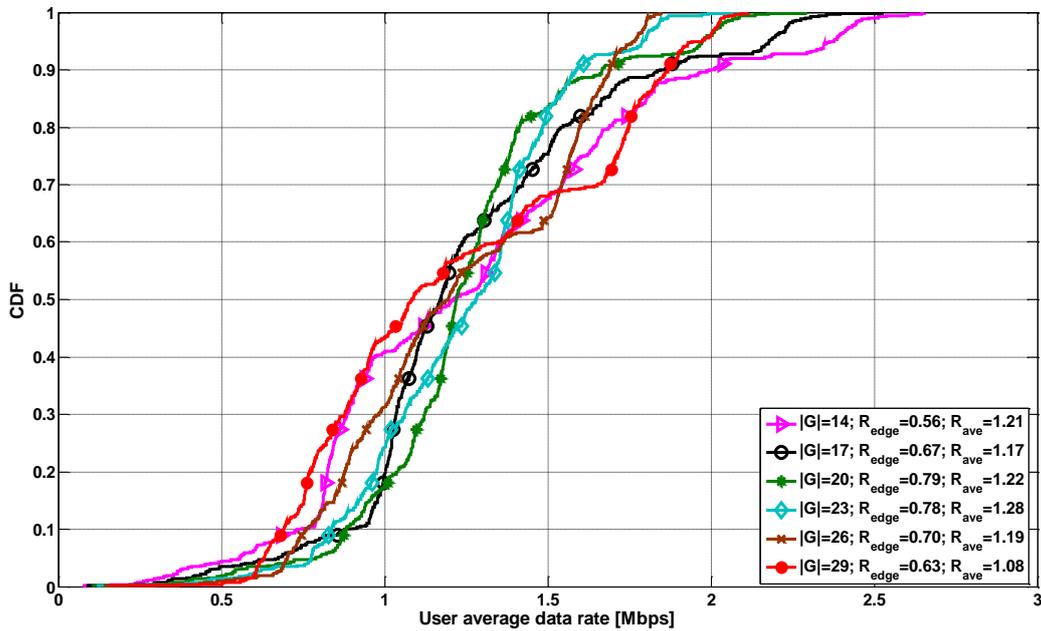


Figure A.60: CDF of user average data rate for FFR-1

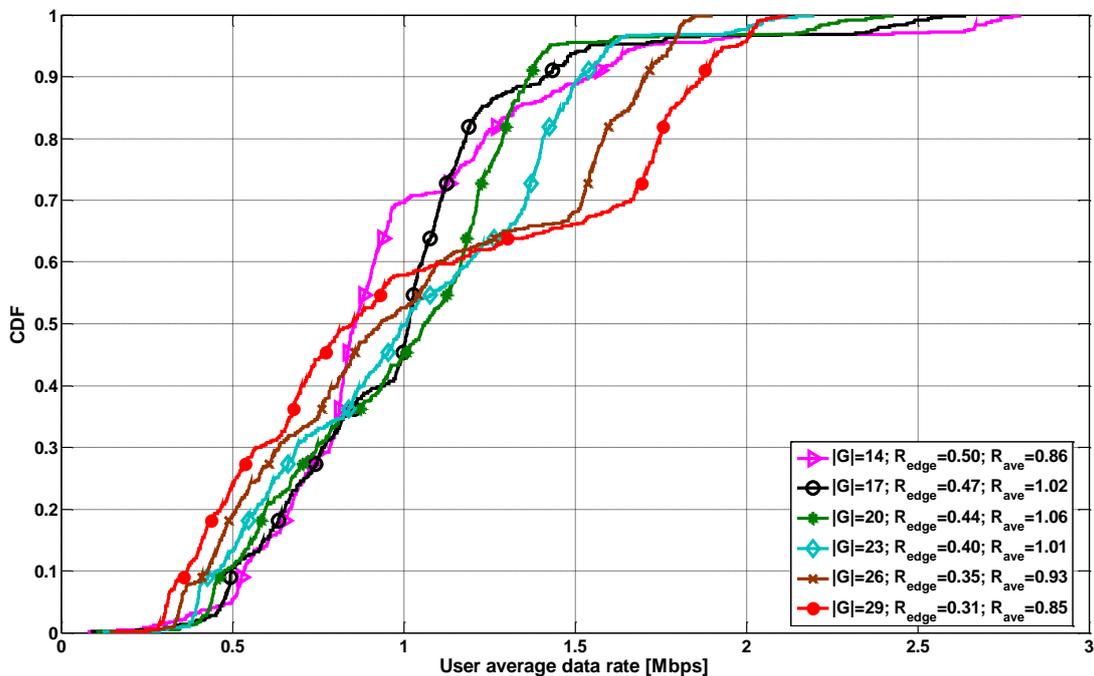
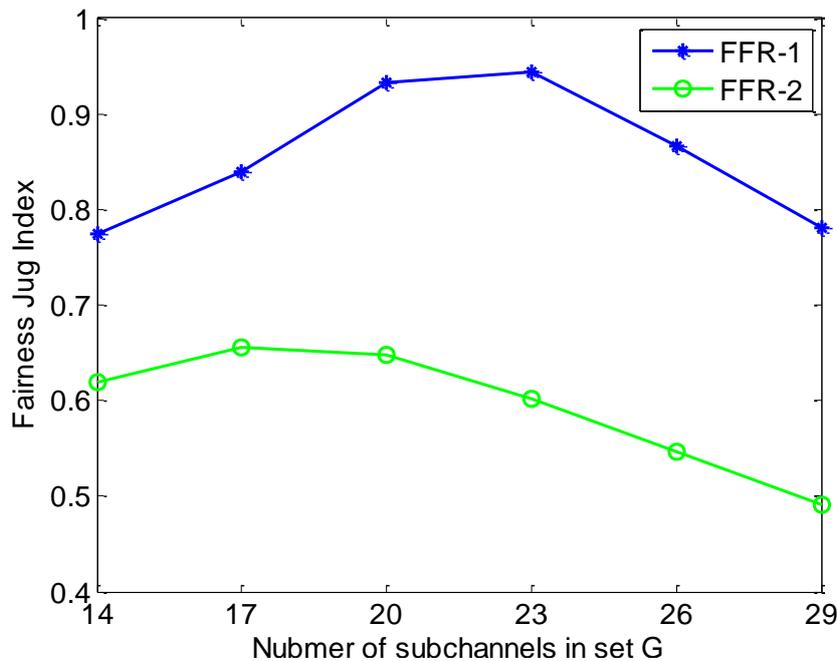


Figure A.61: CDF of user average data rate for FFR-2

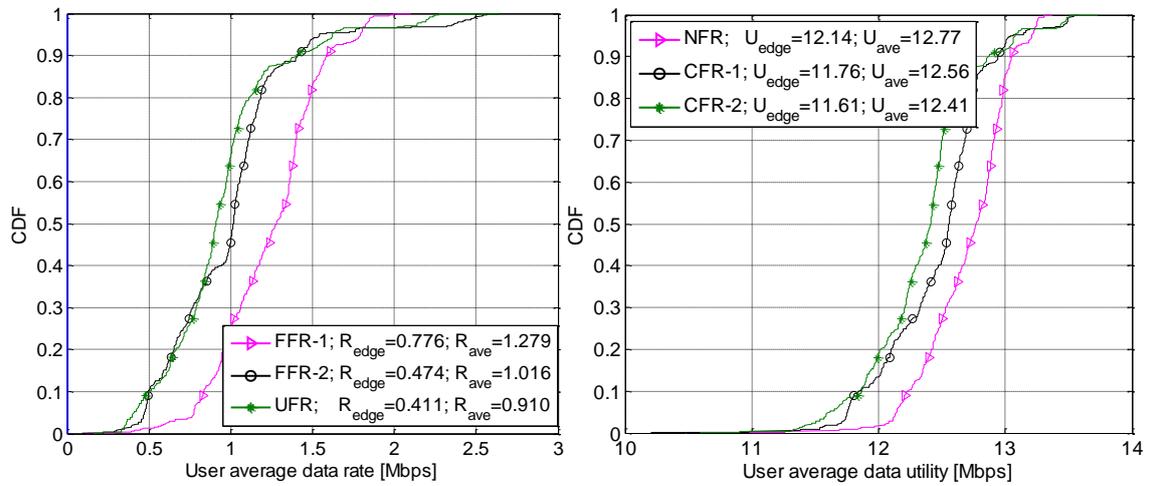
In order to balance the cell-edge and cell-average performance, a fairness jug function with respect to  $|G|$  is defined as  $J(|G|) = \alpha R_{edge}(|G|) + (1 - \alpha) R_{ave}(|G|)$ , where  $\alpha \in [0, 1]$  is a fairness factor reflecting the design objective. As an example,  $\alpha = 2/3$  is selected in this simulation, which means we target the cell-edge performance. The values of fairness jug index  $J(|G|)$  are plotted with different values of  $|G|$  for FFR-1 and FFR-2 in Figure A.62. For each scheme, the user average data rate CDF including all users in the system is first plotted with respect to each

value of  $|G|$ . Then, similar to Figure A.60 and Figure A.61,  $R_{\text{edge}}(|G|)$  and  $R_{\text{ave}}(|G|)$  can be obtained from the corresponding CDF curve. The Fairness Jug Index,  $J(|G|)$ , is finally derived based on the definition. It can be seen from Figure A.62 that the maximum values for both FFR-1 and FFR-2 are obtained when  $|G|=23$ ,  $|F|=9$  for FFR-1 and  $|G|=17$ ,  $|F|=11$  for FFR-2.



**Figure A.62: Fairness jug index with different values of  $|G|$  for FFR-1 and FFR-2**

Finally, we compare these two schemes to the universal frequency reuse (UFR) scheme. In UFR scheme, all subchannels are available for each cluster, and user scheduling is performed independently within each cluster irrespective of the user category (CEU/CCU). Based on the results from Figure A.62, we choose  $|G|=23$  for FFR-1 and  $|G|=17$  for FFR-2. Figure A.63 shows the CDF of the user average data rate (left) and the CDF of the user average utility for these three different schemes (right). Compared with UFR, the cell-edge user data rate of the proposed FFR-2 is improved by 15.41%, while the cell-average user data rate in the FFR-2 is improved by 11.72%. The proposed FFR-1 achieves a much more significant performance improvement compared to the UFR scheme, with about 88.76% increase of cell-edge user data rate and 40.59% increase of cell-average user data rate. Note that although the average data rate and the average utility performance of FFR-2 is close to that of UFR, the complexity of user scheduling and power allocation for FFR-2 is much lower compared with UFR, since only a subset of users is mapped to each subchannel by its fractional frequency reuse rule. In addition, by user partition and frequency partition, the feedback requirements for joint precoding design can be also reduced.



**Figure A.63: CDF of user average data rate for different schemes (left), CDF of user average utility for different schemes (right)**

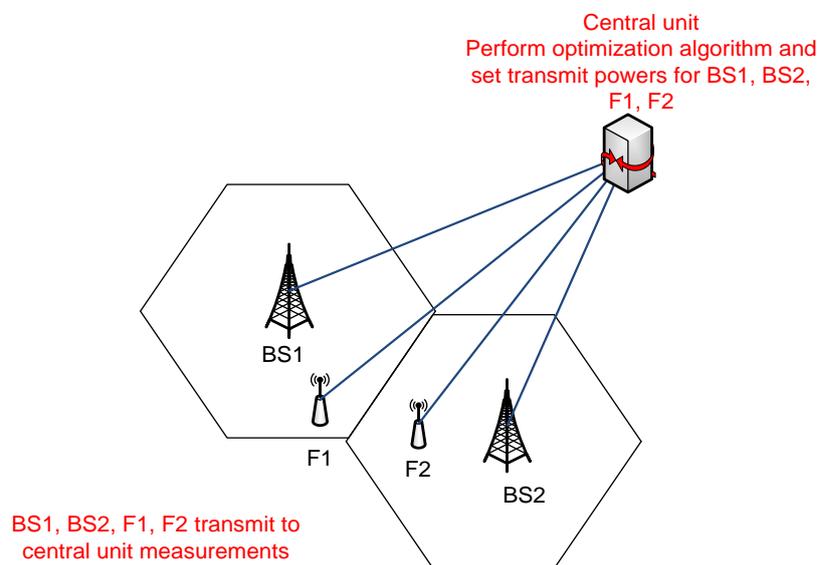
## A6. Inter-Cell Interference Coordination

### A6-1 Coverage control through non linear conjugate gradient optimization

In this section we will introduce and discuss centralized techniques for power setting in the context of heterogeneous networks deployment scenario. More precisely we will focus on campus deployment of femto base stations by considering open and CSG campus deployment.

#### System model:

Consider the general scenario shown in the Figure A.64.



**Figure A.64: general heterogeneous networks scenario**

The macro base stations BS1, BS2 and the femto/relay base stations F1, F2, transmits measurements to the central unit in order to set their powers to optimize the overall coverage of the nodes. The central unit will run optimization algorithm in order to achieve this goal.

#### Optimization problem formulation

The coverage in general is dependent from the overall system parameters, i.e. available PRBs, scheduling algorithms used by the nodes, service type, powers, ..., etc. Coverage control optimization problem can be viewed as the problem of optimizing the worst possible coverage when considering enough PRBs and minimum possible service in each cell of the network.

In this case, the coverage control problem is viewed as the optimization of the cell edge SINRs in the the network such as the minimum service is possible in each cell.

Let us denote the cell edge SINR in the cell  $i$  as

$$SINR(i) = \frac{\alpha_{i,i} P_i}{\sum_{j \neq i} \alpha_{i,j} P_j + N_0} \quad (1)$$

Where:

$P_i$  is the transmit power of the node  $i$

$\alpha_{i,i}$  is the cell edge path gain of the node  $i$

$\alpha_{i,j}$  is the path gain of the interfering node  $j$  over the node  $i$  evaluated in the cell edge of the node  $i$ .

$N_0$  is the additive noise power

Two formulations are possible for the coverage control optimization problem. The first formulation relies on the optimization of cell edge SINRs and the second on the optimization of cell edge spectrum efficiencies that will be approximated by the Shannon capacity formula, i.e. log cell edge SINRs optimization.

#### **Cell edge SINRs optimization problem formulation**

The general formulation of the worst case cell edge SINRs optimization problem is a max-min optimization problem that is given as:

$$\begin{aligned} & \text{maximize } \min_{i \in \{1, \dots, n\}} (SINR(i)) \\ & \text{subject to } P_i \leq P_0 \end{aligned} \quad (2)$$

The non linear minimum function is approximated by generalized mean function as:

$$\min_{i \in \{1, \dots, n\}} (SINR(i)) \approx S_\beta (SINR(1), SINR(2), \dots, SINR(n)) = \left( \sum_{j=1}^n SINR(j)^\beta \right)^{\frac{1}{\beta}} \quad (3)$$

The generalized mean parameter  $\beta$  is considered as low negative parameter and the approximation of the minimum is tight at the limit, i.e.

$$\min_{i \in \{1, \dots, n\}} (SINR(i)) = \lim_{\beta \rightarrow -\infty} \left( \sum_{i=1}^n SINR(i)^\beta \right)^{\frac{1}{\beta}} \quad (4)$$

The Lagrangian of the optimization problem is given as

$$L(P_1, \dots, P_n, \underline{\lambda}) = S_\beta (P_1, P_2, \dots, P_n) - \sum_{i=1}^n \lambda_i (P_i - P_0) \quad (5)$$

The dual parameters  $\lambda_i$  are positive and can be viewed as the average costs associated to the constraints violation. For example when the constraints are not met for the node  $j$ , the

Lagrangian is reduced with  $\lambda_j (P_j - P_0)$  otherwise the Lagrangian is increased. When the constraints are met, the dual parameters are set to zero.

