



Grant Agreement:	247223
Project Title:	Advanced Radio InTerFace Technologies for 4G SysTems ARTIST4G
Document Type:	PU (Public)

Document Identifier:	D1.3
Document Title:	<b>Innovative scheduling and cross-layer design techniques for interference avoidance</b>
Source Activity:	WP1
Editors:	Valeria D'Amico and Jochen Giese
Authors:	Carmen Botella, Loïc Brunel, Cristina Ciochina, Laura Cottatellucci, Valeria D'Amico, Paul de Kerret, David Gesbert, Jochen Giese, Nicolas Gresset, Julien Guillet, Hardy Halbauer, Xiaoran Jiang, Hajer Khanfir, Tilak Rajesh Lakshmana, Bruno Melis, Dario Sabella, Stephan Saur, Tommy Svensson, Patrick Tortelier, Wolfgang Zirwas
Status / Version:	Final / Version 1.0
Date Last changes:	31.03.11
File Name:	D1.3.doc

Abstract:	<p>This document provides an overview of the proposed innovations and activities in Task 1.2 of Work Package 1 (WP1) of the ARTIST4G project, related to interference avoidance.</p> <p>Focus is on the technical approaches applicable at layer 2, which are grouped into five different classes of innovations related to clustering &amp; user grouping, inter-cell interference coordination, coordinated scheduling, scheduling for joint processing and game theory based scheduling.</p> <p>Descriptions of the proposed innovations are given including basic ideas, potential of performance, simulation results, realization options and possible implementation restrictions.</p>
-----------	--

Keywords:	Interference avoidance, Coordinated Multi Point, Coordinated Scheduling, Coordinated Beamforming, Joint Processing, Inter Cell Interference Coordination, Heterogeneous Networks, Clustering, Game theory, ARTIST4G
-----------	---

Document History:	
31.03.2011	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 Generic L2 aspects of interference avoidance schemes .....</b>	<b>9</b>
3.1 Definitions .....	9
3.2 Coordinated multipoint schemes .....	10
3.3 L2 issues of centralized and decentralized architectures .....	13
3.4 Dynamic and semi-static schemes .....	15
3.5 Downlink and uplink specific aspects .....	16
3.6 Impact of duplexing modes .....	17
3.7 Homogeneous and heterogeneous scenarios .....	18
3.8 Key Performance Indicators .....	19
<b>4 Description of specific resource allocation and scheduling algorithms for interference avoidance.....</b>	<b>20</b>
4.1 Clustering and user grouping.....	20
4.1.1 User-centric clustering for partial joint processing .....	21
4.1.2 Clustering based on partial CoMP .....	26
4.1.3 User Pairing algorithms for MU-MIMO .....	32
4.1.4 Optimized resource allocation for SC-FDMA MU-MIMO schemes .....	36
4.2 Inter-Cell Interference Coordination.....	40
4.2.1 Blind and semi-centralised power setting for heterogeneous networks.....	40
4.2.2 Distributed sub carrier allocation for ICIC .....	46
4.3 Coordinated Scheduling .....	50
4.3.1 Dynamic algorithms for coordinated scheduling .....	51
4.3.2 Precoding optimization algorithm for coordinated beamforming.....	55
4.3.3 Distributed scheduling for beam coordination .....	61
4.3.4 Coordinated scheduling based on restriction requests .....	65
4.3.5 Coordinated scheduling for heterogeneous deployments.....	70
4.4 Scheduling for joint processing.....	76
4.4.1 Impact of scheduling on the performance of downlink multicell processing .....	76
4.4.2 Scheduling Aspects of Partial CoMP .....	81
4.5 Game theory based scheduling.....	83
4.5.1 Resource allocation in slow fading interfering channels with partial knowledge of the channels .....	84
<b>5 Conclusions and next steps .....</b>	<b>88</b>

<b>References .....</b>	<b>91</b>
<b>List of acronyms and abbreviations .....</b>	<b>94</b>

## Authors

Name	Beneficiary	E-mail address
Carmen Botella	Chalmers University of Technology	carmenb@chalmers.se
Loïc Brunel	MERCE	l.brunel@fr.mercede.mee.com
Cristina Ciochina	MERCE	c.ciochina@fr.mercede.mee.com
Laura Cottatellucci	EURECOM	Laura.Cottatellucci@eurecom.fr
Valeria D'Amico	Telecom Italia	valeria1.damico@telecomitalia.it
Paul de Kerret	EURECOM	paul.dekerret@eurecom.fr
David Gesbert	EURECOM	David.gesbert@eurecom.fr
Jochen Giese	Qualcomm	jgiese@qualcomm.com
Nicolas Gresset	MERCE	n.gresset@fr.mercede.mee.com
Julien Guillet	MERCE	j.guillet@fr.mercede.mee.com
Hardy Halbauer	Alcatel Lucent	hardy.halbauer@alcatel-lucent.com
Xiaoran Jiang	MERCE	x.jiang@fr.mercede.mee.com
Hajer Khanfir	France Telecom	hajer.khanfir@orange-ftgroup.com
Tilak Rajesh Lakshmana	Chalmers University of Technology	tilak@chalmers.se
Bruno Melis	Telecom Italia	bruno1.melis@telecomitalia.it
Dario Sabella	Telecom Italia	dario.sabella@telecomitalia.it
Stephan Saur	Alcatel Lucent	Stephan.saur@alcatel-lucent.com
Tommy Svensson	Chalmers University of Technology	tommy.svensson@chalmers.se
Patrick Tortellier	France Telecom	Patrick.tortelier@orange-ftgroup.com
Wolfgang Zirwas	NSN	wolfgang.zirwas@nsn.com

## 1 Executive Summary

This document represents an overview of the activities in Task 1.2 of WP1 related to interference avoidance schemes applied at Layer 2. The main idea of interference avoidance is based on the observation that the performance of current cellular mobile communication systems where one user is served by a single access point is limited by the interference caused by the communication of neighbouring access points. Reduction or avoidance of interference is therefore a promising strategy to improve performance.

After the introduction in the section 2, the specific issues on Layer 2 for further improvements together with relevant scenarios and key performance indicators are highlighted in section 3.

The main activities of this task are then described in section 4, structured into five categories:

- **Clustering and user grouping**

In order to avoid interference, different base stations can be considered as a cluster where coordination is used to allow interference avoidance. Similarly, different users can be grouped together to allow coordination of the users' transmission. Several concepts serving as basis for the specific coordination schemes are provided

- **Inter-Cell Interference Coordination**

Two contributions for interference avoidance based on semi-static coordination techniques are described.

- **Coordinated Scheduling**

Interference avoidance based on dynamic coordination of schedulers is presented in five different sections, focusing on the impact of transmission modes, decentralization of the coordination algorithm and the impact of deployment in a heterogeneous network.

- **Scheduling for joint processing**

An assessment of joint processing techniques where data is transmitted to a user from several base stations is provided in two contributions.

- **Game theory based scheduling**

Models of both non-cooperative games as well as coalition games are proposed to study the effect of fading and mutual interference on base station coordination when users do not have complete information on channel states and the states of other base stations.

Conclusions on the presented innovations and an outlook on the next steps to be taken within WP1 are presented in Section 5.

## 2 Introduction

The main objective of the ARTIST4G Work Package 1 (WP1) is to build forward on the 3GPP Long Term Evolution (LTE) Release 8 baseline, proposing a novel fair mobile broadband technological framework in which to design innovative, practical, scalable and cost-effective interference avoidance solutions. Such an approach will enable the identification of optimal strategies also taking into account the practical implications on the real system.

In particular, as specified in the ARTIST4G Description of Work, the specific aim of Task 1.2 of WP1 is:

- to propose and to define innovative scheduling and cross layer design techniques to be applied at the transmitter end of a communication system, in which also a certain level of coordination/cooperation is introduced among different transmission points, in order to achieve interference avoidance;
- to define innovative interference management strategies in heterogeneous deployments, with the aim of avoiding/limiting the mutual interference also considering the case of inter-topology and inter-operator interference.

The performance of multi-cell systems can be substantially increased, if the available resources are assigned in a coordinated manner. To achieve such optimal scheduling decisions in a real system, detailed channel information has to be exchanged between base stations and advanced algorithms need to be applied. This deliverable D1.3 presents the various technical approaches at the Layer 2, which are under investigation within the scope of Task 1.2 of WP1, aiming at taking full advantage of the interference avoidance potential.

In an earlier ARTIST4G deliverable D1.1 [ARTD11], the preliminary requirements of the targeted innovations have been assessed with respect to the impacts on the existing RAN architecture. In the current document a more detailed technical description of these innovations is provided, including basic ideas, potential performance improvements, simulation results, realization options and potential implementation restrictions.

Deliverable D1.3 is related to deliverable D1.2 [ARTD12], which provides the basic physical layer signal processing methods linked to the addressed innovations. D1.2 also describes the practical constraints of the physical layer, to which the advanced resource assignment strategies and scheduling algorithms of D1.3 have to be adapted. In the following the structure and content of deliverable D1.3 is briefly introduced.

In section 3 an overview description of the main generic Layer 2 aspects of interference avoidance schemes is given. First, the definitions of partial Channel State Information (CSI), centralized architecture and distributed optimization algorithms, as they are used within this deliverable, are presented. Then the generic aspects of the two types of coordinated multipoint schemes, joint processing and coordinated scheduling/coordinated beamforming are explained. In contrast to joint processing, coordinated scheduling/beamforming does not require exchange of user data between enhanced Node Bs (eNB), reducing the bandwidth requirements over the backhaul link. Also the impact of centralized versus decentralized scheduling, as well as clustering and user grouping, is addressed. Clustering and user grouping seem to be relevant topics in joint processing and coordinated scheduling/beamforming areas, since they can significantly influence the potentially achievable performance gains.

The choice of centralized and decentralized architectures for interference avoidance schemes depends on the type of deployment, backhaul capacity and latency. For example, intra-eNB Coordinated Multi Point (CoMP) schemes are naturally centralized, whereas inter-eNB CoMP approaches involve the coordination of cells belonging to different eNBs and the exchange of information requires the use of backhaul links. For inter-eNB CoMP schemes, decentralized approaches can be introduced.

Interference avoidance schemes can be also designed as semi-static or dynamic. The basic differences between both approaches are the time scale in which they are defined, and their flexibility. Important Layer 2 (L2) aspects to be considered in this context are the availability of CSI at the eNBs, and the signalling overhead requirements.

In general, downlink (DL) or uplink (UL) assumptions, and the use of Frequency Division Duplex (FDD) or Time Division Duplex (TDD) modes, have a direct impact on the requirements of

interference avoidance schemes. The interference profiles in the DL and in the UL are by essence different, which will lead to different interference management algorithms. The choice of FDD or TDD modes will affect, for example, the availability of CSI in the eNBs. Issues like synchronization requirements and channel reciprocity are highlighted. The impact of a heterogeneous network, i.e. a wireless network where macro eNBs can coexist with relays, pico-cells, and/or HeNBs [ARTD11] is also addressed. In particular, the importance of the definition of a model for the different properties of the interference in homogeneous (macro eNBs only) and heterogeneous networks is discussed. Finally, traffic models and key performance indicators and their relation to 3GPP definitions are introduced.

Section 4 comprises the description of specific resource allocation and scheduling algorithms for interference avoidance. The activities in Task 1.2 are divided into 5 main classes of innovations, which are reflected by the subchapters of section 4: clustering and user grouping, inter-cell interference coordination, coordinated scheduling, scheduling for joint processing and game theory based scheduling.

**Clustering and user grouping** are essential enabling techniques for most of the interference avoidance schemes addressed in Task 1.2, targeting at feasible and practically applicable schemes. We consider a static clustering approach for distributed coordinated scheduling and a semi-static clustering technique where the clusters are updated based on user measurements. Partial CoMP clustering is a user-centric clustering approach proposed to achieve simultaneously high penetration rate and limited backbone and feedback overhead. In the user grouping topic, user grouping for transmission to multiple users based on channel orthogonality metrics is considered, and subcarrier pairing is performed in the framework of single carrier frequency division multiple access (SC-FDMA) to solve impairment problems.

**Inter-Cell Interference Coordination (ICIC)** is a low-overhead cooperation scheme, i.e., only long-term measurements are needed. In the activity described in this section, a graph-based dynamic spectrum allocation scheme for interference coordination between femto cells and macrocells is evaluated. In heterogeneous and femto cell deployments, eNB/Home eNB (HeNB) and HeNB/HeNB interference coordination is also addressed with blind and distributed power control and Radio Resource Management (RRM).

The third main topic of Task 1.2 is **coordinated scheduling**. In this activity, research on a base station coordinated beam selection approach is performed, where the coordination of the selection of beams at neighbouring cells is based on restriction requests or signalling messages. In the distributed scheduling for beam coordination approach, the impact of the antenna downtilt adaptation in combination with a beam coordination algorithm is addressed. Precoding optimization algorithms for coordinated beamforming and algorithms to dynamically select transmission modes, where the proper precoding scheme is decided by the scheduler, are also considered. Finally, the potential gain of coordinated scheduling in heterogeneous scenarios is investigated.

**Scheduling for joint processing** in CoMP systems is the fourth main topic of Task 1.2. Ideally, a scheduler for joint processing should include several dimensions, such as time, frequency, space, and selection of the subset of users to be served. One objective is to investigate the impact of scheduling on the performance of downlink joint processing. A proper Signal to Interference and Noise Ratio (SINR) based scheduler may reduce the interference, decreasing then the impact of the gains related to joint processing. For the partial CoMP clustering technique, the objective is to design a low complexity scheduler. In the user configuration approach, the scheduler decides which users benefit from joint processing and configures the best joint processing scheme to serve the user, from a set of predefined options.

The last main topic covers a very useful and intuitive framework to study coordination of distributed scheduling, the **game theory based scheduling**. The field provides a rich set of tools to not only evaluate performance limits through base-station interaction, but also leads to the development of distributed mechanisms to achieve optimal performance. Models of both non-cooperative games as well as coalition games are proposed to study the effect of fading and mutual interference on base station coordination when users do not have complete information on channel states and the states of other base stations.

Finally, in section 5 an overall view of the preliminary assessment of the described innovations is given and relationships to planned trials [ARTD61] are indicated. Next steps and directions of the work on these innovations within ARTIST4G are pointed out.

## 3 Generic L2 aspects of interference avoidance schemes

The implementation of any wireless communication scheme that is explicitly designed to avoid interference will have to take into account many aspects of the system that were not necessary as important in a standard cellular wireless system without interference avoidance. While generic aspects and specific innovations on the physical layer have been described in [ARTD12], this deliverable is concerned with aspects on the media access control layer. An overview of the most important aspects on this layer will be described in this section. Starting out with a set of definitions in section 3.1 we proceed in section 3.2 to outline the different types of Coordinated Multipoint schemes. The subsequent sections describe Layer 2 issues with respect to centralized and decentralized architectures (section 3.3), dynamic and semi-static schemes (section 3.4), downlink and uplink (section 3.5), the impact of duplexing modes (section 3.6) and homogeneous and heterogeneous network scenarios (section 3.7). Finally section 3.8 elaborates on the traffic models that can be used in the evaluations and on key performance indicators used to assess different schemes.

### 3.1 Definitions

In general, CoMP schemes can be classified as centralized and decentralized and the coordination algorithm can be distributed or non-distributed. This classification is tightly linked to the definition and usage of partial Channel State Information at the Transmitter (CSIT) for each user. We apply the following definitions (cf. [ARTD12]):

#### Partial Channel State Information

There are various forms and definitions of partial CSI. The three most important ones are described below.

- **Partial CSI based on incomplete information:** each eNB acquires only a subset of the coefficients for the global CSI matrix. For instance, the eNB in cell  $i$  obtains CSI for users served by cell  $i$  but not for other users. In another example, the eNB obtains CSI related to the direct channel gains (i.e. to their eNBs) for all network users, but no information related to the channel from a user and the interfering eNBs.
- **Partial CSI based on statistical information:** this scenario is similar to the one above, but some statistical information (mean, variance, correlation coefficients) is added to the partial instantaneous CSI for some of the missing CSI matrix elements. This extra information helps the eNB refine its optimization of the transmission parameters.
- **Partial CSI based on imperfect information:** in this case, the eNB acquires all or a subset of the CSI matrix coefficients, however the coefficients are only imperfectly represented, due either to channel estimation errors or to quantization effects over the feedback channel.

#### Centralized Architecture

In a centralized architecture of multi-cell processing or coordination, the CSI needed to compute the optimal transmission decisions is collected to a single central physical entity (which could be co-located with one of the eNBs or possibly implemented in a separate location of the network). This physical entity is referred to in the following as the Central Coordination Node (CCN). The CCN processes the channel/user information and computes the final decisions which are then distributed to the eNBs involved in the coordination cluster or set of collaborating eNBs. For instance, in Coordinated Beamforming, the CCN collects all CSI and computes all the beamforming weights required to pre-code the data from each of the eNBs. The beamforming coefficients pertaining to a given eNB are then sent to this eNB alone, which exploits them to perform the local beamforming operation.

In the example of Joint Processing a similar centralized architecture can be used. However another variant of a centralized architecture can be envisioned where the CCN in addition to computing the beamforming coefficients, also collects the user data to perform the actual beamforming operation on the data. In this case, the CCN sends the final precoded data to the eNBs. The eNBs can then map the precoded data to the transmit antennas and launch it over the air after some standard upconversion and filtering operations.

### **Decentralized Architecture**

In the decentralized architecture of coordinated scheduling, beamforming or joint processing, there is no CCN. Rather, the computation of the coordinated scheduling or beamforming decisions are carried out individually by each one of the eNBs and implemented locally as well.

### **Distributed optimization algorithms**

A distributed optimization of a coordination or CoMP scheme refers to the capability of computing the transmission decisions (beamforming coefficient, power level, subcarrier usage, scheduler user index, etc.) based on non complete CSI data. Therefore this relates to the mathematical nature of the employed technique rather than where it is physically implemented (in this latter case one will refer to above described centralized vs decentralized architecture).

An example of distributed coordination is illustrated by distributed coordinated scheduling where each eNB makes a scheduling decision primarily based on the link quality and interference information reported by its own cell users, in the absence of link quality information reported by other cell users.

Also, a distributed Joint Processing CoMP scheme refers to a scenario where a eNB computes the beamforming matrix to be used at this eNB alone, based on partial CSI only.

## **3.2 Coordinated multipoint schemes**

### **Joint Processing**

At a first glance Joint Processing (JP) has many similarities with Multi User Multiple-Input Multiple-Output (MU-MIMO) schemes, leading to similar challenges regarding scheduling, optimal resource allocation as well as signalling and feedback design. One of the main challenges known from MU-MIMO is to find suitable pairs of uncorrelated users achieving higher throughput compared to single user transmission, as there is a trade-off between increased usage of spatial resources and performance loss due to mutual interference. JP has an even more severe user grouping challenge due to the typically increased number of simultaneously supported User Equipments (UE).

In addition, investigations so far clearly indicate that advanced JP solutions for cellular mobile radio systems will have to include both intra- as well as inter-site cooperation because without inter-site cooperation cell edge UEs will experience strong interference from adjacent sites that cannot be mitigated. This will involve the backhaul network and will require potentially complex common or aligned scheduling over different sites. The required backhaul links might be bottlenecks for user and control data exchange. They will also include more or less additional delays for these data. As for JP the precoder performance will be affected from inaccurate or missing user data as well as CSI outdated. The scheduler at the Central Unit (CU) of a Cooperation Area (CA) has to take these effects into account for allocation of certain resources to certain sets of cooperating UEs. This might affect the link adaptation, selection of Modulation and Coding Schemes (MCS) and might lead to further signalling overhead.

In particular, JP affects the following L2 issues. Firstly, multi-dimensional scheduling is required in frequency, time, space and groups of users, similar as for MU-MIMO, but for an increased number of UEs compared to typical MU-MIMO scenarios. Basically JP is adding a further dimension to the scheduler, which has now to find the optimum frequency, time, spatial precoding (beamformer plus power) allocation and additionally the optimum set of cooperating users. As the selection of the best performing set of cooperating users is a non-convex optimization problem, no efficient algorithm to achieve the overall upper bound is known, leaving so far only the option of exhaustive search for all possible user groupings. The schedulers become even more challenging in the case in which there are no predefined CAs, i.e. fully adaptive setup of CAs would be allowed. CAs might be defined as either network or user centric, either fixed or adaptive, either for sets of cells or overlapping so that one cell might participate in different CAs. In such a case, exhaustive search over all possible clusters of CAs would be required. That is why useful heuristics have to be found to limit scheduler complexity and to at least approach the overall optimum. Optimum user grouping as well as clustering, i.e. the setup of optimum CAs (which again would need ideally network wide exhaustive search) will be one of the main activities treated in WP1. The selected clustering approach will affect possible system level gains as well as the scheduling decisions.

Typically the number of data streams per Physical Resource Block (PRB) will have to be adapted depending on channel conditions. This is a trade-off between maximum penetration rate of users benefitting from CoMP, optimum usage of available spatial resources (ideally full frequency one reuse), gain per user group due to interference mitigation versus loss due to power normalization. So ideally the scheduler should compare achievable throughput for single cell transmission, single UE transmission per CA (macro diversity) and for a variable number of cooperating UEs. The penetration rate of CoMP users is important as even very large gains for specific user groups will have only marginal effect on system level gains, if these gains can be achieved only for a very low number of UEs. The scheduling of CoMP and non CoMP UEs requires additional careful consideration because typically CoMP UEs have to be scheduled first due to the need of finding common resources on all cooperating cells. This might affect frequency dependent scheduling gains for non CoMP UEs negatively. Depending on traffic model and channel conditions the overall system level performance might be affected.

Moreover, there is an interrelation of the precoding concept with scheduling, channel estimation, reporting and the optimum resource allocation. Singular Value Decomposition (SVD) or unitary precoding will have to adapt MCSs per data stream due to varying inter stream interference conditions which requires additional feedback. For Zero Forcing (ZF) typically UE specific power normalization will be done so that recalculation of the optimum MCSs can be done by the eNB scheduler without further feedback. The switching between CoMP, MU-MIMO, SU-MIMO modes can be transparent or non-transparent and affects the Channel Quality Indicator (CQI) reporting as well as the overall system level performance. A fast transparent switching between different MIMO and CoMP modes might require individual CQI reporting per mode so that the correct MCSs can be selected.

In addition, the handling of inter-CA interference might lead to specific resource allocation or scheduler restrictions, which might be defined over larger areas of the network or even for the full network. This can be seen as specific combination of ICIC mechanisms with CoMP. Backhaul bottlenecks might require to limit cooperation to a subset of UEs, based on different selection criteria from maximum gain to highest priority or lowest feedback overhead. In addition, precoding losses due to X2 delay (the delay on the inter eNB connections) have to be taken into account for the selection of proper MCSs as CSI outdated, i.e. the difference between the CSI at time of measurement and the time of precoding, has a strong effect on achievable JP CoMP gains.

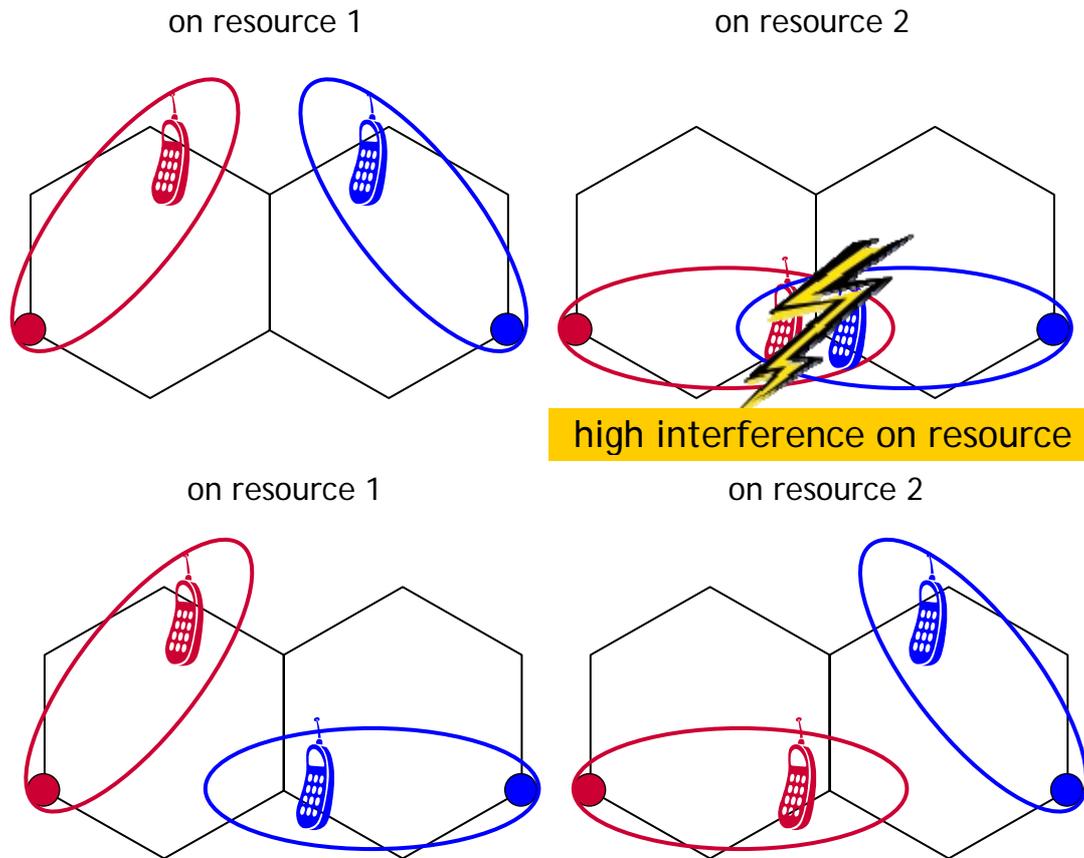
In general, JP can be realized by centralized as well as distributed precoding and/or scheduling. A simple concept is to multicast user data to all cooperating eNBs and to exchange only CSI as well as scheduling information over the X2 interface. Under the assumption that all eNBs calculate exactly the same precoding matrices and scheduling decisions the result is the same as for JP with one single CA.

In summary, the implementation of JP touches on many issues in the design of algorithms on layer 2, some of which have been described in some detail in this section. There are several other issues like optimized timing advance for JP CoMP, Hybrid Automatic Repeat Request (HARQ) optimizations, frequency and time synchronization between cooperating eNBs etc. which might also affect the design of any practical system.

### **Coordinated Scheduling / Coordinated Beamforming**

Interference mitigation by Coordinated Scheduling / Coordinated Beamforming (CS/CB) relies on the appropriate selection of precoding or beamforming weights to exploit the additional degree of freedom introduced by coordination (see following Figure 3-1).

Interference caused to other cells varies with the beamforming or precoding vector used for transmission. For the specific implementation of codebook based precoding or beamforming, as defined in the current Rel-8 of the 3GPP standard, the interference is a function of the assigned Precoding Matrix Index (PMI). In CS/CB an eNB therefore selects a UE and thereby a PMI that shows low mutual interference with UEs scheduled by interfering cells. As an additional degree of freedom, this coordination is applied separately per resource block, where the smallest resource block consists of a pair of PRBs adjacent in time.



**Figure 3-1: Unfavourable scheduling decisions lead to high interference (top). This can be mitigated by coordinating scheduling decisions among base stations (bottom).**

For CS/CB it is required that cooperating eNBs exchange information enabling a resource allocation which minimizes the resulting interference. Depending on the used algorithm this can be scheduling information and/or information on interference caused by the eNBs in other cells. In contrast to joint processing, no exchange of user data is required. This drastically reduces the bandwidth requirements to the backhaul link.

**Additional channel measurements**

To determine beneficial scheduling decisions, the eNBs must be able to predict the interference caused by the targeted resource allocation. This requires knowledge of the inter-cell channel between UEs and their neighbour cell eNB or between the eNB and the UEs of the neighbour cell. This information can be given in the form of, for example, the spatial correlation or the interference per PMI. Therefore, additional inter-cell channel measurements by the UE for DL and by the eNB for UL, respectively, are required and an appropriate measurement procedure needs to be defined. In TDD, the channel reciprocity might be exploited and eNB measurements could to some extent be used for DL, too, saving UE to eNB reporting overhead in UL.

**Centralized and decentralized scheduling**

Based on such interference predictions, and taking fairness constraints into account, the scheduler selects for each time-frequency resource such combinations of UEs that show low mutual interference, improving SINR and therefore increasing throughput. Such coordinated scheduling can be done either centralized or decentralized. Hybrid schemes like centralized within a cluster and decentralized among neighbouring clusters (yielding a 2<sup>nd</sup> level of coordination clusters of clusters) are also possible.

In the centralized coordinated scheduling approach, the CSI in form of channel measurements or already computed interference predictions per schedulable UE is reported from the

coordinated eNBs to the central entity (which is possibly one of these eNBs) along with information describing these schedulable UEs in detail, e.g. load or Quality of Service (QoS). The central entity then carries out scheduling for all cells altogether and distributes its decisions back to the eNBs for execution and actual transmission of data according to these decisions.

In decentralized scheduling approaches, the exchanged information heavily varies with the properties and requirements of the specific scheme. Using a scheme with an approach similar to centralized scheduling, CSI and information about schedulable UEs are exchanged among all eNBs within the cluster, which is then used by each individual eNB to determine its scheduling decisions within the cluster. This requires non-overlapping clusters and the very same well-defined scheduling algorithm to be carried out by each eNB within the cluster, so that the outcome at each eNB is the very same scheduling, from which each eNB only uses those scheduling decisions relevant to its own UEs.

At the other end of distributed scheduling schemes, only final scheduling decisions along with scheduling constraints to other eNBs are exchanged among coordinated eNBs. Such constraints are the outcome of each eNB's local scheduling decisions and are the means by which a reduction in mutual interference and an improved throughput is finally achieved. Such a scheme is well-suited for overlapping clusters.

### **Impact of clustering**

To achieve high gains with CS/CB, most of the strongest interferers should be located within an eNB's coordination cluster. Due to shadowing, the geographically nearest neighbours are not always the strongest interferers. Nevertheless, a smaller cluster size might be desirable to reduce backhaul requirements as well as inter-cell measurement requirements. A special case of non-overlapping clusters is intra-site respectively intra-eNB coordination. This special case doesn't pose any further requirements to the backhaul. Only the air interface specific requirement of additional inter-cell measurements remains.

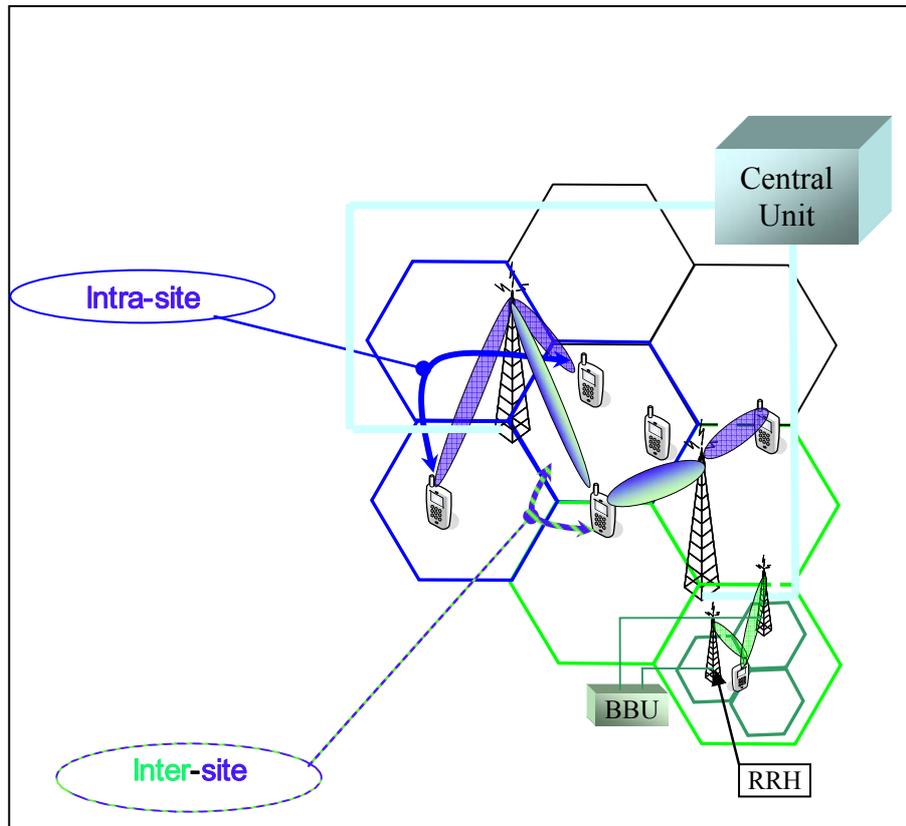
Intra-site coordination has the drawback that only a fraction of the strongest interferers can be coordinated, which accordingly reduces the gains of CS/CB. Non-overlapping clusters in general have the drawback that cells at the cluster border are penalized compared to cells in the cluster centre, since parts of their strongest interferers do not fall within their coordination cluster. At the same time, though, non-overlapping clusters have the advantage that each cell within the cluster has the very same set of cells within its own cluster as any other cell in that cluster. This gives more degrees of freedom to the design of coordination algorithms and enables algorithms which would not be feasible with overlapping clusters. This typically leads to better overall performance, but at the expense of possibly poor cluster edge performance.

Summarizing, CS/CB schemes address L2 aspects with relation to characterization of UEs derived from channel measurements, the coordination of scheduling decisions of a certain number of eNBs and the exchange of control information between these eNBs.

### **3.3 L2 issues of centralized and decentralized architectures**

Downlink CoMP transmission is expected to have a significant impact on the network infrastructure. The choice of centralized or decentralized approaches for coordinated multipoint schemes depends closely on the type of deployment, and the backhaul capacity and latency.

For example, as shown in the following Figure 3-2, for intra-eNB CoMP the coordination is performed by the same eNB, naturally resulting in a centralized architecture. In addition, the exchange of the necessary cell-related information does not involve the backhaul. Two different scenarios fall under the intra-eNB denomination: coordination between the sectors of the same site and Radio Remote Heads (RRH) controlled by the same Baseband Unit (BBU) via Radio over Fiber (RoF). However, inter-eNB CoMP involves the coordination of cells belonging to different eNBs; consequently, the exchange of information involves the backhaul. In that case, a centralized architecture requires a central unit performing the coordination between the different eNBs. This central unit is actually a logical entity that can be located in practice in any of the eNBs.



**Figure 3-2: An illustration of intra-site, inter-site and RRU CoMP architecture**

In the centralized approach, a central unit performs the scheduling and decides on the user precoding jointly for all the UEs in the cluster. This kind of approach can apply both to joint processing and coordinated scheduling/beamforming techniques. Typically, the central unit gathers the CSI of all the users in its area in order to perform joint scheduling and transmit signal processing operations. The main downside of this approach is the vast amount of backhaul required for information exchange between the involved base stations in the case of inter-eNB CoMP.

The centralized approach is challenging in the case of inter-eNB CoMP, because a low-latency backhaul is required in order to convey accurate CSI to the central unit (which has to be located outside - at least some of - the coordinated eNBs) and then to communicate the scheduling and possibly precoding decisions to the coordinated eNBs. In addition, joint processing requires the backhaul to be high-capacity since the serving eNB needs to send the user data to the other eNBs participating in the transmission.

The decentralized approach allows for lower latency backhaul, and for having only a partial or even no central unit. This approach can be adopted for both joint processing and coordinated scheduling or beamforming.

Since less backhaul capacity is required, CS/CB appears more suited to inter-eNB CoMP with moderate backhaul capabilities. One possibility to avoid a central unit is to adopt a Master/Slave approach, where the serving cell chosen as Master communicates to coordinated interfering cells (the slave cells) the time-frequency resources that will be used for transmission to the scheduled UE, together with the constraints for the coordinated cells' schedulers, which should try as much as possible to respect these constraints while performing their own resource allocation. In that case there is no need for a large capacity backhaul but a low latency remains important, although some small delay can be afforded provided the CSI does not change too rapidly (i.e. for low users velocities).