Solving the optimization problem of equation (1) is finding the powers  $P_1, P_2, \dots, P_n$  and the dual parameters that maximize the Lagrangian of the equation (2) and the optimal solution is such that achieves the Karush-Kuhn-Tucker (KKT) conditions [BOY04], i.e. solution for which the gradient of the Lagrangian is minimized and the dual parameters  $\underline{\lambda}$  are zero.

The gradient of the Lagrangian is useful for developing search algorithms for finding the optimal primal/dual parameters.

The gradient of the Lagrangian is given by the following relation:

$$\nabla_{\underline{P}} L = \nabla_{\underline{P}} S_{\beta} - \underline{\lambda}$$

$$\nabla_{\underline{\lambda}} L = -(\underline{P} - P_0)$$

Where the gradient of the generalized mean is given as:

$$\nabla_{\underline{P}} S_{\beta} = \begin{pmatrix} \omega_{\beta,1} & -SINR(2) \frac{\alpha_{2,1}}{\alpha_{2,2}} \omega_{\beta,2} & \dots & -SINR(n) \frac{\alpha_{n,1}}{\alpha_{n,n}} \omega_{\beta,n} \\ -SINR(1) \frac{\alpha_{1,2}}{\alpha_{1,1}} \omega_{\beta,1} & \omega_{\beta,2} & \dots & -SINR(n) \frac{\alpha_{n,2}}{\alpha_{n,n}} \omega_{\beta,n} \\ -SINR(1) \frac{\alpha_{1,3}}{\alpha_{1,1}} \omega_{\beta,1} & -SINR(2) \frac{\alpha_{2,3}}{\alpha_{2,2}} \omega_{\beta,2} & \omega_{\beta,3} & \dots \\ -SINR(1) \frac{\alpha_{1,n}}{\alpha_{1,1}} \omega_{\beta,1} & -SINR(2) \frac{\alpha_{2,n}}{\alpha_{2,2}} \omega_{\beta,2} & \dots & \omega_{\beta,n} \end{pmatrix} \begin{pmatrix} \frac{S_{\beta}}{P_1} \\ \frac{S_{\beta}}{P_2} \\ \dots \\ \frac{S_{\beta}}{P_n} \end{pmatrix}$$

The parameters  $\omega_{\beta,i}$  are given as:

$$\omega_{\beta,i} = \frac{SINR(i)^{\beta}}{\sum_{j=1}^n SINR(j)^{\beta}}$$

The idea of the primal/dual descent algorithms for the solving the optimization problem of the equation (3) is to search in the space of the possible values of the vector of the powers and the dual parameters for the values that maximizes the Lagrangian until optimality condition is achieved, i.e. the vector of the dual parameters is decreased to zero.

The basic search procedure is the projected gradient ascent algorithm that is given by the following update equations:

$$\underline{P}^{(k)} = \left( \underline{P}^{(k-1)} + \mu_1 \nabla_{\underline{P}} L \right)_P$$

$$\underline{\lambda}^{(k)} = \left( \underline{\lambda}^{(k-1)} - \mu \nabla_{\underline{\lambda}} L \right)_+$$

In these equations, the new vector of powers at the iteration  $k$  is chosen to increase the Lagrangian while the dual parameters are chosen in order to decrease the overall cost of the constraints violation in the system.

The parameters  $\mu_1, \mu$  are the steps of the power gradient ascent and the dual gradient descent algorithms. At each iteration  $k$ , the results of the gradient ascent/dual descent algorithms are projected over the domains of the powers and the dual parameters respectively.

### **Cell edge spectrum efficiency optimization problem formulation**

The general formulation of the cell edge spectrum efficiency optimization problem is given as

$$\begin{aligned} & \text{maximize } \min_{i \in \{1, \dots, n\}} \log(1 + \text{SINR}(i)) \\ & \text{subject to } P_i \leq P_0 \end{aligned}$$

Similarly to the cell edge SINR optimization problem, it is possible to use the generalized mean formulation for the approximation of the mean function.

The Lagrangian of the system is written in this case as

$$L(P_1, \dots, P_n, \underline{\lambda}) = \left( \sum_{i=1}^n \log(1 + \text{SINR}(i))^\beta \right)^{\frac{1}{\beta}} - \sum_{i=1}^n \lambda_i (P_i - P_0)$$

The Gradient equations are similar to the case of cell edge SINR optimization and the generalized mean gradient equations are given in this case as

$$\nabla_P S_\beta = \begin{pmatrix} \frac{\text{SINR}(1)}{(\text{SINR}(1)+1)c_1} \omega_{\beta,1} & \frac{-\text{SINR}(2)}{(\text{SINR}(2)+1)c_2} \frac{\alpha_{2,1}}{\alpha_{2,2}} \omega_{\beta,2} & \dots & \frac{-\text{SINR}(n)}{(\text{SINR}(n)+1)c_n} \frac{\alpha_{n,1}}{\alpha_{n,n}} \omega_{\beta,n} \\ -\frac{\text{SINR}(1)}{(\text{SINR}(1)+1)c_1} \frac{\alpha_{1,2}}{\alpha_{1,1}} \omega_{\beta,1} & \frac{\text{SINR}(2)}{(\text{SINR}(2)+1)c_2} \omega_{\beta,2} & \dots & \frac{-\text{SINR}(n)}{(\text{SINR}(n)+1)c_n} \frac{\alpha_{n,2}}{\alpha_{n,n}} \omega_{\beta,n} \\ -\frac{\text{SINR}(1)}{(\text{SINR}(1)+1)c_1} \frac{\alpha_{1,3}}{\alpha_{1,1}} \omega_{\beta,1} & -\frac{\text{SINR}(2)}{(\text{SINR}(2)+1)c_2} \frac{\alpha_{2,3}}{\alpha_{2,2}} \omega_{\beta,2} & \omega_{\beta,3} & \dots \\ -\frac{\text{SINR}(1)}{(\text{SINR}(1)+1)c_1} \frac{\alpha_{1,n}}{\alpha_{1,1}} \omega_{\beta,1} & -\frac{\text{SINR}(2)}{(\text{SINR}(2)+1)c_2} \frac{\alpha_{2,n}}{\alpha_{2,2}} \omega_{\beta,2} & \dots & \omega_{\beta,n} \end{pmatrix} \begin{pmatrix} \frac{S_\beta}{P_1} \\ \frac{S_\beta}{P_2} \\ \dots \\ \frac{S_\beta}{P_n} \end{pmatrix}$$

where the parameters  $c_i$  are the cell edge capacities of the node  $i$ , defined as  $c_i = \log(1 + \text{SINR}(i))$ . For this problem, the  $\omega_{\beta,i}$  parameter is calculated for the capacities

$$\omega_{\beta,i} = \frac{\log(1 + \text{SINR}(i))}{\sum_{j=1}^n \log(1 + \text{SINR}(j))}$$

### **The gradient ascent solutions investigated**

Three solution techniques were investigated in this note, in order to investigate the impact of the generalized mean parameter  $\beta$  and the overall convergence of the solutions. The envisioned solutions are:

- Standard gradient based ascent
- Adaptive Gradient ascent technique
- Conjugate Gradient technique with Polak- Ribiere ascent direction definition and adaptive step

We have used the Wolf- Armijo conditions for the calculations of the adaptive step of the gradient ascent technique, i.e.:

$$\mu_1(k) \geq \frac{(S_\beta(P_{k+1}) - S_\beta(P_k))}{\omega |\nabla S_\beta(P_k)|^2}$$

The function  $S_\beta(P_k)$  is the generalized mean function of the signal to interference ratio SINRs or the capacity. The parameter  $\mu_1(k)$  is the adaptive step of the gradient ascent technique and the parameter  $\omega$  is a relaxation parameter that is set to 0.5 in the simulations.

We have used for the conjugate gradient ascent, the Polak- Ribiere ascent direction calculation. In this case, the ascent equation for the powers are written as the following generalized ascent equation

$$\underline{P}^{(k)} = (\underline{P}^{(k-1)} + \mu_1(k) \underline{d}^{(k-1)})_p$$

Where the ascent direction is evaluated as

$$\begin{aligned} \underline{d}^{(0)} &= -\nabla_{\underline{P}} L^{(0)} \\ \underline{d}^{(k)} &= -\nabla_{\underline{P}} L^{(k)} + \beta_k \underline{d}^{(k-1)} \end{aligned}$$

The Polak- Ribiere step is given as

$$\beta_k = \frac{\|\nabla_{\underline{P}} L^{(k)}\|^2 - (\nabla_{\underline{P}} L^{(k)})^T \nabla_{\underline{P}} L^{(k-1)}}{\|\nabla_{\underline{P}} L^{(k-1)}\|^2}$$

### ***Fairness study of the power setting for the campus solution***

Power setting technique for campus of femto base stations may also be formulated as a mean for increasing the cell edge throughput fairness between the base stations of the campus. In this section we present the optimization problem and gradient based solution for three classical fairness measures:

- Jain index of the cell edge throughputs
- Proportional fairness of the cell edge throughputs
- Generalized proportional fairness of the throughputs

Jain index of the cell edge throughputs is defined as

$$J(r_1, r_2, \dots, r_n) = \frac{\left( \sum_{i=1}^n r_i \right)^2}{n \sum_{i=1}^n r_i^2}$$

Where the variables  $\{r_i\}_{i=1}^n$  are the cell edge throughputs of the base stations of the campus. General formulation of the generalized proportional fairness for the cell edge throughputs, i.e.  $\alpha$  – fairness can be formulated as :

$$F(r_1, r_2, \dots, r_n) = \begin{cases} \sum_{i=1}^n \frac{1}{(1-\alpha)} r_i^{(1-\alpha)} & \alpha > 0 \\ \sum_{i=1}^n \log(r_i) & \alpha = 1 \end{cases}$$

Proportional fairness is obtained for  $\alpha = 0$ , the other values of the parameter  $\alpha$  defines the generalized proportional fairness criterion.

Power setting optimization problem for the maximization of the fairness metric  $F(r_1, r_2, \dots, r_n)$  can be formulated as :

$$(P_1, P_2, \dots, P_n) = \max \{F(r_1, r_2, \dots, r_n)\}$$

#### **Jain index of the cell edge throughputs:**

The fairness metric to be maximized in this case is the Jain index of the cell edge throughputs, i.e.  $F(r_1, r_2, \dots, r_n) = J(r_1, r_2, \dots, r_n)$ . The power setting technique that we consider is a gradient ascent technique. In this technique, the powers are set in the direction of the Gradient of the jain index with respect to the powers.

This Gradient is calculated as the following:

$$\frac{\partial J}{\partial P_q} = \left( \frac{2nJ}{\sum_{i=1}^n r_i} \right) \sum_{i=1}^n \frac{\partial r_i}{\partial P_q} - \left( \frac{2nJ}{\sum_{i=1}^n r_i^2} \right) \sum_{i=1}^n r_i \frac{\partial r_i}{\partial P_q}$$

The cell edge throughputs are considered as  $r_i = \log(1 + SINR(i))$  so the derivatives of the expression above are given as

$$\frac{\partial r_i}{\partial P_q} = \frac{1}{(1 + SINR(i))} \frac{\partial SINR(i)}{\partial P_q}$$

The powers are set in the direction of the Gradient of the Jain index as:

$$\underline{P}^{(k)} = \left( \underline{P}^{(k)} + \delta \nabla J_{\underline{P}^{(k)}} \right)_p$$

So the algorithm sets the transmission powers of the campus in order to maximize the overall Jain index of the campus, i.e. defined previously.  $F(r_1, r_2, \dots, r_n)$ .

#### **Generalized Proportional fairness:**

The powers are set in order to maximize the fairness metric of the campus. So, the powers are set in the direction of the gradient of the fairness metric

$$\frac{\partial F}{\partial P_q} = \sum_{i=1}^n r_i^{-\alpha} \frac{\partial r_i}{\partial P_q}$$

The powers are set in the direction of the gradient as

$$\underline{P}^{(k)} = \left( \underline{P}^{(k)} + \delta \nabla F_{\underline{P}^{(k)}} \right)_P$$

In the simulation results section we present the performance of the power setting techniques based on jain index maximization and generalized proportional fairness maximization.

### Simulation results

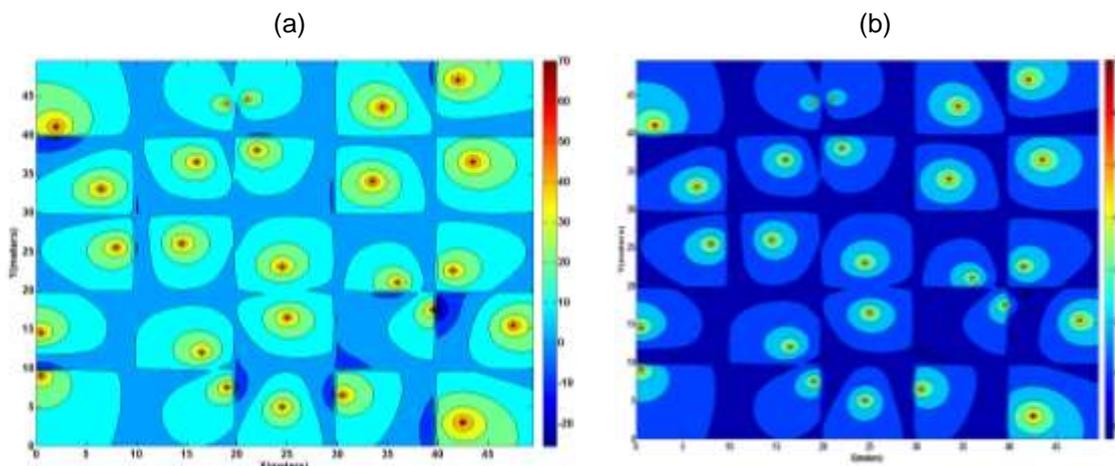
#### **Campus deployment of femto base stations:**

Two scenarios are investigated in this study for the campus deployment of the femto base stations:

- Campus deployment of CSG femto base stations
- Campus deployment of open base stations

For both scenarios, it is assumed a group of 25 femto base stations deployed in indoor, grid environment as shown in the coverage maps below.

The main difference between the two scenarios is that UEs are allowed to connect to the neighbouring base stations for the campus of open femto base station scenario while for the CSG scenario the UEs are only allowed to connect to its own femto base station.



**Figure A.65: typical coverage map for the femto campus scenario: (a) CSG femto base stations, (b) open base stations**

The coverage map of the Figure A.65 (a) is showing coverage holes, i.e. signal to interference ratio below -10dB, while the SINR levels in the coverage map of the open campus, i.e. Figure A.65 (b), are above the minimum SINR level necessary for the UE attachment (-10dB).

The common simulation parameters for the two scenarios are given in the following Table A.18.