An alternative technique aiming at avoiding using a central unit is proposed in [PHG09]. The principle is that all the UEs report their CSI to all the eNBs in the cluster. Provided the eNBs

have the same buffer states and they receive the same CSIs from the UEs, they will make the same scheduling and precoding decisions. Obviously, such an approach is sensitive to different error patterns affecting the feedback links to the various eNBs. Solutions to this drawback were discussed in [SCWY].

Another framework for decentralized approach is described in [3GPP-R1093141] and is based on Over The Air (OTA) signalling. In the proposed scheme there is no direct information exchange between the eNBs. Instead, the users send request messages to all cells in the cluster in order to perform link adaptation.

Alternatively, hybrid approaches exist where some operations (e.g. the scheduling) are performed by a centralized unit, whereas others (e.g. the precoding weights computation) are performed locally [WIN14]. To perform the user scheduling, each eNB transmits to the central unit all the CQI between them and all their UEs. Upon receiving the scheduling decisions from the central unit, all the eNBs estimate the CSI and derive the corresponding precoding coefficients. This hybrid approach limits the transferred information over the backhaul as well as the computational complexity in the central unit.

### 3.4 Dynamic and semi-static schemes

Interference avoidance schemes, e.g. ICIC, coordinated scheduling/beamforming or joint processing, can be semi-static, or dynamic. The basic differences between both approaches are the time scale in which they are defined and their flexibility. Semi-static schemes are defined on long time scale, whereas dynamic schemes are defined on short time scale (see also [ARTD11]) hence, they can adapt to the variability of the channel, the mobility of the users, the variability of the user requirements or the possible changes in the distribution and number of users over the service area.

The implementation of a semi-static or dynamic interference avoidance scheme will have a certain impact on L2 aspects, some of which are described in the following.

#### CSI acquisition

The basic question (in DL), is how the CSI is available at the transmitter node. Clearly, semi-static or dynamic interference avoidance schemes have different requirements in the amount of CSI that is needed. Dynamic schemes naturally increase the complexity of the system, since the CSI is needed in a short time scale.

FDD and TDD modes also have an impact in the requirements of semi-static and dynamic interference avoidance schemes. FDD increases the complexity of the channel estimation and requires the use of feedback mechanisms. In TDD, it is possible to use the reciprocity of the channel, but it should be noted that the interference distribution is not reciprocal.

Semi-static and dynamic schemes also present different levels of robustness with respect to the impact of imperfect CSI. Semi-static schemes are more robust in this respect. Here, the imperfect CSI includes channel estimation errors, feedback errors, backhauling errors, delayed CSI and synchronization errors.

#### Scheduling

In the context of interference avoidance schemes, the scheduling function should consider several aspects, such as the optimum allocation over frequency, time and space resources or the optimum set of users to be served given a certain objective function. Clearly, scheduling decisions in dynamic interference avoidance schemes can improve the performance of the system, but this enhanced performance comes at the cost of a higher overhead and computational complexity.

#### Overhead

The overhead related to a given interference avoidance scheme can be split up into several categories. The basic ones are the signalling overhead, which is related to the CSI acquisition (e.g. number of pilots needed, feedback load) and the backhauling overhead, or how the information is exchanged between the nodes that are involved in the interference avoidance scheme.

Dynamic interference avoidance schemes require a higher and frequent (short time scale) exchange of information between nodes. Hence, the number of cells or users served involved in

the scheme should also be considered as a design constraint, depending on the available/required backhaul. Notice that decentralised solutions may help to alleviate the overhead of dynamic interference avoidance schemes, still providing an improved performance with respect to a semi-static approach.

### **Clustering and user grouping**

Clustering and user grouping are approaches that arise to reduce the overhead of schemes implying some level of coordination/cooperation between several nodes. Clustering and user grouping approaches are discussed in more detail in section 4.1. Basically, semi-static clustering techniques are proposed to reduce the overhead requirements, but introduce inter-cluster interference. Dynamic clustering techniques, where the nodes included in one cluster change over the time, based on different optimization metrics, reduce the problem of inter-cluster interference, but increase the overhead requirements.

In the framework of CoMP schemes, dynamic and semi-static interference avoidance approaches can be characterized with respect to different features. One of them is the feasibility of dynamic and coherent CoMP approaches in realistic scenarios. Another feature is the definition of serving cell which might have to be extended (the user only receives its control channel from a single cell).

In summary, there is a clear trade-off between the performance and complexity of interference avoidance schemes. While in particular dynamic interference avoidance schemes can improve the performance of specific users, still a significant overall gain in relevant system-wide Key Performance Indicators (KPI) is needed to motivate their use.

## **3.5 Downlink and uplink specific aspects**

The interference profiles in the downlink and in the uplink are by essence different, which will lead to different interference management algorithms.

In the downlink, the average SINR observed by a UE decreases with the distance to the cell-centre for two reasons: the path gain to the serving eNB decreases while the path gains to the main interferers increase. Thus, the cell-edge throughput is drastically reduced in interference limited scenarios, where the noise level is negligible with respect to the interference level, and interference management for downlink has been identified as a key point in the early study items of 3GPP-LTE.

In the uplink, the distribution of interference observed at the eNB on top of a UE signal is independent of the UE position in the cell. Thus, only the path gain to the serving eNB varies with the UE position, while the interference level changes from a sub-frame to the other via the randomness of the interfering cells' UL scheduler. The SINR drastically reduces when two close UEs of two neighbouring cells are scheduled on the same time/frequency resource.

These events' occurrence are negligible when the number of UE per cell is large, which explains why UL interference management has not been pointed out as a major research topic for macrocell deployments. However, when a low number of UE request a large throughput (e.g. as in HeNB deployment), and when two close UEs of two different cells are active, they interfere one with each other with a high probability. Thus, UL interference management is particularly useful to preclude long term outage for heterogeneous networks, even if the outage event itself is of low probability.

Of course, the link adaptation, schedulers, and power control strategies have a large impact on the interference distributions in the downlink and in the uplink and should be designed altogether.

If the interference management is done on a long-term basis (larger than several sub-frames), information relative to the average wideband SINR, cell load and power profiles are exchanged. If the interference management is done on a short-term basis (typically each sub-frame), the CSI and scheduling information can be exchanged between base stations.

In downlink, the CSI is measured by UEs and fed back to the base station, while in uplink the CSI is measured by the base station itself. Some interference management algorithms do not need any exchange between base stations, which is particularly useful when no coordination channel is available between two nodes (e.g. no X2 between eNBs and HeNBs).

### 3.6 Impact of duplexing modes

The choice of a duplexing method, namely FDD or TDD, impacts substantially the design and behaviour of a wireless network. Some elements are generic to any wireless system, while certain others are specifically important in a system adopting cooperation techniques at the heart of their operation.

Choice for adopting a specific duplexing method are expected to be made based on the generic aspects detailed in the following.

#### **Flexibility in the design of the uplink-downlink rate asymmetry ration**

TDD allows the system designer to dynamically adjust the UL and the DL bandwidth allocation depending on the traffic requirements. The asymmetry ration can be adjusted over time (slowly) and to some extent over the geographical areas (region wise). In contrast FDD systems operate on paired bands of spectrum of fixed width making it difficult to adjust with respect to traffic evolution.

#### **Frequency planning**

Since TDD systems do not require paired spectrum, it can be deployed in areas where single spectrum bands are available, making spectrum usage more efficient.

#### **Latency requirements**

Because the TDD frame is divided into consecutive alternate UL and DL resource slots, the DL (or UL) transmission is not continuous, as is the case in FDD mode. In certain classes of delay sensitive traffic, this makes it harder to satisfy strict latency requirements.

#### **Guard time**

TDD operation requires guard time to avoid an overlap between transmit and receive periods caused by propagation delay.

#### **Impact of switch on audio systems**

In TDD, the transmitter must quickly switch on and off to give way to receiver operation. The fast switches creates a discontinuity which can hamper certain audio equipments.

The aspects listed in the following are of specific importance in cooperation and interference management.

#### **Synchronization requirements**

TDD often presents the system designer with tighter synchronization requirements between cells. In particular, a lack of UL/DL synchronization between neighbouring cells may result in two eNBs interfering directly with each other. Such interference can be severe due to possible line of sight conditions between the base stations and due to high transmit power levels. This situation cannot occur in properly designed FDD systems where a base station cannot receive on the same spectrum where another one transmits. In FDD this advantage comes at the price of a (usually expensive) duplexer.

#### **Channel reciprocity**

Because TDD systems operate on the same frequency for both UL and DL, the channel state information is identical (up to the coherence time) in the UL and DL. This is not the case in FDD where the frequency separation between UL and DL is enough to fully decorrelate the fading coefficients between these links. Thus in FDD the only way to acquire channel state information at the eNB is via a feedback channel from the user terminals over which the channel information is quantized. In TDD, the channel information is obtained through conventional training methods and reused in the next transmit slot.

The availability of channel information at the transmitter without the quantization loss and delay imposed by a feedback channel further encourages the use of advanced channel-aware transmission and scheduling techniques. This is particularly important in cooperation methods based on joint multi-cell transmission and scheduling which require fast and accurate channel state information.

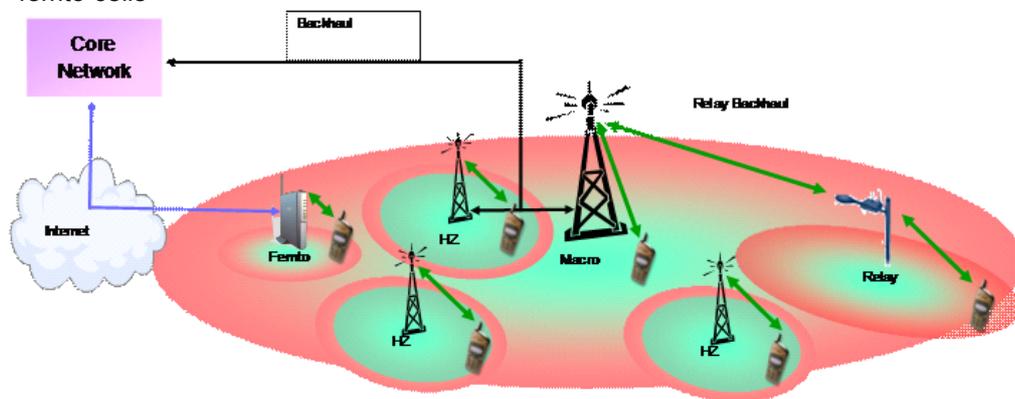
### 3.7 Homogeneous and heterogeneous scenarios

The basic concept of any cellular mobile communication system involves the idea to cover a given service area with a number of immobile radio access points that serve users within the reach of the emitted radio waves. The area covered by a single radio access point is also called the cell size which is a fundamental parameter in the performance assessment and the capacity of the system.

A very simple approach to increase network capacity is the reduction of cell sizes using additional radio access points (each serving a smaller area), which, if regular (also called “macro”) radio nodes are used, is an expensive option. Therefore, various concepts have been proposed to increase network capacity with lower infrastructure requirements, as shown in Figure 3-3.

In particular, the following concepts are currently discussed [3GPP36814]:

- relays
- hotzones (HZ)
- remote radio heads (RRH)
- femto cells



**Figure 3-3: Illustration of a heterogeneous network.**

The additional nodes associated with these concepts differ in e.g. backhaul connection, user access and deployment options. A commonality of these different nodes is the property that transmission power is significantly lower compared to macro nodes. A network with immobile radio nodes with different transmission powers is also referred to as a Heterogeneous Network (HetNet) in contrast to a homogeneous network with only macro nodes. If the low power nodes transmit in the same frequency band as the macro nodes, the network is considered a co-channel HetNet as opposed to a separate-carrier based HetNet where different carriers are used for low power and macro nodes.

Due to the different output powers, potentially unplanned deployment and restrictive user access modes, the characteristic properties of interference caused by neighbouring cells can differ dramatically in a co-channel HetNet compared to those perceived in a homogeneous network. A specific example is a heterogeneous network involving femto cells that are installed in the consumer's home and connected to the network backhaul via high-speed fixed connections, e.g. digital subscriber lines.

The advantages of introducing these femto cells include the very good indoor coverage that results in high user throughputs, the ability to allow for offloading of macro nodes, i.e. the macro nodes can serve a smaller number of users with improved service, and the overall better resource reuse in the network. On the other hand, an unplanned user-installed deployment of femto cells can pose significant challenges in terms of interference caused to out-of-cell users. This can become particularly problematic if the access to a femto cell is restricted to a Closed Subscriber Group (CSG), i.e. a small group of users that, e.g. live in the vicinity of the femto cell or own it.

A femto with a defined CSG serves only those users that belong to its CSG. This restricted access could pose a serious problem to a user who is located very close to a femto but cannot

connect to it because the user is not included in the corresponding CSG. Such a user may experience strong interference from the nearby femto cell resulting in poor downlink reception from the user's serving macro eNB.

The accurate modelling of different network types requires a multitude of assumptions on the parameters of the system. Examples for these assumptions include channel model parameters (e.g. path loss, shadowing, angular spread at transmitter or receiver), placement of mobile users and radio access points, vehicular speeds, data traffic models, transmit power, receiver imperfections, level of detail in modelling, etc. In order to allow for a comparison of different techniques, some agreement on parameters with fundamental impact on performance needs to be achieved. Specific channel model and scenario descriptions [3GPP36814], [3GPP25996], [NGMN08],[FF09] are therefore used to specify in detail placement and properties of radio access points, mobile users, the radio communication channel and data traffic models that need to be simulated. Of specific importance for the evaluation of interference avoidance schemes is the definition of a model that captures in sufficient detail the different properties of the interference in homogeneous and heterogeneous networks. In particular, the relevant ("dominant") interference observed by the mobile user in the downlink can be caused by a smaller number of radio nodes in a heterogeneous network which therefore opens up a better opportunity for interference avoidance schemes because a smaller number of cooperating nodes is involved.

### 3.8 Key Performance Indicators

When assessing advanced interference management schemes, it is of great importance that the modelled system reflects a realistic usage of the mobile network. Within ARTIST4G, some guidelines for this assessment have been defined and presented in [ARTD51] together with performance indicators and evaluation scenarios. Similar to 3GPP [3GPP36814], the following KPIs are to be considered as possible metrics to assess the performance in the presence of full-buffer and models:

- Mean user throughput
- Throughput Cumulative Distribution Function (CDF)
- 5% worst user throughput

In addition, the Jain index can be used to illustrate different levels of fairness in system designs [ARTD51].

The full buffer simulation methodology has been proven as an effective way of estimating the capacity of mobile networks. However, the user data rate derived from full buffer traffic simulations is based on a very high network utilization level which does not always appear realistic. In fact, the user data rates offered by real mobile networks are inherently variable, since they depend on the number of active users, on the user location within the cell and on the network loading conditions. Therefore, the resulting interference conditions of the network will be impacted by the way the user traffic is modelled.

Traffic models for system performance evaluations have been assigned in 3GPP [3GPP36814]. System throughput studies can be assessed using full-buffer traffic model capturing continuous traffic and non-varying interference. But a step further is being considered by means of evaluations with time-varying interference that can be carried out using bursty traffic models. A standard model for bursty traffic simulation (see also [3GPP36814]) defines packet sizes and models the packet arrival times using a Poisson process. If the delay-specific performance characteristics of an innovation are assessed, additional performance indicators like user perceived throughput and its average over several packets can be used.

Finally, for heterogeneous network performance evaluation, the following performance metrics are considered to be of high priority for the assessment of system performance:

- Existing full buffer performance metrics
- Throughput CDFs are for all UEs, i.e., macro UEs and HeNB/pico Ues

Moreover, the following indicators are useful for system evaluation:

- Macro cell area throughput
- Fraction of throughput over low power nodes

- Macro and low power node serving UE throughput ratio

## 4 Description of specific resource allocation and scheduling algorithms for interference avoidance

In this chapter we describe and investigate the performance of specific resource allocation and scheduling algorithms for interference avoidance. We start by investigating different clustering and user grouping techniques. Related to interference control, we look into inter-cell interference coordination schemes and coordinated scheduling techniques, and then we look into scheduling for joint processing. We conclude the chapter by exploring the potential of a game theoretic approach for resource allocation.

### 4.1 Clustering and user grouping

CoMP transmission and reception with a number of base stations is a promising candidate to increase the spectral efficiency of a cellular system as well as its cell-edge user throughput. CoMP transmission and reception techniques have been mainly focusing on coordinated resource allocation and/or user scheduling among several collaborating base stations, exploiting the time, frequency and space domains. Since these schemes require exchange of control data only, such as CSI between collaborating base stations, they are regarded as feasible to implement in the near term. In JP CoMP, multiple base stations can collaborate also on the transmission and reception of user data. Under the assumption of perfect channel knowledge, perfect synchronization among collaborating nodes and negligible delays, the theoretical gains with JP CoMP are substantially larger than with coordinated scheduling.

From a practical point of view, one of the major drawbacks related to the implementation of JP, as the number of users and base stations increases, is the amount of feedback needed from the users (assuming FDD) and the large signalling overhead related to the inter-base information exchange. This last feature opened an active area of research in the topic of rate-constrained links between base stations or between the base stations and a central unit. Therefore, the design of efficient algorithms and principles that could reduce these complexity requirements is of great interest in the field of JP. On the other hand, an interesting trade-off between the performance of a JP CoMP system and the required amount of feedback from the users and backhaul exchange should be pointed out. This trade-off is one of the reasons for restricting the use of JP techniques to a limited number of base stations or areas of the system. The system is typically divided into *clusters* of cells, and the JP schemes are implemented within the base stations included in each cluster.

The cluster formation can be static, or dynamic. The static cluster formation specifies a predefined set of clusters of base stations which do not change in time, whereas the dynamic clustering approaches arise to take into account the changing channel conditions or to optimize a given system metric.

Static cluster formation suffers from inter-cluster interference. Increasing the size of the cluster may reduce this interference, but requires a higher signalling overhead. Hence, CoMP gains may be significantly reduced. Reference limited inter-cluster coordination may be introduced to mitigate the interference suffered by users located at the cluster-edges. However, this approach requires a higher number of spatial degrees of freedom, and the number of users that can be served is reduced. There is a trade-off between increasing the fairness among users (mitigating the interference of cluster-edge users) and improving the sum-rate. The cluster size parameters are analysed based on metrics that consider the fairness-sum-rate trade-off. To overcome the problems related to static clustering techniques, dynamic approaches are proposed.

From the point of view of where the decision is taken, clustering techniques can be divided into network-centric or user-centric. Typically, network-centric clustering techniques divide the network into a set of disjoint clusters of base stations, that is, one base station can belong only to one cluster. In the user-centric clustering approaches, one base station may belong to more than one cluster, depending on the parameter under consideration. From the user point of view, this means that in the cluster area, each user may have a different set of cooperating base stations. Finally, other interesting approaches that imply some clustering of base stations are the virtual/group cell and sliding windows approaches, where user mobility aspects are taken

into account, and the cluster of collaborating base stations is updated and moves over the system in order to keep the user in the centre of the cluster.

The activity of clustering and user grouping is treated in the following sections, focusing mainly on achieving practical and feasible schemes.

#### 4.1.1 User-centric clustering for partial joint processing

Currently, the benefits of Partial Joint Processing (PJP) schemes are being investigated, where different stages of JP between base stations (BSs) are defined based on a user-centric clustering approach. More precisely, the objective is to identify when and where in the cluster area there are gains that motivate the additional system complexity and backhaul load [BSX+10].

The PJP scheme lowers the backhaul load due to a decreased exchange of information in terms of CSI and beamforming weights. The PJP scheme considered here is a particular case of a full JP scheme. The PJP scheme dynamically defines active sets or subclusters of BSs for each user located in a static cluster area. Hence, it results in the formation of overlapping subclusters inside the static cluster of BSs, one per user. Note that these subclusters can be formed by different number of BSs.

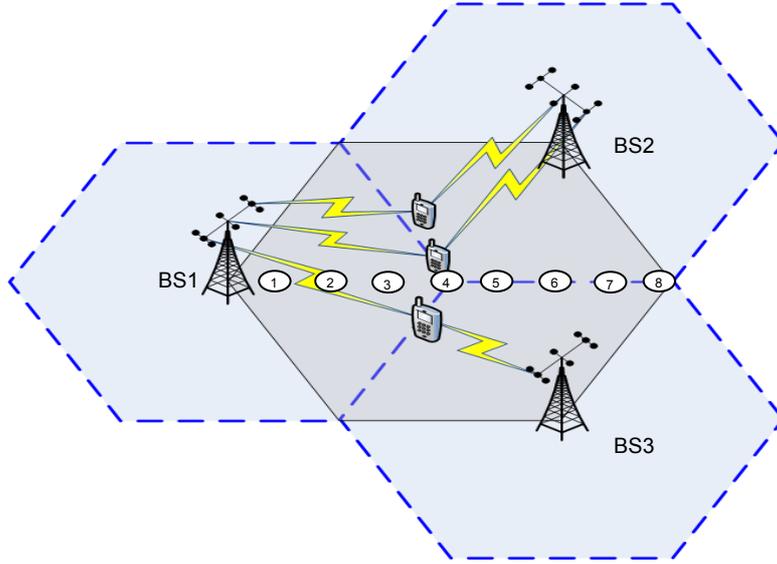
To form the serving subcluster for a given user, a threshold-based approach is used [PBG+04]. The user estimates the gain of the received channels, one from each BS, and defines its reference link or strongest channel, associated to a given BS (usually the serving BS). Then, the user compares the channel gains related to the remaining BSs with the reference link, and includes these BSs in its active set/subcluster only if their channel gains are above a relative threshold, with respect to the strongest channel. The threshold value is specified by the cluster (at CCN level).

From the system point of view, three benefits are provided: feedback reduction (users only feed back channels with an acceptable quality), lower inter-base information exchange (user data is only needed in the BSs included in its active set) and transmit power saving (power is saved from poor quality channels). However, the PJP scheme introduces multi-user interference in the system, since less CSI is available (compared to full JP) at the central unit or CCN to design the linear precoding matrix [ARTD12].

Similar approaches can be found in [PBG+08] [TCJ08]. In fact, these approaches can be seen as an extension to CoMP systems of the idea of [GA04], where users are scheduled depending on a threshold-based mechanism.

#### System description of the innovation

In frequency selective channels, the question of how to perform user-centric clustering techniques remains open. One way is to group the users in a first step, and then allocate each group of users to a given set of resource blocks (RBs). However, the focus is not on user grouping, as a first approach to the problem, the worst case interference scenario is considered where all the users are allocated to all the available RBs in each time slot. The frequency selectivity of the channel is exploited by performing the active set thresholding in every RB. This approach is called Frequency Adaptive (FA) thresholding. But, when the active set thresholding is performed based on the entire channel i.e., an average of all the RBs and applying this threshold to every RB is called Non Adaptive (NA) frequency thresholding [LBS+10]. The focus of the investigation that follows is on the potential gains with FA user-centric clustering approach.



**Figure 4-1: Cluster layout**

Recall the system model considered in section 4.2.1 of [ARTD12] where the framework for JP is defined. In this section, this model is reused with the cluster area as shown in Figure 4-1, where  $M$  users are considered for JP in each of the 8 positions. These users can be served by  $K$  BSs with  $N_t$  antennas each. The active set thresholding process can be represented by a binary matrix  $\mathbf{T}$  of size  $[M \times K]$ , where the  $(m, k)$ th element takes a value '1' or '0' for an active or inactive link between the  $k$ th BS and the  $m$ th user, respectively. Data transmission to a user  $m$  from a BS  $k$  takes place when the  $(m, k)$ th link is marked active. The active link parameter also establishes the amount of feedback (in terms of CSI) that is needed from a user and the backhaul load needed to transmit to the user. The backhaul load is here defined as the information that needs to be exchanged between the BSs and the CCN in a centralized framework. The BSs send the CSI towards the CCN. The CCN performs the precoding (beamforming and power allocation) together with the user data to be transmitted, and sends these values to the corresponding BSs as per the active links. Alternatively, the user data could be multiplied by the precoder at the CCN. In short, the number of active links relates to the system level KPI to minimize the complexity and control channel overhead, as defined in [ARTD51], but also to lower the backhaul load.

In this work, a distinct joint partial zero-forcing beamformer based on the definition of useful and interference channel matrices as in [WWK+09] [ARTD12] is used per RB and suboptimal power allocation is performed [ARTD12]. Under these assumptions, PJP with FA thresholding is performed in every RB and compared with NA frequency thresholding. With FA thresholding, the partial zero-forcing beamformer needs to be performed in every RB based on the useful channel matrix. This depends on the active set matrix  $\mathbf{T}^{FA}$  which can change in every RB compared to NA frequency thresholding, where the active set matrix  $\mathbf{T}^{NA}$  remains the same for all the RBs. The useful channel matrix [ARTD12] is modified to be processed for every RB as  $\mathbf{S}_x = [\mathbf{T}^x \otimes \mathbf{1}_{N_t}] \bullet \mathbf{H}(f_a)$ , where  $x$  denotes if FA or NA is being used,  $\mathbf{T}^x$  is the active set threshold matrix of size  $[M \times K]$ ,  $\mathbf{1}_{N_t}$  is an all ones  $N_t$  row vector,  $\mathbf{H}(f_a)$  is the channel matrix in the  $a$ th RB of size  $[M \times K \cdot N_t]$ ,  $\otimes$  and  $\bullet$  represent the Kronecker product and element-wise multiplication, respectively. The interference introduced due to the transmission to the  $m$ th user from the BSs is captured by the interference channel matrix as  $\mathbf{V}_{x,m} = [\mathbf{T}^{x,m} \otimes \mathbf{1}_{N_t}] \bullet \mathbf{H}(f_a)$ . The useful and the interference channel matrices are used to design the partial zero-forcing beamformer. The average sum-rate per cell (or per BS sector) per RB is given by

$$SR_{RB} = \frac{1}{K} E_H \left[ \sum_{m=1}^M \log_2(1 + \text{SINR}_m) \right], \text{ where } E_H \text{ is the expectation over } \mathbf{H}(f_a), \text{ and } \text{SINR}_m \text{ is the}$$

Signal to Interference Noise Ratio at the  $m$ th user. Refer to section 4.2.1 of [ARTD12] and [LBS+10] for more details regarding the design beamformer and  $SINR_m$  calculations.

In the PJP scheme, the optimal approach would be to perform the active set clustering or thresholding algorithm per RB, i.e., the subcluster of BSs transmitting to a user would be defined within each RB (FA). In this case, the increased granularity of forming the active sets of BSs in every RB, could potentially increase the sum rate or throughput per RB compared to the NA frequency thresholding. This increased granularity/accuracy of the active sets could also optimize the number of active links that needs to be feedback per user, potentially decreasing the CSI that the users need to feedback, and the backhaul load related to the exchange of CSI and precoding weights. However, this would imply that the subclusters of BSs could change in each RB for a given time slot. If the backhaul cannot take advantage of this adaptive channel feedback and beamforming data needs, the user data would need to be available in all the cooperating BSs at all the times. That trade-off is for further study by taking into account a more detailed backhaul network model. Here the focus is on investigating the potential throughput gains and backhaul savings, by not assuming any constraints on the backhaul adaptation capabilities.

### Performance results and future steps

Consider a static cluster consisting of 3 BSs with 3 antennas each with inter-site distance of 433m. The cluster layout as in Figure 4-1 is similar to that used in section 4.2.1 [ARTD12]. There are 8 predefined grid positions along a line from one of the base stations (BS1) towards the cluster center and along the cell-edge of the remaining BSs (BS2 and BS3). 6 single antenna users are uniformly dropped along an ellipse around each of the 8 positions. The system SNR or the reference value of one user located at the cell-edge is fixed at 15dB. The frequency selective nature of the channel is exploited by dividing the channel into various subcarriers. For the WINNER II channel model [WIN2D112], scenario B1, 256 subcarriers are used such that there are 256 RBs with all the users being served in every RB. A range of PJP threshold values are considered, such as 3dB, 5dB, 10dB, 15dB, 20dB and 40dB. The simulation parameters are summarized in Table 4-1.

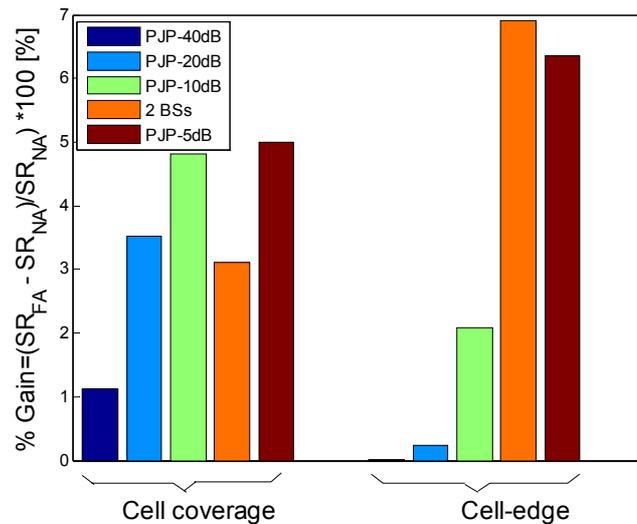
Figure 4-2 shows the gain in average sum rate per cell per RB due to FA thresholding compared to NA frequency thresholding. Refer to Figure 4.18 of [ARTD12] for the average sum rate per cell per RB with NA frequency thresholding. The cell coverage area corresponds to positions 1 to 3 while the cell-edge area corresponds to positions 4 to 8. These grid positions are shown in Figure 4-1. Thus, the FA thresholding outperforms the NA frequency thresholding, with gain in sum-rate per cell per RB. The 2BS case is a special case of PJP, where the best two BSs always serve a user.

Figure 4-3 shows the comparison of the relative average number active links per RB with FA and NA frequency thresholding. The negative values indicate that the average number of active links with FA thresholding is lesser relative to NA frequency thresholding. The FA technique makes the active set threshold more accurate, as they are performed per RB. Due to this, there is lesser number of active links leading to savings both in CSI feedback and backhaul load in terms of CSI and precoding weights, when compared to NA frequency thresholding.

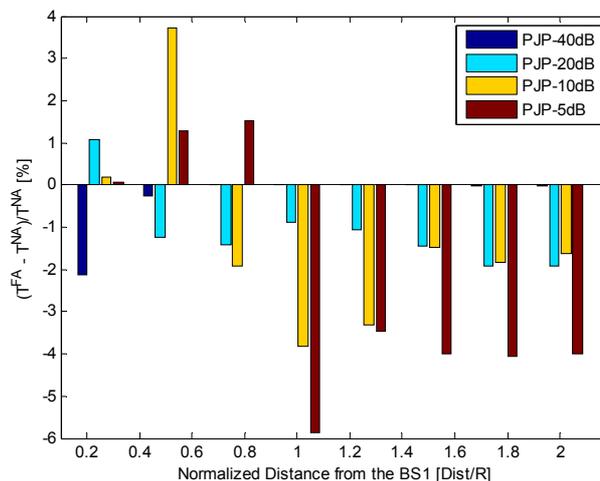
**Table 4-1: Simulation parameters**

Simulation Parameter	Value
Number of base stations, $K$	3
Number of antennas at each base station, $N_t$	3
Slant angle of each dipole	12 degrees
Antenna type at the base station	Uniform Linear Array
Antenna spacing	$4\lambda$
Number of single antenna users, $M$	6
User speed	3 km/h
Number of channel realization at each position	500

Centre frequency	2 GHz
Channel model	WINNER II, scenario B1 (NLOS)
Channel bandwidth	100 MHz
Cell radius	500 m
Inter-base station distance	433 m
Number of resource blocks	256
Cell-edge SNR	15 dB



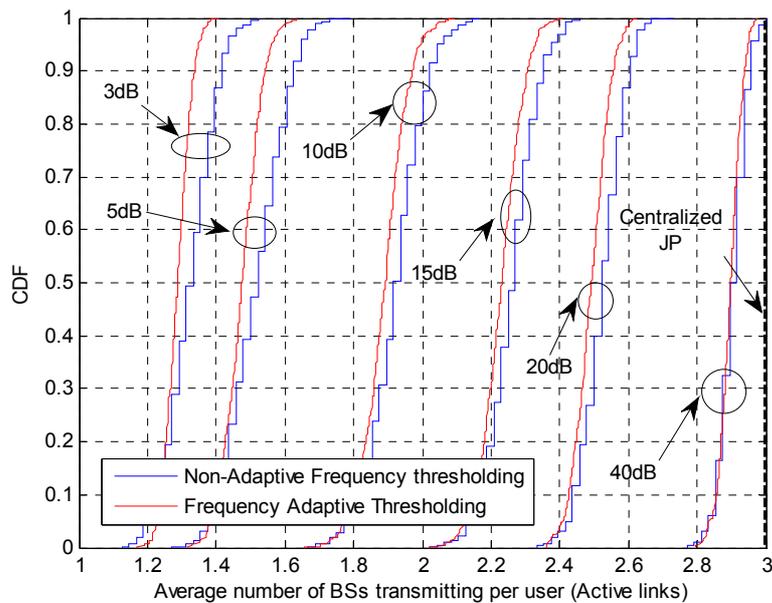
**Figure 4-2: Percentage gain in the average sum-rate per cell per RB when using FA thresholding within the cell coverage area and the cell-edge.**



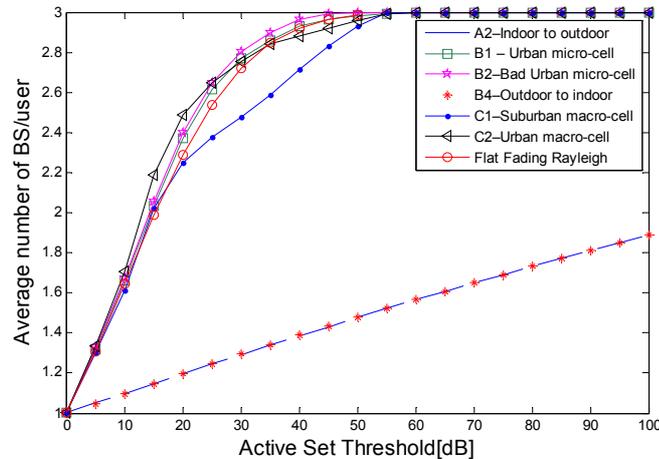
**Figure 4-3: Relative average number of active links of FA thresholding versus NA frequency thresholding per RB.**

Figure 4-4 shows the CDF of the average number of BSs serving a user (or the average number of active links per user) for various PJP active set thresholds at the 8 predefined positions in the cluster area. The number of active links relates to the system level KPI to minimize the complexity and control channel overhead, as defined in [ARTD51]. The active links relate to the backhaul load in terms of CSI feedback and beamformer weights. When the threshold is small, as expected, there is less number of active links. With a larger threshold, there is more number

of active links per user. The smooth red-curves represent the CDF of the active links for FA thresholding while the staircase-blue curves represent the NA frequency thresholding case. The appearance of the staircase-like curve is due to the fact that the NA frequency thresholding technique has a fixed number of active links throughout all the RBs, in a give time slot, while the FA technique is a smooth CDF, since the active set changes in different RBs, in a given time slot. The red curve is placed towards the left of the blue curve, indicating that the FA technique yields lesser number of active links. The difference in the number of active links is more prominent in the lower range of threshold values at the cluster centre, such as 3dB, 5dB, etc. One can also observe that an increase in the threshold also increases the number of active links per user. The PJP algorithm asymptotically reaches the performance of full JP using 3 BSs (Centralized JP). The CDF of the Centralized JP is the dashed line where all the 3 BSs are serving a user. When the users are closer to a BS, the average number active links per user is close to 1 for small threshold values, as the closer BS dominates the active set. The BSs that are far away are not included in the active set. At the cell-edge, there is more likelihood of a user receiving data from more than one BS. It was observed that all the 3 BSs serve a user with greater stability/reliability at the cluster centre for higher threshold values, but when close to the BS, a single BS serves a user more reliably for lower threshold values.