**Table A.18: Common simulation parameters for the campus of femto base stations**

Parameter	Value
Carrier	2GHz
Antenna Gain	5dBi
Body loss	6dB
Powers	Minimum value : -5dBm Maximum value: 20dBm
Pathgain model	Motley Keenan : $PL(dB) = a + b \log_{10}(r)$ $a = 127$ $b = 30$ $r$ : radius in Kms
Walls attenuation	10dB
Noise density	-174dBm/Hz
Noise factor	9dB
Shadowing	No Shadowing

We have simulated for both scenarios the performance of power setting algorithm for maximizing the minimum SINR over the 25 femto base stations in terms of cumulative distribution. This cumulative distribution is calculated over 100 independent Monte- Carlo trials of the femto base stations.

### **CSG campus performance results**

As mentioned previously, in the CSG scenario the UEs are not allowed to connect to the neighbouring femto base stations. So, coverage holes are present in this case as shown in Figure A.65 (a). We have simulated gradient ascent based power setting algorithm where the powers are set in the direction of the gradient  $\nabla_{\mathbf{p}} S_{\beta}$ . The cumulative distribution (CDF) of the worst case cell edge SINR is shown in the Figure A.66 for various values for different optimization criteria. The cumulative distributions are calculated for different trails of the topology of the campus, i.e. for different trials of the positions of the femto base stations.

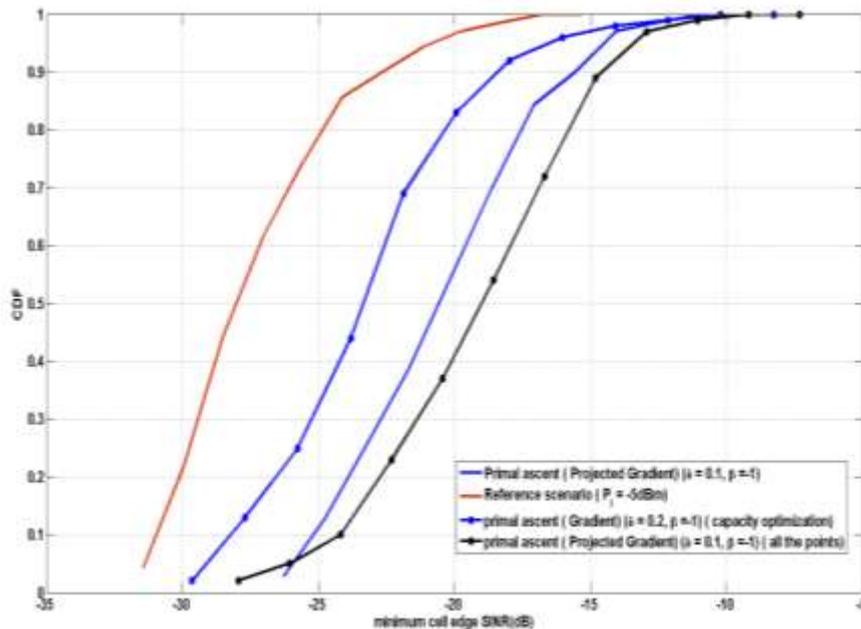


Figure A.66: CDF of worst case cell edge SINR for CSG campus

We have considered several optimization criteria in the figure above:

- Generalized mean of the cell edge SINRs of the base stations of the campus (blue solid line in the figure)
- Generalized mean of the cell edge capacities of the base stations of the campus (circle marked blue solid line in the figure)
- Generalized mean of the *all* SINRs of the campus

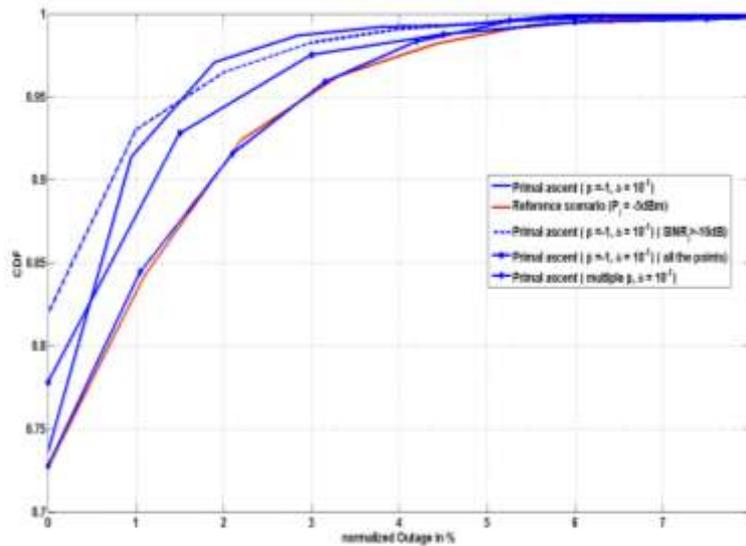
Several observations can be made from the Figure A.66:

The overall SINR level is low for the CSG campus scenario, typically around -28dB for the reference scenario where all the base stations powers are set to -5dBm.

The SINR level is improved by the primal ascent optimization technique and the overall improvement is around 9dB for 50% of the possible topologies of the campus. For this case, the overall cell edge SINR is still low and don't allow the attachment of the UE to the base stations of the campus

The best SINR improvement is obtained when taking into account the overall measured SINRs in the base stations of the campus. It is around -19dB for 50% of the campus topologies.

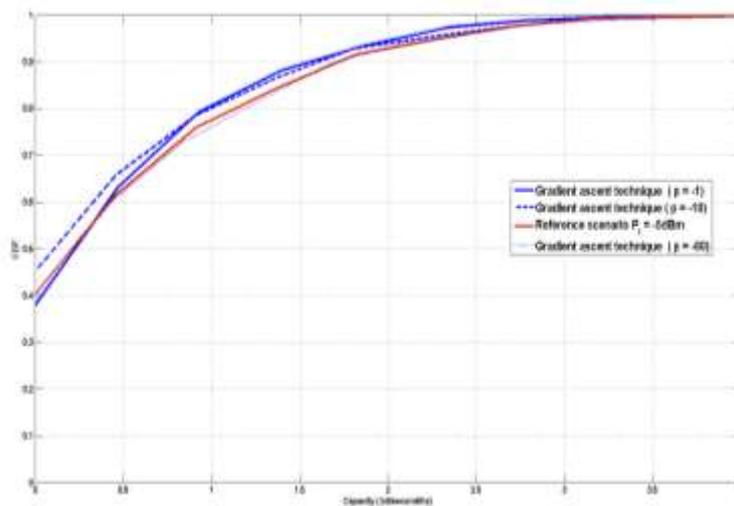
In Figure A.67, we have plotted the CDF with respect to the topologies of the campus of the maximum outage level in the CSG campus. This outage level is defined as the normalized coverage where the measured signal to interference ratio (SINR) is below the UE attachment/synchronization threshold of -10dB.



**Figure A.67: Normalized outage for the CSG campus scenario**

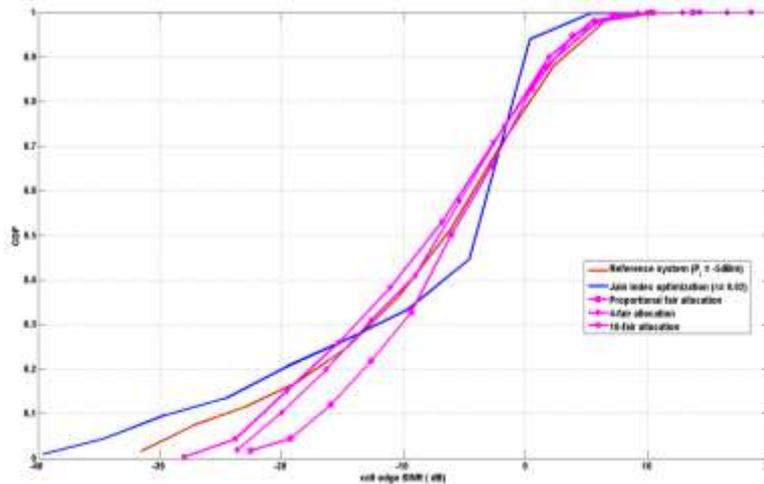
It is seen that the overall coverage in outage is low for the considered scenario. This coverage is around 2% of the coverage of each apartment in the CSG campus, i.e. 2 square meters for an overall coverage of 100 square meters. The proposed coverage control technique improves the overall coverage in outage for the CSG campus since it reduces the normalized coverage in outage with respect to the reference scenario where all the base stations are transmitting at power of -5dBm. The overall improvement is at maximum 50% for the primal ascent power setting technique that is taking into account the SINR's above the attachment threshold of -10dB. In this case, the normalized coverage in outage is below 0.8% for 90% of the tested campus topologies.

In the Figure A.68 we have plotted the cell edge capacity of the CSG campus mainly considering the SINR optimization model and for different values of the generalized mean parameters  $\beta$ .



**Figure A.68: cell edge capacity CDF for gradient ascent technique with different generalized means parameters**

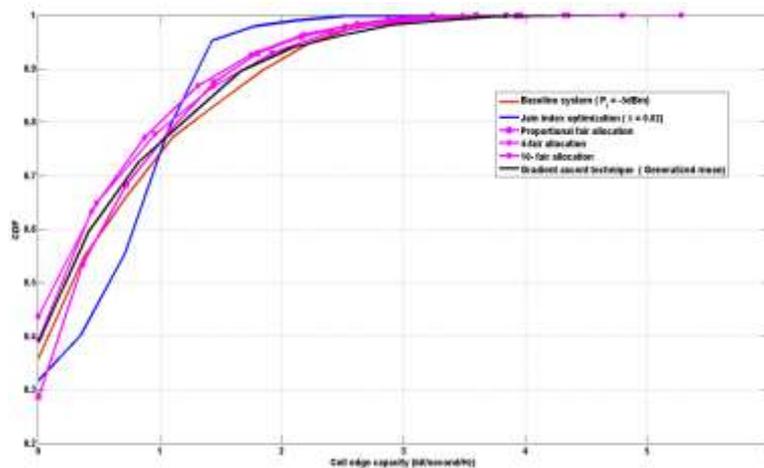
In Figure A.69 we have compared for the CSG campus, the cumulative distributions of the cell edge SINRs for the Jain index maximization and generalized proportional fairness metric maximization.



**Figure A.69: Cell edge SINR for different fairness criteria in CSG campus**

It is seen from the figure that the Jain index maximization provides the best fairness among the different fairness metrics considered since it improves the cell edge SINR in the [-10dB, 0dB] of around 3dB for 50% of the campus topology trials. This improvement comes at the price of a cell edge SINR degradation and a reduction of the maximum range of the cell edge SINRs [0dB, 10dB].

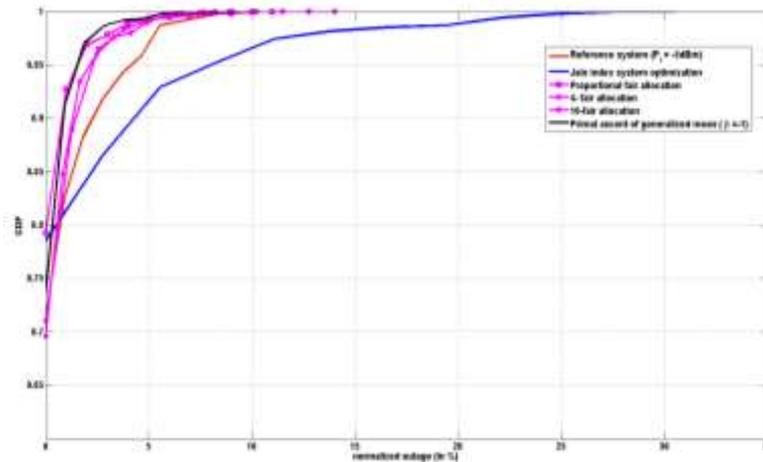
These results are confirmed in the Figure A.70 where we present the corresponding cell edge throughputs.



**Figure A.70: cell edge capacity for different fairness criteria for CSG campus**

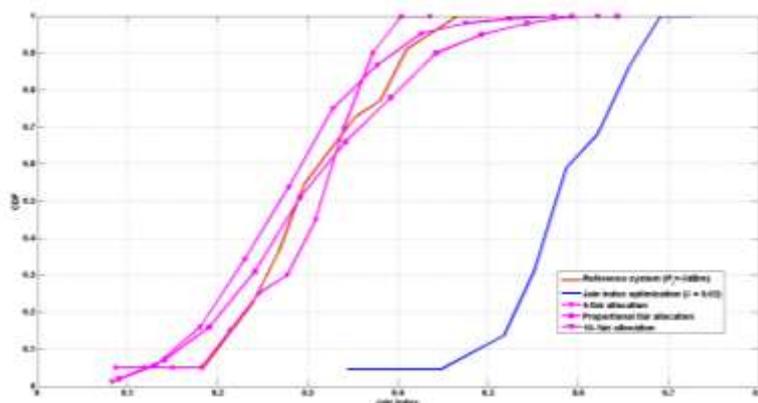
We have recalled in the figure, the results of the generalized mean optimization. It is seen that these results are similar to the results of the 4-fair power setting algorithm and that the Jain index maximization outperforms the different schemes in terms of throughput fairness. This improvement is around 0.32 bit/second/Hz in the lower region of the cell edge spectrum efficiencies (below 1 bit/second/Hz).

In Figure A.71 we present the normalized coverage in outage for CSG campus for the different power setting techniques based on fairness metric maximization.



**Figure A.71: Normalized coverage in outage**

When compared to most of the outage CDFs and to the reference system, it is seen that the Jain index maximization power setting increases the coverage in outage. This is obvious since the cell worst case cell edge SINRs for the Jain index maximization is reduced. Another observation is that the normalized coverage in outage for the proportional fair power setting is similar to the performance of the generalized mean maximization. As a summary, the Figure A.72 presents the Jain index cumulative distribution function for the different fairness metrics as well as for the baseline system.



**Figure A.72: Jain index CDF for different fairness metrics for CSG campus**

It is seen that the Jain index of the generalized proportional fairness metric is similar to the Jain index of the reference system. Jain index maximization technique improves the overall Jain index of the system of 0.2 for 50% of the topologies.

### Open campus performance results

In the Figure A.73, we present the minimum cell edge SINR results of the power adjustment technique based on gradient ascent of the generalized mean of the cell edge SINR's for the open campus. In the open campus scenario each UE is allowed to connect to the best femto base station in the campus coverage area.

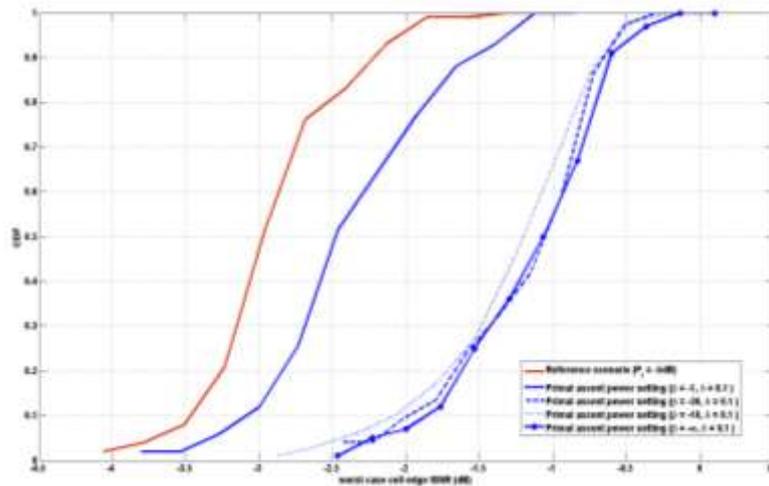


Figure A.73: CDF of worst case cell edge SINR for open campus

The worst case cell edge SINR performance results of the open campus shows better results when compared to the CSG campus cell edge SINR. The reference performance for 50% of the topology configurations of the open campus is around -3dB. The performance improvement of the power setting technique is around 2dB, i.e. the best minimum cell edge SINR for the open campus is -1dB for the primal ascent power setting technique.

In Figure A.74, we plot the CDF of the minimum cell edge capacities of the open campus showing an overall improvement of 43% over the minimum cell edge capacities of the reference scenario. This improvement is around 0.24 bit/second/Hz for worst case spectrum efficiency

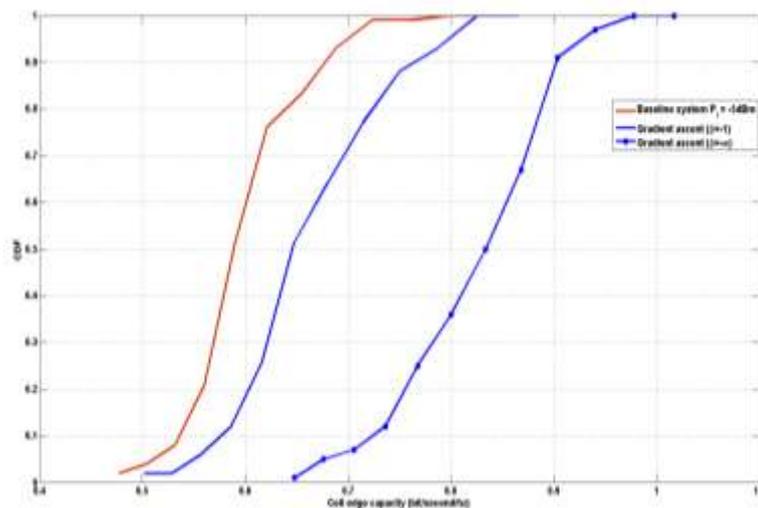
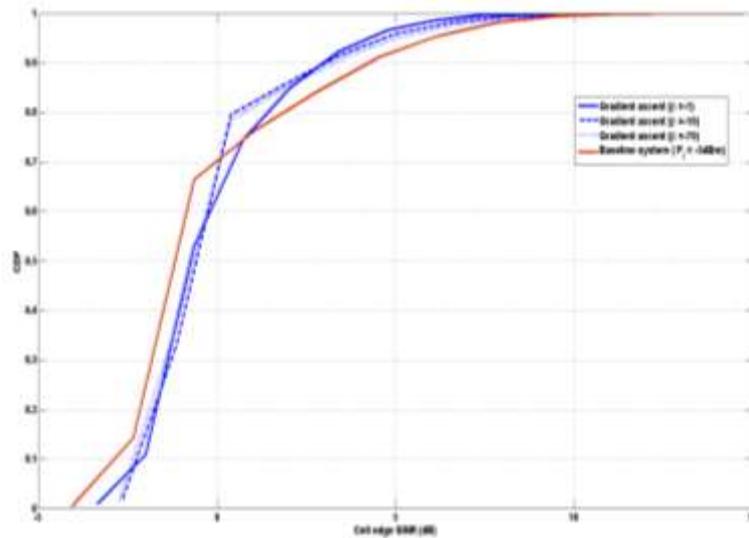


Figure A.74: CDF of worst case cell edge capacities for open campus

In the Figure A.75 we have plotted the cumulative distribution of the cell edge signal to interference ratio of the open campus for different generalized mean parameters  $\beta$ .

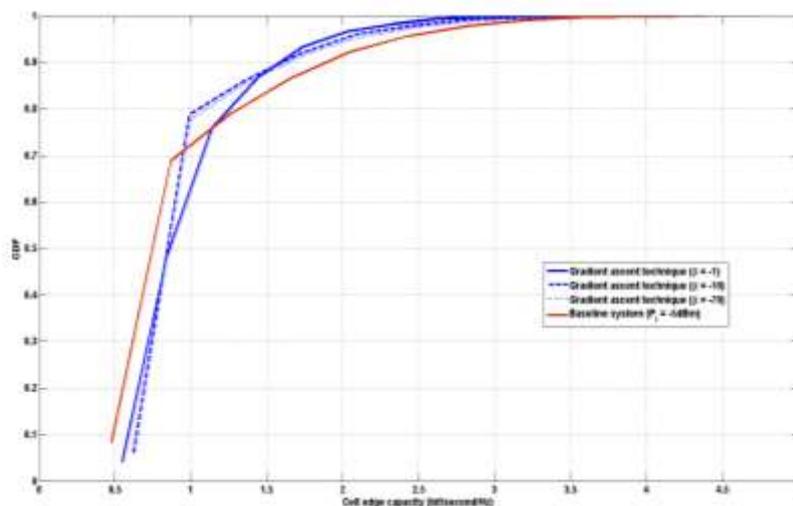


**Figure A.75: CDF of cell edge SINR for the open campus**

In this figure it is seen that power setting technique, based on gradient ascent improves the low range of cell edge SINR's for 70% of the topologies of the open campus scenario. The power setting technique improvement is around 43% for cell edge SINR's below 0dB, for example, the cell edge SINR is improved from -1.51 dB for the baseline scenario to -0.88dB for the power setting technique. It is seen that the SINR improvement increases with the parameter  $\beta$ .

The high range of SINR is reduced of around 1dB for the high range of SINRs, i.e. the SINRs are above 1dB. It is shown that this SINR reduction is increased with the parameter  $\beta$ .

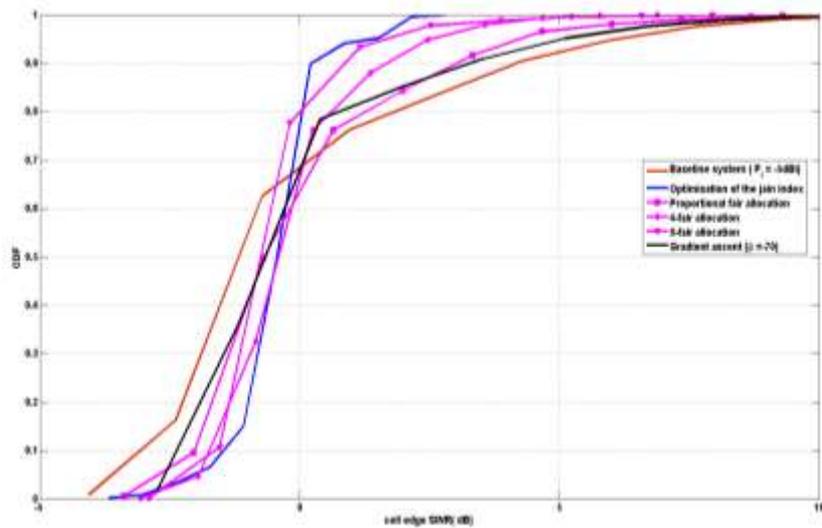
In Figure A.76 we have plotted the CDF of the cell edge capacities of the open campus for different values of the generalized mean parameters.



**Figure A.76: CDF of cell edge capacities for the open campus**

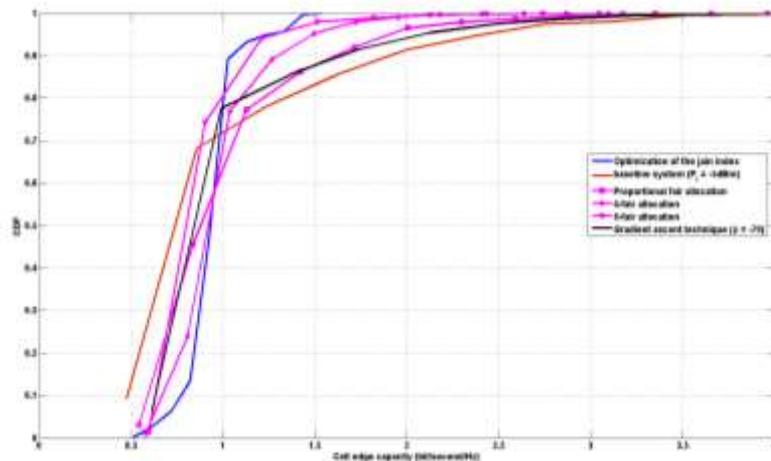
Open campus cell edge capacity results in Figure A.76 are showing similar behaviour to the SINR results of the Figure A.75. The power setting technique improves the low range cell edge capacities, i.e. below 1 bit/second/Hz. For example, the capacity is improved of 26% for 20% of the topology realizations of the open campus. The relative improvement drops to 14% for 50% of the open campus topology realizations and initial capacity of 0.5 bit/second/Hz. The high range of the cell edge capacities is reduced, i.e. the capacity is reduced from 2 bit/sec to 1.5 bit/second/Hz.

As for the CSG campus case, we will present in the following figures the performance of the power setting techniques based on fairness metric maximization. In Figure A.77 we show the cumulative distribution of the cell edge signal to interference ratios for the different power setting algorithms based on the maximisation of the different fairness metrics defined previously for CSG campus.



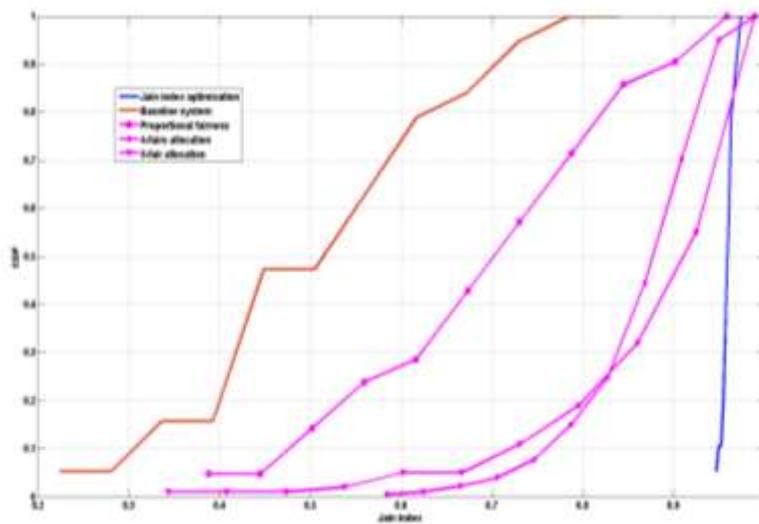
**Figure A.77: Cell edge SINR for different fairness criteria in open campus**

The figure shows that the power setting algorithm based on Jain index outperforms the different fairness metrics considered since it improves the overall cell edge SINR of around 2dB in the lower range of cell edge SINRs, i.e. the [-4dB 0dB] range. The  $\alpha$  – fairness optimization approaches the performance of the Jain index maximization for 8-fair power setting system and the performance of the generalized mean of the SINRs is similar to the proportional power setting algorithm. In Figure A.78 we have plotted the cell edge capacity CDF for the open campus with the different fairness maximization strategies.



**Figure A.78: cell edge capacity for different fairness criteria for open campus**

In this figure we see that the Jain index maximisation based power setting algorithm outperforms the other power setting techniques, i.e. around 0.3 bit/second/Hz for 20% of the topologies and a baseline performance of 0.5 bit/second/Hz. Similarly to the previous results, the proportional fair power setting technique exhibits similar performance as the generalized mean of the SINRs. Finally, the Figure A.79 shows the CDF of the Jain index for the different power setting techniques.



**Figure A.79: Jain index CDF for different fairness metrics for open campus**

It is seen from the figure that the Jain index based power setting improves the overall Jain index of the open campus of about 90% ( from 0.5 to 0.92). The Jain index is increased with the parameter  $\alpha$  of the generalized proportional fair metric.

As a final note it is seen that the proposed power setting technique improves worst cell edge capacities of the campus at the expense of reducing the high range cell edge capacities. So it is clear that the proposed technique introduce some fairness between the base stations of the campus with respect to the baseline system.

In Table A.19 we have evaluated the Jain fairness index [VAN00] for both cell edge SINRs and cell edge capacities for various parameters of the generalized mean.

**Table A.19: Jain index evaluation for the open campus scenario**

$\beta$	Cell edge SINRs		Cell edge capacities	
	Reference system Jain index	Power setting system Jain index	Reference system Jain index	Power setting system Jain index
-1	0.4596	0.6667	0.8144	0.8834
-10	0.4378	0.5449	0.7926	0.8621
-70	0.4589	0.5035	0.8013	0.8439

Several observations can be made from Table A.19:

The proposed power setting technique improves the Jain index of the system for the different values of the generalized mean parameters both for cell edge SINRs and cell edge capacities. The maximum improvement of the Jain index is obtained for low values of the generalized mean parameter, i.e. the power setting with low generalized mean parameter  $\beta$  increases the fairness among the base stations of the open campus when compared to higher values of the generalized mean parameter and the reference system.

The Jain index for the cell edge capacities is high when compared to the Jain index of the cell edge SINRs and the similar improvement and behaviour is observed with respect to the generalized mean parameter.

## A7. Coordinated Scheduling

### A7-1 Performance evaluations of the interference management concept

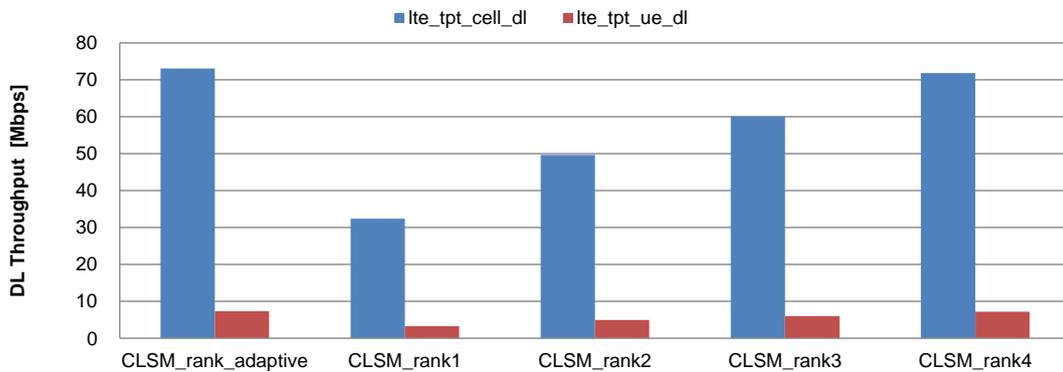
In this appendix section performance evaluations are presented by means of system level computer simulations, in order to assess the behavior of different transmission modes in the view of the final design of an overall L1/L2 interference management scheme concept.

In particular, system level simulations (see Table A.20 for detailed assumptions) are relative to a system with MIMO 4x4 antenna configuration in different propagation scenarios.