**Figure 4-4: CDF of the average number of base stations serving a user, the smooth red curves represent the FA thresholding and the staircase blue curves represent NA frequency thresholding technique. PJP threshold values shown are 3dB, 5dB, 10dB, 15dB, 20dB and 40dB.**



**Figure 4-5: Average number of BSs serving a user (or active links) in the cluster area for various WINNER II scenarios.**

The choice of the active set threshold depends on the scenario. Figure 4-5 shows the average number of BSs serving a user in the cluster area for various threshold levels with NA frequency thresholding. It can be observed that with a PJP active set threshold value greater than 40dB, all the BSs are involved in serving the user, showing the asymptotic behaviour of PJP towards full JP. More importantly, this is not true for the case of A2 and B4 WINNER II scenarios [WIN2D112], as these scenarios have no LOS and NLOS dominates these channels causing a severe degradation of the channel. The penetration loss due to the walls and the non-frequency selective nature of the channel with fewer multipaths compared to scenario B1, renders the need for higher active set threshold. Thus, the threshold value for PJP should be chosen carefully depending on the scenario/channel conditions. The flat fading Rayleigh channel performs similar to most of the WINNER II scenarios considered, showing that a flat fading Rayleigh can be assumed for theoretical work for PJP.

### Conclusions and future steps

In PJP, the frequency adaptive thresholding technique has fewer active links, due to which CSI feedback is further reduced compared to the non-adaptive frequency thresholding. There is gain in the average sum rate per cell per RB with frequency adaptive thresholding compared to non-adaptive frequency thresholding. With frequency adaptive thresholding, the backhaul load consisting of the exchange of CSI feedback and precoding weights is lower compared to non-adaptive frequency thresholding.

The performance results presented show that the evaluated schemes can improve the average sum-rate in the cell edge, thus achieving a higher uniformity of the performance over the served area. However, users located close to a base station may benefit from conventional single base station transmission, which would contribute to increase the uniformity of the average sum-rate per cell over the considered area. In future works, first the Jain Index as one of the main ARTIST4G KPIs will be evaluated [ARTD51]. Then, instead of allocating all the users in all the RBs, user grouping can be performed. This should decrease the interference and improve the backhaul load. Alternatively, a hybrid two-step thresholding technique can reduce the backhaul cost with some performance degradation. Finally, we intend to investigate the actual backhaul savings in the context of a more detailed backhaul model.

#### 4.1.2 Clustering based on partial CoMP

Cellular radio systems like 3GPP LTE suffer from strong inter cell interference, which is for example obvious by comparing the spectral efficiency of a single cell with that of the multi cell environment [FKV06]. For the here assumed cooperative transmission techniques several cells or eNBs are doing joint precoding of Tx-signals to several UEs, thereby cancelling interference between the cooperative UEs.

Under ideal assumptions like user centric setup of cooperation areas and perfect channel knowledge even for few cooperating cells significant capacity and coverage gains are possible. See for example system level simulations in [TSH+2009] as well as measurement results in [JFJ+10].

In 3GPP LTE Release 10 a so called CoMP study item investigated the potential of CoMP under practical restrictions and limited to only intra- and no inter-site cooperation to avoid any backbone traffic. In the end performance gains were either limited to a few percentage [3GPP-R1101431] or the overall system concept very complex, requiring cooperation over many – potentially far off – cells. It was decided that CoMP needs further evaluation before it might be getting part of any LTE standard, leading to a further CoMP SI that has started in beginning of 2011. To make CoMP a success high performance, practical, robust and easy to standardize concepts have to be found.

Regarding system level performance advances for the setup of CA – i.e. clustering of cells - and quite related the optimum user grouping within CAs seems to be the most pressing and challenging issue. For JP CoMP the CAs have to include the  $x$  strongest cells –  $x$  in the range of 2 to 5 - for all cooperating UEs as otherwise one stronger interferer will spoil most of the potential gains. Unfortunately in non line of sight (NLOS) scenarios UEs see a high variety of different cells as strongest interferers spread over a large area. As a further challenge each UE might need a different setup of CAs leading for the direct approach to very small penetration rates of best served UEs. Optimization algorithms as proposed in the literature [PGH08] trade one setup of CAs versus another with the goal to minimize the performance loss, but gains are typically limited as improving JP gain for one UE is paid by loss for another one.

The goal for the here proposed novel clustering scheme is to find a more fundamental solution serving at least most of the UEs – e.g. 90% - user centric. As will be shown this requires a new dimension for optimization, which we call in the following cover shifts ( $c_s$ ). The  $c_s$  means that each eNB is part of more than one CA allowing the eNB to schedule UEs into the individually best fitting CA/  $c_s$ .

This innovation targets the most advanced CoMP scheme, i.e. JP CoMP for intra- plus inter-site cooperation is really allowing to cancel interferers and at least theoretically is known to provide large performance gains. It is the most complex scheme, requiring tight time and frequency synchronization, very fast exchange of control and user data between all cooperating eNBs, highly accurate and frequency selective channel estimation in a multi cellular environment for a high number of channel components and – in case of FDD - fast feedback of CSI information to limit precoding errors due to outdated channel information. In addition new features should be as backward compatible as possible to LTE Release 8 and fit well into existing LTE standardization strategies.

Minor system level gains in the range of few percentage will make it difficult to justify the high complexity for JP CoMP. Without a convincing clustering solution progress for JP CoMP might be significantly delayed.

### System description of the innovation

Probably the most important aspect regarding the future of CoMP is to find practical concepts generating significant CoMP gains in the order of 100% or even more. JP CoMP gains for different size of CA from 2 to 10 have been simulated. In case of optimum user centric cooperation - meaning that each UE is served by its e.g. 3 strongest cells – encouraging performance gains of more than 100% have been reported.

In reality user centric cooperation is a real challenge as it requires that all UEs of a CA have exactly the same set of strongest cells, i.e. are being served by the same set of eNBs. For realistic radio channel conditions with strong NLOS probability, finding of such sets of UEs is extremely seldom, leading to a very low penetration rate of so called 'happy' users. Figure 4-6 illustrates the main issue of clustering as well as the basic idea of the partial CoMP concept. Visible are several sites  $s$  with 3 cells each - i.e. the 3 sectors of a site. In case of intra site cooperation (thick arrows) the CAs are defined by the 3 cells of one single site. For line of site (LOS) conditions significant performance gains are achievable as all UEs  $k_u$  of a CA  $u$  are served from their nearest – and therefore strongest – cells. This is the important general condition for user centric clustering:

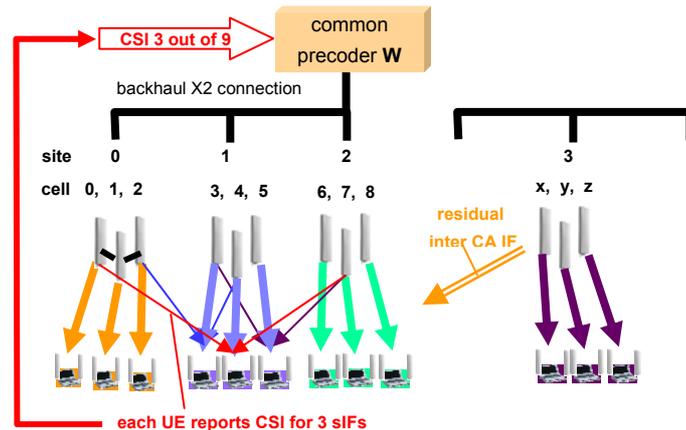
$$RSRP(n_u, k_u) \geq RSRP(n, k_u); \forall n \neq n_u,$$

where the Reference Symbol Received Power (RSRP) as defined in 3GPP LTE is a measure of the Rx power or equivalently - for fixed Tx power - of the path loss.  $n_u$  are the cells of CA  $u$  while  $n$  is the general cell index.

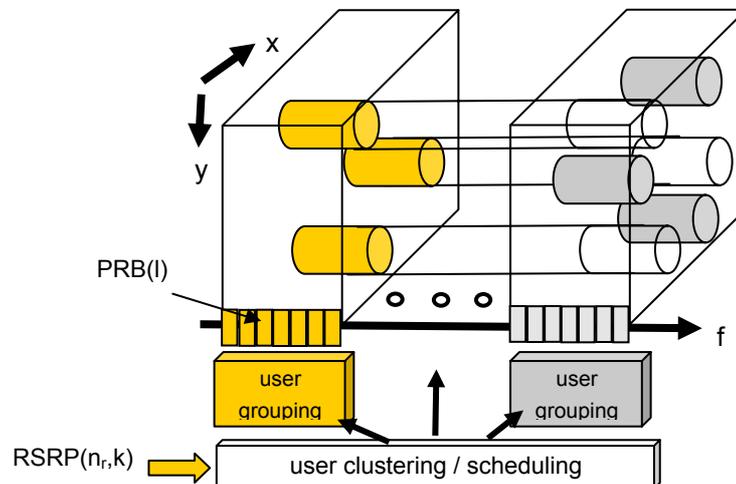
Under more realistic channel conditions with strong shadowing and many NLOS components the general condition for user centric clustering will be unsatisfied for intra site cooperation, indicated by the narrow inter site arrows in Figure 4-6. Such a scenario is requiring a new clustering and a common JP precoder  $\mathbf{W}$  over several sites  $s$ . Optimized clustering is a hot research topic and often well-known optimization algorithms are adapted accordingly, for example relying on Graph colouring schemes [PGH08].

*Partial CoMP* has a fundamentally different and more structured approach. In a first step the size of the CAs is expanded from 3 cells to e.g. 3 sites comprising 3 cells each, i.e. to overall 9 cells. From a practical point of view beneficially this requires only two inter site backhaul connections limited to adjacent sites, while reporting of channel components for 9 cells still would be a challenge. For that reason as a further means only the e.g. 3 strongest from the overall 9 cells of a CA are reported by the UEs. This *partial* reporting is the motivation to call the overall scheme *partial CoMP*.

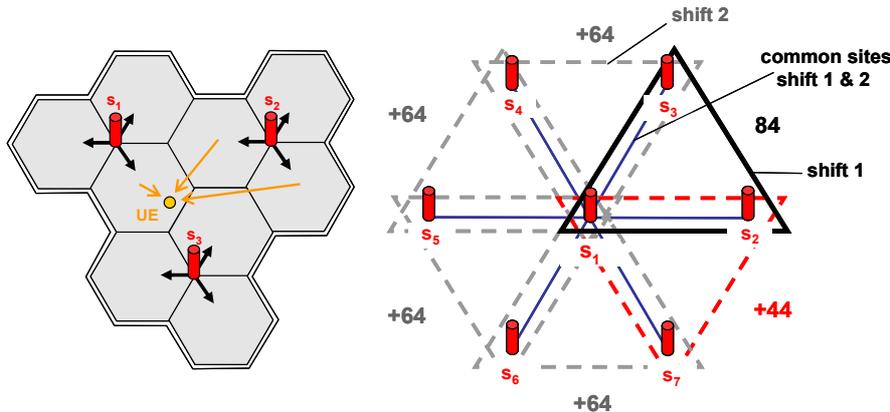
The unreported low power channel components will generate intra CA precoding errors, but these errors will be typically relatively small as they are per definition of low power components.



**Figure 4-6: Partial CoMP over 3 sites / 9 cells and UEs reporting their 3 strongest interferers each.**



**Figure 4-7: Resource allocation for partial CoMP.**



**Figure 4-8: Left: single CA of 3 sites. Right: 6 cover shifts for  $s_1$ .**

While the penetration rate is increased there is still significant inter CA interference – indicated by double line arrows - and unfortunately this easily spoils most of the JP performance gains. Therefore we have to open up a new dimension for optimization.

Under the assumption of a relative high number of users we can group UEs into different resources like different subbands and/or subframes, as in the cover shifts  $c_s$ . Into each cover shift only CA centric UEs will be scheduled, i.e. those UEs having all their e.g. 3 strongest interferers within the CAs of  $c_s$ . As illustrated in Figure 4-7 in case of sufficiently high number of cover shifts  $c_s$  all or at least most of the UEs should be servable without inter CA interference, at least with respect to the e.g. 3 strongest interferers. Depicted are two different cover shifts/frequency subbands (orange and light grey) over frequency  $f$  for a certain spatial area defined by  $x$ - and  $y$ -coordinates. Important is that per cover shift the CAs – or more precise the CA centric users - are spatially separated and that UEs not served in one cover shift will be CA centric within one of the other cover shifts. In the end almost all UEs can be served CA centric. The proposed concept fulfils the general condition for user centric clustering (see above) for user centric cooperation for most of the UEs.

The UEs  $k$  are scheduled into their best fitting  $c_s$  based on the reported  $RSRP(n_r, k)$  values, where  $n_r$  are the indices of the reported cells, e.g. the 3 strongest cells seen by the UE. For the UEs of  $c_s$  the scheduler has to find the best fitting groups of UEs per PRB I maximizing the multi user (MU) throughput, basically defined by the mutual spatial orthogonality of the UEs.

Figure 4-8 helps to understand the much higher penetration rate  $P$  of user centric served UEs for *partial CoMP* compared to conventional setup of CAs. The main difference is that conventionally the goal is to find 3 UEs simultaneously being served by 3 cells, i.e. there have to be 3 UEs seeing the same 3 eNBs as strongest interferers.

For *partial CoMP* beneficially all UEs with 3 strongest interferers out of 9 cells can cooperate. According to equation (2) this is fulfilled for  $N_c(1)=84$  cell combinations for one  $c_s$  (the 1 in brackets indicates one single cover shift). Figure 4-8 right illustrated 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 different CAs per cover shift indicated by according triangles. Let us take the solid line black triangle as basis CA in  $c_s=1$  comprising sites  $s_1, s_2$  and  $s_3$ .  $c_1$  covers  $N_c(1)=9!/(6! 3!)=84$  different combinations of 3 strongest cells. Overall we have 6 cover shifts to serve an UE, but between two adjacent cover shifts there are always two common sites. Therefore all cell combinations 3 out of 6 for these two sites are common between adjacent cover shifts. To calculate the overall possible cell combinations  $N_c(1...6)$  for all 6 cover shifts serving an UE user centric with 3 strongest interferers, one has to subtract these common cell combinations  $N_{c,overlap}(n_{c-1}, n_c)$ . The last cover shift has common sites with the left and right cover shift and therefore has to subtract  $N_{c,overlap}$  two times.

$$N_c(1) = \frac{9!}{6! 3!} = 84;$$

$$N_{c,overlap}(n_{c-1}, n_c) = \frac{6!}{3! 3!} = 20;$$

$$N_c(n_c) = N_c(1) - N_{c,overlap}(n_{c-1}, n_c) = 64; \quad n_c = 2, \dots, 5;$$

$$N_c(6) = N_c(1) - 2N_{c,overlap}(n_{c-1}, n_c) = 44;$$

$$N_c(1 \dots 6) = \sum_{n_c} N_c(n_c) = 384;$$

$$N_{c,IRC}^{60}(1) = \frac{9! * 60}{7! 3!} = 720;$$

In the above equations,  $N_c(1 \dots 6)$  has been calculated to 384. This leads to a very high penetration rate  $P$  of CoMP UEs (UEs gaining by JP CoMP), as each UE can be served by 384 different cell combinations, including with high likelihood the combination of 3 strongest interferers. At the same time the sites of the network are connected only between direct neighbours (blue lines) avoiding very complex network structures. Specifically 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 [MZ10]. This might be further improved by an interference rejection combining (IRC) receiver at the UE. In that case one of the interferers might esteem from any cell in the network. Assuming a maximum reach of 60 cells than  $N_{c,IRC}^{60}(1)$  - the number of possible user centric cell combinations with IRC - will be already 720 for a single shift and for all shifts even higher. Alternatively one can use the IRC to cancel 4 instead of 3 strongest interferers with according higher performance (providing about 3dB signal to interference ratio gain).

In reality there will be load variations for different cover shifts, leading potentially to underutilization 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 computed as follows

$$\begin{aligned} P_c(1s) &= k_1 / N_c(1) = 3/84 = 3.6\%; & P_{ch}(1s) &\approx 40\%; \\ P_c(2s) &= k_2 / N_c(1) = 40/84 = 47\%; & P_{ch}(2s) &\approx 30 - 40\%; \\ P_c(3s) &= k_3 / N_c(1) = 41/84 = 49\%; & P_{ch}(3s) &\approx 20 - 30\%. \end{aligned}$$

$P_c(1s)$  is for the most valuable intra eNB UEs, i.e. those which can be scheduled into any cover shift, but is with about 3.6% very low. Fortunately in real macro networks the collocation of cells at one site in combination with distance dependent path loss leads to much higher penetration rates  $P_{ch}(1s)$  of about 40% as known from many system level simulations.  $P_{ch}(2s)$  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 than some few UEs will have to be scheduled into second best cover shifts.

Note instead of frequency subbands any other set of orthogonal resources like time domain or a combination of frequency and time domain is possible. Further note that this is a full frequency reuse 1 system.

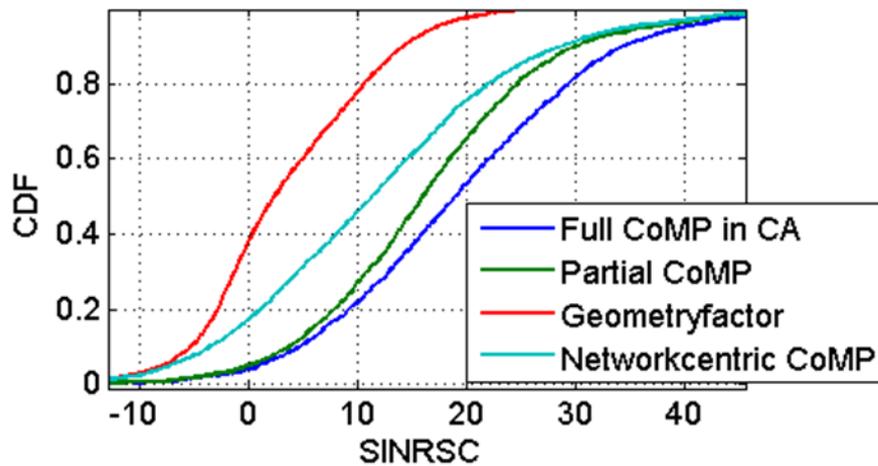
### Performance results and future steps

In [MZ10], system level simulations have been done for the proposed scheme and have been compared with the geometry factor as reference, network centric or intra site CoMP as well as full cooperation, meaning precoding based on full reporting for all 9 cells of a CA. Without going

into details it is obvious that specifically cell edge UEs profit from user centric *partial* CoMP compared to network centric cooperation. For the 5 percentile users the signal to interference and noise ratio (SINR) increases from about -5 to 0 dB for *partial* over network centric CoMP. As expected *full* CoMP over 9 cells is superior to *partial* CoMP, but mainly in the high SINR region above 20dB, which is of small interest as it is beyond the highest possible modulation and coding scheme 64QAM5/6 of LTE. At the same time the improvement of the cell edge performance is quite similar for partial and full CoMP as can be concluded from Figure 4-9.

**Table 4-2: Simulation parameters**

Simulation Parameter	Value
Number of eNBs:	57
Number of Cellsites:	19
Cells (Sectors) per Cellsite:	3
Sector width:	120 deg
Number of Subcarriers:	32
Bandwidth per sub carrier	180 kHz
TxAEs, RxAEs:	1
Algorithm for CJP:	ZF
Channelmodell:	SCME
InterCellSiteDistance:	500m
Antenna Tilting:	12 deg
Penetration Loss:	No
CSI:	Ideal



**Figure 4-9: CDF of SINR for different cooperation schemes and geometry factor as reference**

The proposed *partial* CoMP clustering concept is already very promising allowing for user centric CAs with limited feedback overhead and simple backhaul connections just between adjacent sites. It leads to high penetration rates by the increased size of the CAs including e.g. 3 sites or 9 cells and the introduction of different cover shifts per eNB. The proposed concept is somewhat specific, but generally the right way for clustering seems to be to setup more than one CA per cell or site. Further site conditions like minimum backbone overhead, minimizing CA sizes or keeping most of multi user scheduling gains might lead to different setup of CAs.

The system level simulations so far verify the expected performance gains at least for the single antenna case per eNB and UE. Including more antenna elements – i.e. potentially up to 8

for LTE Advanced – will lead to further degrees of freedom for optimization and naturally for accordingly higher performance gains.

*Partial* CoMP includes intra CA interference, which needs further careful analysis as it might be enlarged by the CA common precoder so that even small unreported channel components might lead to reasonable intra CA interference. Precoder design is highly related to user grouping mandating for an according inter-layer analysis.

#### 4.1.3 User Pairing algorithms for MU-MIMO

In multiple antenna broadcast channels the system capacity can be increased by means of Space Division Multiple Access (SDMA) serving multiple users simultaneously over the same transmission resources. This method of transmission is denoted in literature as Multi-User MIMO (MU-MIMO). This technique is of particular interest for the objective of the ARTIST4G project due to its ability to increase the total cell throughput rather than acting solely on the peak data rate of users positioned in very limited parts of the cell.

It is well known from the theory that MU-MIMO requires some level of channel knowledge at the transmitter and that, differently from Single-User MIMO (SU-MIMO), the achievable capacity is highly dependent on the level of CSI knowledge at the transmitter. Moreover it is of crucial importance to identify an efficient strategy to select the pair of users to be served on the same resources. The paired users should be in good radio channel conditions and also be sufficiently separated in space to minimize the multi-user interference.

The contribution proposed in this section is focused on the investigation and evaluation of an appropriate user-pairing algorithm to be used in a MU-MIMO system configuration, while maintaining a reasonable level of complexity. It should be stressed that particular focus will be dedicated to identifying an appropriate selection strategy, rather than providing a benchmark of the overall system performance.

#### System description of the innovation

In this section we analyse some user pairing metrics that can be exploited by the scheduler to perform the optimal user pairing for MU-MIMO transmission. We first describe the possible foreseen approaches and then provide a first comparison of the obtainable system performance expressed in terms of achievable system and user throughput, obtained by means of computer simulations. It should be noticed that, in order to guarantee the fairness between different pairing algorithms, only MU-MIMO Transmission Mode (TM) has been used for all the users present in the system, even if this is not necessarily the optimum choice in a real system that is capable to select the best TM as a function of the instantaneous channel conditions. The purpose of this study is to perform a relative comparison of the different considered pairing strategies, regardless of the absolute system performance.

In future work, the comparison among different transmission modes and the related achievable system performance that maximizes the system spectral efficiency will be performed, leading to the benchmarking of the overall system throughput in a more realistic scenario. In fact, in order to decide if the user pairing is effective in terms of spectral efficiency, the selection algorithm should compare the estimated sum-rate of the two paired users with the achievable rates of the two users if they would be scheduled separately, using SU-MIMO, over the considered data region. Clearly, the user pairing can be considered effective only if the achievable sum-rate using MU-MIMO increases the maximum achievable rate of the two users considered separately.

We here consider the following four user pairing strategies, based on different metrics:

- Random pairing;
- Best CQI pairing;
- CQI and channel orthogonality pairing;
- Exhaustive search pairing.

The applicability of these metrics is strongly related to the level of CSI knowledge at the transmitter. In a first phase the followed approach will be based on a complete CSI knowledge

at the transmitter. In a subsequent step of the analysis, left as future work, some form of limited feedback condition, with partial CSI knowledge at the transmitter, could also be considered. The complete CSI corresponds to a full knowledge of the channel matrix  $\underline{\mathbf{H}}$  at the transmitter. This assumption is rather ideal, but it is useful in the analysis in order to derive the upper bound performance of the user pairing and resource allocation algorithm. This level of knowledge can be achieved for example with TDD duplexing or, in case of FDD duplexing, by performing the scalar quantization of the channel matrix at the receiver. The quantized matrix has to be compressed by means of suitable algorithms and then sent to the transmitter over the reverse link. In case of complete CSI knowledge the information available at the transmitter is the channel matrix  $\underline{\mathbf{H}}$  complemented by the CQI that indicates the level of interference and thus the SINR experienced by the receiver before pairing.

### Random pairing

The proposed random pairing strategy represents the reference scenario with minimum complexity. This approach simply pairs users randomly among the available set of users and over the available frequency subbands.

### Best CQI pairing

The proposed best CQI pairing strategy is based on the CQI measurement. In particular, the CQI can be exploited to select a sub-set of users that experience the best channel conditions in terms of SINR, for which it is worthwhile in terms of spectral efficiency to perform Multi-User MIMO transmission.

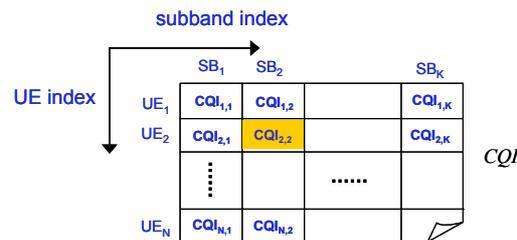


Figure 4-10: CQI information used for user pairing.

In particular the selection algorithm considers the CQI values of all the UEs among all the possible subbands and then selects the UE corresponding to the maximum CQI value. This UE will then be paired with a second UE, that corresponds to the next best CQI value that appears over the same sub-band. At this point the algorithm removes from the list the index of the considered subband and of the allocated users. The algorithm iterates such approach until all users have been paired.

### CQI and channel orthogonality pairing

The proposed CQI and channel orthogonality pairing strategy is based on the CQI measurement complemented by a second suitable metric that accounts for the level of channel orthogonality of the two users. In particular, the CQI can be exploited to select a sub-set of users that experience the best channel conditions in terms of SINR, for which it is worthwhile to perform Multi-User MIMO transmission. In particular the proposed algorithm considers, as suitable candidates for pairing, those users within the selected sub-set that possibly have comparable CQI (i.e. balanced SINR) and experience the maximum channel orthogonality. The channel orthogonality can be measured by means of several metrics such for example the ones described in [XWT+08].

An example of orthogonality metric is the Orthogonal Deficiency (OD) defined in [3GPP-R1083774]. The orthogonal deficiency is comprised in the range  $OD \in [0 ; 1]$ . In particular, OD is equal to zero if the user channel vectors are orthogonal. In general, the lower the OD is, the better is the channel orthogonality between the paired users. It follows that a pairing criterion can be to pick up the users that minimize the OD.

The proposed MU-MIMO user selection algorithm selects the first active user based on the maximization of the average CQI for user over the considered data region, and then selects the second user based on the minimum value of OD. We exploit the combination of CQI pre-selection with a second metric that captures the spatial properties of the radio channel. The CQI pre-selection allows to select only the users in good channel conditions, which are the most suitable for using MU-MIMO as transmission mode, and to reduce the number of potential pairing combinations. The second metric, like the orthogonal deficiency, allows to determine the subset of users that are most orthogonal and thus spatially separable using correlated antennas.

#### Exhaustive search pairing

The proposed exhaustive search pairing strategy is the most complex one in terms of computational complexity and scans all the possible pairs of UEs in the system, selecting the one pair that maximizes the achievable system performance in terms of sum capacity after pairing.

#### Performance results and future steps

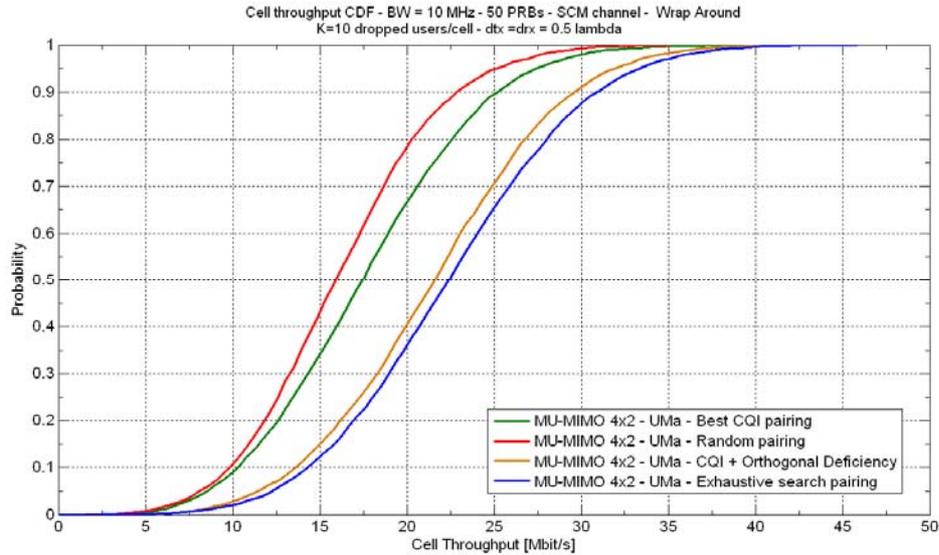
The Table 4-3 below lists the parameters used in the system level simulations. A hexagonal cellular layout of 57 cells with wrap around has been simulated. In each cell 10 users are randomly dropped and each user is camped on the best serving cell characterized by the largest average SINR. Each user can be allocated over one sub-band formed by 10 adjacent RBs. Each eNB is equipped with four vertically polarized antennas spaced by half wavelength, while each UE is equipped with two vertically polarized antennas spaced by half wavelength. The channel knowledge at the transmitted is assumed ideal (i.e. without errors or delays) with a frequency granularity equal to one RB.

**Table 4-3: Simulation parameters**

Simulation Parameter	Value
Simulated scheme	Single Cell MU-MIMO
UE CSI reporting	Ideal (channel matrix known for each RB at the transmitter)
Cellular Layout	Hexagonal grid, 19 sites, 3 sectors per site
Number of dropped users per cell	10
Number of PRB allocated to each user	10
Simulated Link	Downlink
Deployment scenario	Urban Macrocellular and Microcellular (UMa, UMi)
Traffic model	Full buffer
User Pairing	CQI pre-selection + Best secondary user
Bandwidth	10 MHz (50 PRB)
Channel model	Spatial Channel Model (SCM)
Number of antenna elements (eNB, UE)	(4, 2)
Antenna separation (eNB, UE)	$(\lambda/2, \lambda/2)$ - Vertically polarized
Link to system interface	MIESM
HARQ	Realistic (embedded into link level performance curves)
UE channel Estimation	Realistic (embedded into link level performance curves)

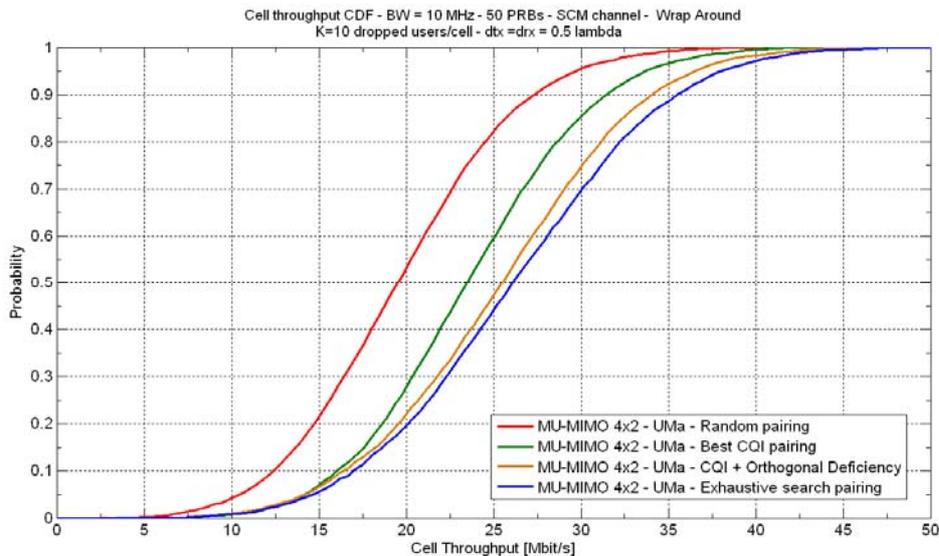
The Figure 4-2 shows the comparison of the different user pairing strategies in case of urban macrocellular (UMa) environment. The beamforming weights are calculated using the method based on the concept of signal leakage, as described in [STS+07]. In particular the beamforming coefficients are calculated in order to maximize the Signal to Leakage plus Noise Ratio (SLNR) of the paired users. In case of Figure 3-2 the beamforming weights are calculated

by using a channel matrix averaged over the sub-band, formed by 10 adjacent RBs. The beamforming weights calculated in this way are then kept constant over the sub-band. The results show that a strategy based on the usage of the CQI (before pairing) for the primary user and a channel orthogonality metric for the secondary user is able to grab a significant part of the channel capacity achievable with an exhaustive search over the population of served user.



**Figure 4.2: Cell Throughput distribution for UMA (sub-band precoding granularity)**

In Figure 4-3 it is investigated the impact of the CSI granularity at the transmitter. In this case the beamforming weights are calculated for each RB using the channel matrix supposed available at the transmitter. The availability of a finer CSI at the transmitter shows a clear improvement, regardless of adopted user pairing strategy, in the order of 10-20% for the median cell throughput. Clearly the optimum precoding granularity must be designed to obtain a reasonable trade-off between performance and corresponding uplink signalling overhead.



**Figure 4.3: Cell Throughput distribution for UMA (RB precoding granularity)**

In this contribution different user pairing approaches are compared, ranging from the trivial random pairing that provides a lower performance bound up to the exhaustive search pairing that provides an upper performance bound. A good performance-complexity trade-off is

identified when using the algorithm that adopts a user pairing approach based on both CQI and orthogonal deficiency metrics. The results show that a strategy based on the usage of the CQI (before pairing) for the primary user and a channel orthogonality metric for the secondary user is able to grab a significant part of the channel capacity achievable with an exhaustive search over the population of served user. Further studies will be dedicated to the fine tuning for optimal utilization of these metrics.

In future work, the comparison among different transmission modes and the related achievable system performance that maximizes the system spectral efficiency will be performed, leading to the benchmarking of the overall system throughput in a more realistic scenario.

#### 4.1.4 Optimized resource allocation for SC-FDMA MU-MIMO schemes

Single-Carrier Space Frequency Block Coding (SC-SFBC) is an innovative mapping scheme suitable for implementing transmit diversity in Single-Carrier Division Multiple Access (SC-FDMA) systems [CCM+09]. The main advantage of SC-SFBC is that it preserves the low envelope variations of SC-FDMA, which is particularly interesting for the uplink of wireless communications systems. This transmit diversity technique, originally designed for a user equipped with 2 transmit antennas, has been combined with spatial multiplexing and extended to a single-user multiple-input multiple-output (SU-MIMO) scenario for users equipped with 4 transmit antennas. This was presented in [ARTD12]. Here, we apply the SC-SFBC concept in a multiuser (MU)-MIMO scenario. We introduce a novel algorithm allowing the optimization of the parameters of SC-SFBC in order to enable low-complexity decoding at the receiver side and to maximize the overall spectral occupancy in MU-MIMO SC-FDMA systems, and we show the good performance of the proposed MU scheme.

##### System description of the innovation

We consider that several users, disposing of at least 2 transmit antennas each, are managed by the same base station eNB in an SC-FDMA system. The eNB tries to map the uplink signals of these users in a given limited bandwidth in an optimal manner. Each such user implements SC-SFBC as a transmit diversity scheme, using the scheme in Figure 4-11, where the  $M$ -sized vectors  $\mathbf{s}^{\text{Tx}_n(t)}$ ,  $n=0\dots1$ , represent the frequency-domain samples (after  $M$ -sized DFT-precoding and before  $N$ -sized IDFT subcarrier mapping) on each of the 2 transmit antennas of a SC-FDMA/SC-SFBC signal occupying  $M$  (out of maximum  $N$ ) subcarriers. This was thoroughly described in [ARTD12]. The  $SC_M^p$  operation consists in taking the complex conjugates of vector  $\mathbf{s}$  in reversed order, applying alternative sign changes and then cyclically shifting down its elements by  $p$  positions. Alamouti-precoded pairs appear on couples of non-adjacent subcarriers  $(k_0, k_1 = (p - 1 - k_0) \bmod M)$ . The maximum separation between subcarriers carrying frequency samples precoded together is  $\max(p, M-p)$  and is thus controlled by the parameter  $p$ . Distant subcarriers might experience different or even uncorrelated channel realizations, which generates some interference within the Alamouti-precoded pair. The optimum value of  $p$ , minimizing the maximum distance between subcarriers carrying Alamouti pairs is the even integer closest to  $M/2$ ,  $p_{\text{opt}} = 2\text{floor}(M / 4)$ .

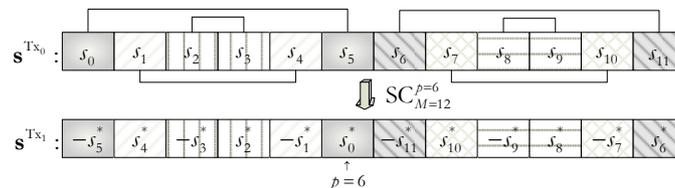
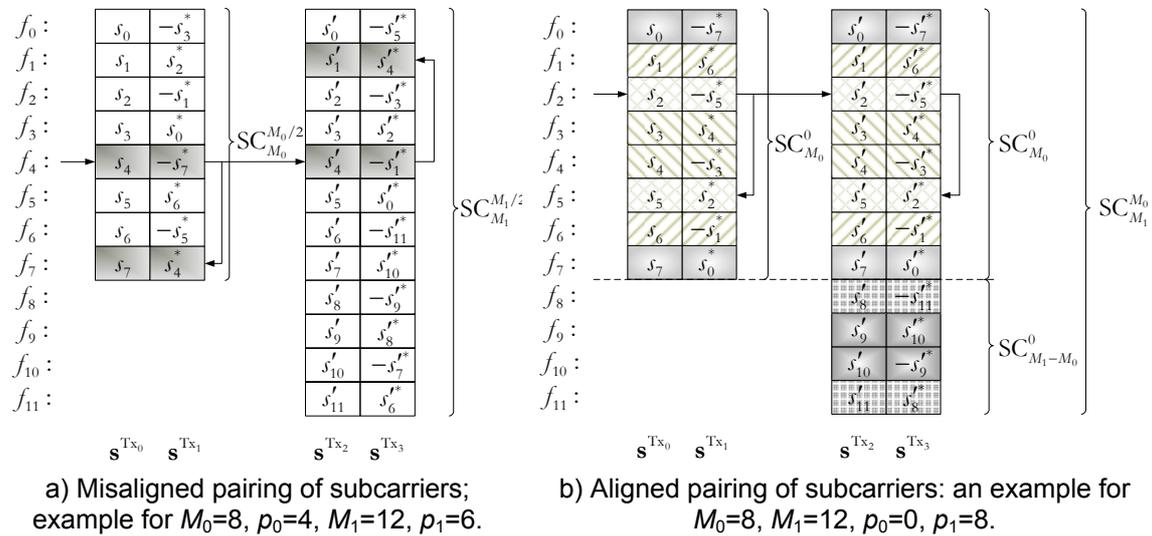


Figure 4-11 SC-SFBC precoding; example for  $M=12$ ,  $p=6$ .

According to the desired throughput, to the capabilities of each mobile station and to the corresponding channel quality, the scheduler at the eNB will decide the spectral allocation and the MCS of each user. To optimize the spectral occupancy and increment the throughput, it is interesting to allow some spectral superposition between users having either the same or different spectral allocations.

Let us assume that the scheduler allows at most two users ( $UE_0$  and  $UE_1$ ) to share simultaneously groups of subcarriers corresponding to all or to part of the subcarriers allocated to each user. Each user is employing transmit diversity techniques, e.g. SC-SFBC, independently from the other users. For the part of the spectrum where the two users transmit simultaneously, this is equivalent to applying a MIMO scheme that combines spatial multiplexing and SC-SFBC in a MU context. The MU-MIMO channel has  $N_{Tx}$  transmit antennas, split into two non-collocated groups,  $N_{Tx} = N_{Tx_0} + N_{Tx_1}$ .

The spectral allocation decided by the scheduler consists in computing the number of subcarriers  $M_i$ , as well as the starting position  $n_i$  of the portion of spectrum allocated to each  $UE_i$ . When SC-SFBC is used, to minimize the maximum distance between subcarriers coded together, the best strategy is to employ  $SC_M^{p=2\text{floor}(M/4)}$ . In a MU-MIMO context, double SC-SFBC might have some pairing incompatibility problems. Indeed, let us analyse the situation depicted in Figure 4-12a), where  $UE_0$  is allocated  $M_0=8$  subcarriers and  $UE_1$  is allocated  $M_1=12$  subcarriers.



**Figure 4-12 MU Double SC-SFBC with different types of subcarrier pairing.**

The portions of spectrum occupied by the 2 UEs start with the same spectral position,  $n_0 = n_1 = 0$ , which means that the first occupied subcarrier by each UE is the one with index 0, denoted  $f_0$ . Therefore,  $UE_0$  uses  $SC_8^4$  and  $UE_1$  uses  $SC_{12}^6$ . Subcarriers with indexes  $(k_0, k_1 = (p-1-k_0) \bmod M)$  contain Alamouti pairs. Each UE uses its optimum  $p$  parameter, respectively  $p_0=4$  and  $p_1=6$  in this example. On the 5-th occupied subcarrier  $f_4$  for example,  $UE_0$  transmits frequency samples  $s_4$  and  $-s_7^*$  onto its two transmit antennas respectively. Next,  $f_4$  is paired with  $f_7$ , onto which  $UE_0$  transmits frequency samples  $s_7$  and  $s_4^*$ , respectively. On the same subcarrier  $f_4$ ,  $UE_1$  transmits frequency samples  $s_4'$  and  $-s_1'^*$ , respectively, onto its two transmit antennas. Since  $UE_1$  uses  $SC_{12}^6$ ,  $f_4$  is paired with  $f_1$ . As a result, the pairing of subcarriers is not compatible between  $UE_0$  and  $UE_1$ . Because of this incompatibility, this structure does not correspond to a double SC-SFBC construction and the conventional MMSE simplified detector cannot be employed anymore. A joint MMSE detection over all the bandwidth containing cross-codes subcarriers is necessary. For the example in Figure 4-12a), this would involve inverting a matrix of order  $M_0 + M_1 = 20$  instead of 2 matrices of order 4 and 2 matrices of order 2, as it would have been the case if the two UE were correctly aligned to form double Alamouti pairs on the overlapping subcarriers, and simple Alamouti pairs on the remaining subcarriers. In practice, the complexity of this scheme is a real issue. Indeed, the receiver must

be designed for the worst-case scenario and must thus be able to invert matrices of rank hundreds or thousands.

To show how this incompatibility problem can be avoided, let us notice that any  $SC_M^p$  operation can be decomposed into the juxtaposition of  $SC_p^0$  and  $SC_{M-p}^0$  operations. This is a direct result of the very structure of SC-SFBC. Let us denote the number of subcarriers simultaneously used by two UEs by  $M_{\text{overlap}}$ . To avoid any pairing incompatibility, the 2 UEs need to transmit the same symbol structure over the overlapping spectral portion. Based on the property stated above, when the two UEs have strictly different spectral allocations, the only valid option is to chose  $p$  parameters  $p_i$  such that the overlapping portion has a structure based on  $SC_{M_{\text{overlap}}}^0$ .

The case where the two UEs have the same number of allocated subcarriers  $M_0=M_1$  and share the same bandwidth has already been treated in [ARTD12]. We only treat here the case of different spectral allocation  $M_0 \neq M_1$ , let us assume for example  $M_0 < M_1$ .

A solution is given in Figure 4-12b). We need to impose UE<sub>0</sub> to use  $SC_{M_0}^{p_0=0}$  and UE<sub>1</sub> to use  $SC_{M_1}^{p_1=M_0}$ . The  $SC_{M_1}^{p_1=M_0}$  can be seen as the juxtaposition of two SC-like operations:

- $SC_{M_0}^0$  to match the configuration of UE<sub>0</sub>; on this part of the spectrum, double SC-SFBC transmission can thus be employed;
- The remaining  $SC_{M_1-M_0}^0$  corresponds to a simple SC-SFBC transmission and keeps an overall SC-type signal to be transmitted by UE<sub>1</sub>.

Hence, it is no longer possible to use a default value for the  $p$  parameter for all the system (highest even integer is inferior to the half of the respective number of allocated subcarriers), but double SC-SFBC potential is kept at the expense of a modification of the  $p$  parameter, *i.e.*, some performance degradation as the maximum distance between subcarriers that are jointly precoded is increased. But complexity is strongly reduced: only two matrices of order 4 and two matrices of order-2 need to be inverted during MMSE decoding for the example in Figure 4-12b), while for the structure in Figure 4-12a) an inversion of an order 20 matrix was required. It should also be noted that additional signaling is necessary to indicate the values of  $p$  to be used by each UE in this case.

### Performance results and future steps

We have to investigate thus the performance of two schemes:

Misaligned double SC-SFBC: Each user is allowed to egoistically use its own optimum  $p$  parameter in order to try and improve its individual performance by minimizing the interference within Alamouti pairs. High complexity detection is needed.

Aligned double SC-SFBC: For each user, the  $p$  parameter is optimized from a global point of view, taking into account the characteristics of other users that might overlap in the frequency domain. Low complexity detection is needed, but some performance loss might occur due to higher interference level within the Alamouti pairs of each user.

Let us investigate from a performance point of view the performance-complexity trade-off between Aligned and Misaligned double SC-SFBC.

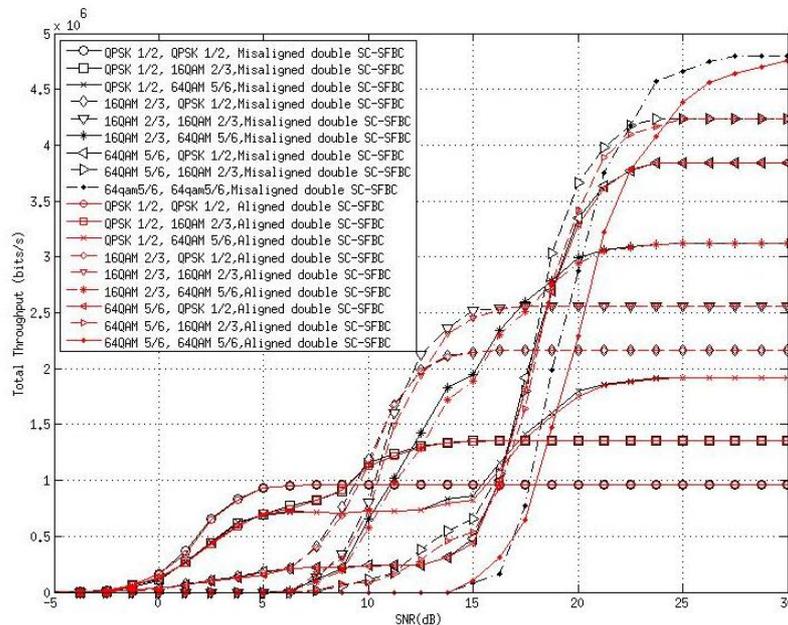
Simulation parameters are summarized in the table below:

**Table 4-4: Simulation parameters**

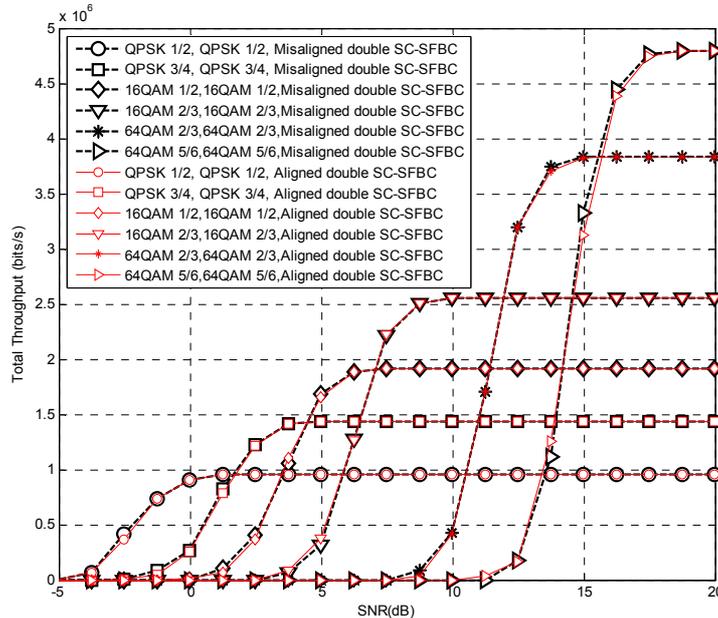
Simulation Parameter	Value
Transmission bandwidth	5MHz
Number of subcarriers ( $N_{\text{FFT}}$ )	512
Maximum number of data subcarriers ( $M_{\text{max}}$ )	300
Number of allocated subcarriers ( $M$ )	60, 20
Signal mapping	QPSK, 16QAM, 64QAM
FEC coding	Turbo code 1/3, 2/3, 3/4, 5/6

Number of transmit antennas ( $N_{Tx}$ )	4
Number of receive antennas ( $N_{Rx}$ )	2, 4
Channel type	3GPP Typical Urban
Channel estimation	Ideal
Velocity	120kmph
Detection	MMSE Successive Interference Cancelling (SIC)

First, we consider two users occupying  $M_0=60$  and respectively  $M_1=20$  subcarriers in an SC-FDMA system with 512 subcarriers out of which only 300 are active data carriers to fit within a bandwidth of 5MHz. Different symbol mapping (QPSK, 16QAM, 64QAM) and turbo code with different coding rates are employed. Each user performs SC-SFBC-based transmit diversity. The KPI used here is the throughput as defined in [ARTD51].



**Figure 4-13. Performance comparison between Misaligned and Aligned double SC-SFBC,  $M_0=60$ ,  $M_1=20$ ,  $N_{Rx}=2$ .**



**Figure 4-14. Performance comparison between Misaligned and Aligned double SC-SFBC,  $M_0=60$ ,  $M_1=20$ ,  $N_{R_x}=4$ .**

Figure 4-13 and Figure 4-14 show the comparative performance of Misaligned and Aligned double SC-SFBC. When only two receive antennas are employed, Misaligned double SC-SFBC has a slight advantage on the Aligned scheme when high order modulation (64QAM) is employed (2.8dB at a throughput of 450kbps). This is due to the different interference profile within an Alamouti pair. As explained in the previous section, this interference is stronger in the case of the Aligned scheme, who allowed precoding between distant subcarriers in order to reduce the detection complexity. The effect is more visible on high order modulations, more sensitive to interference, and employing higher coding rates.

In practice, the base station will be equipped with 4 or more receive antennas. In this case, results in Figure 4-14 show that both techniques display similar performance. The lower complexity of the Aligned scheme is to be preferred.

We investigated specific techniques of parameter optimization for MU-MIMO SC-SFBC. We assessed the performance of two different double SC-SFBC schemes in MU-MIMO configurations: Misaligned double SC-SFBC and Aligned double SC-SFBC. The performances of these two schemes are similar, but the second one demands much lower decoding complexity. With the aligned scheme, the system performances in terms of throughput with different MCSs and pathloss were evaluated. The simulation results showed the way how MCS used by one can impact the performance of the other user.

## 4.2 Inter-Cell Interference Coordination

In this section, we address inter-cell interference coordination for heterogeneous deployments involving eNBs and HeNBs. First, a power setting strategy is presented so as to limit the impact of the in-band massive deployment of femto cells on the macro network. Then, the HeNB/eNB or HeNB/HeNB interference can be further reduced by using a dynamic resource allocation scheme.

### 4.2.1 Blind and semi-centralised power setting for heterogeneous networks

In current mobile cellular networks, like 3GPP-LTE networks, heterogeneous deployments mixing deployment of macro base stations (eNBs) and deployment of home base stations (HeNBs) are foreseen as an effective way to ensure both mobility within a large geographical area and high data throughput, comparable to wireless LAN, at home.

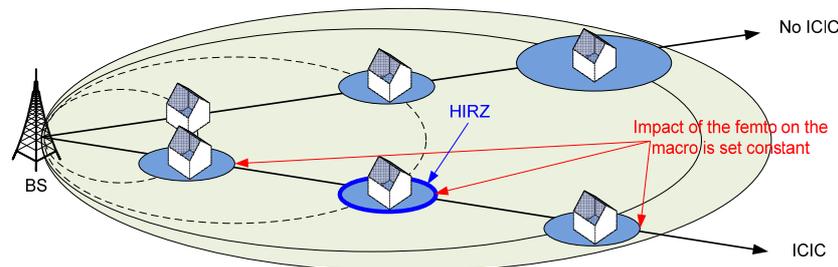
Besides, due to the constant cell size reduction and spectral efficiency increase, inter-cell interference has become a main issue. During the standardization phase of 3GPP-LTE, inter-cell interference coordination (ICIC) techniques have been extensively discussed and UE reports to its serving eNB and eNB-to-eNB messages have been standardized in order to allow performing efficient ICIC. However, the eNB-to-eNB messages will be conveyed through a direct eNB-to-eNB logical link, the X2 interface, which might not exist between an eNB and a HeNB. Indeed, a massive deployment of HeNBs may prevent from having an X2 interface between an eNB and all HeNBs within its coverage area.

These HeNBs strongly interfere with eNB and even create coverage holes. In order to secure the operator macro traffic, priority should be put on minimizing the interference created by HeNBs on eNBs. However, the HeNB throughput inside home should at least achieve the performance of WiFi. Furthermore, the interference mitigation should work without eNB-to-HeNB X2, i.e., without a fast cooperation channel. Blind ICIC in downlink and semi-centralised power control in uplink satisfy these requirements.

## System description of the innovation

### Blind ICIC in downlink

In downlink (DL), the interference impact depends on the path-gain between each Macro-UE (MUE) close to the HeNB and its serving eNB. The lower the average path-gain between MUEs close to a HeNB and eNB, the larger the area in which MUEs are strongly interfered by the HeNB. Figure 4-15 illustrates the distance-dependent coverage of HeNBs in a case without shadowing. The downlink blind ICIC scheme described in this section aims at making the impact of the HeNB on the neighbouring UEs as constant as possible.



**Figure 4-15: Downlink ICIC.**

Having the knowledge of the eNB-to-HeNB path gain or the Received Signal Strength (RSS) from the eNB, the HeNB can independently set its transmit power. A HeNB may have UE receiver capabilities and measure its environment when switched on and then continue to do so periodically. Thus, it is able to determine the RSS from eNB, including shadowing and distance-dependent path loss effects. This HeNB measurement is representative of what neighbouring MUEs experience if the eNB shadowing correlation distance is much higher than the home buildings size. The wall penetration loss can also be compensated by the HeNB.

In a heterogeneous deployment, involving several eNBs and HeNBs, interference from other eNBs and AWGN must be taken into account. With strong interference from other eNBs, the impact of HeNB interference on MUEs becomes small. Thus, the degradation of MUE performance due to a HeNB is not only linked to the serving eNB received power but also to the interference plus noise level.

In order to evaluate the impact of the HeNB on neighbouring MUEs, we define a high interference reference zone (HIRZ), e.g., a ring around the HeNB building as shown in Figure 4-15, which represents the location of MUEs highly impacted by the HeNB transmission. We also define as a function representative of the MUE performance degradation the ratio of the MUE SINR with HeNB interference and the MUE SINR without HeNB interference  $x$ :

$$x = \frac{SINR_{withHeNB}}{SINR_{withoutHeNB}} = \frac{P_{I,M-MUE}}{P_{I,M-MUE} + P_{I,H-MUE}}$$

where  $P_{I,M-MUE}$  is the interference experienced by MUEs from neighbouring eNBs plus AWGN and  $P_{I,H-MUE}$ , the interference experienced by MUEs from the HeNB. Due to shadowing, there may be different reception conditions in HIRZ: the performance degradation is a random variable. Thus, we define a degradation threshold  $x_{th}$  and an outage probability  $O_p$  and the HeNB transmit power  $P_{t,HeNB}$  is set in order to satisfy the outage probability in HIRZ:

$$\Pr(x \leq x_{th} | HIRZ) = O_p$$

The transmit power  $P$  which results in a degradation  $x_{th}$  is such that

$$P = \frac{(x_{th}^{-1} - 1)P_{I,M-MUE}}{P_{G,H-MUE}}$$

where  $P_{G,H-MUE}$  is the path gain between HeNB and MUEs, which is a random variable depending on shadowing. The higher the transmit power of the HeNB  $P_{t,HeNB}$ , the higher the degradation  $x$ . We compute the CDF of the random variable  $P$  and set  $P_{t,HeNB}$  to the value of  $P$  for which the CDF equals  $O_p$ .

$$P_{t,HeNB} = Q_P(O_p)$$

where  $Q_P(O_p)$  is the quantile of  $P$  at  $O_p$ . Thus, in a fraction  $O_p$  of random variable events,  $P_{t,HeNB}$  will be too high and will result in a value of  $x$  lower than  $x_{th}$ , i.e., in too high a degradation. The probability of having  $x$  lower than  $x_{th}$  will be  $O_p$ , which is our criterion. In practice, the quantile is computed by considering a log-normal distribution for  $P_{G,H-MUE}$ .

### Semi-centralised ICIC in uplink

In uplink, the higher the average path-gain between HeNB UEs (HUEs) and eNB, the higher the interference created by HUEs on all MUEs served by eNB. Besides, MUEs also interfere on the HUEs. The lower the average path-gain between eNB and MUEs at the neighbourhood of the HeNB, the higher the interference created by MUEs on HUEs served by HeNB, due to MUE power control.

Since an X2 connection is not available between eNBs and HeNBs, we consider a power control solution, where the HUE and MUE transmit power is set in order to optimize the HUE and MUE SINR at low eNB-HeNB signalling cost. The power control solution is applied for each cell of each eNB independently and involves a coordinator. The interactions between the eNB, the HeNBs and the coordinator are described in Figure 4-16. MUEs (resp. HUEs) report path gains from the serving eNB (resp. HeNB) and interfering HeNBs (resp. eNB). The HeNB-to-HeNB interference is not considered here. Each eNB (resp. HeNB) builds from MUE (resp. HUE) reports statistics of the useful path gain and interfering path gains. The statistics are sent from the eNB and HeNBs to the coordinator and the coordinator optimises the power control for the eNB and HeNBs. The power control rule is a function  $f$  of the pair (useful path gain, interfering path gain). The eNB and each HeNB receive the power control information, i.e., the function  $f$ , and apply a UE-specific power control according to the function  $f$  taking as arguments the measurement pairs (useful path gain, interfering path gain).

Here, we assume that the macro power setting is already available and we want to define a global HeNB power setting rule, i.e., a function  $f$  common to all HeNBs.

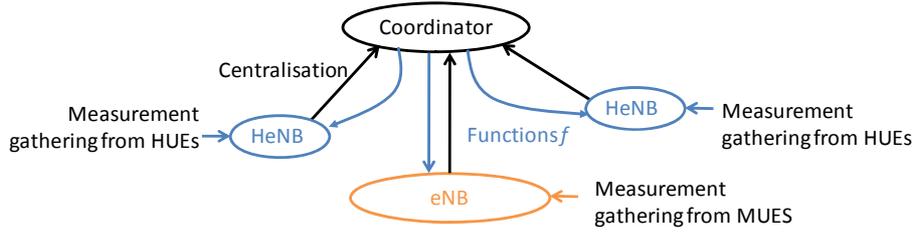
In a first step, we consider the following structure for function  $f$ :

$$f(P_{G,HUE-M}) = \frac{\beta}{P_{G,HUE-M}}$$

where  $P_{G,HUE-M}$  is the interference path gain, i.e., the path gain between a HUE and the interfered eNB. The transmit power  $P_{t,HUE}^j$  of a HUE  $j$  is set according to the interfering path gain

$P_{G,HUE-M}^j$  measured by the HUE:

$$P_{t,HUE}^j = f(P_{G,HUE-M}^j)$$



**Figure 4-16: Interactions in the semi-centralized UL ICIC.**

The higher this path gain, the lower the HUE transmit power. This solution is optimal from the eNB performance perspective. The parameter  $\beta$  is computed by the coordinator based on feedback from eNB and HeNBs.

In order to further simplify the algorithm implementation, we assume that the HeNB  $i$  has UE receiver capability. It can measure  $P_{G,H-M}^i$ , the path gain between the eNB and itself, which is a good approximation of  $P_{G,HUE-M}^j$ . Therefore, the HeNB assigns the same transmit power to all its HUEs based on this single measurement and parameter  $\beta$ . Thus, HUE-specific power control, which would require one measurement per HUE, is not performed at the initial setup. This additional measurements could be used in order to refine the HUEs power control afterwards. In order to control the eNB performance degradation due to HUE interference, we define it as the ratio  $\alpha$  of mean interference level on eNB due to HUEs in the eNB coverage and mean interference plus AWGN level on eNB without HUE interference ( $E[N^M]$ ). With  $N_f$  HeNBs and a system load of  $0 \leq \rho_i \leq 1$  for HeNB  $i$ , the  $\beta$  value is

$$\beta = \frac{\alpha E[N^M]}{\sum_{i=1}^{N_f} \rho_i} \leq \frac{\alpha E[N^M]}{N_f}$$

In a second step, we introduce HUE-specific power control by also considering the HUE useful path gain  $P_{G,HUE-H}^j$  of HUE  $j$  as argument of function  $f$ :

$$P_{t,HUE}^j = f(P_{G,HUE-M}^j, P_{G,HUE-H}^j) = \min \left( \frac{\beta_1}{P_{G,HUE-M}^j}, \frac{\beta_1}{E \left[ \frac{P_{G,HUE-M}^j}{P_{G,HUE-H}^j} \right]} \frac{\beta_2}{P_{G,HUE-H}^j} \right) = \frac{\beta_1}{P_{G,HUE-M}^j} \min(1, \beta_2 z)$$

where  $E[x]$  denotes the mean of  $x$  and

$$z = \frac{P_{G,HUE-M}^j}{P_{G,HUE-H}^j} E \left[ \frac{P_{G,HUE-M}^j}{P_{G,HUE-H}^j} \right]^{-1}$$

With this function  $f$ , an appropriate balancing between useful pathgain dependent power control and interference dependent power control is performed. Based on HeNBs' feedback, the coordinator optimises parameters  $\beta_1$  and  $\beta_2$ , i.e., finds values such that the eNB performance degradation remains  $\alpha$  while maximizing the cell-edge Shannon capacity  $C_{CE}$  of the HeNBs:

$$(\beta_1, \beta_2) = \arg \max_{\beta_1, \beta_2} C_{CE} \left| \sum_{i=1}^{N_f} \rho_i E [P_{G,HUE-M}^j f(P_{G,HUE-M}^j, P_{G,HUE-H}^j)] = \alpha E[N^M] \right.$$

The coordinator assumes that random variables have log-normal distribution, due to shadowing, and performs numerical optimisation. The eNB has to report  $E[N^M]$  and the HeNBs the means and variances of path gains  $P_{G,HUE-H}$ ,  $P_{G,HUE-M}$  in dB and of the noise plus interference level  $N_i^H$  at HeNB  $i$ . The HeNBs also report their load  $\rho_i$ . Finally, after optimisation, the coordinator

broadcasts  $\beta_1$  and  $\beta_2$  to all HeNBs under the eNB coverage, which then apply function  $f$  for UE-specific power control.

### Performance results and future steps

The ICIC schemes are evaluated in a static system level simulator with perfect scheduling, i.e., assuming that each UE is always allocated the PRB maximizing the considered metric, on a 5-MHz bandwidth. The MUE and HUE throughput based on an outage capacity metric is evaluated, taking into account inter-cell interference, in a heterogeneous network comprising 19 tri-sectorized eNBs in an hexagonal grid. The inter-site distance is 1732m and the propagation model is the LTE Case 3 propagation model [3GPP25814]. The simulation scenario is summarized in Table 4-5.

**Table 4-5: Simulation parameters**

Simulation Parameter	Value
Simulated scheme	Multi-Cell macro + femto, SU-SIMO
Frequency	2 GHz
Cellular Layout	Hexagonal grid, 19 sites, 3 sectors per site
Number of dropped users per cell	Full system load
Simulated Link	Downlink/Uplink
Deployment scenario	Urban Macrocellular 3GPP case 3
Femto deployment	Circular femto buildings uniformly deployed
Femto path gain model	Indoor : 3GPP LTEA femtocell Indoor to outdoor : Attenuation coefficient of 3GPP microcell NLOS
Link to system interface	LUT of SINR vs. spectral efficiency (b/s/Hz)
Bandwidth	5 MHz (25 PRBs)
Small-scale Channel model for macro BS	ITU-TU6
Small-scale Channel model for femto BS	ITU-InH-NLOS
Number of antenna elements (BS, UE)	(1,2)
Number of PRB allocated to each user	1

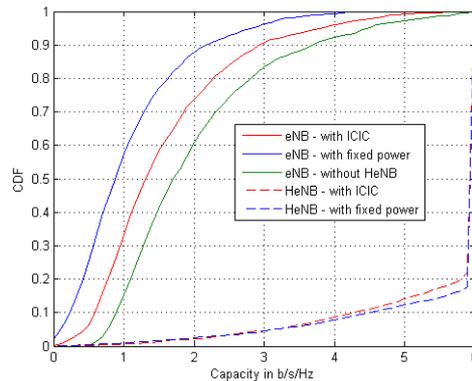
The throughput of one user as a function of the SNR is defined by

$$s(SNR) = \max_{R_c, m} (R_c \cdot m (1 - p_{out}(R_c, SNR, m))) \quad \text{where } 0 < R_c < 1 \quad \text{and} \quad m = \{2, 4, 6\}$$

where  $R_c$  is the code rate and  $m$  is the number of bits per QAM modulation symbols. The outage probability  $p_{out}(R_c, SNR, m)$  is computed from the Gaussian input capacity of the system.

### Downlink ICIC

Figure 4-17 shows the CDF of DL outage capacity for MUEs close to a HeNB and HUEs with shadowing and a 20dB wall penetration loss. Here, without ICIC, the fixed transmit power is set to 5dBm. There is a 50% MUE performance improvement towards performance without HeNB at the expense of a negligible HUE performance degradation.



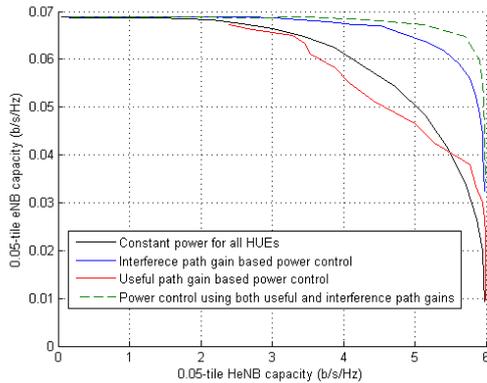
**Figure 4-17: DL capacity CDF with shadowing and 20-dB wall penetration loss.**

### Uplink ICIC

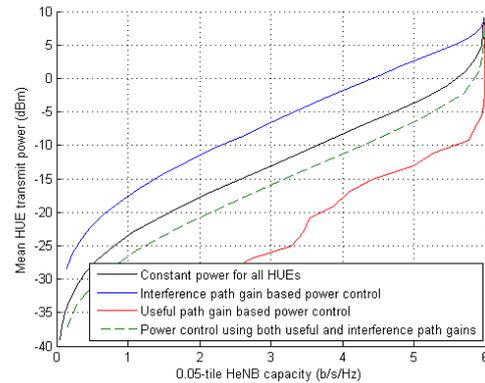
We evaluate a deployment of 50 HeNBs per eNB sector. Figure 4-18 and show the HeNB and eNB cell-edge capacity gains and average HUE power reduction achieved by performing joint power control over HeNBs in order to mitigate inter-cell interference compared to no power control or HeNB useful path gain based power control.

We observe from Figure 4-18 and Figure 4-19 that useful path gain based power control minimizes the mean HUE transmit power among all HeNBs. Compared to having constant HUE transmit power (no power control), it achieves similar performance trade-off with lower average HUE transmit power. With the very simple non-UE-specific interference path gain based power control, the eNB and HeNB performance trade-off is improved at the price of an increased HUE average power.

By performing a more advanced UE-specific power control combining useful path gain and interference path gain interference, performance trade-off is further improved while limiting the HUE average power increase compared to useful path gain based power control.



**Figure 4-18: 5%-ile eNB capacity vs. 5%-ile HeNB capacity.**

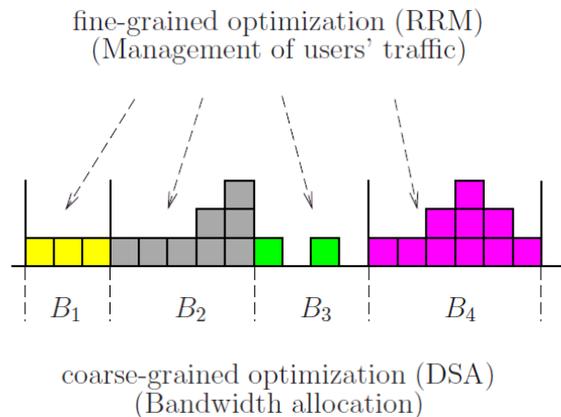


**Figure 4-19: Mean HUE transmit power.**

Thanks to DL blind ICIC, the eNB performance can be improved without reducing the HeNB performance significantly. Principles of an UL ICIC requiring reduced backhaul messages have been described and two approaches have been compared. By sending simple statistics like means and variances to the coordinator, more efficient UL power control preserving MUEs can be achieved. The next step of this study will include further evaluation of the DL and UL algorithms.

#### 4.2.2 Distributed sub carrier allocation for ICIC

This section addresses Dynamic Spectrum Allocation (DSA) schemes for a better efficiency of the spectrum use, taking advantage of the time and spatial variations of the traffic (and therefore of spectrum needs). This approach targets clearly licensed spectrum management and can be of interest for the interference coordination between femto and macro cells in a multicarrier context (LTE). It fits well with an automatic optimization of the network to cope with space and time variations of the traffic as well with the expected behaviour of femto cells (new femto eNB appear at random). DSA should be distinguished from RRM techniques because it addresses a different time-scale: RRM is about managing user's traffic within a given spectrum allocation and is a fine grained optimization of the resources. When it is no longer possible to face the increase of traffic, new spectrum resources must be allocated to the different Radio Access Technologies (RATs) with a coarser time-scale (a coarse-grained optimization) as sketched in Figure 4-20.



**Figure 4-20: RRM vs. DSA: two levels of optimization**

The main problem is to manage interference. In a first, centralized approach, known positions of transmitters and path loss models allow us to compute co-channel interference (spatial reuse of a same frequency band) or adjacent channel interferences coming from imperfect filtering of transmitters and/or receivers (leakage, selectivity). This minimization problem can be tackled by means of combinatorial optimization schemes because we have a finite number of frequency bands and a finite number of transmitters. Typical solutions are Tabu search algorithms, Simulated Annealing or genetic algorithms. Nevertheless these methods are not convenient to implement in a distributed way (we mean that each node would run its own problem). We describe here a distributed version without a centralized entity.

### System description of the innovation

We shall suppose that the spectrum is divided in a number  $N$  of sub-bands. A number of transceivers (macro cells eNB, or femto eNB) operate in the given spectrum, each with a given demand (the required amount of spectrum given in number of subcarriers); the problem is, given an available amount of spectrum, to allocate each transmitter its required number of subcarriers and at the same time managing interferences in order to satisfy quality of transmissions. This is a Min-interference allocation problem as described in [PGB07] [PGD07] [PAG+08]. We can give a graphical description of interference with a graph where two nodes will be connected by an edge if their transmissions cause mutual interference when using a same frequency band: with this definition all nodes would be connected to all other nodes (the graph would be a clique). If we remind that attenuation increases greatly with distance, "neighbour" nodes will contribute to the interference more heavily than distant nodes. To capture this phenomenon we introduce a threshold (in dB) and say that two nodes will be connected by an edge if the path loss (in dB) between them is less than this threshold. Each edge of this graph is weighted by the value  $g_{i,j}$  (in a linear scale) of the attenuation between nodes.