**Table A.20: Simulation assumptions.**

Simulation Parameters	Value
Simulated scheme	LTE 10MHz; MIMO 4x4
Cellular Layout	Indoor Hotspot environment
Simulated Link	Downlink
Traffic model	Full buffer
Packet Scheduling	Proportional Fair, with and without rank adaptation.
RI period	5 TTI
Feedback latency (*)	6 TTI
Antenna separation (BS, UE)	(10 $\lambda$ , $\lambda/2$ ); co-polarized antennas
(*) feedback generated in the TTI n is applied in TTI n+6	

System performances are depicted in Figure A.80 and show the average user and cell throughput for an Indoor Hotspot scenario. The considered scenario (Indoor Hotspot) permits to better appreciate performances gains obtained with TM switching, due to more favourable interference conditions. In particular, the impact of UE feedback is analyzed when considering a packet scheduler with fixed rank or with rank adaptation, and average system performances show benefits in adapting to RI reported by the terminal (under realistic feedback assumptions, see Table A.20).

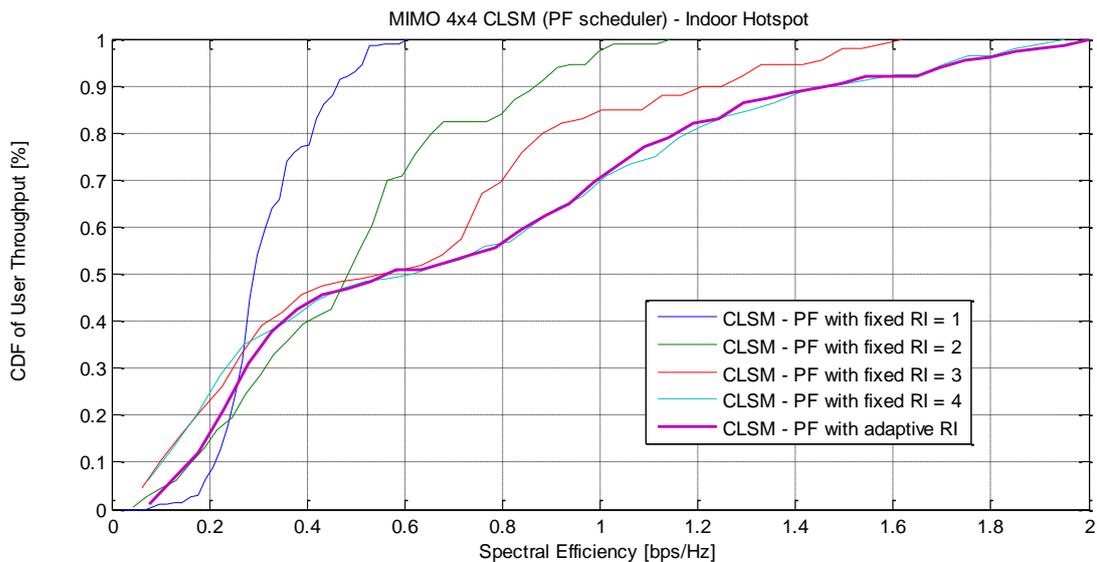


**Figure A.80: LTE MIMO 4x4 average system performances (Indoor Hotspot).**

More in detail, system performances are also analyzed in terms of CDF distribution of user throughput, in order to assess the behavior of the scheduler with respect to cell edge users, or in general as an indication of the fairness of the system. Simulation results show that rank adaptation implemented at the scheduler side is beneficial not only in terms of cell capacity (see Figure A.81) but also in terms of user throughput, even if a different behavior can be highlighted depending on the channel conditions:

For higher CDF values (best channel conditions, mainly center cell users), rank adaptation gives the best performances (maximum rank).

For low CDF values (poor channel conditions, mainly cell edge users), rank adaptation performances are affected by feedback latency (realistic working conditions).



**Figure A.81: LTE MIMO 4x4 system; CDF of user throughput (Indoor Hotspot).**

In summary, these simulations (conducted in a fully loaded system) show that the performances of a rank adaptation implemented in the scheduler depend on the feedback latency. Under realistic assumptions, a MIMO switch between different ranks is anyway beneficial in terms of overall system capacity and especially for centre cell users (better channel conditions and more reliable feedback), while for cell edge users feedback latency limits the performances of such adaptation.

## A8. Scheduling for joint processing

### A8-1 Impact of scheduling on the performance of downlink multicell processing

#### System Model and Opportunistic Schedulers

We consider a multicell cellular network and we focus on the eNBs taking part in a cooperation cluster of  $n_{comp}$  eNBs. Each eNB is equipped with one antenna and transmits to only one UE. There are  $K$  UEs in each of the cells, also equipped with a single antenna. The noise at the UE is a zero mean AWGN of variance  $\sigma^2$ . Moreover, each eNB transmits with its maximal power  $P$  (unless otherwise stated), and we consider a Rayleigh fading channel with a long term path loss effect where only the first ring of interferers is assumed to emit significant interference. The channel between eNB  $i$  and UE  $k$  in the cell of the eNB  $j$  is denoted as  $h_{ij}^{(k)}$ . The cooperation cluster is denoted as  $\wp$  and the elements in it are assumed to be ordered from 1 to  $n_{comp}$ . When considering the eNB  $j$ , the set of all the neighbouring eNBs is denoted as  $N_j$ . The interference is divided into two parts,  $I_{int,j}^{(k)}$  and  $I_{ext,j}^{(k)}$ , which represent the interference coming from the eNBs inside and outside the cooperation cluster, respectively. Thus, they can be written as

$$I_{int,j}^{(k)} = \sum_{i \in N_j \cap \wp} |h_{ij}^{(k)}|^2, I_{ext,j}^{(k)} = \sum_{i \in N_j \setminus (N_j \cap \wp)} |h_{ij}^{(k)}|^2.$$

For JP-CoMP, the matrix  $\mathbf{H}_{comp}^{(\mathbf{k})} \in \mathbb{C}^{n_{comp} \times n_{comp}}$  represents the multi-user channel between the cooperating eNB and the scheduled UEs. The multi-index  $\mathbf{k} \in N^{n_{comp} \times 1}$  is made of the  $n_{comp}$  indices of the UEs inside the  $n_{comp}$  cells of the coordination cluster. If it is written with a star symbol, it indicates that it is the optimal index according to some given scheduler.

The division of the interference between intra- and extra- cluster interference is useful only in the case of JP-CoMP where the intra-cluster interference are reduced via the precoding. In the case of single cell processing, the cooperation cluster has no meaning and only the sum of both the intra- and the extra- cluster interference matters.

In a first step we will discuss algorithms maximizing the performances without any consideration on fairness. We denote these algorithms as *unfair* (based on their expected highly unfair behaviour), while the modified versions of these algorithms in order to improve the fairness between the UEs are called the *Opportunistic Round Robin* [KR03] algorithms.

#### Unfair Distributed Schedulers:

Single cell processing with distributed scheduling: In that case, the eNBs do not cooperate with each other and each eNB selects individually the UE to transmit to. The first and main focus of our work is the *max-SINR scheduler*, which consists in selecting the UE with the maximal SINR. It is the most interesting distributed scheduler since it increases the gain of the direct link and reduces the interference at the same time.

For each eNB  $j$  in the cooperation cluster  $\wp$ , the index of the scheduled UE is given by

$$k_{\max-SINR, j}^* = \arg \max_{k \in \{1, \dots, K\}} SINR_{j, (k)} = \arg \max_{k \in \{1, \dots, K\}} \frac{|h_{jj}^{(k)}|^2}{\sigma^2 + I_{\text{int}, j}^{(k)} + I_{\text{ext}, j}^{(k)}}, \forall j \in \mathcal{J}.$$

The sum rate is then computed by summing the rates achieved by the scheduled UEs across the cells of the cooperation cluster:

$$R_{\max-SINR} = \sum_{j \in \mathcal{J}} \log(1 + SINR_{j, (k_{\max-SINR, j}^*)})$$

We will also consider the performances of a less elaborate distributed scheduler which only selects the UE with the largest SNR for the direct link and is denoted as the *max-SNR scheduler*. The index of the scheduled UE is then obtained from the optimization

$$k_{\max-SNR, j}^* = \arg \max_{k \in \{1, \dots, K\}} \frac{|h_{jj}^{(k)}|^2}{\sigma^2}, \forall j \in \mathcal{J}.$$

JP-CoMP with distributed scheduling: In that scheme, the scheduler is also distributed and selects the UE with the maximal SINR. However, the intra-cluster interference is not taken into account because it will be removed via the ZF precoder. In any case, simulations also show that the intra-cluster interference in the scheduler does not have any significant impact. The index of the scheduled UE is then obtained as

$$k_{\text{comp}, j}^* = \arg \max_{k \in \{1, \dots, K\}} SINR_{j, \text{comp}}(k) = \arg \max_{k \in \{1, \dots, K\}} \frac{|h_{jj}^{(k)}|^2}{\sigma^2 + I_{\text{ext}, j}^{(k)}}, \forall j \in \mathcal{J}.$$

Once the scheduled UEs are chosen via the distributed scheduler, the matrix  $\mathbf{H}_{\text{comp}}^{(k_{\text{comp}}^*)} \in \mathbb{C}^{n_{\text{comp}} \times n_{\text{comp}}}$  is computed from the given choice of UEs. A precoder is then computed to remove all the intra-cluster interference, i.e., to diagonalize the effective channel between the eNB and the UEs. Water-filling is then applied using the diagonalized effective channel obtained after using the derived precoder. Finally, the power allocation obtained from water-filling is normalized so as to fulfill the power constraint per eNB, and the sum rate is computed from the SINRs as previously given for the single cell processing.

No-interference upper bound: To evaluate the influence of the interference on the performance, we consider an upper bound presented in the introduction and consisting in removing the interference from the surrounding cells. A simple *max-SNR distributed* scheduler is then applied, which is actually optimal because of the absence of interference.

### Opportunistic Round Robin Schedulers:

A fairer alternative to these schedulers is called Opportunistic Round Robin (ORR) scheduler [KR03]. The principle of the ORR scheduler is to remove the UE from the set of UEs once it has been scheduled. The set of possible UEs is reduced by one and for the next time slot the scheduler is applied on this remaining set of UEs. This continues until all the UEs have been scheduled and served once. It has for consequence that in  $K$  time slots, each UE is scheduled one and only one time. Thus, the position of the UEs does not bring any diversity gain and the only multiuser diversity gain is obtained from the Rayleigh fading. Indeed, once a UE is

scheduled, while the position of the UEs is kept, new realizations are taken for the Rayleigh fading since it corresponds to a different time slot and the Rayleigh fading is short term fading. ORR can then select a UE with a good Rayleigh fading realization and a diversity gain is obtained. ORR achieves the same fairness as a more usual random round robin, but with the advantage of exploiting some of the multiuser diversity available at the same time.

The principle of ORR scheduling does not state which figure of merit is used to select the UE, and we will in fact have an ORR version of each of the previously described schedulers (*max-SINR*, *max-SNR*, *JP-CoMP*, and *no-interference upper bound*).

### Simulations Results

The parameters of the simulations are given in Table A.21 and correspond to a cellular multicell network. The UEs considered are located in a polygon whose corners are defined by the cooperating eNBs, i.e., in our cases, the triangle made of the three cooperating eNBs. It corresponds to a meaningful clustering and is assumed without loss of generality. Note that we do not consider the whole cell because of the constraint to let the UE be inside the polygon made of the cooperating eNBs, so that what we call the number of UEs per cell is in fact the number of UEs for the part considered in the cell.

**Table A.21: Simulation parameters**

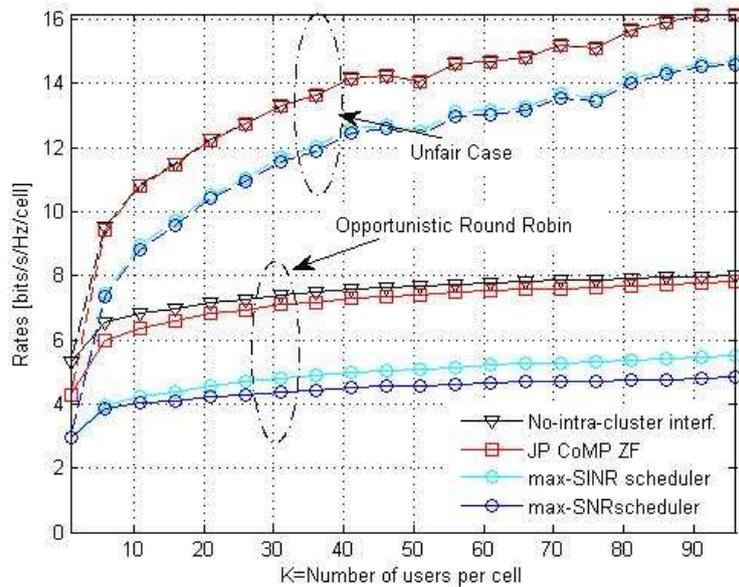
Simulation Parameter	Value
Transmission bandwidth	$\Delta=5$ MHz
Frequency	$f=2$ GHz
Radius of cell	$R=2$ Km
Minimal distance from UE to eNB	$\rho_{min}=20$ m
Path Loss Model	Hata model, with $h_{eNB}=15$ m, $h_{UE}=1.6$ m
Fading Law	Rayleigh Fading
Maximal power per eNB	$P=1$ W
Noise variance	$\sigma^2=8.3 \cdot 10^{-14}$ W
eNB antenna Gain	$G_{ant}=20$ dB
UE antenna Gain	$G_{UE}=5$ dB
Number of UEs per cell	$K=100$
Number of cooperating eNBs	$n_{comp}=3$

The simulations results have been provided in Section 4.2.3.1 and are recalled here for the sake of completeness.

In Figure A.82, the average rate is plotted for  $K$  time slots, so that each UE can be served when the ORR algorithms are used. We observe that for both the unfair and the fair versions, JP-CoMP ZF achieves an average rate very close to the average no-intra-interference rate, while the schemes with single cell processing are characterized by larger losses. Still, distributed processing allows to achieve good performance with much lower requirements for the system.

We also note that the *max-SNR scheduler* introduces little loss compared to the *max-SINR scheduler* in both cases. However, the difference is much smaller in the case without fairness constraint. In that case, it seems that the two distributed schedulers have the same scaling in

terms of the number of UEs. This is to put in relation with the fact that the scheduled UEs are then the UEs located very close to the eNBs, so that the interference power is very small. Consequently, the max-SNR scheduler and the max-SINR scheduler are likely to schedule the same UEs.



**Figure A.82: Average rate per cell and time slot.**

With our simulations, we have analysed in a realistic environment the impact of the scheduler on the interference in a multi-cell scenario. We can see that distributed opportunistic scheduling leads to good performance compared to the ideal case without interference and is an efficient tool to manage interference without much requirement on the architecture. Yet, it cannot by itself manage interference and it has to be complemented by other methods reducing interference, like for example JP-CoMP. An analytical analysis of the rate of convergence and the rate loss due to distributed scheduling can be found in [KG11].

## A8-2 Scheduling Aspects of Partial CoMP

[The results in this section are discussed in Subsection 5.2.]

The IMF-A framework as being introduced in chapter 5 includes the cover shift, tortoise and partial reporting concepts as described in A5-1. These help to localize the overall optimization problem of user grouping and scheduling to single cooperation areas (CA). A CA comprises for example nine cells with four antenna ports (AP) each and serves a single cover shift of e.g. 18 physical resource blocks (PRB). The goal is to find the optimum scheduler achieving the highest spectral efficiency  $SE_{\max}$  by serving simultaneously up to 27 UEs per PRB – i.e. 3 UEs per cell – under the constraint of at least proportional fairness.