Furthermore Automatic Neighbour Relation (ANR) is an automatic setting of the neighbours list of a cell, and thus it is an information which is the equivalent of the interference graph used in the DSA algorithm, with the notable difference that the whole graph is not necessarily stored in a central place, but all nodes have a list of their neighbours (which could be extended with the associated path-losses).

We introduce some notations:

- $\nu$  is the number of nodes (transmitters)
- $V(i)$  is the neighbourhood of the node  $i$ , that is to say the list of nodes  $j$  such that there is an edge  $(i, j)$  in the interference graph. An edge means that two neighbouring nodes should not use a same frequency (subcarrier).
- $n$  is the number of subcarriers and  $\mathbf{f}_i = (f_{i,1}, f_{i,2}, \dots, f_{i,n})$  is a binary vector which indicates frequencies in use at node  $i$ ,  $f_{i,k} = 1$  if frequency  $k$  is used at node  $i$ , otherwise  $f_{i,k} = 0$ .

- The demand at node  $i$  is known,  $d_i = \sum_{k=1}^n f_{i,k} \in \{1, 2, \dots\}$

- There is a conflict between nodes  $i$  and  $j \in V(i)$  when a same frequency  $k$  is in use at both nodes. With this notation the number of conflicts between nodes  $i$  and  $j$  is simply

given by a scalar product  $\mathbf{f}_i \cdot \mathbf{f}_j = \sum_{k=1}^n f_{i,k} f_{j,k}$  and the total number of conflicts at node  $i$  is

$$C(i) = \sum_{j \in V(i)} \sum_{k=1}^n f_{i,k} f_{j,k} = \sum_{k=1}^n f_{i,k} \left\{ \sum_{j \in V(i)} f_{j,k} \right\}$$

We shall denote  $w_k = \sum_{j \in V(i)} f_{j,k}$  which is the number of neighbours of node  $i$  using the frequency

$k$ . The problem at node  $i$  is thus to minimize the function  $C(i) = \sum_{k=1}^n f_{i,k} w_k$  under constraint

$\sum_{k=1}^n f_{i,k} = d_i$ . This looks like a linear programming problem except that all variables are binary; furthermore we have in fact  $\nu$  such coupled optimization problems.

We introduce a relaxation of variables  $f_{i,k}$  which can now take their values in  $[0,1]$ . The **local** problem at node  $i$  becomes:

$$\begin{aligned} &\text{minimize } C(\mathbf{f}_i) = \sum_{k=1}^n f_{i,k} w_k, \text{ with } w_k = \sum_{j \in V(i)} f_{j,k} \\ &\text{such that } \sum_{k=1}^n f_{i,k} = d_i \\ &\quad f_{i,k} \in [0,1] \quad \forall k \end{aligned}$$

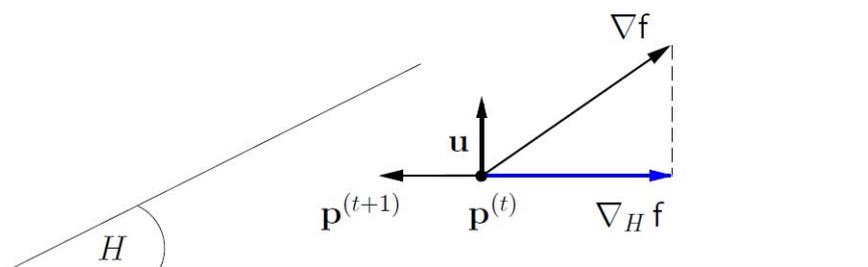
Each node  $i$  can implement a gradient projection method: it is a well-known technique [Boyd04, Ros60] whose main characteristic (as in the steepest descent method) is to follow a direction corresponding to a minimization of the objective function. The difference with steepest descent method is the presence of constraints (here the demand  $\sum_{k=1}^n f_{i,k} = d_i$ , which is a hyper plane) and we use the projection of the gradient on the hyper plane.

Denoting  $\mathbf{u} = (1,1,\dots,1)$  the orthogonal vector to the hyper plane defined by  $\sum_{k=1}^n f_{i,k} = \mathbf{u} \cdot \mathbf{f}_i = d_i$ , the projection of a gradient  $\nabla f$  on this plane is given by

$$\nabla_H f = \nabla f - \frac{\nabla f \cdot \mathbf{u}}{\|\mathbf{u}\|^2} \mathbf{u}$$

For the local problem at node  $i$  the gradient is  $\nabla f = (w_1, w_2, \dots, w_n) = \mathbf{w}$  and the projection is given by

$$\nabla_H f = (w_1 - \bar{w}, w_2 - \bar{w}, \dots, w_n - \bar{w}), \quad \bar{w} = \frac{1}{n} \sum_{i=1}^n w_i$$



**Figure 4-21: gradient projection on hyper plane H**

Each node  $i$  performs a step in the opposite direction to the projected gradient, that is to say:

$$\mathbf{f}_i^{(t+1)} = \mathbf{f}_i^{(t)} - \lambda \nabla_H C(\mathbf{f}_i), \quad \lambda \geq 0$$

We can easily check that  $\nabla_H f \cdot \mathbf{w} \geq 0$  so that any move  $-\lambda \nabla_H f, \lambda > 0$  will cause a decrease of the objective function  $C(\mathbf{f}_i) = \sum_{k=1}^n f_{i,k} w_k$ .

The problem can be easily modified to take into account the propagation attenuations between nodes with the introduction of the gain  $g_{i,j}$  (the path-loss between nodes  $i$  and  $j$  in a linear scale) so that the coefficients  $w_k$  are now given by:

$$w_k = \sum_{j \in V(i)} g_{i,j} f_{j,k}$$

The projection itself and the choice of the parameter  $\lambda$  must be done in a way that ensures that the (soft) variables  $f_{i,k}$  remain in the interval  $[0,1]$ . The result is a generalization of the initial projection operator  $\mathbf{P}_0$  (which does not account for bounds on variables  $f_{i,k}$ )

$$\nabla_H f = \nabla f - \frac{\nabla f \cdot \mathbf{u}}{\|\mathbf{u}\|^2} \mathbf{u} = \left[ I - \frac{1}{\|\mathbf{u}\|^2} \mathbf{u}^t \mathbf{u} \right] \cdot \nabla f = \mathbf{P}_0(\nabla f)$$

Projection  $\mathbf{P}_0$  becomes now:

$$\mathbf{P}_1 = \mathbf{I} - \mathbf{A}(\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t$$

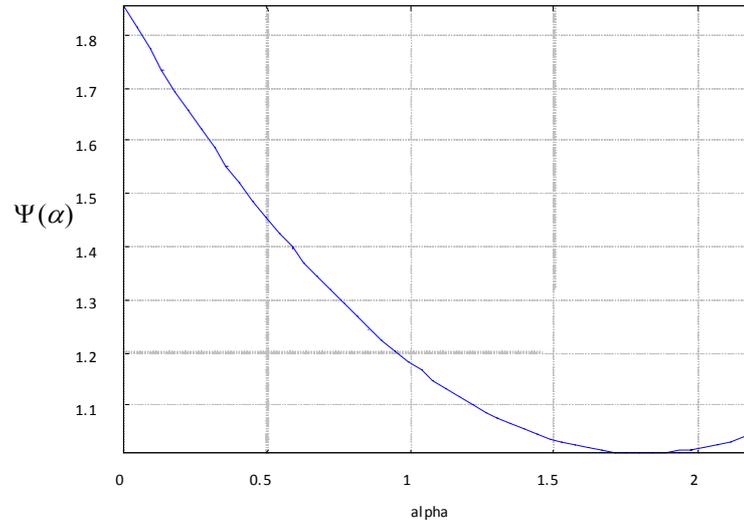
where  $\mathbf{A}$  is a rectangular matrix whose columns are  $\mathbf{u} = (1,1,\dots,1)$  and unit vectors  $\mathbf{e}_{i1}, \mathbf{e}_{i2}, \dots, \mathbf{e}_{in}$  indicating active coordinates (they must remain constant after projection). For example, if  $n = 6$ , the projector orthogonal to  $\mathbf{u} = (1,1,\dots,1)$  and leaving unchanged coordinates 2 and 5 is given by the matrix

$$\mathbf{P}_1 = \begin{pmatrix} 3/4 & 0 & -1/4 & -1/4 & 0 & -1/4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1/4 & 0 & 3/4 & -1/4 & 0 & -1/4 \\ -1/4 & 0 & -1/4 & 3/4 & 0 & -1/4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1/4 & 0 & -1/4 & -1/4 & 0 & 3/4 \end{pmatrix}$$

We observe that, as expected, the components 2 and 5 of the projected gradient  $\mathbf{P}(\nabla f)$  are equal to 0, so that the same components of the solution  $\mathbf{f}_i^{(t+1)}$  remain unchanged after a move  $\mathbf{f}_i^{(t+1)} = \mathbf{f}_i^{(t)} - \lambda \nabla_H C(\mathbf{f}_i)$ ,  $\lambda \geq 0$ . The aim was to prevent coordinates of a new solution being less than 0 or greater than 1.

#### Putting together all local problems:

- we start from an initial random allocation  $\mathbf{f}_i^{(t=0)} = (f_{i,1}, f_{i,2}, \dots, f_{i,n})$  for all nodes  $i \in \{1 \dots \nu\}$  such that all demands are satisfied :  $\sum_{k=1}^n f_{i,k} = d_i$ .
- All nodes  $i$  compute estimations of  $\Delta \mathbf{f}_i = (w_1, \dots, w_n)$ ,  $w_k = \sum_{j \in V(i)} g_{i,j} f_{j,k}$  and the projection  $\mathbf{P}(\Delta \mathbf{f}_i)$  of these gradients accounting for bounds constraints.
- All nodes  $i$  compute the greatest value  $\alpha_i$  such that the vector  $\mathbf{f}_i^{(t)} - \alpha \mathbf{P}(\Delta \mathbf{f}_i)$  is feasible, where 'feasible' means that all its components are in  $[0,1]$ .
- In the next step all nodes  $i \in \{1 \dots \nu\}$  perform a small change in the opposite direction to the projected gradient:  $\mathbf{f}_i^{(t+1)} = \mathbf{f}_i^{(t)} - \alpha \mathbf{P}(\Delta \mathbf{f}_i)$  with  $0 \leq \alpha \leq \max(\alpha_1, \dots, \alpha_\nu)$ , the value of  $\alpha$  being such that the criterion  $\Psi(\alpha) = \sum_i C[\mathbf{f}_i^{(t)} - \alpha \mathbf{P}(\Delta \mathbf{f}_i)]$ ,  $0 \leq \alpha \leq \max(\alpha_1, \dots, \alpha_\nu)$  is minimum, for generally  $\Psi(\alpha)$  exhibits a minimum for some  $\hat{\alpha}$  as depicted below ( $\hat{\alpha} \approx 1.8$ ).



**Figure 4-22: Typical behavior of the criterion  $\Psi(\alpha)$**

This step is not fully distributed and corresponds to the present status of our work.

(e) Update current solution at each node:  $\mathbf{f}_i^{(t+1)} = \mathbf{f}_i^{(t)} - \lambda \nabla_H C(\mathbf{f}_i)$ ,  $\lambda \geq 0$  and go to step (b).

### Performance results and future steps

We are still in the development phase of our algorithm, simulation results will be available in a next version of the deliverable.

We initially intended to implement a distributed version of a centralized heuristic algorithm, Tabu Search (TS), available from a previous work. Unfortunately TS is not convenient for this purpose: the elementary step of TS involves the choice of a single node  $i$ , and then the choice of two frequencies to be swapped. This can be easily done when a central entity performs all computations, but is not well suited for a distributed scheme. So we switched to another scheme based upon a relaxation of the initial problem.

We have already implemented steps (a) to (d) aforementioned and are still working on step (e) because simulations we have done show that projected gradients become smaller and smaller while we are not near an optimum.

### 4.3 Coordinated Scheduling

As aforementioned, there are two families of techniques for downlink CoMP: CS/CB and JP. In the first category data is only available at the serving cell but user scheduling decisions are made with coordination among cells. For JP, the transmission to a single UE is simultaneously performed from multiple cells. These two approaches aim at improving user throughputs, in particular at the cell-edge, leading to higher system efficiency.

Among CB/CS and JP, CB/CS turns out to be less sensitive to CSI accuracy than JP, thus being more robust in particular in the case of non-pedestrian UE velocities [WIND 18]. Even if it appears that CB/CS is less restrictive than JP, several aspects of CB/CS still need to be clarified. In this section we present the coordinated scheduling based on three aspects: cross-layer approach for advanced system, coordinated beams approach and heterogeneous deployment approach.

The first approach, presented in section 4.3.1 and section 4.3.2, proposes to resolve two optimization problems. The first problem consists in maximizing the user rates over all the different transmission modes such as the SU-MIMO and the MU-MIMO and over the frequency/time resources where the scheduler is based on the proportional fair algorithm. Scheduling algorithm coordinated between the eNBs able to dynamically manage different transmission modes is designed. The second problem maximizes the min SINR margin over all

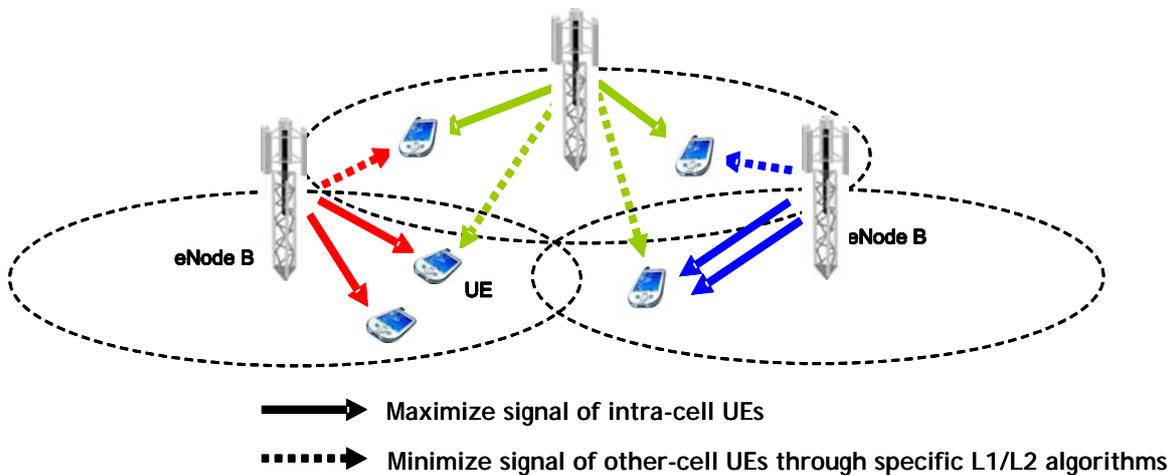
possible beamformers and transmission powers under a maximal transmission power available at each cell. The solution of this optimization problem is an iterative algorithm based on the uplink-downlink duality.

The second approach is based on beam collision avoidance. In fact, standard approaches are based on coordination of channel accesses either in the frequency dimension (fractional frequency reuse) or the time dimension using some form of time division duplex transmission. If the transmitter (the eNB, in the case of downlink transmission) is equipped with several antennas, the possibility of beamforming allows the steering of the radiated signal in specific directions, thereby opening up the spatial dimension for interference coordination. If a transmitting cell uses a beam directed towards a cell-edge user, it may cause significant interference to a nearby user that is served from a different cell.

The contributions in sections 4.3.3 and 4.3.4 require feedback information from the interfered UEs. As for the first one this information is the worst companion beam combined with downtilt information. For the second the RESTRICTION REQUEST information is exchanged between the macro and femto cells. In section 4.3.5 four basic coordinated femtocells schemes are considered: interference nulling, joint superimposed transmission, joint orthogonal repetition and joint orthogonal transmission.

### 4.3.1 Dynamic algorithms for coordinated scheduling

This work is focused on the study of a Multi Cell Layer 1 (L1) and L2 Interference Control scheme, with particular reference to the aspects related to the activation of appropriate L2 mechanisms able to dynamically realize an efficient resource allocation. Figure 4-23 depicts the overall system, in which the presence of L1/L2 mechanisms are designed with a cross-layer approach, and in particular at L2 the packet scheduler has in charge the exploitation of the radio resources, while taking into account also the level of interference (also given by L1 interference rejection schemes) in order to optimize system performances.



**Figure 4-23: Multi Cell L1/L2 Interference Control scheme.**

From this point of view, the scheduling algorithm performances are strictly linked to the L1 interference rejection mechanisms studied in Task 1.1 (and described in deliverable D1.2 [ARTD12]) and user pairing activities carried out in Task 1.2 (and described in section 4.1.3).

### System description of the innovation

Current definition of LTE technology related to packet scheduling operation foresees the usage of different degrees of freedom during transmission, mainly related to choice of different modulation and coding schemes, antenna mapping and transmission modes selection, resource allocation and power assignment, etc. In particular, as a first step, the selection of a proper transmission mode should be exploited in a multi-user environment in which each user is

characterized by time and frequency varying channel conditions, and these conditions also differ from user to user.

According to this need a TM Switching (TMS) algorithm has been designed, in order to first perform single cell SU-MIMO TM selection among Transmit Diversity (TxD) and open and closed loop Spatial Multiplexing (SM) managed through control channels in order to enable the adaptation to fast fading characteristics. Moreover, the addition of single cell MU-MIMO mode is the second step, necessary to exploit MU-MIMO in a multi cell scenario, also in conjunction with interference rejection techniques performed at L1. Finally, the formulation of the proposed Multi Cell L1/L2 Interference Control scheme will be based on scheduling coordination between eNBs: in this case a minimal set of information exchange between nodes through control plane will be introduced (e.g. the set of scheduled users and corresponding resource allocation decisions provided by the other nodes).

Thus, as a first step a dynamic switch between different TMs is envisaged. In particular this can be signalled in two ways:

- by means of a proper Radio Resource Control (RRC) message (LTE Rel-8): for example the switching between SU-MIMO and MU-MIMO can be realized by changing TM4 and TM5 involving RRC signalling, with a latency that makes it unfeasible to follow fast fading characteristics;
- by exploiting a fast signalling exchange over the control channel (LTE-Advanced and beyond), potentially in real-time and with a fixed latency of 1 Time Transmission Interval (TTI).

In this contribution a proper scheduling algorithm able to manage different transmission modes in a dynamic way has been designed. The TMS is performed through control channels, thus by enabling the adaptation with fast fading characteristics.

The starting point of the scheduler design is the well-known Time Domain Packet Scheduling (TDPS) Proportional Fair (PF) algorithm, where users are scheduled according to PF policy and just by taking into account the time dimension. This algorithm has been extended for the SU-MIMO TDPS / Frequency Domain Packet Scheduling (FDPS) scheduler by taking into account the extra frequency and spatial dimensions. Let's first define the PF metric for  $c$ -th PRB,  $j$ -th Transmission Mode and time  $t$ :

$$\lambda_{i,j}^c(t) = \frac{\hat{r}_{i,j}^c(t)}{R_i(t)}$$

where  $\hat{r}_{i,j}^c(t) = r_{i,j}^c(t) \cdot (1 - p_{e,j}(SNR))$  is the effective data rate estimated at the transmitter side, for the  $i$ -th user and  $j$ -th Transmission Mode (where  $r_{i,j}^c(t)$  is the channel rate and  $p_{e,j}(SNR)$  is the error probability) and  $R_i(t)$  is the long term service rate:

$$R_i(t+1) = (1 - \alpha) \cdot R_i(t) + \alpha \cdot \sum_c \sum_j x_{i,j}^c(t) \cdot \hat{r}_{i,j}^c(t)$$

where  $\alpha$  is a constant typically of the order of  $1/1000$ , the variable  $x_{i,j}^c(t) \in \{0,1\}$  indicates whether or not RB  $c$  is assigned to user  $i$  with transmission mode  $j$  at time instant  $t$ . Thus, if  $x_{i,j}^c(t) = 1$  then user  $i$  is scheduled and has an effective data rate of  $\hat{r}_{i,j}^c(t)$  for RB  $c$  with transmission mode  $j$  at time instance  $t$ .

According to conventional PF formulation, a first version of the PF objective function at time instant  $t$  could be defined as follows:

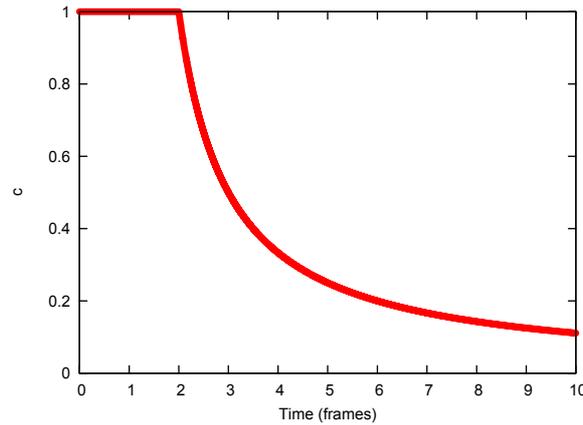
$$\max \sum_i \sum_c \sum_{j \in M} x_{i,j}^c(t) \cdot \lambda_{i,j}^c(t)$$

where  $M$  is the set of available transmission modes, and additional constraints are considered e.g. in order to use only one MIMO mode per user per time instance.

The basic algorithm builds on some simplified assumptions that do not take into account practical implementation aspects. In our studies two additional aspects have been addressed:

1. we extended the approach by considering realistic scenarios such as realistic feedbacks (reporting aligned with Rel-8 specifications), MAC overhead, traffic models, propagation scenarios.
2. we defined a heuristic formulation of the algorithm, since optimal solution is unfeasible with algorithms running on polynomial time.

In particular, regarding the feedback management, it is at most defined per sub-band; moreover it may be not be available for all sub-bands in every TTI for all UEs; finally, the UEs can only indicate the CQI for one single transmission mode, which is the one currently employed. Thus a smart feedback filtering has been realized by means of a proper confidence function (in the range [0,1]) according to reliability of the reported measure: the closer is  $c_{i,j}^s(t)$  to 1 (where s is the sub-band), the more “confident” become the corresponding  $r_{i,j}^s(t)$ . Figure 4-24 shows an example of confidence function used to weight reporting messages, where in an initial time interval the value is set to 1 and in subsequent frames is decreased in order to take into account the degradation of the feedback reliability due to its obsolescence. The definition of an appropriate function for computing  $c_{i,j}^s(t)$  is left for further study.



**Figure 4-24: Example of confidence function used to weight reporting messages.**

Regarding the heuristic formulation of the algorithm, for each queued user the best transmission mode is then selected in the following way:

$$\tau = \arg \max_{j \in M} \sum_{s=1}^{n_{SB}} f(\lambda_{i,j}^s(t), c_{i,j}^s(t))$$

where  $n_{SB}$  is the number of sub-bands in the system,  $f(\lambda_{i,j}^s, c_{i,j}^s)$  is the new PF metric (here called *summary value*) defined in a first step as the product of the two parts (further and more advanced mathematical definitions of this function are left for future study).

Finally, the current formulation of the TM switching algorithm consists (for each TTI) in the following steps:

1. **Feedback processing:** for each sub-band s calculate the couple  $\hat{r}_{i,j}^s, c_{i,j}^s$  then the summary  $f(\lambda_{i,j}^s, c_{i,j}^s)$ ;

2. **Transmission mode selection and user scheduling:** for each queued user  $i$  choose the TM  $\tau$  and add for each sub-band the corresponding t-uple  $\langle s, i, \tau \rangle$  in a sorted list  $L$  in decreasing order of the corresponding summary value;
3. **Post-processing phase:** for each tuple  $\langle s, i, \tau \rangle$  in the sorted list  $L$  allocate the backlogged bytes of user  $i$  and start filling the corresponding sub-band  $s$  (when a sub-band  $s$  is already fully allocated by other users, then the relative tuple is skipped because it doesn't correspond to a feasible scheduling decision).

This formulation can be extended with the addition of MU-MIMO mode (for the exploitation of this new degree of freedom), and also by studying the relationship of the scheduling operation with L1 interference rejection techniques and user grouping, in order to correctly manage cell resources during transmission.

### Performance results and future steps

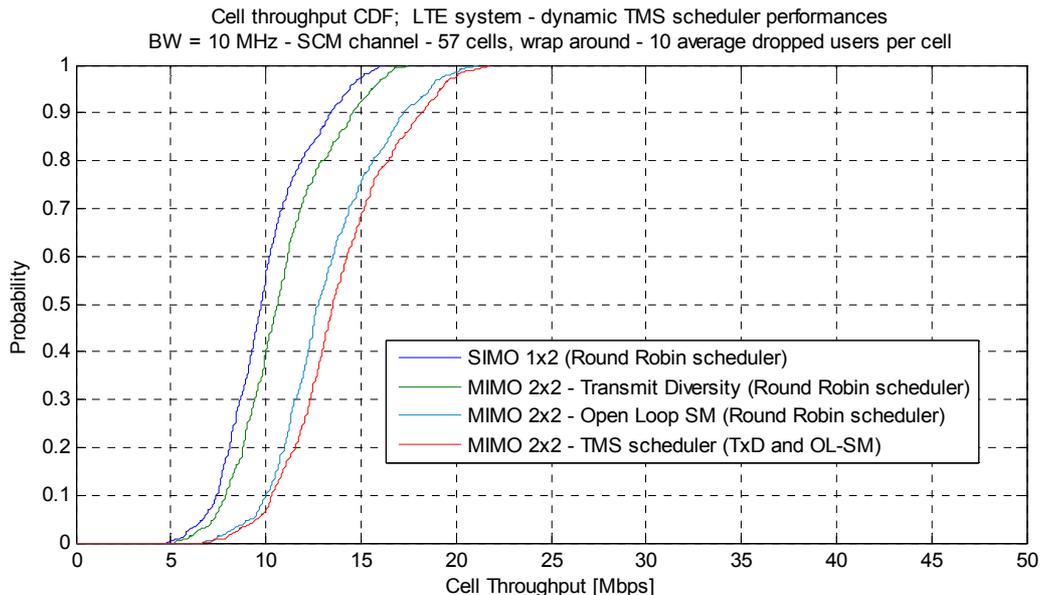
Simulation results have been obtained with a dynamic system level simulator, implementing the well-known "snapshot method": during each snapshot a certain number of users is dropped in the system at random positions; each snapshot is simulated for a time duration that has been verified to be long enough in order to produce stable results; the network is supposed to be stable (no birth/dead or handover events occurs since the time granularity of the simulation is relative to traffic sessions of a certain number of active users).

The table below summarizes the main parameters used in system level simulations:

**Table 4-6: Simulation parameters**

Simulation Parameter	Value
Simulated scheme	SIMO 1x2; MIMO 2x2; Transmission Mode Switching
Cellular Layout	Hexagonal grid, 19 sites, 3 sectors per site
Number of average dropped users per cell	10
Simulated Link	Downlink
Deployment scenario	Urban Macrocellular (UMa)
Traffic model	Full buffer
Packet Scheduling	TMS (Transmission Mode Switching), Round Robin.
Adaptive Modulation and Coding	3GPP LTE standard transport formats; AMC $PER_{target} = 10\%$
Bandwidth	10 MHz (50 PRB)
Channel model	Spatial Channel Model (SCM)
Interference	Explicit (the 9 strongest interference cells are considered)
Number of antenna elements (BS, UE)	(2, 2)
Antenna separation (BS, UE)	$(\lambda/2, \lambda/2)$
Link to system interface	Mutual Information Effective SNR Mapping (MIESM)
HARQ	Stop and wait; synchronous adaptive
Number of HARQ processes	8
Retransmission interval	8 ms
Maximum number of retransmissions	3 (corresponding to a maximum of 4 transmissions)
CQI reporting	Wideband CQI, no PMI on PUCCH (mode 1-0)
UE channel Estimation	Realistic (embedded into link level performance curves)

Preliminary simulations have been performed in order to make a comparison between a SIMO 1x2 system and a MIMO 2x2 system in which transmission modes are statically chosen (Transmit Diversity or Spatial Multiplexing) or dynamically managed by means of the TMS scheduler described in the above. Results depicted in Figure 4-25 show the benefit obtained when using the TMS scheduler, thus improving the cell throughput distribution with respect to Spatial Multiplexing performances.



**Figure 4-25: Cell Throughput distribution for UMa (TMS scheduler with MIMO 2x2).**

In a first step TMS algorithm performances have been assessed by considering Single User MIMO TM selection by means of control channels, enabling the adaptation to fast fading characteristics. In next steps Single cell MU-MIMO will be added as an additional transmission mode and taken into account by the scheduler (simulation assessment will be conducted in a multi-cell interference scenario). Exploitation of MU-MIMO in a multi cell scenario will be considered for scheduling decisions in relationship with interference rejection techniques performed at L1.

Further improvement of the study is based on scheduling coordination between eNBs: in this case a minimal set of information exchange between nodes through control plane will be introduced (e.g. the set of scheduled users and corresponding resource allocation decisions provided by the other nodes).

#### 4.3.2 Precoding optimization algorithm for coordinated beamforming

The coordination between the different transmission points in the network can be seen as an additional system resource regarding the classical resources (space, bandwidth and power). Substantial performances gains can be achieved, depending on the number of the coordinated points, their corresponding transmit processing schemes and the exchanged information. However, these gains come at the cost of an increased backhaul load. Thus, CB/CS is a potential solution.

For single cell multiuser MIMO systems, there are two closely related optimization problems to derive the beamforming coefficients: the first one is based on the maximization of the joint achievable SINR under a total power constraint, and the second one is based on the minimization of the total transmission power while satisfying a set of SINR constraints.

These two problems have been addressed in both downlink and uplink. Downlink joint power control and beamforming, is more complicated than the uplink one because the received user SINRs are coupled by the transmit beamforming vectors and powers which must be jointly optimized.

When non orthogonal transmission is allowed, which is the case of our study, the choice of one UE's beamforming vector may affect the crosstalk experienced by other UEs. For fixed SINR targets, the downlink joint power control and beamforming problem has been solved in [SB04, FLT98]. A key technique of the solution is based on a duality between the multiuser downlink and a virtual multiuser uplink [SB04, VM99]. The duality theorem states that, under the same sum power constraint, both the downlink and uplink have the same achievable SINR region [SB04]. Moreover these targets can be achieved by the same set of beamforming vectors. A second class of algorithms has been developed for the multiuser beamforming approach, which is based on Semi Definite Programming (SDP) [BO99].

In a coordinated multi-cell system, an iterative algorithm based on uplink-downlink duality was introduced in [WIND18, DY08], where the downlink beamformers are designed locally based on the reciprocal uplink channels and virtual uplink powers of all users. This also allows for a distributed implementation, where virtual uplink powers are exchanged between eNBs in a coordinated manner. In our study, a joint power control and beamforming is addressed in a coordinated beamforming multipoint system.

It has been shown that the downlink transmit channel covariance matrix can be obtained from the uplink Rx channel covariance matrix by means of channel reciprocity, even for an FDD system [3GPP-R1100853]. Our study is based on the availability of long-term downlink transmit channel covariance matrices. We propose an algorithm that jointly finds a set of feasible transmit beamforming weight vectors and downlink transmit power allocations such that the SINR at each link is greater than a target value. A user selection step is also added to the initial algorithm to improve the cell-edge performance.

### System description of the innovation

Consider  $N_B$  cells in the network transmitting to  $K$  users. Each UE is equipped by  $N_r = 1$  receive antenna and each cell is equipped with  $N_t$  transmit antennas.

The received signal at UE  $k$  is expressed as

$$\mathbf{r}_k = \sqrt{p_k} \mathbf{h}_k \mathbf{w}_k s_k + \sum_{\substack{j=1 \\ j \neq k}}^K \sqrt{p_j} \mathbf{h}_k \mathbf{w}_j s_j + n_k$$

Where  $\mathbf{h}_k = [\mathbf{h}_{k,1}^T, \dots, \mathbf{h}_{k,n}^T, \dots, \mathbf{h}_{k,N_B}^T]^T$ ,  $\mathbf{h}_k$  is the channel vector of dimension  $(1 \times N_t N_B)$  between the  $k^{th}$  UE and all the  $N_B$  transmit cells. The precoding vectors  $\mathbf{w}_k = [\mathbf{0}^T, \dots, \mathbf{w}_{k,n}^T, \dots, \mathbf{0}^T]^T \in C^{N_t N_B \times 1}$  precode the data streams intended for the  $k^{th}$  UE at all the cells. The vector  $\mathbf{w}_{k,n} \in C^{N_t \times 1}$  is the precoding vector which is used at cell  $n$  corresponding to its active user  $k$ . Notice that each cell can transmit to multiple users: in such case we have the multiuser MIMO scheme.

The scalar  $n_k$  is zero-mean additive white Gaussian noise with unit variance at the  $k^{th}$  UE. The average pre-detector SINR of  $k^{th}$  UE conditioned to  $\mathbf{W}_k$  is given by:

$$SINR_k^{DL} = \frac{p_k \mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k}{\sum_{\substack{j=1 \\ j \neq k}}^K p_j \mathbf{w}_j^H \mathbf{R}_k \mathbf{w}_j + \sigma_k^2}$$

Where  $\mathbf{R}_k = \text{diag}(\mathbf{R}_{k,1}, \mathbf{R}_{k,2}, \dots, \mathbf{R}_{k,N_B})$  and  $\mathbf{R}_{k,n} = E\{\mathbf{h}_{k,n}^H \mathbf{h}_{k,n}\}$  is the covariance matrix of channel  $\mathbf{h}_{k,n}$  between the  $k^{th}$  UE and the  $n^{th}$  cell.

Each user has a SINR target  $\gamma_k$  which must be reached in order to have a reliable transmission. The following information is assumed to be available at all eNBs:

- the SINR targets  $\gamma_k$  of each user in the network.
- covariance matrix  $\mathbf{R}_{k,n} = E\{\mathbf{h}_{k,n}^H \mathbf{h}_{k,n}\}$ ,  $n \in \{1, \dots, N_B\}$  of the channel from cell  $n$  (the interfered cell) to the  $k^{\text{th}}$  UE, obtained via long-term uplink-downlink channel reciprocity at cell  $n$ .

Each user in the network transmits its data using power  $p_k$ , but this power is under a total power constrained such that:  $\sum_{k \in S_n} p_k \leq P_{\max}$ , where  $S_n$  is the set of active users in cell  $n$  and

$P_{\max}$  is the maximum transmit power at each cell  $n \in \{1, \dots, N_B\}$ .

All the target thresholds  $\gamma_1, \dots, \gamma_k, \dots, \gamma_K$  can be achieved simultaneously if and only if

$\min_{1 \leq k \leq K} \frac{SINR_k}{\gamma_k} \geq 1$ . The problem will be then a joint power and beamforming optimization which can be expressed as follows:

$$\begin{cases} C^{DL}(\mathbf{W}, P_{\max}) = \max_{\mathbf{W}, P} \min_{k=1..K} \frac{SINR_k(\mathbf{W}, P)}{\gamma_k} \\ \text{constr: } p_n = \sum_{k \in S_n} p_k \leq P_{\max} \quad \forall n = 1, \dots, N_B \end{cases}$$

Where  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_k, \dots, \mathbf{w}_K]$ ,  $P = [p_1, \dots, p_k, \dots, p_K]^T$  and  $p_n$  is the transmit power used by the cell  $n$ .