In chapter 5 eq 5.2.1  $SE_{\max}$  has been derived as:

$$SE_{\max} = \arg(\max(\sum_{n_{DS}} \sum_{N_{PRB}} GP(CA_b^{opt}, UE_{1..K}^{opt}, \mathbf{W}^{opt}, \mathbf{F}_{1..K}^{opt}, PRB_i) / N_{cell}),$$

leading to a complex optimization problem including the optimum number of data streams  $n_{DS}$ , finding the optimum setup of CAs, user grouping per CA, precoder  $\mathbf{W}$ , UE receive filters  $\mathbf{F}$  and allocation of UEs to PRBs.

Figure A.83 contains a high level block diagram illustrating our proposed 2-stage scheduling strategy for achieving  $SE_{\max}$  assuming one single localized CA as described above.

Starting point is the time domain scheduler selecting 10 UEs per cell for the next subframe, summing up for 9 cells to 90 active UEs per CA. Each per cell scheduler uses the channel state information (CSI) provided by the 10 active UEs to form sets of 3 potentially simultaneously served UEs in this cell, which we call ‘overbooking’. Overbooking means that at least for some of the UE sets the condition of the channel matrix  $\mathbf{H}_i$  will be very poor in the end suggesting to serve only two or one UE of this set. Overall from 10 UEs 120 different sets  $\mathcal{L}_i$  of 3 UEs  $\mathcal{L}_i = \{UE_x, UE_y, UE_z\}$ ;  $x \in 1 \dots 120$ ;  $y \neq x \in 1 \dots 120$ ;  $z \neq x, y \in 1 \dots 120$ ; can be formed. For each set  $\mathcal{L}_i$  and for the whole subband  $f_{SB}$  comprising for example 18 physical resource blocks (PRB) the power normalization loss (PNL) is being calculated in case of simultaneous transmission of this set of users. This PNL can be found by a testwise precoding with  $\mathbf{W}_{t,i} = \text{pinv}(\mathbf{H}_{t,i})$  – with  $\text{pinv}(\cdot)$  being the pseudo inverse of a matrix - for the per cell test channel matrix  $\mathbf{H}_{t,i} \in \mathbb{C}^{3 \times 4}$  containing all channel components  $h_{km}$  for the UEs of set  $\mathcal{L}_i$ . A typical result can be seen in Figure A.83 left middle indicating a strong PNL variation between few to 20 or more dB varying for different  $\mathcal{L}_i$  differently. An iterative score based scheduler [JST+09] is applied to the sets  $\mathcal{L}_i$  of UEs. In each step the set  $\mathcal{L}_i$  containing the UE with the lowest score is selected. For this set the PRB from the not yet scheduled PRBs is taken exhibiting the smallest PNL for the selected set. Afterwards for all UEs of set  $\mathcal{L}_i$  the score is increased by one. These iterations are continued until all PRBs have been scheduled.

The overall effort for this scheduling is quite high, but here we benefit from the per cell scheduling strategy limiting the number of potential sets and leading to a small precoding matrix  $\mathbf{W}_{t,i}$  of size three times four. Anyway so far system level gains have been the main focus, but for an implementation a number of simplifications should be possible. For example the number of sets can be easily reduced from 120 to e.g. 20 without losing too much performance gain. Instead of doing the testwise precoding, measures with less computational complexity might be used for testing of the mutual orthogonality of a certain UE set.

The per CA scheduler as depicted in Figure A.83 right collects the scheduling results of all nine cells, forms a common CA wide channel matrix  $\mathbf{H}$  and the according CA wide zero forcing precoder  $\mathbf{W} = \text{pinv}(\mathbf{H})$  for all PRBs  $\text{PRB}_i$ . The achieved SINR CDF for the  $3 \times 9 = 27$  UEs can be found in Figure A.84 and indicates that about 70% of UEs experience a SINR greater than 20dB, but at the same time about 20% of UEs suffer from a PNL of more than 10dB. Under the assumption of ideal CSI knowledge the PNL is not an issue, but for real CSI estimation an increasing PNL leads to accordingly larger accuracy requirements for the channel estimation. For that reason the per CA scheduler has to do some fine tuning by taking out those UEs generating very high PNL. Interestingly the PNL per CA for a certain UE is most of the time lower compared to that of the per cell precoding and the reason is the well known rank enhancement for dislocated Tx points [JJT+09]. In about 50% of the time the PNL is even below 4.7dB, which would be the power required to serve 3 fully orthogonal UEs. For more details see [ZMK12].

Figure A.85 contains a typical outcome of the per cell (left) and the per CA scheduler (right). The achieved SINRs for the sets of 3 UEs are colour coded for the per cell scheduler and indicate a very good fairness level as almost all UEs have at least on some resources a very high SINR. A similar observation can be made from the results for the 27 UEs of the overall CA.

The high SINRs are transferred by the SINR-to-datarate-mapping table in Table A.22 into throughput values for each UE allowing the calculation of an overall spectral efficiency. Under ideal conditions it is close to 6.5bit/s/Hz as already mentioned in chapter 5. For that result an overhead of 43% has been taken into account so the gross SE without any overhead would be over 11 bits/s/Hz/cell.

As there is a close interaction between scheduling decisions with requirements for channel estimation for example over the power normalization loss and/or the number of relevant channel components these two areas have to be optimized together. Further research is needed to find the best overall interference mitigation setup taking realistic channel estimation into account.

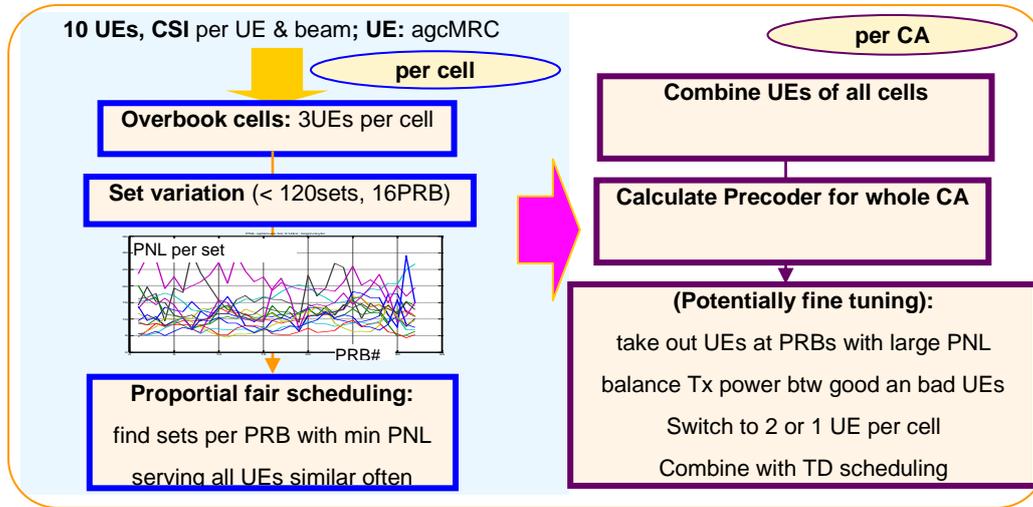


Figure A.83: high level block diagram of per cooperation area and per cell scheduler

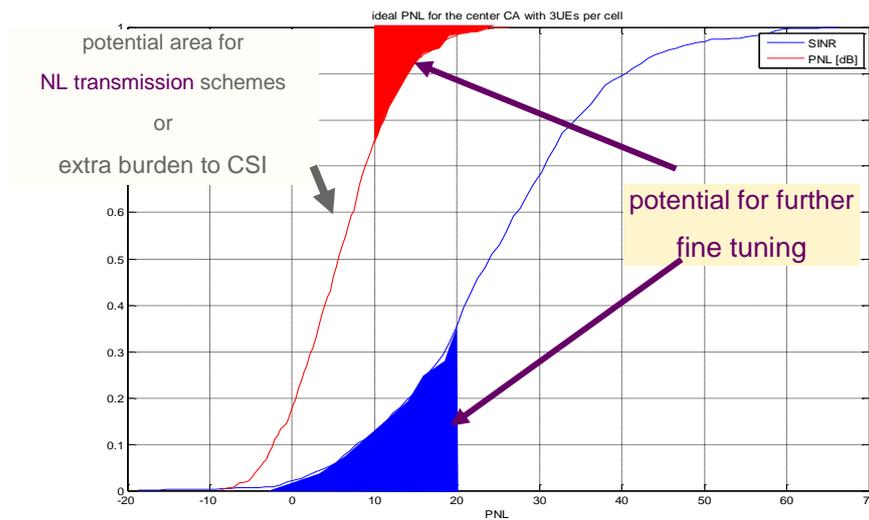
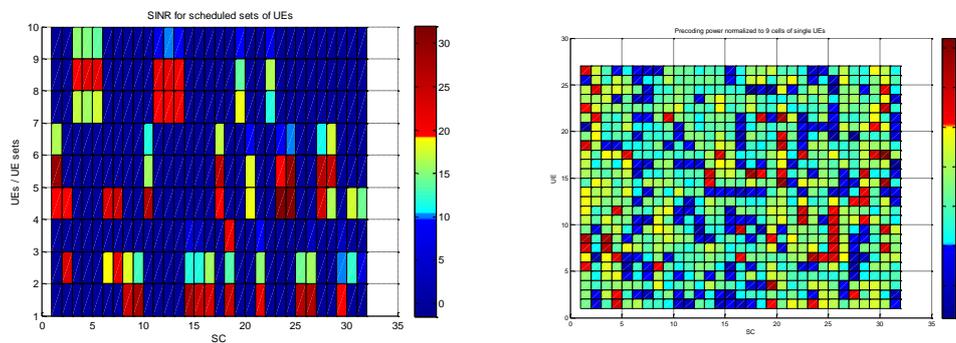


Figure A.84: red: power normalization loss (PNL) and blue: SINR of the scheduler above serving 27 UEs per CA

**Table A.22: SINR to data rate mapping table**

CQI	Mod	CR	Bits per RE	SINR [dB]
0		-		
1	QPSK	0.08	0.15	-8
2	QPSK	0.12	0.23	-6
3	QPSK	0.19	0.38	-4
4	QPSK	0.30	0.60	-2.5
5	QPSK	0.44	0.88	-1
6	QPSK	0.59	1.18	1.5
7	16QAM	0.37	1.48	3.8
8	16QAM	0.48	1.91	4.5
9	16QAM	0.60	2.4	7
10	64QAM	0.45	2.73	8
11	64QAM	0.55	3.32	11
12	64QAM	0.65	3.90	12.5
13	64QAM	0.75	4.52	16.5
14	64QAM	0.85	5.12	18
15	64QAM	0.9258	5.5547	19



**Figure A.85: typical per cell (left) and per CA (right) scheduling results. The SINR is color coded in dB. The per cell scheduler allocates 3 UEs per SC and the CA scheduler 27 UEs. Left/right: red colour indicates SINR > 20/15dB and blue SINR < 10/0dB.**

## References

- [3GPP1] 3GPP LTE Release 11. DL MIMO IRC.
- [3GPPTR1] 3GPP TR 36.819 V11.0.0 (2011-09), 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Coordinated multi-point operation for LTE physical layer aspects (Release 11).
- [3GPPTR2] 3GPP TR 36.814 V9.0.0 (2010-03), 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Further advancements for E-UTRA physical layer aspects (Release 9).
- [3GPP25814] 3GPP TR 25.814, "Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA)".
- [3GPP25996] 3GPP, "3GPP TR25.996 V9.0.0 (2009-12) Spatial channel model for Multiple Input Multiple Output (MIMO) simulations (Release 9)," 2010.
- [3GPP25996b] 3GPP, "3GPP TR25.996 V10.0.0 (2011-03) (2009-12) Spatial channel model for Multiple Input Multiple Output (MIMO) simulations (Release 10)," 2011.
- [3GPP36211] 3GPP TS 36.211, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation (Release 10)".
- [3GPP36420] 3GPP TS 36.420, "X2 general aspects and principles (Release 10)", 2010.
- [3GPP36814] 3GPP TR 36.814 v9.0.0, Further Advancements for E-UTRA Physical Layer Aspects, 2010. 3GPP Ftp Server (<http://www.3gpp.org/ftp>), 2010.
- [3GPP36942] 3GPP TR 36.942 V9.1.0, Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios, 2010.
- [3GPP-R1083774] 3GPP, NXP Semiconductors and Philips, "Feedback and Precoding Techniques for MU-MIMO for LTE-A", R1-083774.
- [3GPP-R1090777] 3GPP, Alcatel-Lucent, "UE PMI feedback signalling for user pairing/coordination," R1-090777.
- [3GPP-R1093141] 3GPP, Qualcomm Europe, "Signaling for spatial coordination in DL CoMP," R1-093141.
- [3GPP-R1100853] 3GPP, Ericsson, "Channel reciprocity in FDD systems including systems with large duplex distance," R1-100853.
- [3GPP-R1101431] 3GPP, Nokia Siemens Networks, "CoMP performance evaluation," R1-101431.
- [3GPP-R1105081] R1-105081: Summary of the description of candidate eICIC solutions, 3GPP TSG-WG1 #62, Madrid, Spain, August 23rd – 27th, 2010.
- [3GPP-R1112339] 3GPP, LG electronics, "CoMP performance evaluation under low-capacity and high-latency backhaul,," R1-112339.
- [AM79] B.D.O. Andersson and J.B. Moore, Optimal Filtering, Prentice-Hall, 1979.
- [Aro11] D. Aronsson, "Channel estimation and prediction for MIMO OFDM systems – Key design and performance aspects of Kalman-based algorithms", Ph.D Thesis, Dept. of Engineering Sciences, Signals and Systems Group, Uppsala University, March 2011.
- [ARTD11] Artist4G, "D1.1 – Definitions and architecture requirements for supporting interference avoidance techniques" Jan. 2011.
- [ARTD12] ARTIST4G consortium, D1.2 – "Innovative advanced signal processing algorithms for interference avoidance", *ARTIST4G technical deliverable*, 2010.

- [ARTD13] ARTIST4G consortium, "D1.3 – Innovative advanced signal processing algorithms for interference avoidance," project report, March, 2011.
- [ARTD25] ARTIST4G consortium, "D2.5 – Flexible interference control – Evaluation," Project report, June 2012.
- [ARTD35] ARTIST4G consortium, "D3.5 – Performance evaluations of advanced relay concepts," Project report, June 2012.
- [ARTD42] ARTIST4G consortium, "D4.2 – Feedback from RAN constraints" Project report, Oct 2011.
- [ARTD43] ARTIST4G consortium, "D4.3 – Final conclusions on RAN architecture," Project report, June 2012.
- [ARTD51] ARTIST4G consortium, "D5.1 – Scenarios, KPIs and Evaluation Methodology for Advanced Cellular Systems", project report, March 2012.
- [ARTD61] ARTIST4G consortium, "D6.1 – First feedback on implementation aspects connected to the selected innovations," project report, March 2011.
- [ARTD62] ARTIST4G consortium, "D6.2 – Laboratory and field trial results connected to the first set of innovations," project report, Sept. 2011.
- [ASA12] R. Apelfröjd, M. Sternad and D. Aronsson, "Measurement-based evaluation of robust linear precoding for downlink CoMP", *IEEE ICC'12*, Ottawa, Canada, June 2012.
- [BH06] F Boccardi and H Huang, Zero-forcing precoding for the MIMO broadcast channel under per-antenna power constraints, in Proc. of the 7th IEEE Workshop on Signal Processing Advances in Wireless Communications, 2006.
- [Bon04] T. Bonald, "A score-based opportunistic scheduler for fading radio channels," *Proc. Of European Wireless*, 2004.
- [BOY04] Stephen Boyd and Lieven Vandenberghe, "Convex Optimization", Cambridge University Press, 2004.
- [Brä11] L.-J. Brännmark, "Robust sound field control for audio reproduction: A polynomial approach to discrete-time acoustic modelling and filter design," Ph.D Thesis, Dept. of Engineering Sciences, Signals and Systems Group, Uppsala University, 2011.
- [BS06] H. Boche and M. Schubert. Duality theory for uplink downlink multiuser beamforming. In *Smart Antennas - State-of-the-Art*. EURASIP, Hindawi Publishing Corp., 2006.
- [BSX+10] C. Botella, T. Svensson, X. Xu, and H. Zhang, "On the performance of joint processing schemes over the cluster area," *IEEE Vehicular Technology Conference*, 2010.
- [C07] J.M. Cioffi, "Advanced Digital Communications, EE379c", Stanford University Course Notes, 2007-2008, <http://www.stanford.edu/class/ee379c>
- [C75] A. Carleial, "A case where interference does not reduce capacity (corresp.)," *Information Theory*, *IEEE Transactions on*, vol. 21, no. 5, pp. 569 – 570, Sep 1975.
- [C78] A. Carleial, "Interference channels," *IEEE Transactions on Information Theory*, vol. 24, p. 60, Jan. 1978.
- [C83] Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, 1983, pp. 439-441.
- [CJ08] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the K -user interference channel," *IEEE Transactions on Information Theory*, vol. 54, pp. 3425–3441, Aug. 2008.
- [CL08] D. Chizhik, J. Ling, "Propagation over Clutter: Physical Stochastic Model", *IEEE Trans. On Antennas and Propagation*, Vol. 56, No. 4, April 2008.

- [D03] Devaney, Robert L. "Introduction to Chaotic Dynamical Systems", Westview Press.
- [DHG+12] M. Danneberg, J. Hofeld, M. Grieger, M. Amro, G. Fettweis, „Field Trial Evaluation of UE Specific Antenna Downtilt in an LTE Downlink“, Workshop on Smart Antennas (WSA 2012), March 7-8, 2012, Dresden, Germany.
- [DY10] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp.1748-1759, May 2010.
- [Ekm02] T. Ekman, "Prediction of mobile radio channels: Modeling and Design." Ph.D Thesis, Dept. of Engineering Sciences, Signals and Systems Group, Uppsala University, 2002.
- [Eng05] A. P. Engelbrecht, "Fundamentals of computational swarm intelligence", *John Wiley & Sons*, pp.171-172, 2005.
- [EO07] T. Eriksson and T. Ottosson, "Compression of feedback for adaptive transmission and scheduling," *Proceedings of IEEE*, vol 95, no. 12, pp 2314 – 2321, Dec. 2007.
- [ESB+98] O. Edfors, M. Sandell, J.J. van de Beek, S.K. Wilson and P.O. Börjesson, "OFDM channel estimation by singular value de-composition", *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 931-939, July 1998.
- [FF12] Fritzsche, R.; Fettweis, G.; "Robust MMSE Precoding with General Power Constraints", accepted for publication at *IEEE ISWCS 2012*.
- [FKV06] G. Foschini, K. Karakayali, and R. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," *IEEE Proceedings-Communications*, vol. 153, pp. 548-555. Aug. 2006.
- [FOF12] Fritzsche, R.; Ohlmer, E.; Fettweis, G.; „Where to Predict the Channel in Cooperative Cellular Networks?“, submitted to IEEE SCC 2013.
- [GK11] D. Gesbert and M. Kountouris, "Rate Scaling Laws in Multicell Networks Under Distributed Power Control and User Scheduling," *IEEE Transactions on Information Theory*, vol. 57, no.1, pp. 234-244, Jan. 2011.
- [GMK09] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "Interference alignment for the K-user MIMO interference channel," *CoRR*, vol. abs/0909.4604, 2009.
- [Gol05] A Goldsmith, *Digital Communications*, Cambridge University Press 2005.
- [GSC09] A. García Armada, M. Sánchez-Fernández, and R. Corvaja, "Waterfilling Schemes for Zero-Forcing Coordinated Base Station Transmission", *IEEE Global Communications Conference (GLOBECOM 2009)*, Honolulu, Hawaii, December 2009.
- [GUI11-1] J. Guillet, L. Brunel, N. Gresset, "Downlink Femto-Macro ICIC with Blind Long-Term Power Setting", *PIMRC'11*, Toronto, Canada, September 2011.
- [GUI11-2] J. Guillet, L. Brunel, N. Gresset, "Uplink femto-macro ICIC with semi-centralized power control", *PIMRC'11*, Toronto, Canada, September 2011.
- [HE06] P. Hammarberg and O. Edfors, "A comparison of DFT and SVD based channel estimation in MIMO OFDM systems," *IEEE PIMRC 2006*.
- [HSK+12] H. Halbauer, S. Saur, J. Koppenborg, C. Hoek, "Interference Avoidance with Dynamic Vertical Beamsteering in Real Deployments," *4G Mobile Radio Access Networks Workshop at IEEE Wireless Communications and Networking Conference (WCNC 2012)*, April 1, Paris France.
- [HT11] Harri Holma, Antti Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, 2nd Edition ISBN: 978-0-470-66000-, 2011.

- [HZH08] K. Jee Hyun, W. Zirwas, and M. Haardt, "Adaptive Codebooks for Efficient Feedback Reduction in Cooperative Antenna Systems," IEEE Global Telecommunications Conference, 2008.
- [J05] N. Jindal, "MIMO Broadcast Channels with Finite Rate Feedback," in IEEE Global Telecommunications Conference (GLOBECOM), 2005.
- [JJJ+09] S. Jaeckel, L. Jiang, V. Jungnickel, L. Thiele, C. Jandura, G. Sommerkorn, and C. Schneider, 'Correlation properties of large and small-scale parameters from multicell channel measurements', European Conference on Antennas and Propagation (EuCAP 2009), Berlin, Germany, 2009.
- [JJT+09] V. Jungnickel, S. Jaeckel, L. Thiele, L. Jiang, U. Krüger, A. Brylka, C. von Helmolt, Capacity Measurements in a Cooperative MIMO Network, IEEE Transactions on Vehicular Technologies, VOL. 58, NO. 5, JUNE 2009.
- [JKG+02] Joham, M.; Kusume, K.; Gzara, M.H.; Utschick, W.; Nossek, J.A.; "Transmit Wiener filter for the downlink of TDDSS-CDMA systems," *IEEE Seventh International Symposium on Spread Spectrum Techniques and Applications, 2002*.
- [JST+09] Jungnickel, V., Schellmann, M., Thiele, L., Wirth, T., Haustein, T., Koch, O. Zirwas, W. Schulz, E., Interference-Aware Scheduling in the Multiuser MIMO-OFDM Downlink, IEEE Communications Magazine, June 2009, Vol. 47, No.6, pg 56-66.
- [JTS+08] V. Jungnickel, L. Thiele, M. Schellmann, T. Wirth, W. Zirwas, T. Haustein, E. Schulz, "Implementation Concepts for Distributed Cooperative Transmission ", 42nd Asilomar Conference on Signals, Systems and Computers, Monterey, USA, Oct. 2008.
- [JWS+08] V. Jungnickel, T. Wirth, M. Schellmann, T. Haustein, and W. Zirwas, "Synchronization of cooperative base stations", ISWCS '08, Reijavik, Island, Oct. 2008.
- [KE95] J. Kennedy, R. C. Eberhart, "Particle swarm optimization", *Proc. IEEE International Conference on Neural Networks*, pp.1942–1948, 1995.
- [KF10] K. Kumar and X. Feng, "An iterative algorithm for joint signal and interference alignment," in Proc. IEEE International Symposium on Information Theory Proceedings (ISIT), 2010, Austin, TX, 2010, pp. 2293–2297.
- [KFO6] K.M. Karakayali, G.J. Foschini and R.A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications*, vol. 13, no.4 Aug. 2006, pp. 56-61.
- [KG11a] P. de Kerret, D. Gesbert, "The Asymptotic Limits of Interference in Multicell Networks with Channel Aware Scheduling", Proc. International Workshop on Signal Processing Advances in Wireless Communications, San Francisco, USA, 2011.
- [KG11b] P. de Kerret, D. Gesbert, "The Multiplexing Gain of a Two-Cell MIMO Channel with Unequal CSI", Proc. International Symposium on Information Theory, Saint-Petersburg, Russia, 2011.
- [KHS+12] J. Koppenborg, H. Halbauer, S. Saur, C. Hoek, "3D Beamforming Trials with an Active Antenna Array", Workshop on Smart Antennas (WSA 2012), March 7-8, 2012, Dresden, Germany.
- [KPK+09] Koivisto, T., Pedersen, K.I., Kovacs, I.Z., Raaf, B., Pajukoski, K., Rinne, M.J., LTE-Advanced: The path towards gigabit/s in wireless mobile communications, Wireless VITAE 2009. May 2009.
- [LBS+10] T.R. Lakshmana, C. Botella, T. Svensson, X. Xu, J. Li, and X. Chen, "Partial joint processing for frequency selective channels," in *Proc. IEEE Vehicular Technology Conference, Ottawa, Canada 2010*.

- [LBS12a] T. R. Lakshmana, C. Botella, T. Svensson, "Partial Joint Processing with Efficient Backhauling in Coordinated MultiPoint Networks," in *Proc. IEEE Vehicular Technology Conference*, Yokahama, Japan May 2012.
- [LBS12b] T. R. Lakshmana, C. Botella and T. Svensson, "Partial joint processing with efficient backhauling using particle swarm optimization", *EURASIP Journal on Wireless Communications and Networking* 2012, 2012:182 doi:10.1186/1687-1499-2012-182.
- [LBS12c] J. Li, C. Botella and T. Svensson, Resource allocation for clustered network MIMO OFDMA systems, *EURASIP Journal on Wireless Communications and Networking* 2012, 2012:175 doi:10.1186/1687-1499-2012-175.
- [LCB+12] J. Li, X. Chen, C. Botella, T. Svensson and T. Eriksson, Resource allocation for OFDMA systems with multi-cell joint transmission, in *Proc. IEEE SPAWC'12*, Çeşme, Turkey, June 2012.
- [LH04] D. J. Love, R. W. Heath Jr., "Grassmannian Beamforming on Correlated MIMO Channels," *Proc. Globecom*, Dallas, 2004.
- [LMF10] B.K. Lau, J. Medbo and J. Furuskog, "Downlink cooperative MIMO in urban macrocell environments". *2010 IEEE International Symposium on Antennas and Propagation (ARSURSI)*.
- [LPA+12] J. Li, A. Papadogiannis, R. Apelfröjd, T. Svensson, M. Sternad, "Performance Analysis of Coordinated Multi-Point Transmission Schemes with Imperfect CSI", *PIMRC'12*, Sydney, Australia, Sept. 2012.
- [LSB+11] J. Li, T. Svensson, C. Botella, T. Eriksson, X. Xu, and X. Chen, "Joint scheduling and power control in coordinated multi-point clusters", *IEEE VTC'11*, San Francisco, USA, September 2011.
- [LXC+10] J. Li, X. Xu, X. Chen, X. Tao, T. Svensson, and C. Botella, Downlink radio resource allocation for coordinated cellular OFDMA networks, *IEICE Transactions on Communications*, vol. E93.B, no. 4, pp. 3480-3488, 2010.
- [LZX+10] J Li, H Zhang, X Xu, X Tao, T Svensson, C Botella and B Liu, A novel frequency reuse scheme for coordinated multi-point transmission, in *Proc. IEEE VTC2010-Spring*, Taipei, Taiwan.
- [MF11] Marsch, P.; Fettweis, G., "Uplink CoMP under a Constrained Backhaul and Imperfect Channel Knowledge," *IEEE Transactions on Wireless Communications*, vol.10, no.6, pp.1730-1742, June 2011.
- [MF11] P. Marsch, G. P. Fettweis, "Coordinated Multi-Point in Mobile Communications Subtitle: From Theory to Practice, chapter 9.1", 2011, Cambridge University Press, ISBN 978-1-107-00411-5.
- [MSK+09] J. Medbo, I. Slomina, A. Kangas, and J. Furuskog, "Propagation channel impact on LTE positioning accuracy – A study based on real measurements of observed time difference of arrival". *IEEE Personal, Indoor and Mobile Radio Communications (PIMRC) 2009*, Sept. 13-16, 2009, Tokyo, Japan.
- [MZ11] W. Mennerich and W. Zirwas, "Reporting effort for cooperative systems applying interference floor shaping," *IEEE PIMRC 2011*, Toronto Canada, Sept. 2011.
- [MZ12] W. Mennerich, W. Zirwas, "User Centric Coordinated Multi Point Transmission", *Proc. IEEE VTC-fall*, 2010.
- [NSN11] Nokia Siemens Networks, '3GPP White paper 2020: Beyond 4G Radio Evolution for the Gigabit Experience', 2011-08-24.
- [P12] A. Pascht, "LightRadio™- A Ground-Breaking Mobile Technology", *International Multi-Conference on Systems, Signals and Devices*, March 20 – 23, 2012, Chemnitz, Germany.