Following the same problem resolution developed by [SB04] based on the uplink and downlink duality, for a given beamforming matrix  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_k, \dots, \mathbf{w}_K]$  we can derive the downlink power allocation that maximizes the worst case SINR margin:

$$\begin{cases} C^{DL}(\tilde{\mathbf{W}}, P_{\max}) = \frac{SINR_k(\tilde{\mathbf{W}}, \tilde{P})}{\gamma_k} \\ \text{constr: } \tilde{p}_n = P_{\max} \quad , \quad \forall n = 1, \dots, N_B \end{cases} \quad , \quad 1 \leq k \leq K$$

This means that we must find the set of cells that can transmit at the maximum transmit power. At least we can have only one cell  $n^*$  which satisfies this constraint. The power allocation problem is solved by the resolution of the following eigensystem derived from the above simplified optimization problem:

$$\Gamma(\tilde{\mathbf{W}}, P_{\max}, n^*) \tilde{\mathbf{p}}_{ext} = C^{DL}(\tilde{\mathbf{W}}, P_{\max})^{-1} \tilde{\mathbf{p}}_{ext}$$

$$\text{with: } \Gamma(\mathbf{W}, P_{\max}, n) = \begin{bmatrix} \mathbf{D}(\mathbf{W})\Psi(\mathbf{W}) & \mathbf{D}(\mathbf{W})\boldsymbol{\sigma} \\ \frac{1}{P_{\max}} \mathbf{1}_{n^*}^T \mathbf{D}(\mathbf{W})\Psi(\mathbf{W}) & \frac{1}{P_{\max}} \mathbf{1}_{n^*}^T \mathbf{D}(\mathbf{W})\boldsymbol{\sigma} \end{bmatrix} \quad , \quad \tilde{\mathbf{p}}_{ext} = \begin{bmatrix} \tilde{\mathbf{p}}^T & 1 \end{bmatrix}^T$$

$\tilde{\mathbf{p}}_{ext}$  is a  $(K+1) \times 1$  extended power vector of  $\tilde{\mathbf{p}}$ .  $\mathbf{1}_{n^*}^T$  is  $(K \times 1)$  vector having ones only in the position of the scheduled users of the cell  $n^*$ .

$$\mathbf{D}(\mathbf{W}) = \text{diag} \left( \frac{\gamma_1}{\mathbf{W}_1^H \mathbf{R}_1 \mathbf{W}_1}, \dots, \frac{\gamma_K}{\mathbf{W}_K^H \mathbf{R}_K \mathbf{W}_K} \right)$$

$$[\Psi(\mathbf{W})]_{j,k} = \begin{cases} \mathbf{W}_k^H \mathbf{R}_k \mathbf{W}_k & , j \neq k \\ 0 & j = k \end{cases} \quad \text{and} \quad \boldsymbol{\sigma} = [\sigma_1^2, \dots, \sigma_K^2]^T$$

$C^{DL}(\tilde{\mathbf{W}}, P_{\max})^{-1}$  is the eigenvalue of the nonnegative extended coupling matrix  $\Gamma$  and  $\tilde{\mathbf{p}}_{ext}$  is its eigenvector corresponding to the optimum power allocation.

One of the most important results of the uplink-downlink duality is that under the given constraints both uplink and downlink have the same SINR achievable region. Moreover the targets can be achieved by the same set of beamforming vectors [SB04]. This result holds for equal receiver noises or by using a scaled covariance matrix such us:  $\mathbf{R}'_k = \frac{\mathbf{R}_k}{\sigma_k^2}$  and noise

variances  $\sigma_k^2 = 1$ ,  $1 \leq k \leq K$ . The SINR for the uplink is given by:

$$SINR_k^{UL} = \frac{q_k \mathbf{w}_k^H \mathbf{R}'_k \mathbf{w}_k}{\mathbf{w}_j^H \left( \sum_{\substack{j=1 \\ j \neq k}}^K q_j \mathbf{R}'_k + I \right) \mathbf{w}_j}$$

Where  $q_k$  is the transmit power of UE  $k$ . The same targets achieved in the downlink with the same fixed beamforming vectors can be achieved in the uplink by an uplink power allocation  $\tilde{q}_k$ , so we can write:  $SINR_k^{UL}(\tilde{\mathbf{W}}, \tilde{q}) = \gamma_k C^{DL}(\tilde{\mathbf{W}}, P_{\max})$

Using the expression of  $SINR_k^{UL}$  and the matrix notation we obtain the uplink power allocation:

$$\tilde{q}_k = \left( C^{DL}(\tilde{\mathbf{W}}, P_{\max})^{-1} D(\tilde{\mathbf{W}})^{-1} - \Psi^T(\tilde{\mathbf{W}}) \right)^{-1} \mathbf{1} \boldsymbol{\sigma}$$

where  $\mathbf{1}$  is the  $(K+1)$  vector with only ones in the position of the scheduled users, in the coordinated cells.

Since the power allocation is determined we can proceed to find the precoding vectors (downlink vectors are the same as the uplink ones) solved by the dominant generalized eigenvectors of the matrix pairs  $(\mathbf{R}'_k, \mathbf{Q}_k(q_{ext}))$ :

$$\tilde{\mathbf{w}}_k = \arg \max_{\mathbf{w}_k} \frac{q_k \mathbf{w}_k^H \mathbf{R}'_k \mathbf{w}_k}{\mathbf{w}_j^H \mathbf{Q}_k(q_{ext}) \mathbf{w}_j}, \quad \text{where } \mathbf{Q}_k(q_{ext}) = \sum_{\substack{j=1 \\ j \neq k}}^K q_j \mathbf{R}'_k + I, \quad \text{for } 1 \leq k \leq K$$

These power allocation and beamforming steps could be repeated iteratively until convergence, the following algorithm summarizes what we have already described above:

### The algorithm:

**Step1:** All the cells in the network estimate the  $K$  downlink covariance matrices  $\mathbf{R}_{k,n}$  from the uplink channel estimation.

**Step2:** Each cell transmits the covariance matrices to a central scheduler (could be one of the eNB in the network), where the channel covariance matrices for each UE are constructed  $\mathbf{R}_k = \text{diag}(\mathbf{R}_{k,1}, \mathbf{R}_{k,2}, \dots, \mathbf{R}_{k,NB})$  for  $1 \leq k \leq K$ .

**Step3:** Initialisation at time  $t$ :  $it=0$   $\hat{\mathbf{q}} = [0, \dots, 0]^T$   $\mathbf{R}'_k = \frac{\mathbf{R}_k}{\sigma_k^2}$  for  $1 \leq k \leq K$

**Step4: Repeat it  $\leftarrow it+1$**

Calculation of the precoding vectors:  $\tilde{\mathbf{w}}_k = \arg \max_{\mathbf{w}_k} \frac{q_k \mathbf{w}_k^H \mathbf{R}'_k \mathbf{w}_k}{\mathbf{w}_j^H \mathbf{Q}_k(q_{ext}) \mathbf{w}_j}$  for  $1 \leq k \leq K$

**Step5:** Find the cell which satisfies the power constraint

Select the cell  $n^*$  from  $\{1, \dots, NB\}$  which has the maximum cell-edge users.

Put  $p_{n^*} = P_{\max}$

Solve  $\Gamma(\tilde{\mathbf{W}}, P_{\max}, n^*) \tilde{\mathbf{p}}_{ext} = C^{DL}(\tilde{\mathbf{W}}, P_{\max})^{-1} \tilde{\mathbf{p}}_{ext}$

Check if  $p_{n^*} = \sum_{k \in S_n} p_k \leq P_{\max} \quad \forall n = 1, \dots, N_B$

**else if** reduce the number of the coordinated cells and return to step5 **else not continue**

**Step6: until**  $C^{(it)} \approx C^{(it-1)}$  **else compute**  $\tilde{q}_k = \left( C^{DL}(\tilde{\mathbf{W}}, P_{\max})^{-1} D(\tilde{\mathbf{W}})^{-1} - \Psi^T(\tilde{\mathbf{W}}) \right)^{-1} \mathbf{1}\sigma$

and **go to** step4.

### Performance results and future steps

In this section we provide some numerical results to illustrate the performance of the proposed algorithm compared to the coordinated beamforming scheme proposed in section 2.2.3 [WIND18].

The proposed algorithm has been simulated with the hexagonal deployment with 19 sites and 57 cells each one equipped with 4 transmit antennas. The number of users is 10/cell on average. We consider as a baseline the no-coordinated cells scheme using the LTE codebook with 16 precoding vectors.

**Table 4.1: Simulation parameters**

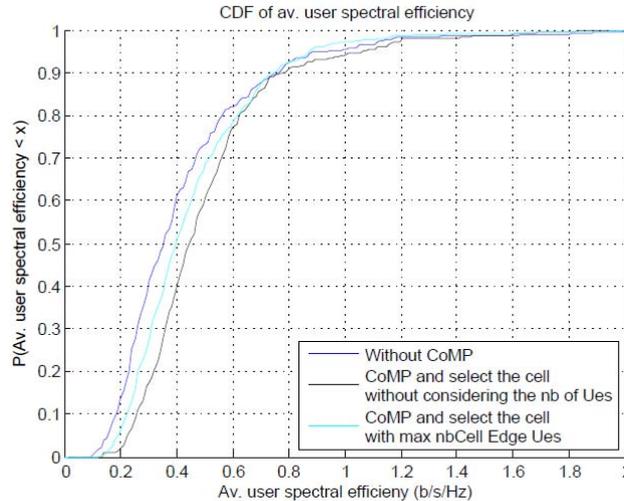
Simulation Parameter	Value
Number of base stations	3
Number of users	19 sites, 57 cells
Number of antennas at each base station	10 Users / cell
Number of antennas at each user	4
Antenna type at the base station	1
Antenna spacing	3GPP path loss and shadowing models
Inter site distance (ISD)	0.5λ Uniform Linear Array
Minimum distance between UE and cell	500m
Number of channel realizations at each position	35m
Centre frequency	1000
Channel Model	2 GHz
Cell radius	one OFDM subcarrier, uncorrelated Rayleigh fading, perfect link adaptation, PF scheduler, MMSE receiver
	500 m

Figure 4-26 shows the performance of the proposed coordinated scheduling algorithm compared to the non-coordinated scheme. The proposed algorithm is also evaluated on the scheduling criteria. In fact, the algorithm is first considered with the constraint on the choice of the selected cells (CoMP1) and second when this constraint is lifted (CoMP2).

The performance of the proposed algorithm when the chosen cell is selected regardless the number of its cell-edge users is shown by the black curve.

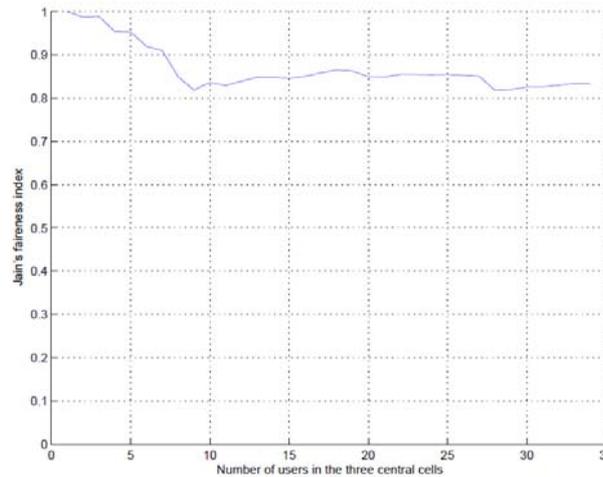
When the selected cell is the one which has the maximum number of cell-edge users, the algorithm remains better than the no-CoMP but falls below the CoMP1 since the other cells are constrained to transmit with lower powers in order to maximize the worst SINR margin of the

selected cell having greater number of cell-edge users. Moreover the power distributed for the cell-edge users of the selected cell comes from the users far from the cell-edge which explains the loss shown from the 0.7 (b/s/Hz) at the average user spectral efficiency.



**Figure 4-26: Performance of the iterative optimization algorithm.**

We have plotted in Figure 4-27 the Jain index obtained for the iterative algorithm CoMP1. We can see that the Jain index decreases slightly as the number of UEs increases. The values of the Jain index don't exceed 0.8 which leads to consider the proposed algorithm as fair for the different users in the network.



**Figure 4-27: Jain index performance.**

We have proposed an iterative algorithm to jointly optimize the beamformers and transmission powers in multi-cell network where each user receives its data from a single cell. This algorithm maximizes the worst SINR margin under per-cell power constraints which is helped by user selection procedure. Here, we have assumed a single receiver antenna at the user side. But the proposed solution can easily be generalized to multiple receive antennas. In addition, we will evaluate this solution under limited feedback links in the future. Finally, the performance of this algorithm will be assessed on a system level simulation in the future.

### 4.3.3 Distributed scheduling for beam coordination

It has already been shown that advanced 3D beamforming can provide substantial gain even without coordination of the scheduling decisions [ARTD12]. However, it is expected that with appropriate beam coordination additional performance improvements can be achieved [ALUD10, DJG+09]. Beam coordination among neighbour eNBs of a multicell wireless system aim at avoiding collisions of the beams assigned to simultaneously scheduled UEs of adjacent cells, as exemplarily indicated in section 3.2. A full optimization of the scheduling decisions in combination with the applied horizontal beamforming weights of all eNBs would require either a centralized entity using system-wide information on scheduling requirements at each eNB, or a decentralized algorithm using the same information but running on each eNB. Both approaches are basically able to derive the global optimum according to criteria like overall spectral efficiency or cell edge throughput maximization. But both approaches require the exchange of all relevant scheduling information either with a central processing unit or with any other eNB. This implies extensive signalling overhead and latency constraints. In combination with 3D beamforming this overhead may even increase due to the additional degrees of freedom and the potentially increased number of scheduling combinations to be taken into account.

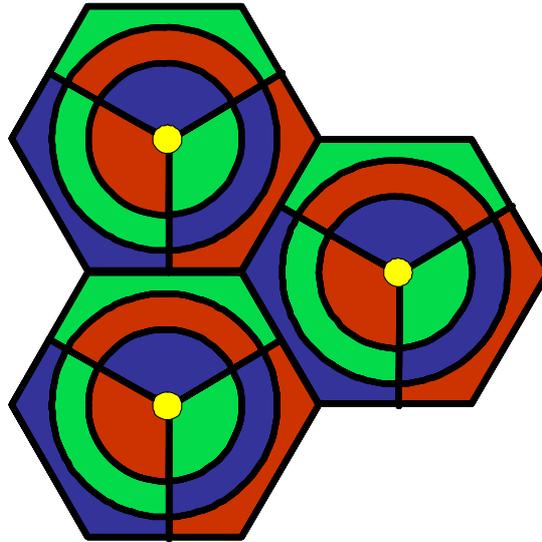
With respect to complexity reduction and practical feasibility, two promising approaches have been identified and analysed. The first approach is an implicit coordination method without explicit information exchange between eNBs. Explicit information exchange here means e.g. the usage of the X2 interface as specified in [3GPP36420]. The location of the UEs within a cell, and their pathloss or received interference level is captured. In combination with an a-priori known network-wide resource configuration, an optimized UE-individual resource assignment can be derived.

The second approach relies on the use of distributed algorithms operating on coordination areas. These coordination areas comprise a relatively small number of eNBs with a limited amount of information exchange between them. For the pure horizontal beamsteering the benefits of coordinated scheduling with such a distributed approach have already been proven [ALUD10]. Within ARTIST4G, this approach will be extended and adapted towards 3D beamforming. The aim is to exploit the additional degree of freedom for resource allocation introduced through vertical beamforming. Therefore new scheduling concepts for interference avoidance, making use of the vertical beamforming capability, will be derived. Different parameter settings and algorithm constraints are considered. The existence of the expected additional gain could already be verified and further improvement is under investigation.

## System description of the innovation

### Implicit coordination

For implicit coordination the schedulers work without control information exchange between eNBs. Coordination of resource assignment is achieved through an appropriate mapping of time and/or frequency resources to specific beams or dedicated sectors within a cell. This could be considered as “location based scheduling”. The basic principle is shown in Figure 4-28. The cell is divided into different areas (3 areas per cell in the example of Figure 4-28), to each area specific time/frequency resources are assigned. The configuration of this resource assignment is a parameter, which is assumed to be known in the network. Based on this knowledge, the scheduler of each cell can assign a predefined resource to a UE located in a specific area. The configuration of the present assignment is supporting interference avoidance in particular at cell edge.



**Figure 4-28: Resource assignment for implicit coordination**

In combination with 3D beamforming, the assignment of resources is complemented by a simultaneous adaptation of the antenna downtilt according to the area where the UE is located.

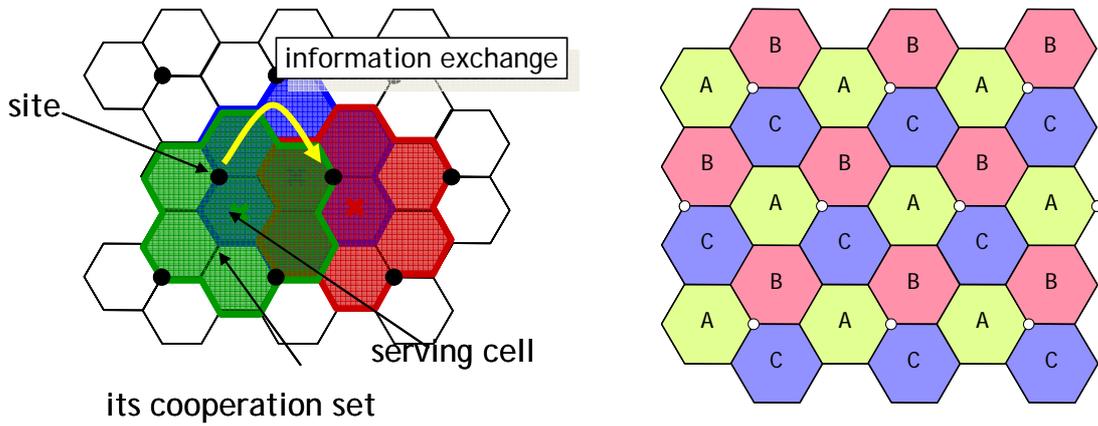
### **3D Beam Coordination**

This coordination scheme requires a limited amount of scheduling information exchange among eNBs. The concept is based on an algorithm derived for horizontal beam coordination and relies on a specific UE feedback [3GPP-R1090777]. For beamforming a specific closed-loop linear precoding with a codebook with 8 entries is assumed. This is appropriate for the 4-element half-lambda antenna array. It is obvious that the appropriate selection of the codebook depends on the used antenna type.

The basic algorithm is as follows:

The UEs measure the channel from their serving cell and report the best beam index (preferred rank 1 PMI) and CQI for their serving cell. The UEs further measure the channels from a set of dominant interfering cells. The UEs report Worst Companion Indicator (WCI) PMIs and the resulting CQI improvements for the case that the WCI is not used (Delta CQI). When a UE is scheduled taking Delta CQI into account, the usage of the associated WCIs by interfering cells must be prevented. This is the actual coordination process. Interfering cells are informed about the WCIs that must not be used. Furthermore, the PMI used in the transmission to the UE and the expected gain is communicated to allow the interfering cell to consider WCI constraints of its own UEs as well as to decide whether obeying to a given WCI constraint at all.

To guarantee a fair distribution of constraints, i.e. to define which eNB may put constraints to others, and which eNB has to accept constraints, a cyclically prioritized scheduling scheme is applied (see Figure 4-29).



**Figure 4-29: Coordination area of a cell, and prioritization of cells**

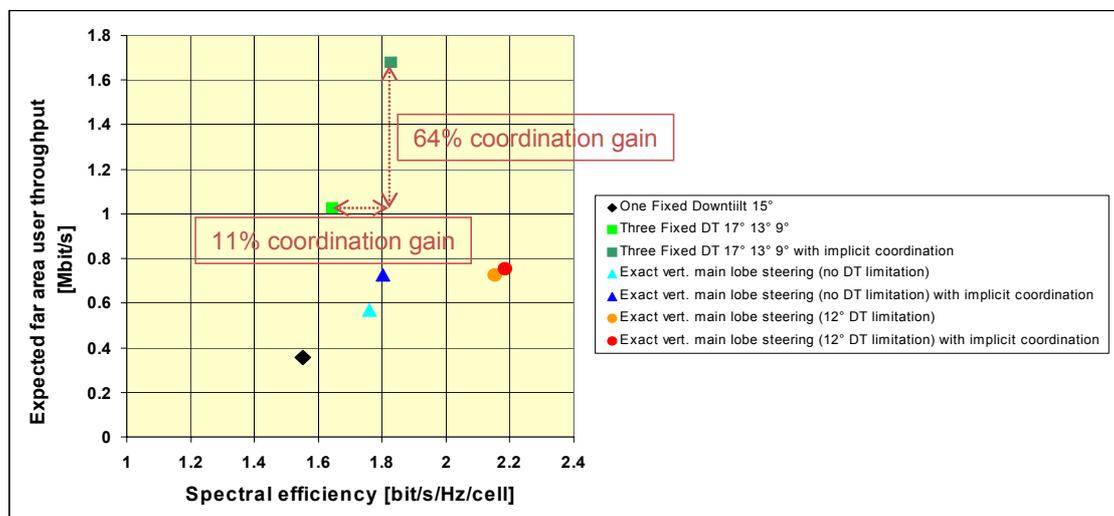
For each cell a coordination area is defined, comprising all cells around it. Each cell's scheduling takes into account only scheduling decisions of cells that fall within their respective coordination area. Conflicting scheduling decisions are prevented through prioritization that serializes scheduling within a coordination area. Depending on the priority A, B or C of the serving cell, it either schedules without constraints and puts constraints on cells with priority B and C, respects constraints from cells with priority A and puts constraints on cells with priority C, or respects all constraints from cells with priority A and B. Fairness is restored by cyclically shifting priority to cell assignments over time and/or frequency, i.e. each cell alternately obtains priorities A, B and C. Within these constraints, each scheduler assigns the resources to the UEs according to a proportional fair scheduling.

A first approach for taking vertical beamforming into account selects a downtilt per UE, which can be based on location and/or other classification criteria (e.g. path loss). The feedback information according to the coordination algorithm is then generated for this assigned downtilt. As in the original algorithm, each scheduler takes into account a configurable number of constraints (e.g. up to 3) for its scheduling decision.

### Performance results and future steps

#### Implicit Coordination

For a specific evaluation scenario some first results are available. In each of the 3 areas per cell 4 UEs are located. The horizontal pattern is a sector pattern, so only the downtilt variation is influencing the results. In Figure 4-30 the spectral efficiency versus the expected throughput of the far area, which not necessarily complies with cell edge user throughput as per definition, is shown. Two different realization options are compared. The case of three fixed downtilts assumes 17° downtilt for the near area, 13° downtilt for the center area, and 9° downtilt for the far area. When using the exact vertical main lobe steering to the UE, either no limitation or a limitation of the smallest downtilt to 12° is applied.



**Figure 4-30: Performance results for implicit coordination**

The case of three fixed downtilts shows best performance gain between the uncoordinated case and implicit coordination. In spectral efficiency an improvement of 11% is achieved, whereas in far area throughput 64% gain results. The same basic behaviour, but with lower gains, can be seen for the exact vertical main lobe steering to the UE with and without limitation of the smallest downtilt. All cases show better performance than the baseline performance with one fixed downtilt of 15°.

### **3D Beam Coordination**

The proposed algorithm has been simulated with a hexagonal deployment with 21 sites. At the eNBs a linear codebook based precoding with 8 beams is assumed, the number of constraints (“worst companion” PMIs) has been set to 0 (no coordination) up to 3. Two different implementation options for vertical beamforming have been investigated: two fixed downtilts assigned to a near and a far area, and exact main lobe steering with a limitation of downtilt to > 11°. The other simulation parameters are given in Table 4-7.

**Table 4-7: Simulation parameters**

Simulation Parameter	Value
Channel Model	3GPP case 1 (SCME, 3D antenna model)
ISD (inter-site distance)	500 m
Velocity	3 km/h
eNB antenna	4 antennas, 0.5 $\lambda$ spacing
UE antenna	1 antenna
Channel estimation	ideal
System bandwidth	10 MHz
Duplex method	FDD
Traffic model	Full buffer
Number of cells	21 (7 sites with 3 cells each), wrap around
Number of UEs per cell	15 (average)

Some results are shown in Figure 4-31. The case with fixed downtilt of 12° and no coordination is the baseline. The performance of fixed downtilt and coordination with 3 constraints (dark green dashed curve) and the performance of two fixed downtilts without coordination (0 constraints) are almost identical and provide 6% gain in spectral efficiency or about 19% gain in cell edge throughput. If applying coordination to the case of two fixed downtilts an additional improvement according to the number of considered constraints (1, 2 or 3 constraints) is feasible. For 3 constraints (light green solid curve) an additional gain of 6.5% in spectral efficiency or 17% in cell edge throughput is achieved. A slightly higher performance is possible when using optimum downtilt to the UE (light green dashed curve).

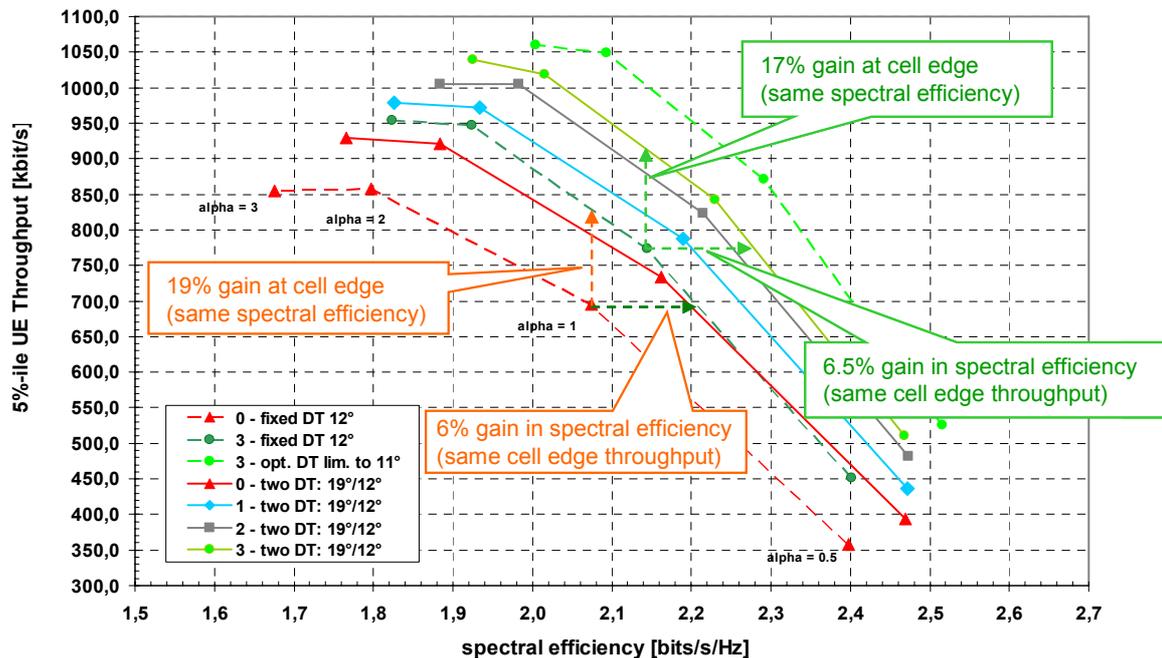


Figure 4-31: Performance of 3D beam coordination versus number of constraints.

## Conclusions and Future Steps

### Implicit coordination

Implicit coordination can provide significant performance gains, but is rather sensitive to the specific configuration and procedure. So, three fixed downtilts perform much better for far area UEs than exact main lobe steering, which primarily privileges near sector UEs, especially if downtilt limitation is applied. For further study some different alternative strategies of location based resource assignment will be investigated, which are better adapted to realistic UE distributions. Therefore also flexible configuration definitions will be assessed.

### 3D Beam Coordination

Gain of 3D beamforming without coordination and gain of pure horizontal beamforming with coordination are independent effects which can be combined. Combining both effects leads to almost an addition of the gains, even with the suboptimum algorithm approach and the simplified implementation options. So it has been proven that the gain due to 3D beamforming can be further enhanced with beam coordination. As next step, an improvement of the coordination algorithm, taking into account a full optimization over PMI and downtilt combinations, promises further gain enhancements in cell edge throughput and spectral efficiency. Also additional implementation options will be investigated.

#### 4.3.4 Coordinated scheduling based on restriction requests

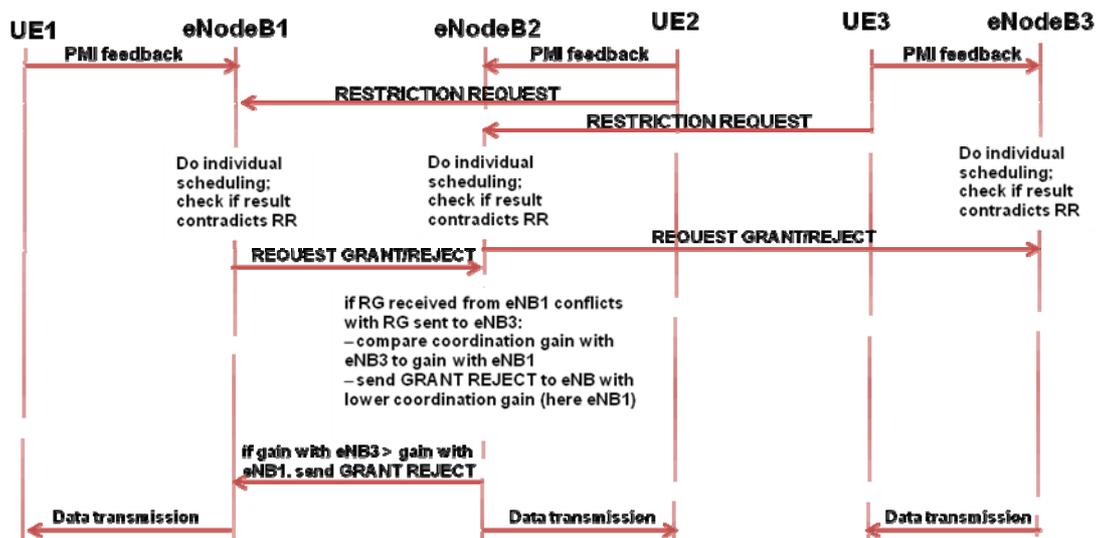
A CB scheme called Base Station Coordinated Beam Selection (BSCBS) was presented in [GA10] in order to mitigate femto cell interference on macro cell users. The obtained results were based on upper bounds referred to actual system performance and therefore did not allow predicting realistic performance gains accurately. In this contribution, the required data exchange protocol between cooperating nodes is presented together with numerical results obtained using full-blown system level simulations and without resorting to bounds. In the following, the CB algorithm is described together with details on the simulation assumptions and an evaluation of gains in an example setup.

## System description of the innovation

The main idea of the so-called BSCBS approach is to coordinate selection of beams at neighbouring cells to avoid beam ‘collision’ when two nearby cell-edge users in different cells are served using the same time-frequency resources.

In order to coordinate the selection of beams at cooperating cells, a distributed and iterative approach has been devised for networks involving an arbitrary number of base stations. The control signalling between base stations to coordinate their scheduling decisions together with the appropriate timeline has also been defined. The coordination is based on feedback from UEs including not only CQI, Rank Indicator (RI) and PMI, but also different types of additional messages to allow the cooperation. For example, the so-called RESTRICTION REQUEST feedback message from the UE contains information about those precoding matrices that, if not used at an interfering cell, would result in reduction of interference at the UE.

An example for a timeline for inter-site coordination between three eNBs, with two UEs requesting PMI restrictions at eNodeB1, is shown in Figure 4-32. The figure shows the additional feedback messages from the UEs and message exchange between eNBs at different steps of the coordination.



**Figure 4-32: Coordination timeline along with message exchange between eNBs.**

The main steps of the coordination algorithm are as follows:

- 1) As a first step, the coordinating cells make tentative scheduling decisions without considering any RESTRICTION REQUEST messages. This scheduling is then the same as it would be done in a LTE Release 8 system.
- 2) In the next step, the scheduler at each cell compares the precoding matrices identified in the received RESTRICTION REQUEST messages with the precoding matrices selected for tentatively scheduled UEs. In case of valid precoding matrix restrictions, the scheduler decides about accepting or rejecting the request after comparing a given UE utility metric. In case the precoding matrices are restricted at the cell, the UE(s) that requested the restriction in the first place would see a gain in their utility metrics. On the other hand, restriction of most suitable precoding matrices results in a utility loss for a UE served by the cell. To compute the utility gains, the required information is reported by the UE in the corresponding RESTRICTION REQUEST. Based on the result of a utility comparison, the following two cases are possible.
  - a) If the overall utility gain in the other coordinating cell is larger than the utility loss in the own cell, the cell revises its tentative scheduling decision and restricts the precoding matrices in question. It informs the serving cell of the UE from which

it received the corresponding RESTRICTION REQUEST by sending a REQUEST GRANT message.

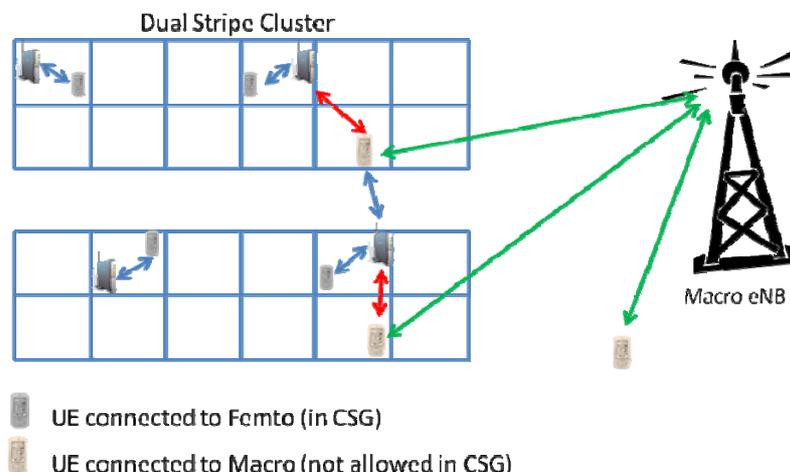
- b) If the overall utility gain is smaller than the utility loss, a REQUEST REJECT message is sent to the serving cell. In this case, the scheduler keeps its tentative scheduling decisions.
- 3) In case the REQUEST GRANT messages sent and received by a cell contradict each other, a conflict resolution is done based on a comparison of net utility gains of the contradicting grants. This can happen if a cell granted a restriction request to avoid using a specific precoding matrix and receives a grant from a neighbouring cell for a restriction request of a UE that is to be scheduled with the same precoding matrix. Depending on the result of the comparison of net utility gains, the cell either revokes its own grant and goes back to using the original scheduling decision or rejects the received grant by sending a GRANT REJECT message to the source cell.

The overall goal of the coordination approach is to maximize the sum of given utility metrics over all participating UEs in the coordinating cells. The maximization is done over different valid combinations of precoding matrices in each PRB.

The presented coordination approach is limited by the speed and payload of messages exchange between femto eNBs due to network infrastructure constraints. The delay of available CQI reports leads to uncertainties in the calculation of the utility metric.

### Performance results and future steps

Since BSCBS coordinates a single interferer that is perceived by the UE as strongest, it is clear that BSCBS is most beneficial if a small number or, in particular, a single strong interferer is present that impairs the communication link between the UE and the serving cell. Such a situation is very likely to appear in scenarios where femto cells with closed subscriber group (CSG) functionality are deployed: If a UE outside the subscriber group is in the vicinity of a femto cell that is not serving it, the interference caused by this femto cell can cause degraded SINR at the out-of-CSG macro UE and possibly lead to outage. A model for this kind of situation is a heterogeneous network with a hexagonal layout of macro cells and dual-stripe clusters of femto cells. Each femto cell serves at least a single UE within its CSG. In order to simulate the case that a macro user is in the vicinity of a femto cell, the placement of the macro users is non-uniform: 80% of the macro users are placed *within* the dual stripe cluster and 20% are placed uniformly within the cell. Therefore, a large number of UEs can be expected to be in poor interference situations. A graphic illustration of the resulting scenario where macro UEs are inside a dual stripe cluster is given in Figure 4-33. In our simulations, two different scheduling algorithms were investigated:



**Figure 4-33: Simulation scenario where users served by the macro cell are strongly impaired by CSG femto cells**

- a) The scheduler at each cell accepts a restriction request if the utility gain in the other cell is larger than the utility loss in the own cell if the request is accepted.
- b) The scheduler accepts any restriction request.

Algorithm a) is trying to maximize sum utility in both concerned cells. Algorithm b) puts more emphasis on supporting cell-edge users that do transmit restriction requests. In order to reduce scheduling complexity, only UEs that are connected to a macro cell (MUE) and are in very bad interference conditions are allowed to transmit restriction requests. The criterion used in the simulations was the geometry (sometimes termed 'wideband SINR') that an MUE sees: if it is above a threshold of -3dB, no restriction request message can be sent. Therefore, three different types of UEs are considered in the simulations:

- 1) Restriction Request (RR)-UEs are MUEs that are allowed to transmit RESTRICTION REQUESTs;
- 2) non-RR-UEs are MUEs but not allowed to transmit RESTRICTION REQUESTs;
- 3) CSG-UEs are connected to a femto cell within a CSG and are not allowed to transmit RESTRICTION REQUESTs.

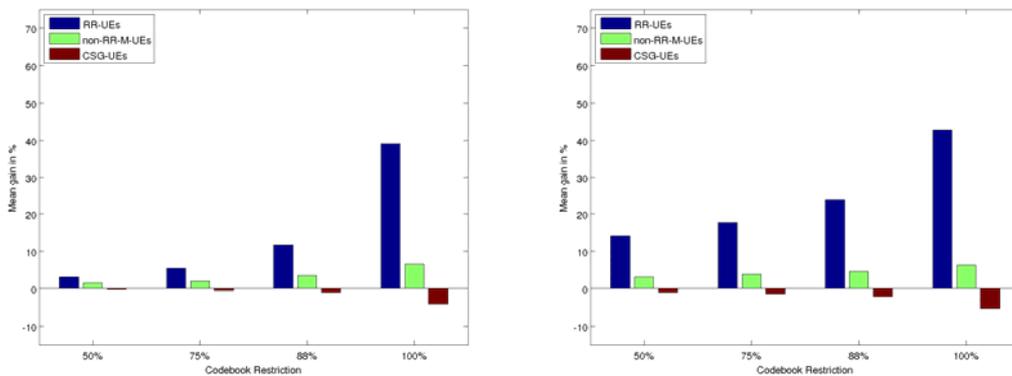
A summary of the relevant simulation assumptions is stated in Table 4-8.

**Table 4-8: Simulation parameters**

Simulation Parameter	Value
Channel Model	SCME urban macro, 1.4MHz bandwidth
Macro eNB antenna configuration	4 antennas. $\lambda/2$ spacing, vertically polarized, sectorized (3 sectors)
Macro user distribution	10 UEs per macro cell
Scheduler, traffic model	Proportional fair, full buffer
UE equalizer	MRC
Coordination threshold	Only UEs below -3dB geometry use coordination
Link adaptation	Ideal, no HARQ
CQI Feedback delay	7ms
X2 message delay	$\ll$ 1ms
Allowed codebook restriction:	Restrict 32 (50%), 48 (75%), 56 (88%) or 64(100%) PMIs from the LTE Rel. 8 codebook with 64 entries for 4Tx antennas
Femto layout	Dual Strip, 1 cluster per macro cell, urban deployment
Node type	Femto (PL and SF according to [3GPP36814]. urban deployment, closed access)
Femto antenna configuration	4 antennas, $\lambda/2$ spacing, vertically polarized, isotropic
UE antenna configuration	2 antennas, $\lambda/2$ spacing, vertically polarized, isotropic
Femto deployment ratio (probability that an apartment contains a femto)	0.05
Probability of indoor macro user (close to femto)	0.8
Femto user distribution	1 CSG user per femto

Cluster for message exchange	1 macro cell with all femto cells placed within the sector associated to that macro cell
Algorithm version	a) w/ utility comparison b) w/o utility comparison (accepting all RRs)

Performance results in terms of throughput are given in Figure 4-34 as a comparison of the mean gain with respect to the reference case for all three UE types. All figures compare the performance if 50%, 75%, 88% and 100% of the codebook of the strongest interferer is restricted.

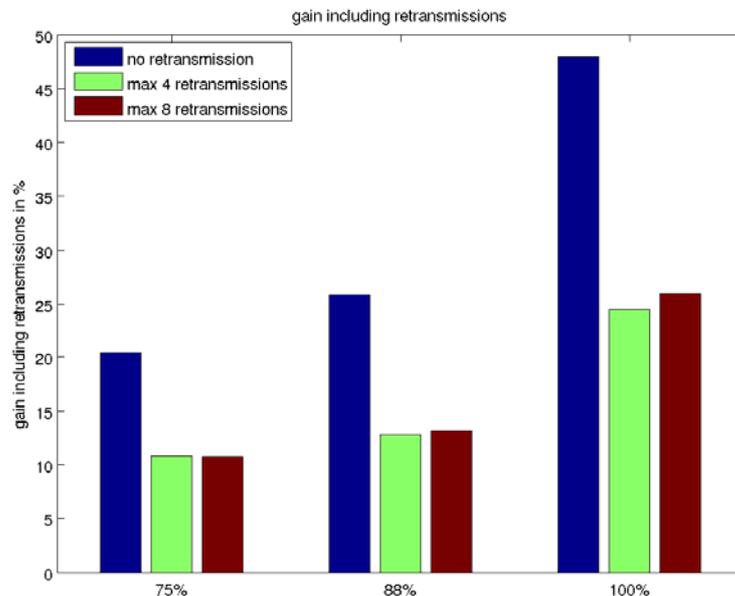


**Figure 4-34: Comparison of gains over LTE Rel. 8 for three different levels of codebook restriction. Left: Algorithm a). Right: Algorithm b)**

With respect to Figure 4-34, it is apparent that a mean gain of more than 40% is available in both algorithms for the RR-UEs if the UEs are allowed to request the strongest interferer to be silenced, i.e., to avoid 100 % of its codebook. This comes at a cost of throughput reduction for the CSG-UEs which lose on average less than 5%. This price to pay is comparatively small, because the restriction requests are only impairing the throughput of the CSG-UEs when the RR-UEs are scheduled. Since the macro cell serves 10 MUEs on average (as stated in the simulation assumptions), only a small fraction of the allocated resources to the CSG-UE is involved. The gain observed through partial restriction is significantly smaller compared to the gain due to silencing, but also leads to smaller losses in the femto cells. It is interesting to note that also the non-RR-UEs connected to a macro node experience a gain when RR are transmitted to the femto cells. If a femto cell is not transmitting at all due to granting an RR, the interference received by non-RR-UEs from these cells is reduced as well.

Moreover, it is apparent from the results that partial restriction is more effective to raise the throughput for RR-UEs using algorithm b) where all restriction requests are accepted, albeit at the cost of slightly higher throughput reduction for the CSG-UEs. It is a matter of future research to find out under which circumstances algorithms a) and b) should be chosen for best performance experience.

The results above were obtained based on the assumption that no HARQ method is used. Thus, when deciding on granting or rejecting a RESTRICTION\_REQUEST, no scheduling rule on how to prioritize possible retransmissions needed to be modelled. However, any practical implementation will require the usage of retransmissions. As a first step towards this goal, we implemented the scheduling rule that any RESTRICTION\_REQUEST is rejected if retransmissions are pending and we allowed a variable number of maximum retransmissions. In this sense, allowing retransmission with higher priority than initial transmissions with coordinated interference had an impact on overall coordination gain. Simulation results for 0, 4 and 8 maximum retransmissions are presented in Figure 4-35 for UEs that are allowed to transmit RESTRICTION REQUESTS. The example is shown for algorithm variant b).



**Figure 4-35: Simulation results for varying number of maximum retransmissions**

In particular, the gain with respect to no coordination is illustrated for 75%, 88% and 100% codebook restriction of the coordinated interferer in order to show the impact of the algorithmic extension with retransmissions on coordination gain. Apparently, retransmission having higher priority than initial transmissions (that are possibly coordinated) reduce the gain due to coordination drastically (e.g., the gain for 75% coordination shrinks from roughly 20% to 10%). It is a matter of future research in how far more advanced scheduling criteria can be developed that avoid this gain reduction.

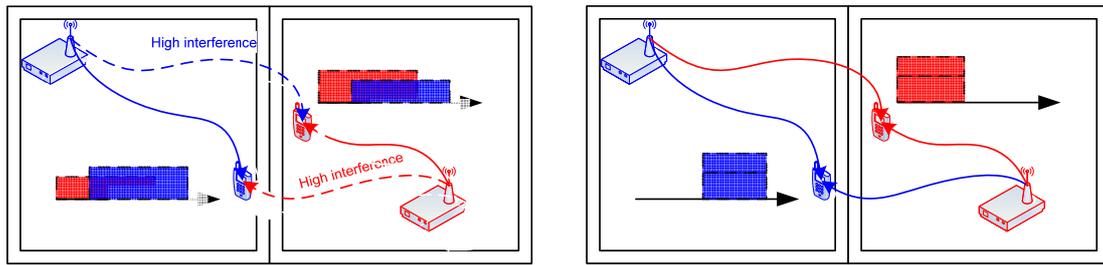
In summary, a specific downlink coordinated scheduling algorithm was presented based on LTE Rel. 8 precoding matrices. Coordination was achieved by restricting the choice of available precoding matrices at the interfering transmitters. The details of this coordination method were outlined and two variants of the algorithm were evaluated using full system-level simulations. Significant gains in the order of 40% were achievable for users that suffered from HeNB interference. Future steps include a thorough investigation of other antenna configurations, advanced scheduling criteria taking retransmissions into account and the overhead involved in the message exchange that is required to achieve the coordination within a group of transmitters. Moreover, if no X2 interface is available between involved (H)eNBs, the impact of an exchange of messages using OTA methods with a mobile station as a relay node is of high interest for practical implementations.

#### 4.3.5 Coordinated scheduling for heterogeneous deployments

In this section we will evaluate the impact of basic CoMP involving two femto base stations deployed in a campus environment, and highlight the impact of the cell load on the performance. We compare the interference nulling scheme to a joint processing scheme and illustrate that most of the gain is brought by interference nulling. A particular aspect of HeNB network is the low mobility of the UEs and the high stationary conditions of the wireless channels. A UE being in a high interference zone might experience bad performance for a long time. Thus, it is particularly relevant to set up an advanced interference avoidance scheme in that case.

#### System description of the innovation

In Figure 4-36, we observe on the left a typical downlink transmission without CoMP, where two cell edge UEs receive a signal from their serving HeNB and a high level of interference generated by the signal sent to the other UE by the other HeNB.



**Figure 4-36: DL transmission without CoMP (left) and with CoMP (right).**

We also observe on the right part of the figure that by scheduling the two UEs from the two different cells with disjoint resource allocation, the SINR observed at each UE is drastically improved. Indeed, the main (and most often only) interference contribution is shut down to zero. Furthermore, the same or additional data can be sent from the other HeNB to further improve the signal strength or multiplexing gain. In some cases, it is better to transmit on less resource with a better SINR than transmitting on more resource with a very low SINR. Furthermore, the lower the SINR is, the higher the improvement by reducing the interference level is. Thus, CoMP techniques mainly target cell-edge throughput improvement.

A CoMP UE  $k$  will experience an average spectral efficiency  $S_k$  which can be expressed as

$$S_k = \beta f(\rho_1, \rho_2)$$

where  $\rho_1$  denotes the SINR for the HeNB1-to-UE link, and where  $\rho_2$  denotes the SINR for the HeNB2-to-UE link. If the  $k$ -th UE is not a CoMP UE, then  $\rho_2=0$ . Finally,  $\beta$  is a scaling factor depending on the CoMP-cost sharing strategy in the cell. This variable also depends on the cell load  $CL$ , which is the number of scheduled resource with respect to the overall resource amount (i.e., a load of  $CL=50\%$  means that only 1 PRB out of 2 actually carries information). We consider two strategies:

**The resource dedicated for CoMP is limited to the free resource in the cell:**

In that case, the alternative HeNB2 indicates to the anchor HeNB1 the amount of free resource  $(1-CL)$ , and only the UEs taking the best benefit from CoMP are selected. This technique allows not to degrade the performance of the alternative HeNB2 UEs ( $\beta=1$ ). However, if one cell load is 100%, no cooperation is possible with that HeNB.

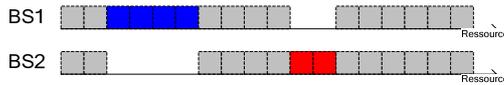
**The cost of resource needed for CoMP is shared between the cell UEs:**

In that case, all UEs taking benefit from CoMP are helped by the alternative HeNB2. Thus, if  $\gamma$  is the portion of resource needed for cooperation purpose in a given cell, then

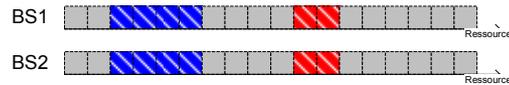
$$\beta = \min(1, 1/(\gamma + CL)),$$

which means that the CoMP cost is divided between all UEs of the cell. Thus, if the free resource is not sufficient for supporting CoMP, all UEs of the alternative cell are impacted. Since this is a cooperative exchange between neighbouring HeNBs, this will mainly impact high spectral efficiency users, as all cell-edge users of both cells take benefit from CoMP. Other strategies could be studied such as only sharing the cost of CoMP among CoMP users.

We consider two basic CoMP schemes: interference nulling and joint superimposed transmission. The last scheme needs the User Plane exchange between two cooperating nodes.



**Figure 4-37: CoMP Interference nulling**



**Figure 4-38: CoMP Joint superimposed transmission**

**Interference nulling:**

Two HeNBs exchange scheduling information on CoMP UEs, such that no data transmission is scheduled in the second HeNB on the same resource as a CoMP UE of the first HeNB. The peak data rate of CoMP UEs is not improved, but the overhead of the technique is low. The interference nulling is illustrated in Figure 4-37. Let  $P_1$  and  $P_2$  denote the signal power received from HeNB1 and HeNB2, respectively. Let  $P_i$  be the sum of noise and interference from other HeNBs. Let  $s(SINR)$  denote the average spectral efficiency for a HeNB to UE downlink transmission at a given SINR. Thus, a UE can take benefit from joint scheduling as soon as

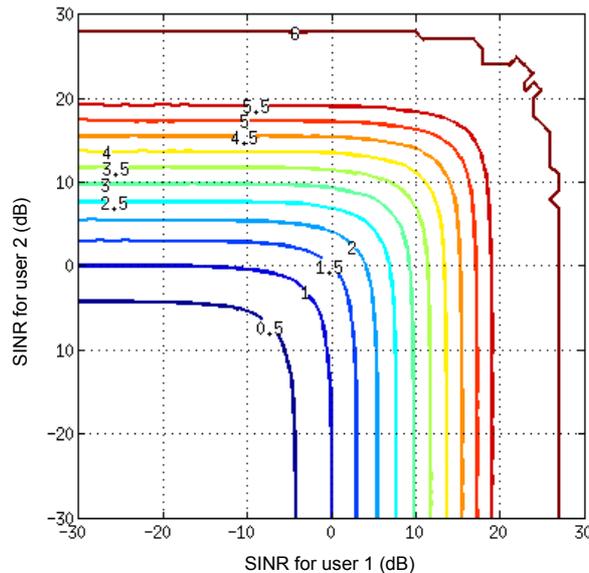
$$s(P_1 / (P_2 + P_i)) < \beta \cdot s(P_1 / P_i).$$

This implies that if the resource dedicated for CoMP is limited to the free resource in the cell, a UE always take benefit from joint scheduling.

**Joint superimposed transmission:**

Two HeNBs exchange scheduling information and data for CoMP UEs, such that the same data transmission is scheduled in the second HeNB on the same resource as a CoMP UE of the first HeNB. The peak data rate of CoMP UEs is not improved and the overhead of the technique is higher than for the joint scheduling case since user plane exchange is needed between two neighbouring nodes. This is illustrated in Figure 4-38. Define  $SINR_1 = P_1 / P_i$  and  $SINR_2 = P_2 / P_i$ . Let  $f_s(SINR_1, SINR_2)$  denote the average spectral efficiency for a superimposed transmission from two HeNBs to one UE, as illustrated in Figure 4-39. We assume that the scheduling is done on the  $SINR_1$  basis, thus, the second link does not take benefit from the whole frequency diversity order. Thus, a UE can take benefit from joint superimposed transmission as soon as

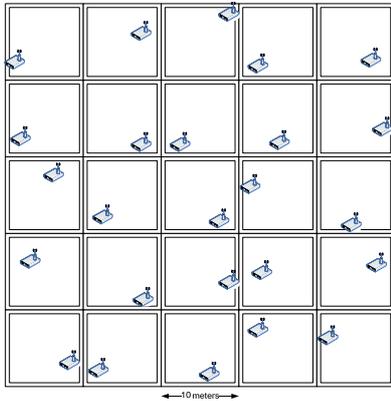
$$s\left(\frac{P_1}{P_2 + P_i}\right) < \beta f_s\left(\frac{P_1}{P_i}, \frac{P_2}{P_i}\right).$$



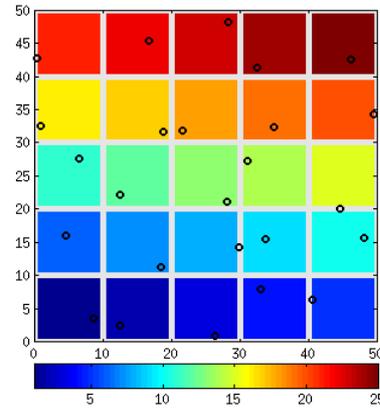
**Figure 4-39: Spectral efficiency of CoMP with joint superimposed transmission.**

### Performance results and future steps

We consider a deployment of HeNBs in a set of 10m x 10m apartments placed in grid, as illustrated in Figure 4-40. For the sake of illustration, each HeNB is allocated a colour as in Figure 4-41. We consider the parameters listed in Table 4-9 for system level evaluations.



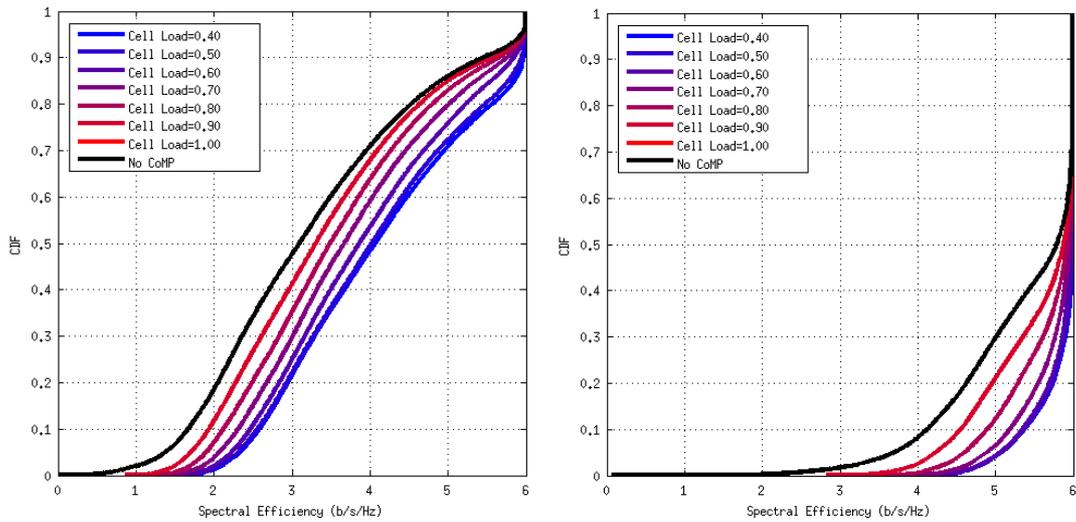
**Figure 4-40: Example of deployment of HeNBs in a 5x5 grid of 10mx10m apartments.**



**Figure 4-41: Cell Identity represented as colours, for the connectivity and cooperation graph**

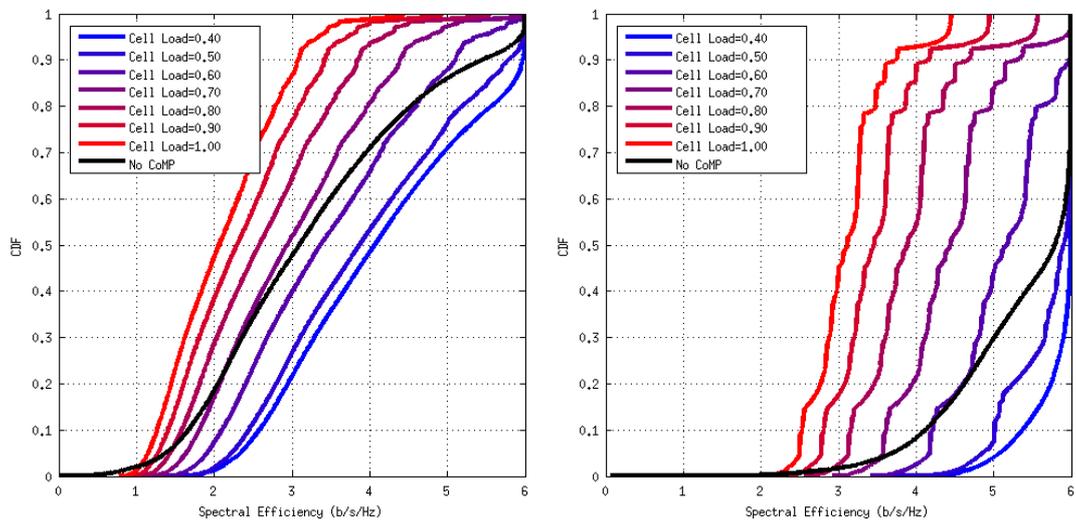
**Table 4-9: Simulation parameters**

Simulation Parameter	Value
<b>HeNB Parameters</b>	One sector. Each UE lying in the same square of 10mx10m as the HeNB is attached to the HeNB. 25 HeNBs are deployed in a 5x5 grid. The antenna has a gain of 5dB with omni directional diagram and is at a height of 1.5m. The transmit power is 20 dBm
<b>UEs Parameters</b>	Deployment by space quantization (evaluate the spectral efficiency on each place of the cell) (100 000 UEs are uniformly sampled in the deployment grid). The antenna gain is 0dB, with an omni directional diagram, at a height 1.5m, with 2 receive antennas. The thermal noise power of -112.5 dBm. The UE to HeNB attachment is based on a geographical criterion..
<b>Propagation models</b>	No Shadowing, Motley Keenan path loss model with a wall attenuation of 10 dB. Multipath channel= InH 17-tap indoor channel with 0,225 us delay spread. 5 MHz bandwidth OFDM channel.
<b>PHY and MAC layers models</b>	One PRB per slot is allocated to one UE, and perfect resource fairness is experienced by the UEs. Each UE is always scheduled on the best PRB. We consider for a given SNR the spectral efficiency $s(SNR)$ such that $s(SNR) = \max_{R_c, m} (R_c \cdot m (1 - P_{out}(R_c, SNR, m))) \text{ where } 0 < R_c < 1 \text{ and } m = \{2, 4, 6\}$ and where $P_{out}(R_c, SNR, m)$ is the $2^m$ -QAM-input outage probability for a data rate $R_c \cdot m$ . This metric is related to the performance with ARQ.



**Figure 4-42: Spectral Efficiency CDF with DL CS CoMP-Joint scheduling with interference nulling (Left: Wall Attenuation=10dB, Right: Wall Attenuation=20dB). The resource dedicated for CoMP is limited to the free resource in the cell.**

The CDF of the average per-UE spectral efficiency with joint scheduling when the resource dedicated for CoMP is limited to the free resource in the cell, and with variable cell load and wall attenuation, are illustrated in Figure 4-42. We observe that in any case, the spectral efficiency is improved (as expected). The gain is more important when the wall attenuation is lower, as more interference is experienced. When the cell load tends to 1, CoMP is almost not used and we observe the poor gain on the CDF.

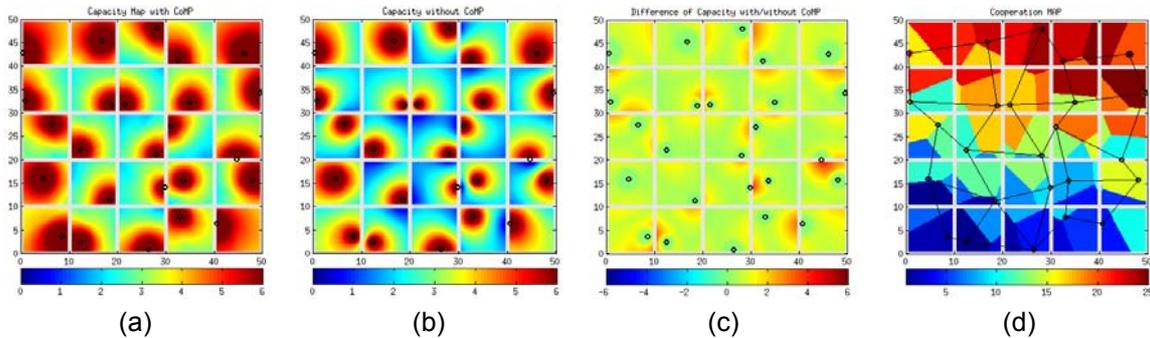


**Figure 4-43: Spectral Efficiency CDF with DL CoMP-Joint scheduling (Left: Wall Attenuation=10dB, Right: Wall Attenuation=20dB). The cost of resource needed for CoMP is shared between the cell UEs.**

The CDFs of the average per-UE spectral efficiency with joint scheduling when the cost of resource needed for CoMP is shared between the cell UEs, and with variable cell load and wall attenuation, are illustrated in Figure 4-43. Only cell edge UEs (taking benefit from CoMP) see a performance improvement as the gain brought by CoMP is higher than the reduction due to the shared cost of CoMP. The performance of other UEs is degraded.

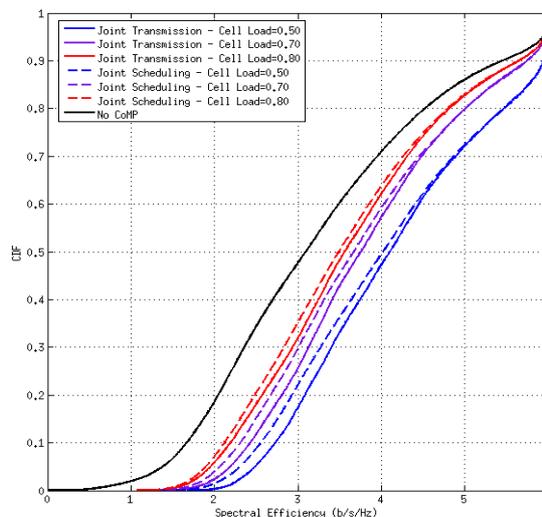
In Figure 4-44, we observe several maps obtained with DL CoMP-Joint scheduling when the resource dedicated for CoMP is limited to the free resource in the cell, for a wall attenuation of 10dB and for a cell load of 50%. Figure 4-44 (a) illustrates the coverage map with DL CoMP-Joint scheduling. Figure 4-44 (b) illustrates the coverage map with no DL CoMP. Figure 4-44 (c) illustrates the map of the difference of the spectral efficiency with and without CoMP. We clearly see that up to 4 b/s/Hz gain can be observed for highly interfered cell-edge UEs, such as the (45,20) region where a neighbouring HeNB is lying just behind the wall. Figure 4-44 (d) illustrates the connectivity and cooperation graph. The edge between the nodes illustrates that at least one UE is using CoMP between the two nodes. The colour of a point on the map illustrates which HeNB is the alternative node, the colour code being illustrated in Figure 4-41. If a UE does not use CoMP, the colour of its anchor node is used.

From the cooperation map, we see that each HeNB has to setup X2 connections with at maximum 4 of its neighbours in order to guarantee CoMP coverage for its UEs in the apartment. Straightforward optimization can be done by evaluating the proportion of the test apartment surface associated to each possible CoMP connection and maintaining the X2 connections that cover proportion of the surface of the test apartment.



**Figure 4-44: Coverage maps improvement with DL CoMP-Joint scheduling. The resource dedicated for CoMP is limited to the free resource in the cell.**

The CDFs of the average per-UE spectral efficiency are illustrated in Figure 4-45 with joint superimposed transmission when the resource dedicated for CoMP is limited to the free resource in the cell with variable cell load and wall attenuation. We observe that cell-edge throughput is improved by the joint transmission with respect to joint scheduling, as expected. However, the improvement is limited with respect to the overhead increase due to User Plane forwarding through a X2 cooperation channel.



**Figure 4-45: Spectral Efficiency CDF comparison of joint scheduling with interference nulling and joint superimposed transmission. The resource dedicated for CoMP is limited to the free resource in the cell, the wall attenuation is 10dB.**

Other simulations would show that only few users take benefit from CoMP-orthogonal repetition and CoMP-joint orthogonal transmission. The orthogonal repetition or orthogonal transmission techniques has the important advantage to receiving the sent symbol on orthogonal resource, which allows a constructive combination of the SINRs (in the joint superimposed transmission, the signals are not added coherently, and a power loss is observed). Unfortunately, because of the resource wasting, they are highly non-efficient from a system point of view. However, they lead us to a more efficient solution considering the multiple user dimensions and the variance of the user performance for cancelling the drawbacks of orthogonal transmission techniques.

We have shown the improvement of femto cell–edge for a campus deployment using the most basic CoMP technologies. At this step of the analysis, most of the gain is provided by the joint scheduling with interference nulling approach.

#### 4.4 Scheduling for joint processing

By exploiting the multi-user diversity dimension, coordinated or non coordinated multi-cell scheduling techniques have the potential of having a strong impact on the distribution of the inter-cell interference seen by the physical layer. Since the spatial beamforming methods studied in WP1 also tackle the avoidance of interference, it is essential to study the combined effects of resource allocation and spatial processing (e.g. JP CoMP). In this section we look at the combination of scheduling and JP CoMP and address some issues related to scheduling in the context of the Partial CoMP presented earlier in this document.

##### 4.4.1 Impact of scheduling on the performance of downlink multicell processing

The rate increase from using JP-CoMP compared to the single cell processing comes from the reduction of the interference received at the UEs from the eNBs in the cooperation cluster and from the beamforming gain. For a given choice of UEs, this rate increase is easily quantified and has been studied in many works, first with perfect CSI, then with a level of detail in realistic modelling of the CSI error. In practical systems, a scheduler is used to select the UE to transmit to and chooses the UE so as to improve some figure of merit. This modifies the distribution of the effective channel and of the effective interference at the scheduled UEs. For example, distributed scheduling based on the SINR maximization, is known to be a good way to mitigate the effects of interference. 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 *max-SINR scheduler* is the same as the scaling law of an upper bound corresponding to a *max-SNR scheduler* with no interfering cell. Mathematically, it means that the average rates achieved in the no-interference upper bound is equivalent to the average rate achieved with the *max-SINR scheduler* when the number of UEs per cell grows large, i.e.,  $E[R_{\max-SINR}] \sim E[R_{no-int\,erf}] \sim f(n)$ , where  $n$  is the number of UEs per cell. Therefore, if  $n$  tends to infinity, the distributed *max-SINR scheduler* leads to a transmission scheme achieving most of the rate obtained in the no-interference upper bound with no-interfering cell, without the need for more advanced MIMO processing. Therefore, it leads us to think that the impact of JP-CoMP would be much reduced when used in addition to a distributed scheduler with many UEs.

Still, some issues remain to be considered before being able to confirm such an affirmation. First, the results are for an infinite number of UEs and state nothing about the speed of convergence or about the performances at realistic number of UEs. Second, the scaling of the rate in terms of the number of UEs is not accurate enough and a more precise description of the average rate needs to be derived. Indeed, the scaling law gives only an asymptotic equivalent but any finite difference, or even any difference growing to infinity at a lower rate than the first term is neglected. Finally, no fairness is considered when using the *max-SINR scheduler* and only UEs located very close to the eNB are scheduled. This is not acceptable in a cellular network, where the operator cannot serve only the UEs around the eNBs and some fairness considerations have to come into play.

We will consider these issues in the following analysis. Particularly, we will compare the performances of JP-CoMP and single cell processing, both applied after a distributed scheduler. We will also consider different distributed schedulers and discuss the impact of the scheduler's choice, particularly on fairness.

## System description of the innovation

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  $\mathbf{H}_{comp}^{(k)} \in \mathbb{C}^{n_{comp} \times n_{comp}}$  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

$$k_{\max-SINR,j}^* = \arg \max_{k \in \{1, \dots, K\}} SINR_j(k) = \arg \max_{k \in \{1, \dots, K\}} \frac{|h_{jj}^{(k)}|^2}{\sigma^2 + I_{int,j}^{(k)} + I_{ext,j}^{(k)}}, \forall j \in \wp.$$

The sum rate is then computed by summing the rates achieved by the scheduled UEs across the cells of the cooperation cluster:

$$R_{\max-SINR} = \sum_{j \in \wp} \log(1 + SINR_j(k_{\max-SINR,j}^*))$$

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 zero-forcing 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

$$k_{comp, j}^* = \arg \max_{k \in \{1, \dots, K\}} SINR_{j, comp}(k) = \arg \max_{k \in \{1, \dots, K\}} \frac{|h_{jj}^{(k)}|^2}{\sigma^2 + I_{ext, j}^{(k)}}, \forall j \in \mathcal{J}.$$

Once the UEs being scheduled are chosen via the distributed scheduler, the matrix

$\mathbf{H}_{comp}^{(k_{comp}^*)} \in \mathbb{C}^{n_{comp} \times n_{comp}}$  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 fulfil 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 interferences.

#### 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*).

#### Performance results and future steps

The parameters of the simulations are given in Table 4-10 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.

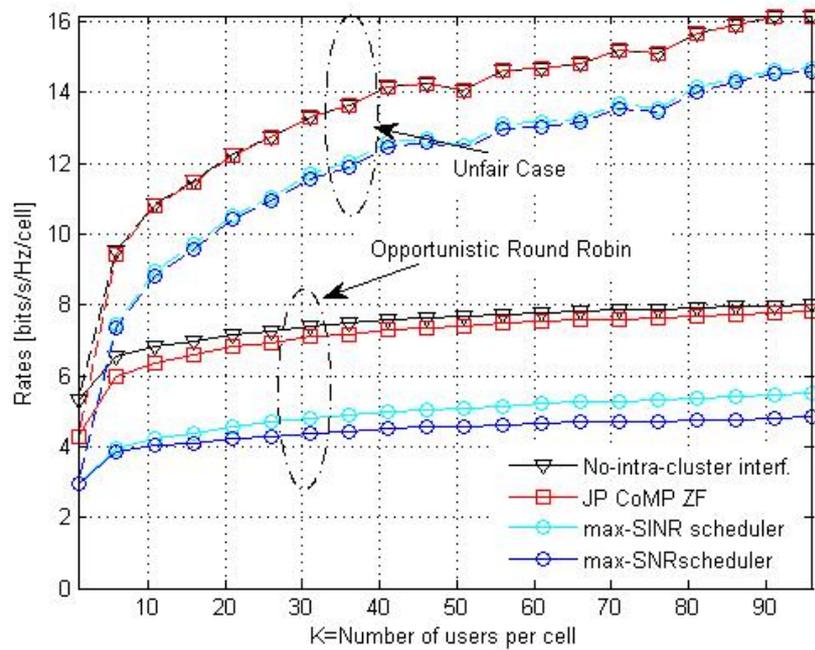
**Table 4-10: Simulation parameters**

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	$n_{comp}=3$

In Figure 4-46, the average rate is plotted for  $K$  time slots, so that each UE can be served when the ORR algorithms are used. Moreover, the simulations are averaged over 100 Monte-Carlo realizations, i.e., 100 random generations of the  $K$  UEs inside the parts of the cell considered.