- [PA10] A. Papadogiannis and G. C. Alexandropoulos, The value of dynamic clustering of base stations for future wireless networks, in Proc. IEEE World Congr. Comput. Intell., Barcelona, Spain, Jul. 2010.
- [PBG+08] A Papadogiannis, H Bang, D Gesbert, E Hardouin, "Downlink overhead reduction for multicell cooperative processing enabled wireless networks", *IEEE Personal, Indoor and Mobile Radio Communications*, 2008.
- [PBG11] A. Papadogiannis, H.J. Bang, D. Gesbert, E. Hardouin, "Efficient selective feedback design for multicell cooperative networks," *IEEE Trans. on Vehicular Technology*, vol. 60, no. 1, pp. 196-205, Jan. 2011.
- [PGH08] A. Papadogiannis, D. Gesbert, and E. Hardouin, "A dynamic clustering approach in wireless networks with multi-cell cooperative processing," in Proc. IEEE ICC'08, Beijing, China, 2008.
- [PH09] S. Peters and R. Heath, "Interference alignment via alternating minimization," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, 2009. ICASSP 2009, April 2009, pp. 2445–2448.
- [PH11] S. Peters and R.W.Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 206–218, Jan 2011.
- [PHG08] A. Papadogiannis, E. Hardouin, and D. Gesbert, "A framework for decentralising multi-cell cooperative processing on the downlink," in Proc. IEEE GLOBECOM Workshops, 2008.
- [PW12] Palleit, N., Weber, T., Channel Prediction in Multiple Antenna Systems. International ITG Workshop on Smart Antennas (WSA'12), S. 1-7, Dresden, 2012.
- [RHG06] Roemer, F., Haardt, M., Del Galdo, G., Higher Order SVD Based Subspace Estimation to Improve Multi-Dimensional Parameter Estimation Algorithms, ACSSC '06, Oct. 29 2006.
- [SA93] M. Sternad and A. Ahlén, "Robust filtering and feedforward control based on probabilistic descriptions of model errors," *Automatica*, vol. 29, pp. 661-679, 1993.
- [SB04] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18-28, Jan. 2004.
- [SB05] M. Schubert and H. Boche. Iterative multiuser uplink and downlink beamforming under SINR constraints. *IEEE Trans. Signal Processing*, 53(7):2324- 2334, July 2005.
- [SD08] M.B. Shenouda and T.N. Davidsson, "On the design of linear transceivers for multiuser systems with channel uncertainty," *IEEE Journal on Selected Areas of Communication*, vol. 26, no.6, pp. 1015-1024, Aug. 2008.
- [SFS+05] M. Sternad, S. Falahati, T. Svensson and D. Aronsson, "Adaptive TDMA/OFDMA for wide area coverage and vehicular velocities," *IST Summit*, Dresden, June 19-23, 2005.
- [SGA+12] M. Sternad, M. Grieger, R. Apelfröjd, T. Svensson, D. Aronsson and A. Belén Martinez, "Using 'predictor antennas' for long-range prediction of fast fading of moving relays," *IEEE WCNC12 Workshop on 4G Systems*, Paris, France, April 2012.
- [SGS+09] U. Salim, D. Gesbert, D. Slock, Z. Beyaztas, Hybrid Pilot/Quantization based Feedback in Multi-Antenna TDD Systems, *GLOBECOM 2009*, IEEE, Nov. 2009.
- [SH09] Bin Song, Haardt, M., "Achievable throughput approximation for RBD precoding at high SNRS" *ICASSP*, 2009

- [SH11] S. Saur, H. Halbauer, "Exploring the Vertical Dimension of Dynamic Beam Steering," 8th International Workshop on Multi-Carrier Systems and Solutions, May 3-4, 2011, Herrsching, Germany.
- [SM03] Samardzija, D.; Mandayam, N., "Multiple antenna transmitter optimization schemes for multiuser systems," IEEE 58th Vehicular Technology Conference, VTC 2003-Fall. 2003, 6-9 Oct. 2003.
- [SSO+07] M. Sternad, T. Svensson, T. Ottosson, A. Ahlén, A. Svensson and A. Brunstrom, "Towards systems beyond 3G based on adaptive OFDMA transmission," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2432-2455, Dec. 2007.
- [SSV+08] Shuying Shi; Schubert, M.; Vucic, N.; Boche, H.; "MMSE Optimization with Per-Base-Station Power Constraints for Network MIMO Systems," *IEEE International Conference on Communications, 2008. ICC '08*, 2008.
- [STS07] M. Sadek, A. Tarighat, and A. Sayed, "A Leakage-Based Precoding Scheme for Downlink Multi-User MIMO Channels," *IEEE Transactions on Wireless Communications*, 2007.
- [SVH+11] N. Seifi, M. Viberg, R.W. Heath, J. Zhang and M. Coldrey, "Multi-mode transmission in network MIMO downlink with incomplete CSI," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, Article ID 743916, doi:10.1155/2011/743916.
- [TCJ08] A. Tölli, M. Codreanu and M. Juntti, Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas, *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1428-1440, 2008.
- [TGR09] R. Tresch, M. Guillaud, and E. Riegler, "On the achievability of interference alignment in the K-user constant MIMO interference channel," in *Proc. IEEE/SP 15th Workshop on Statistical Signal Processing*, 2009. SSP '09, sept 2009, pp. 277–280.
- [TJ08] Lars Thiele, Volker Jungnickel, Adaptive Transmission in a Realistic Multicell Scenario, *Proceedings of the 2008 IEEE Radio and Wireless Symposium (RWS 2008)*, Jan. 2008.
- [TJ10] G. Tiangao and S. Jafar, "Degrees of freedom of the K user  $M \times N$  MIMO interference channel," *IEEE Transactions on Information Theory*, vol. 56, p. 6040, Dec. 2010.
- [TSS05] A. Tarighat, M. Sadek, and A. Sayed, "A Multi User Beamforming Scheme for Downlink MIMO Channels Based on Maximizing Signal-to-Leakage Ratios," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, IEEE, 2005.
- [TWS+09] L. Thiele, T. Wirth, M. Schellmann, Y. Hadisusanto, and V. Jungnickel, "MU-MIMO with Localized Downlink Base Station Cooperation and Downtilted Antennas," *2009 IEEE International Conference on Communications*, IEEE, 2009, pp. 1-5.
- [VAN00] Bobby Vandalore, Sonia Fahmy, Raj Jain, Rohit Goyal, Mukul Goyal, "General Weighted Fairness and its Support in Explicit Rate Switch Algorithms," *Computer Communications*, Vol. 23, No 2, January 2000, pp. 149-161.
- [VENU12] Karrthik, Venu, *SP Wi-Fi Update and Next Generation Hot Spot*, Cisco, 2012.
- [WES06] A. Wiesel, Y.C. Eldar, and S. Shamai, "Linear precoding via conic Optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol 54, no 1, pp.161-176, Jan. 2006.

- [WLS+00] Wacker, A.; Laiho-Steffens, J.; Sipila, K.; Heiska, K., The impact of the radio network planning and site configuration on the WCDMA network capacity and quality of service, Vehicular Technology Conference Proceedings, VTC 2000-Spring Tokyo, 2000.
- [WMM+05] T. Weber, I. Maniatis, M. Meurer, and W. Zirwas, "Performance Investigation of Improved Channel Estimation Exploiting Long Term Channel Properties," *Conference on Multi-Carrier Spread Spectrum Technologies*, 2005.
- [YL06] W. Yu and R. Lui, Dual methods for nonconvex spectrum optimization of multicarrier systems, *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310-1322, 2006.
- [YS10] H. Yu and Y. Sung, "Least squares approach to joint beam design for interference alignment in multiuser multi-input multi-output interference channels," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4960–4966, sept 2010.
- [YTJK08] C. Yetis, G. Tiangao, S. Jafar, and A. Kayran, "Approaching the capacity of wireless networks through distributed interference alignment," in *Proc. IEEE Global Telecommunications Conference*, 2008. GLOBECOM 2008, New Orleans, LO, Dec. 2008, pp. 1–6.
- [YXL09] A. Liu, Y. Liu, H. Xiang, and W. Luo, "On the rate duality of MIMO interference channel and its application to sum rate maximization," in *Proc. IEEE Global Telecommunications Conference*, 2009. GLOBECOM 2009., Honolulu, HI, Dec. 2009, pp. 1–6.
- [ZCA+09] J Zhang, R Chen, J G Andrews, A Ghosh, and R W Heath, Networked MIMO with clustered linear precoding, *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1910-1921, 2009.
- [ZG10] R. Zakhour, D. Gesbert, "Team Decision for the Cooperative MIMO Channel with Imperfect CSIT Sharing", Invited Paper, The Information Theory and Applications (ITA) Workshop, San Diego CA., February 2010.
- [ZMK12] Wolfgang Zirwas, Wolfgang Mennerich, Aneeq Khan, Main Enablers for Advanced Interference Mitigation, *European Transactions on Telecommunications*, to be published 2012.
- [ZMS+09] W. Zirwas, W. Mennerich, M. Schubert, L. Thiele, V. Jungnickel, and E. Schulz, "Cooperative transmission schemes," *Long Term Evolution: 3GPP LTE radio and cellular technology*, B. Furht and S.A. Ahson, Auerbach Publications, 2009, pp. 213-263.
- [ZPO08] X. Zhang, D.P. Palomar and B. Ottersten, "Statistical robust design of linear MIMO transceivers", *IEEE Transactions on Signal Processing*, vol 57, no. 8, pp 3678 – 3689, Aug. 2008.
- [ZWZ+05] Jinfan Zhang; Yongle Wu; Shidong Zhou; Jing Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Communications Letters*, vol.9, no.11, pp. 991- 993, Nov. 2005.
- [ÖAS95] K. Öhrn, A. Ahlén and M. Sternad, "A probabilistic approach to multivariable robust multivariable filtering", *Automatic Control, IEEE Transactions on*, vol 40, No 3, pp 405-418, March 1995.

## List of acronyms and abbreviations

3D	3-Dimensional
3GPP	3rd Generation Partnership Project
4G	4th Generation
AAS	Active Antenna Systems
ABS	Almost Blank Sub-frame
AE	Antenna Element
AMC	Adaptive Modulation and Coding
AP	Access Point
AR	Auto Regressive
ARQ	Automatic Repeat request
ARTIST4G	Advanced Radio Interface Technologies for 4G Systems
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BLAST	Bell Labs Layered Space-Time
BF	BeamFormer
BPC	Binary Power Control
BS	Base Station
BVDM	Building Vector Data Map
CA	Cooperation Area
CBS	Cooperative BS Set
CC	Channel Component (Downlink channel from one beam/antenna)
CCN	Central Coordination Node
CCU	Cluster-Center Users
CDD	Cyclic Delay Diversity
CDF	Cumulative Distribution Function
CE	Channel Estimation
CEU	Cluster-Edge Users
CIR	Channel Impulse Response
C-JT	Centralized Joint Transmission
CoMP	Coordinated Multi Point
CP	Cyclic Prefix; Control Plane
CQ	Channel Quality
CQI	Channel Quality Indicator

CRS	Common Reference Signal
CS/CB	Coordinated Scheduling / Coordinated Beamforming
CSG	Closed Subscriber Group
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
CTF	Channel Transfer Function
CU	Coordination Unit (for centralized joint processing)
CVX	Convex Optimization
D-BLAST	Diagonal-BLAST
DCW	Double Codeword
DD	Distributed Design
DDD	Dependent Distributed Design
DFT	Discrete Fourier Transform
DJP	Distributed Joint Processing
D-JT	Distributed Joint Transmission
DL	DownLink
DoA	Direction of Arrival
DPC	Dirty Paper Coding
DT	Down Tilt
ECSI	Effective Channel State Information
ECSIT	Effective Channel State Information at the Transmitter
eNB	enhanced Node B
FB	Feedback
FBH	Full BackHaul
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FER	Frame Error Rate
FFB	Full FeedBack
FFR	Fractional Frequency Reuse
FFT	Fast Fourier Transform
FMT	HeNB-UEs
FTB	Front-To-Back
GB	GigaBit
GMMSE	Generalized MMSE

GP	Goodput (Throughput of correctly received segments)
GPS	Global Positioning System
HARQ	Hybrid Automatic Repeat Request
HeNB	Home eNB
HIZ	High Interference Zone
HPA	High Power Amplifier
HPBW	Half Power Beam Width
HSPA	High Speed Packet Access
IC	Interference Cancellation
ICIC	Inter-Cell Interference Coordination
ID	IDentity
IDD	Independent Distributed Design
IDFT	Inverse Discrete Fourier Transform
IFA	InterFerence Alignment
IMF-A	Interference Mitigation Framework - Advanced
IPG	Interpolation Gain
IRC	Interference Rejection Combining
ISD	Inter-Site Distance
ITU	International Telecommunications Union
JD	Joint Design
JP	Joint Processing
JS	Joint Scheduling
JT	Joint Transmission
JT-CoMP	Joint Transmission Coordinated Multipoint
KKT	Karush-Kuhn-Tucker
KPI	Key Performance Indicator
L1	Layer 1
L2	Layer 2
LBH	Limited BackHaul
LFB	Limited FeedBack
LO	Local Oscillator
LOS	Line of Sight
LTE	Long Term Evolution
LTE-A	LTE Advanced
MAC	Medium Access Control
MBCP	Model Based Channel Prediction

MC	Multi Carrier
MCS	Modulation and Coding Scheme
MIESM	Mutual Information Effective SINR Mapping
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MMSE	Minimum Mean Square Error
MMT	eNB-UEs
MNO	Mobile Network Operators
MPC	Multipath Components
MRC	Maximum Ratio Combining
MS	Mobile Station
MSE	Mean Square Error
MU	Multi User
MU-BF	Multi-User Beamforming
MU-MIMO	Multi User Multiple-Input Multiple-Output
MU-MISO	Multi User Multiple-Input Single-Output
MWF	Modified Waterfilling
NCIC	Number of Coordinated Interfering Cells
NCJT	Non-Coherent Joint Transmission
NCJT-BPC	Non-Coherent joint transmission with BPC
NL	Non-Linear
NLOS	Non Line Of Sight
NMSE	Normalized complex channel Mean Square prediction Error
OAM	Operation and Maintenance
OCI	Other-Cell Interference
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
ORR	Opportunistic Round Robin
OSWF	Optimal Scaled Wiener Filter
PAPR	Peak to Average Power Ratio
PCI	Physical Cell Identity
P-CSI	Perfect CSI
PDP	Power Delay Profile
PHY	Physical layer
PJP	Partial Joint Processing
PL	(outdoor-to-indoor) Penetration Loss

PMI	Precoding Matrix Index
PNL	Power Normalization Loss
PR	Penetration Rate (fraction of users that use JT CoMP)
PRB	Physical Resource Block
PSO	Particle Swarm Optimization
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RAN	Radio Access Network
RI	Rank Indicator
RLC	Radio Link Control layer
RLP	Robust Linear Precoder
RLP-MSSE	Robust Linear Precoder - Minimum Mean Square Error solution
RLP-ACFF	Robust Linear Precoder – Automatic Control robust FeedForward solution
RMMSE	Robust MMSE
RR	Round Robin (scheduling)
RRH	Remote Radio Head
RS	Reference Signal
RS SINR	Signal to Interference Ratio of Reference Signal
RSRP	Reference Symbol Received Power
RSRQ	Reference Signal Received Quality
RTxWF	Robust TxWF
Rx	Receiver
SB	Score Based (scheduling)
SC	Single Carrier; Singel Cell
SC FDMA	Single Carrier Frequency Division Multiple Access
SC SFBC	Single Carrier Space-Frequency Block Code
SCM	Spatial Channel Model
SCME	Spatial Channel Model Extended
SCW	Single Codeword
SDMA	Space Division Multiple Access
SE	Spectral Efficiency
SFBC	Space-Frequency Block Code
SFR	Soft Frequency Reuse
SI	Study Item
SINR	Signal to Interference and Noise Ratio

SIR	Signal to Interference Ratio
SISO	Single-Input Single-Output
SLNR	Signal to Leakage plus Noise Ratio
SLR	Signal to Leakage Ratio
SM	Spatial Multiplexing
SMSE	Sum Mean Square Error
SNR	Signal to Noise Ratio
SOCP	Second Order Cone Program
SS	Synchronization Signal
STBC	Space-Time Block Code
STTD	Space Time Transmit Diversity
SU	Single User
SU-MIMO	Single User Multiple-Input Multiple-Output
SVD	Singular Value Decomposition
SWF	Scaled Wiener Filter
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
TV	TeleVision
Tx	Transmitter
TxD	Transmit Diversity
TxWF	Transmit Wiener Filter
UE	User Equipment
UFR	Universal Frequency Reuse
UL	UpLink
ULA	Uniform Linear Array
UMTS	Universal Mobile Telecommunications System
UP	Uniform Power; User Plane
UP-COOP	User Plane Cooperation
V-BLAST	Vertical-BLAST
WF	Waterfilling
WLAN	Wireless Local Area Network
WP1	Work Package 1
WSR	Weighted Sum Rate
ZF	Zero Forcing
ZF-OPA	Zero-Forcing joint transmission with Optimal Power Allocation