We observe that for both the unfair and the ORR algorithms, JP-CoMP ZF achieves an average rate very close to the average no-intra-interference rate, while the single cell processing schemes with distributed schedulers achieve significantly lower rates. Even though it is proven that the rate obtained with the max-SINR scheduler converges to the no-intra interference rate when the number of UEs tends to infinity, the rate difference between the two remains significant even for a large but finite number of UEs. Actually, the proof was for the case of equal path loss between the UEs, but intuitively it should also hold for a fair opportunistic scheduler. Indeed, it seems that the difference decreases (very slowly), as the number of UEs increases.

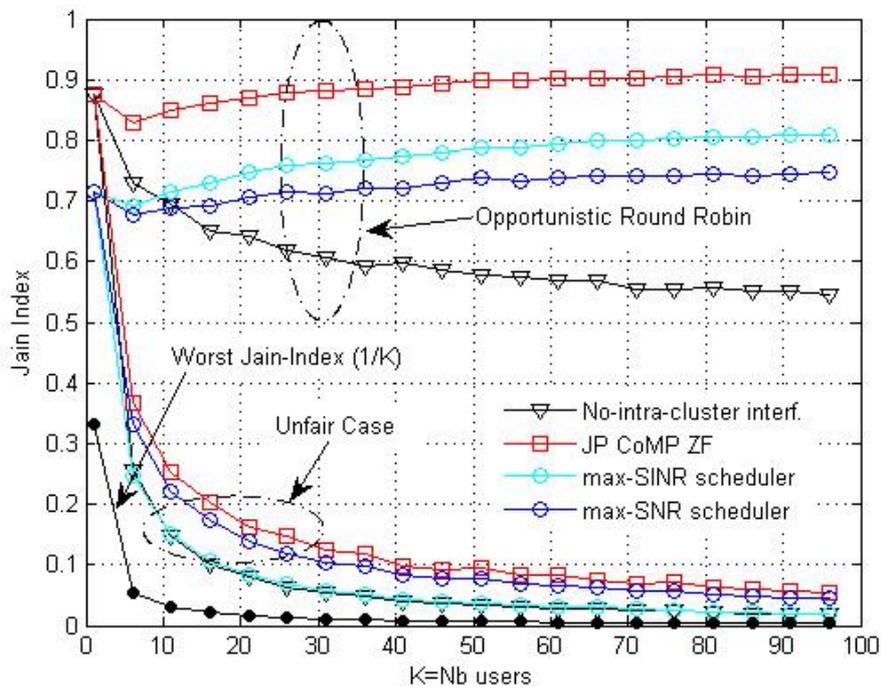
We also note that the *max-SNR scheduler* introduces very little loss compared to the *max-SINR scheduler* in both cases. However, the difference is much smaller in the case without fairness. In that case, it seems that the two distributed schedulers have the same scaling in terms of the number of UEs. This is to put in relation to the fact that the scheduled UEs are located very close to the eNBs, if no fairness is considered, so that the interference power is very small. It also explains why the difference between JP-CoMP and the single cell transmission is more significant (particularly in relative difference) with the ORR algorithms where the scheduled UEs are located in average further away from the eNB.



**Figure 4-46: Average rate per cell and time slot.**

This leads us to the more detailed discussion of the fairness between the UEs. We have plotted in Figure 4-47 the Jain index obtained for the unfair and the ORR algorithms. We can see that the Jain indices obtained with the unfair algorithms tend to zero as the number of UEs increases, which is very unfair. We have plotted the curve representing the worst Jain index possible and we can see that the Jain indices for the unfair algorithms are close to it, which is completely unacceptable in a cellular mobile network.

On the opposite, the ORR algorithms achieve Jain indices between 0.7 and 0.95, which is much better. Finally, we can also observe that the JP-CoMP ZF achieves a better Jain index in both cases, which is a consequence of the better rate achieved by the cell edge UEs.



**Figure 4-47: Average Jain index per cell and time slot.**

With our simulations, we have analysed in a realistic environment the impact of the scheduler on the improvement brought by doing JP-CoMP zero-forcing with water-filling. We have first considered algorithms maximizing only the average performance, which have proven to be very unfair, and then Opportunistic Round Robin versions of these algorithms. The ORR algorithms achieve a good fairness between the UEs at the cost of significant rate losses compared to the previous unfair algorithms. However, it has to be kept in mind that the unfair algorithms serve only the UEs located extremely close to the eNBs, which is completely unacceptable and it is obvious that it will be impossible to achieve a reasonable fairness without significant rate losses compared to these algorithms. Thus, the ORR versions appear quite attractive, with quite good performance and good fairness between the UEs.

We can observe that in spite of the asymptotic theoretic analysis, JP-CoMP brings significant improvement for the average cell capacity, and increases particularly the rate achieved by the cell edge UEs. The difference is higher when some fairness is considered because of the relative stronger interference which is a consequence of the scheduled UEs being distributed uniformly in the cell instead of being located very close to the eNB. This is interesting since it will clearly be impossible to neglect the fairness considerations and serve only the UEs close to the eNB. As a conclusion we can state that the interference cannot be managed only by scheduling with a realistic number of UEs per cell, and JP-CoMP (or other interference reducing methods) needs to be used. The use of a distributed scheduler to select the UEs has no significant effect on the increase brought by using JP-CoMP, which remains important even with a large number of UEs per cell.

#### 4.4.2 Scheduling Aspects of Partial CoMP

The *partial* CoMP concept as an advanced clustering concept has been introduced in detail in chapter 4.1.2 in a comprehensive manner including already some first scheduling aspects. Hence the concept should not be reintroduced once again in this chapter. Instead a short list of specific scheduling aspects with respect to *partial* CoMP will be shortly addressed.

#### System description of the innovation

The partial CoMP concept has been introduced in section 4.1.2. From the high level block diagram of a *partial* CoMP scheduler in Figure 4-7 of section 4.1.2 it should be clear that there

are some specific issues compared to a conventional scheduler. For example RSRP measurements as reported from UEs for different cells have to be used to estimate the best fitting cover shift per UE and to group the UEs accordingly, which is a challenge due to possibly inaccurate RSRP measurements.

Before this can be done, the setup of cover shifts comprising either certain resources like frequency subbands – i.e. set of PRBs –, subframes or a combination thereof is needed. To work properly this setup has to be network wide or at least for a whole hot spot area using JP CoMP. Therefore the simplest way is probably to allocate to each cover shift the same amount of resources, either statically or maybe semi statically, as reallocation of cover shifts over large areas will be difficult to handle. Anyway optimization to specific load conditions in one area might easily lead to degradations in other areas.

In addition the aspect of unequal load conditions for different cover shifts has been already investigated in chapter 4.1.2. and it has been demonstrated that a high number of UEs can be scheduled into different or even any cover shift allowing to balance resource usage.

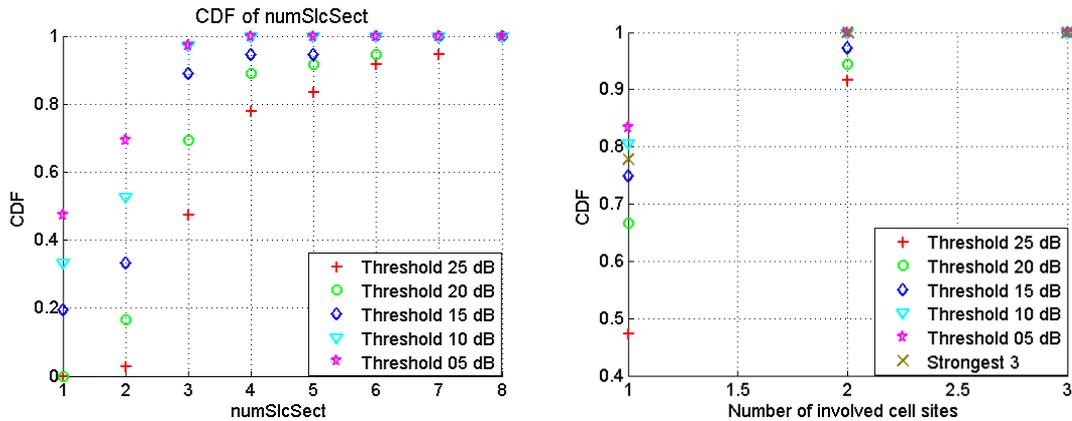
After UEs have been allocated to specific cover shifts the UEs per cover shift have to be scheduled for simultaneous data transmission. This means finding the best set of UEs per PRB providing the highest overall throughput. Unfortunately this is a non-convex optimization problem needing some heuristics. The partial CoMP aspect is that UEs per cover shift are already a reduced number of UEs so that potential multi user scheduling gains for the full pool of all UEs might be missed. Further in case of using frequency subbands as cover shift resources the bandwidth – and as a result the frequency aware scheduling gains – might be reduced. For overall large bandwidth this effect might be small nonetheless with respect to this time domain allocation of cover shifts, it seems to be superior, specifically as conventional schedulers include a time domain scheduler anyway. The increased delay of few ms for e.g. 6 cover shifts for the time domain scheduler is probably tolerable, but a combination of time and frequency might be even the best choice.

The user grouping itself per cover shift is quite similar to conventional JP CoMP schemes. The main issue is to find sets of mutually orthogonal UEs to avoid strong power rise for the zero forcing like precoders  $\mathbf{W}$ . The main benefit of *partial CoMP* is that almost all UEs can be used for cooperation and the reason is that almost all UEs are served user centric ensuring cooperation gains at least in case of well-conditioned precoding matrices.

### Performance results and future steps

As mentioned already in chapter 4.1.2 the freedom of the scheduler to put certain UEs into more than one cover shift – either a frequency subband or a subframe – depends on the number of involved sites. For example UEs with their 3 strongest cells belonging to the same site (intra eNB cooperation) can be scheduled into any cover shift and are for that reason very valuable to overcome unequal load conditions between cover shifts. To get a feeling about the typical percentages of UEs connected to one, two or three sites SL simulation have been analyzed for the scenario as described in 4.1.2.

Reporting thresholds TH have been varied between 5 to 25dB, i.e. a UE reports all cells with a RSRP value higher than the threshold TH, where TH is defined with respect to the strongest cell RSRP value. Additionally reporting of the 3 strongest interfering cells has been simulated as baseline scheme. From Figure 4-48 it can be concluded that with increasing threshold the number of involved cells and sites increase as well and in case of TH=25dB about 40 to 50% of UEs are served by only 1 and a further 40% by 2 sites (see the red crosses for 25 dB threshold in Figure 4-48 right for 1 site at .48 and for two sites at 0.9 of the CDF → 48% of UEs served from one site and 0.9-0.48=40% served from 2 sites). This result is promising and provides quite a high degree of scheduling freedom at least for the investigated scenario.



**Figure 4-48: Number of involved cells (left) and sites (right) for the partial CoMP concept.**

For the time being the work concentrates on analysis of *partial CoMP* performance itself, while more detailed investigations of scheduling issues and optimum user grouping per cover shift will follow later on. For the time being it is important to notice that despite the separation of the resources into several disjoint cover shifts, the scheduling flexibility seems to be only marginally reduced compared to a single cell scheduler.

The first analysis has been limited to full buffer traffic, but for meaningful performance results it would be of great interest to investigate more realistic scenarios with different traffic mixes. Other issues might be: a) better understanding of backhaul traffic overhead and inter node connections, b) optimized schemes for distribution of user data for joint precoding per cover shift over several sites, etc.

#### 4.5 Game theory based scheduling

Sharing a common pool of frequency resources among network nodes characterized by fluctuating levels of the channel quality, requests of services, and traffic, allows for an effective exploitation of the large benefits promised by user, space, and time diversity. *Cross-layer design* approaches benefit at the best from such different kinds of diversity since it jointly optimizes their use. CoMP schemes offer a technical framework/architecture to support such a resource sharing. However, the provided coordination has a cost which could be very relevant in terms of additional frequency band required for information exchange among multiple nodes. Cross layer design, requiring joint optimization, has also a relevant cost in terms of computational complexity.

Cross-layer design approaches have been studied in single-cell systems, i.e. for networks with frequency reuse where the system resources 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 the available results to a CoMP system would require a complete sharing of data among coordinated nodes on the data plane and additional data exchange on the control plane. This later information sharing implies exchange of instantaneous information about both the instantaneous CSI at the transmitters and instantaneous occupation of the queues. This exchange grows exponentially with the number of coordinated nodes. Additionally, optimum available control mechanisms based on cross-layer design such as joint rate and power allocation, scheduling and admission control have a complexity which scales exponentially in the number of coordinated nodes. Thus, CoMP schemes based on full sharing at the level of data and control planes is intrinsically characterized by a limited scalability. Our objective is to investigate the trade-off between performance and level of shared (data and control) information in a cross layer design framework for CoMP.

In order to design and analyse cross-layer design algorithms with a low level of coordination among nodes and then a higher level of scalability, we adopt a *game theoretical approach*. Game theory is a mathematical framework for multiple decision 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 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 peer 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.

A specific branch of game theory 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. Since we are interested in defining algorithms for system with distributed intelligence among nodes and limited amount of CSI, Bayesian game theory offers the most suitable game theoretical framework to design cross-layer algorithms based only on locally available measurements.

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

We aim to study cross-layer design approaches for joint power and rate allocation, packet transmission scheduling, and packet admission control in a CoMP system with different levels of coordination and information sharing among multiple points. Note that the kind of cross layer design presented in this section involves different kinds of control functions compared to the cross layer design scheduling considered in the previous sections. There, the key functionality to be coordinated was beamforming. The scheduling aimed at selecting the receiving nodes which minimized the inter-beam interference thanks to the characteristics of their physical channel. The queue of each receiver was supposed to be saturated, i.e. there were always data available for transmission toward the considered receivers. In the framework considered in this section the traffic is at burst and the queues may be empty. Then, the scheduling in this section accounts also for the state of the queues.

In the system at hand, diversity throughout users, space, and time are realistically modelled in the system by the assumptions that (i) each user is endowed with a finite buffer queue for packet storage, which can be, eventually, empty; (ii) requests of packet transmissions arrive according to Poisson distributions; (iii) channels are varying in time according to a Markov distribution.

We assume that no exchanges occur on the data plane and a very limited amount of information is exchanged at the control plane such that very low rate channels are required among transmitters. More specifically, we assume that the transmitters share information only about statistics of their channels and incoming packets but not instantaneous information about their realizations. Thus, the only instantaneous information about the system available to each transmitter is the information that can be measured locally, i.e. the number of packets in the queue and the state of the links from the transmitters to the receivers. This last assumption

underlies the use of a protocol like TDD for which the reciprocity principle holds and the state of a link in one direction can be inferred from measurements and estimates on the opposite direction.

In this section we present algorithms based on Bayesian game theory to perform jointly rate and power allocation, scheduling and admission control with a level of coordination among nodes very low. The proposed algorithms differ for the assumption made at the receivers, which could perform single user decoding treating the received interference as Gaussian noise or successive interference cancellation decoding. In such a latter case the decodable interference is decoded and cancelled out from the received signal before decoding the information of interest. This improves the performance of the system. Furthermore, the proposed algorithms differ from each other also in the objectives of the transmitters. One class of algorithms assumes that each transmitter aims to maximize its own throughput without considering the effects of its strategy on the other receivers' performance. We dubbed this class of algorithms as non-cooperative games. Another class assumes that the transmitters aim to maximize the global throughput of the network and cooperate. They are referred to as cooperative games. This latter class of algorithms has higher complexity compared to the class of non-cooperative algorithms.

### System description of the innovation

We consider a system consisting of  $N$  source-destination pairs sharing the same frequency band. For example, we may have  $N$  eNB nodes serving  $N$  different user terminals. The time is uniformly slotted. The channel is block fading with duration of a block equal to a time slot. Furthermore, codewords are completely transmitted during a single time slot. The channel in time slot  $t$  is described by a square matrix  $\mathbf{Y}(t)$  whose elements are the power attenuations of the channel between transmitter (associated to rows) and receivers (associated to columns). We dub them as the Channel States (CS). The row  $i$  includes the states of the channels from the transmitting node  $i$  to all the destination nodes. This is the vector of known CS information at node  $i$  and it is denoted by  $\mathbf{y}_i(t)$ . The  $j$ -th column includes the states of the channels from all the transmitting nodes to the receiver  $j$ . This is the column vector denoted by  $\mathbf{y}^j(t)$ . It contains all the CS information necessary to determine the signal to interference and noise ratio (SINR) at the destination node  $j$  at time slot  $t$ . Furthermore, each power attenuation is modelled as an ergodic Markov chain taking values in a discrete set and described by the transition matrix  $\mathbf{T}(i,j)$ . The steady CS probability distribution of the channel between transmitter  $i$  and destination  $j$  is given by the column vector  $\boldsymbol{\pi}(i,j)$ .

At each node, packets arrive from the upper layer according to an independent and identically distributed arrival process  $\zeta_i(t)$  with arrival rate  $\lambda_i$ . Here,  $P(\zeta_i(t))$  is the probability of receiving  $\zeta_i(t)$  packets at time instant  $t$ . The packets have constant length. Each transmitter is endowed with a buffer of finite length. We denote by  $B_i$  the maximum length of the buffer at node  $i$  and by  $q_i(t)$  the number of queuing packets at the beginning of slot  $t$ . In the following, we dubbed the variable  $q_i(t)$  as the queue state (QS). In a given time slot we assume that all the arrivals from the upper layer occur after transmission of packets to the network.

In each time slot, on the basis of the available information at time  $t$  transmitter  $j$  decides:

- (a) the transmission power level  $p_i$ ;
- (b) the number of packets to transmit  $\mu_i$  (determining the transmission rate);
- (c) to accept or reject new packets arriving from upper layers. We denote with  $c_i=1$  and  $c_i=0$  the decision of accepting and rejecting the packets, respectively.

Therefore, the action of the node  $j$  at time slot  $t$  is described by the triplet  $\mathbf{a}_i(t) = (p_i(t), \mu_i(t), c_i(t))$ .

The information available at node  $i$  at time  $t$  is given by the pair  $\mathbf{x}_i(t) = (\mathbf{y}_i(t), q_i(t))$ , i.e. the CSs from transmitter  $i$  to all receivers and the number of the packets in the queue at the beginning of time slot  $t$  (QS). We refer to the pair  $\mathbf{x}_i(t)$  as the transmitter state. Additionally, each transmitter knows the statistics of the other channels and the statistics of the arrival process in the other nodes' buffers.

However, the performance of the receiver  $i$  depends on the state of the links from all the transmitters to the receiver  $i$ , i.e.  $\mathbf{y}^i(t)$ . Therefore, it is convenient to introduce the receiver state (RS) as  $\mathbf{x}^i(t) = (\mathbf{y}^i(t), q_i(t))$ .

The signal of the user of interest is impaired by the interfering signals and additive white Gaussian noise with variance  $\sigma^2$ . When the power level choices of the active transmitters are  $\mathbf{p}=(p_1,p_2,\dots,p_N)$ , and the receiver performs single user decoding, the maximum instantaneous achievable rate for the  $i$ -th communication pair depends on the receiver state  $i$ ,  $x^i(t)$ . We denote it by  $r_i^{\text{SU}}(x^i(t),\mathbf{p})$ . It is given by

$$r_i^{\text{SU}}(x_i(t),\mathbf{p}) = \log_2 (1 + \text{SINR}_i^{\text{SU}}(x_i(t),\mathbf{p}))$$

where  $\text{SINR}_i^{\text{SU}}(x^i(t),\mathbf{p})$  is the signal to interference and noise ratio at receiver  $i$  given by

$$\text{SINR}_i^{\text{SU}}(x^i(t),\mathbf{p}) = \begin{cases} \frac{y_i^i(t)p_i(t)}{\sigma^2 + \sum_{j \neq i} y_i^j(t)p_j(t)} & q_i(t) > 0 \\ 0, & q_i(t) = 0 \end{cases}.$$

If the receiver performs successive interference cancellation (SIC) decoding and, additionally, knows the transmission rate of the decodable interferes the maximum instantaneous achievable rate for the  $i$ -th communication pair is given by

$$r_i^{\text{SIC}}(x_i(t),\mathbf{p}) = \log_2 (1 + \text{SINR}_i^{\text{SIC}}(x_i(t),\mathbf{p}))$$

where

$$\text{SINR}_i^{\text{SIC}}(x^i(t),\mathbf{p}) = \begin{cases} \frac{y_i^i(t)p_i(t)}{\sigma^2 + \sum_{\substack{j \neq i \\ q_j(t) > 0 \\ j \text{ not decodable}}} y_i^j(t)p_j(t)} & q_i(t) > 0 \\ 0, & q_i(t) = 0 \end{cases}$$

At each time slot, a node chooses its action without having a global view of the channel states and the other users' interference. There is no coordination among transmitters' actions and only local information is available at each node. Therefore, for any choice of the transmission power and rate  $(p_i,\mu_i)$ , there is no guarantee that the  $\mu_i$  transmitted packets can be received correctly when the channel state at the transmitter is  $x_i$ . Let  $R$  be the rate required to transmit a packet in a time slot. The probability that  $\mu_i$  packets can be transmitted successfully in a time slot  $t$  by source  $i$  is given by the probability that the maximum instantaneous achievable rate is higher than the transmission rate utilized for transmitting  $\mu_i$  packets, i.e.  $R \mu_i$ . This probability is given by

$$\Pr \{r_i(x_i(t),\mathbf{p}) \geq \mu_i(t)R\}.$$

In such scenario, it is interesting to maximize the throughput, i.e. the average number of packets successfully received by the destination. Additionally, for physical and QoS reasons we need to consider that the transmitters are subjected to constraints on the average transmitted powers, on the average queue length, and eventually on the maximum outage probability.

Then, a formal statement of the problem is as follows. Each transmitter  $k$  (e.g. eNB) maximizes the throughput

$$\max_{(p_k,\mu_k)} E[ \Pr \{r_i(x_i(t),\mathbf{p}) \geq \mu_i(t)R\} \mu_k(t)R ]$$

Subject to constraints on:

- Average power:  $E[ p_k(x_k(t))] \leq \bar{p}_k$
- Average buffer length:  $E[ q_k(t)] \leq q_k$
- (Eventually) Probability of outage at the steady state:

$$\Pr \{r_i(x_i(t),\mathbf{p}) \geq \mu_i(t)R\} \leq \bar{P}_k^{\text{out}}$$

In this study we will consider two different approaches:

(A-self)  $\rightarrow$  each user independently optimizes its strategy to maximize its own throughput (selfish game);

(A-coop)  $\rightarrow$  each user independently from the others optimizes its strategy to maximize the joint throughput of the whole network (team game).

Each approach will be investigated for two different kinds of receivers:

- (a) receivers performing single user decoding;
- (b) receiver performing SIC decoding.

Approach A-x, being x=self or x=coop and decoding d=SU or d=SIC is addressed as A-x-d.

We formulate the above described problem as a stochastic N-player game with constraints on the average transmit powers, and average occupancy of the queues. The solutions of the games can be obtained via successive best responses of one transmitter to the policies of the remaining transmitters. The algorithm boils down to successive linear programming problems. Interestingly, simulations showed that the algorithm converges very quickly.

### Performance results and future steps

For performance analysis and evaluation we considered two scenarios with parameters detailed in Table 4-11. All the links from a transmitter to a receiver are independent and identically distributed. Each of them is described by a Markov chain with transition probabilities described in Figure 4-49 (left).

We perform a two level admission control. One is performed by our offline algorithm that, for each state of the transmitter, indicates whether to accept or reject the incoming packets in a certain time slot. However, since the game based algorithm indicates only whether to accept or not incoming packets independently of their number, it may happen that the available space in the buffer is not sufficient to store all the arrived packets. Therefore, a second (realtime) control mechanism is needed in order to drop the packets when the queue is full.

**Table 4-11: Simulation parameters for stochastic game**

Simulation Parameter	Value Setting 1	Value Setting 2
Number of Communications	2	3
Buffer Length	5	5
Channel State Cardinality	3	3
Maximum Number of Packets Transmittable in a Time Slot	5	5
Maximum Average Transmit Power	1.5	1.5
Maximum Average Buffer Occupancy	3	3
Average rate of the Poisson distributed packet arrival process	1	1

From simulations we could verify that the best response algorithm converges to a single solution in the A-self-SU model while for the A-self-SIC model two distinct solutions are obtained.

In the following, we compare the performance of such strategies in the network.

The performance measures are:

- Throughput, i.e. the number of packets per time slot correctly decoded by the receiver,
- Outage rate, i.e. the fraction of transmitted packets which cannot be decoded correctly,
- Drop rate, i.e. the fraction of arriving packets from upper layer which are rejected due to the admission control.

Table 4-12 compares the performance of the policies obtained for different kinds of receivers and a selfish or a cooperative approach.

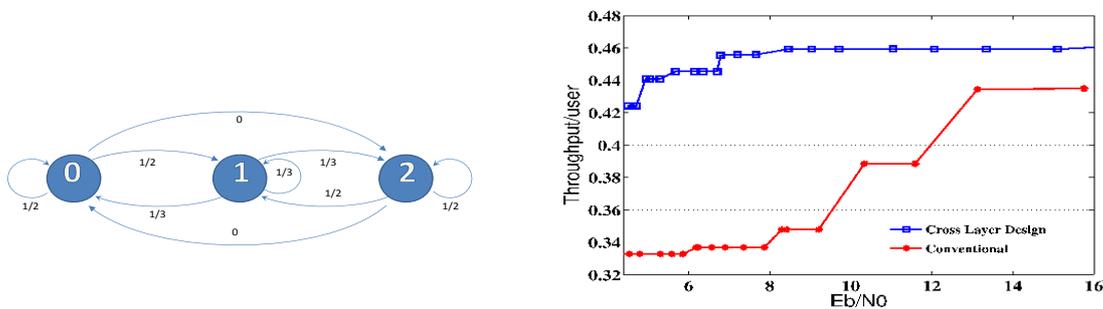
**Table 4-12: Performance of different utility functions and receivers**

A-{self,coop}-{SU,SIC}	Throughput	Outage Rate	Drop Rate
A-self-SU	0.49	0.42	0.15

A-self-SIC (1st equilibrium)	0.64	0.24	0.16
A-self-SIC (2nd equilibrium)	0.69	0.19	0.15
A-coop-SU	0.5	0.4	0.16

From Table 4-12 it is apparent that a selfish or a cooperative approach based on a competitive or a team approach do not affect substantially the performance while the receiver capabilities to perform successive interference cancellation decoding has a remarkable impact on both the throughput and the outage rate.

The advantages offered by the proposed cross-layer design approach compared to the conventional resource allocation, which does not take into account the state of the queue, are apparent from Figure 4-49 (right) where the average throughput per user versus the rate between the transmitted power per information bit and the noise variance  $E_b/N_0$  are shown for the conventional and the cross-layer design approach.



**Figure 4-49: Markov chain modeling a link (left) and Throughput vs  $E_b/N_0$  (right)**

Our analysis shows that, even with a very limited amount of exchanged information among transmitters sharing the same frequency band, it is still possible to design resource allocation, scheduling, and admission control policies able to guarantee a reasonable level of communications. A key point to improve the system performance when the transmitters do not share data and control information, is to enhance the interference mitigation capabilities of the receivers. Therefore, saving of frequency band for the feedback channels can be compensated, to some extent by the use of more costly and powerful equipments at the receiver side.

The next steps of the study will be focused on the design and analysis of algorithms for joint rate and power allocation, scheduling and admission control with increasing levels of channel knowledge at the transmitters. The final goal is an analysis of the trade-off between gain in throughput on the data plane and bandwidth needed on the control plane for control information exchange.

## 5 Conclusions and next steps

Resource allocation and cross-layer design provide important degrees of freedom for managing inter-cell interference in a way which is at the same time different and much complementary to the signal processing and coding approaches investigated in WP1 (and reported under D1.2). A fundamental aspect of multi-user communications in random fading channels, typically encountered in real-life contexts, is that the interference signals are subject to the same kind of variability due to random path gain and fading which affects the main information signal. Clustering exploits this by grouping certain users belonging to different cells together when they are compatible from an interference point of view. Resource allocation schemes further assign them to certain resource blocks, carefully selected in time and frequency. Doing so, the system can exploit the multi-user diversity so as to obtain desirable interference patterns. The obtained benefits are shown to vary depending on whether the clustering strategy simply relies on slow-varying path gain information or also includes fast fading information. Our results indicate that significant interference reduction gains can be obtained from a number of various techniques summarized in the following. The resource allocation schemes can be used on their own as a means to mitigate interference or can be exploited in conjunction with the physical-layer

oriented JP CoMP and beamforming methods. A number of different scenarios were considered.

In section 4.1, clustering and user grouping techniques have been addressed, targeting at feasible and practical schemes. We have considered a static clustering approach for distributed coordinated scheduling and a semi-static clustering technique where the clusters are updated based on user measurements. The performance results show that the evaluated schemes can improve the average sum-rate in the cell edge, thus achieving a higher uniformity of the performance over the served area. Partial CoMP user-centric clustering approach was also proposed to achieve both high penetration rate of users that benefit from CoMP and limited backbone and feedback overhead. The system level simulations verify the expected performance gains for the single antenna case per eNB and UE. Including more antenna elements will lead to further degrees of freedom for optimization and higher performance gains.

In the contributions related to user grouping, transmission to multiple users based on channel orthogonality metrics was considered. Different user pairing approaches were compared, ranging from the trivial random pairing that provides a lower performance bound up to the exhaustive search pairing that provides an upper performance bound. A good performance-complexity trade-off was identified when using the algorithm that adopts a user pairing approach based on both CQI and orthogonal deficiency metrics. Subcarrier pairing is also considered in the framework of single carrier frequency division multiple access (SC-FDMA) to solve impairment problems. The performance assessment of two different double SC-SFBC schemes in MU-MIMO configurations was carried out: misaligned double SC-SFBC and aligned double SC-SFBC. The performances of these two schemes are similar, but the second one demands much lower decoding complexity.

In section 4.2, inter-cell interference coordination techniques for heterogeneous deployments have been developed in order to minimize the impact of a massive deployment of HeNBs on the eNB network. A blind power setting algorithm is derived for the downlink while the uplink requires a minimal exchange of information through a central entity. The system performance can be further improved by the use of this central entity that computes an optimization of the spectrum allocation in the aim of further reducing the eNB/HeNB and HeNB/HeNB interference.

In section 4.3, coordinated scheduling was considered. The coordinated scheduling aims at improving the cell edge throughput and spectral efficiency by exchanging only information on the scheduled resources between the coordinated nodes. One procedure is to use optimization methods to derive algorithms in order to maximize throughput. In this context two algorithms have been derived. The first algorithm is a scheduling algorithm coordinated between the eNBs able to dynamically manage different transmission modes (section 4.3.1). This algorithm is the solution of a maximization problem of the objective function representing the effective data rate of each user and its confident function. This confident function aims to reflect the realistic scenarios such as the realistic feedbacks, Mac overhead, traffic models and propagation scenarios. The second algorithm (section 4.3.2) proposes to resolve downlink joint power control and beamforming problem in a multi-cell context. This algorithm jointly finds a set of feasible transmit beamforming weight vectors and downlink transmit power allocations such that the SINR at each link is greater than a target value. A user selection step is also added to the initial algorithm to help satisfy the per-cell power constraint. This algorithm shows a significant gain at the cell edge. It is assumed a single receiver antenna at the user side but the proposed solution can easily be generalized to multiple receive antennas. In future work, this solution will be evaluated under limited feedback links. Other solutions proposed in section 4.3.3 and section 4.3.4 for the CS, are based on the restriction on the use of resources on some cells. In section 4.3.3 two approaches have been proposed the implicit coordination method and the 3D beam coordination. The second one outperforms the first one at the cost of increased exchanged information between the coordinated areas. Nevertheless the achieved gains through additional information on the downtilt are very promising. In section 4.3.4 the coordination was achieved by restricting the choice of available precoding matrices at the interfering transmitters. Two ways are investigated, in the first one the restriction is applied if the gain in the requesting cell is greater the loss in the requested cell. In second way the restriction is always applied. Those two algorithms applied in heterogeneous scenario provide a gain up to 40% on the non-requested scheme. In section 4.3.5, two basic coordinated femtocells schemes are considered: interference nulling and joint superimposed transmission. Most of the gain is provided by the

interference nulling in a femtocell campus scenario. The future investigations will be more realistic by considering traffic models and resource allocation.

In Section 4.4, we investigated the effect of scheduling on the reduction of interference and its impact on the overall benefits of JP CoMP. It was seen that depending on the type of scheduler, the impact was widely different. In the case of rate maximizing schedulers without fairness constraints, the scheduler ultimately tends to reduce the interference to a small level when compared to the direct channel gains. A scheduled user then benefits little from an extra layer of interference mitigation via JP CoMP. In contrast, with a fairness oriented scheduler, the importance of spatial domain interference avoidance using JP CoMP remains obvious.

In Section 4.5, game theory, and more specifically Bayesian games, have been adopted as a mathematical framework in order to develop distributed algorithms for cross layer design. Our analysis shows that with a very limited amount of exchanged information among transmitters sharing the same frequency band, it is still possible to design resource allocation, scheduling, and admission control policies being able to guarantee a reasonable communication quality. However, in this case of very limited information exchange, the analysis shows that a cooperative behaviour of the transmitter does not provide a substantial performance improvement compared to a selfish behaviour while the complexity increases considerably. Nevertheless in both cases, the system performance brought by the game theoretic transmitter algorithm can be further enhanced by allowing successive cancellation instead of single user decoding, as developed under ARTIST4G's WP2. The next steps of the study will be focused on the design and analysis of algorithms for joint rate and power allocation, scheduling and admission control with increasing levels of channel knowledge at the transmitters. The final goal is an analysis of the trade-off between gain in throughput on the data plane and bandwidth needed on the control plane for control information exchange.

Such studies indicate in general that the impact of scheduling, resource allocation scheme and cross layer design on the system performance for an interference-prone wireless network is substantial.

The strategies offered to the system designer are diverse in nature, with two clearly leading concepts however: (i) user grouping, (ii) coordinated multicell scheduling and power control

Nevertheless the precise effect of resource allocation schemes on the interference distribution highly depends on certain critical parameters that merit further investigations or specifications. Among these, we note: (i) The number of simultaneously active users available as "degrees of freedom" to the Layer 2 protocols, as this determines the level of multi-user diversity, (ii) the nature and performance characteristics of the information exchange mechanism between the cells engaged in the cooperative resource allocation protocols.

## References

- [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.
- [3GPP36420] 3GPP TS 36.420, "X2 general aspects and principles (Release 10)", 2010
- [3GPP36814] 3GPP, "TR 36814-900 Further advancements for E-UTRA - Physical Layer Aspects," 3GPP Ftp Server (<http://www.3gpp.org/ftp>), 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.
- [ALUD10] "Co-Scheduling", Alcatel-Lucent contribution to EASY-C Public Workshop April 16, 2010
- [ARTD11] D1.1 – "Definitions and architecture requirements for supporting interference avoidance techniques," *ARTIST 4G technical deliverable, August 2010*.
- [ARTD12] D1.2 – "Innovative advanced signal processing algorithms for interference avoidance", *ARTIST4G technical deliverable, 2010*.
- [ARTD51] D5.1 – "Scenarios, Key Performance Indicators and Evaluation Methodology for Advanced Cellular Systems," *ARTIST4G technical deliverable, June 2010*
- [ARTD61] D6.1 – "First Feedback on Implementation Aspects Connected to the Selected Innovations," *ARTIST4G technical deliverable March 2011*
- [BO99] M. Bengtsson and B. Ottersten, "Optimal Downlink Beamforming Using Semidefinite Optimization," in Proc. 37th Annual Allerton Conf. Control Computing, Monticello, Sep. 1999,.
- [Boyd04] S.B. Boyd, L. Vandenberghe: Convex optimization, chapter 9, (<http://www.stanford.edu/~boyd/cvxbook/>)
- [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-spring, 2010.
- [CCM+09] C. Ciochina, D. Castelain, D. Mottier and H. Sari, "New PAPR-Preserving Mapping Methods for Single-Carrier FDMA with Space-Frequency Block Codes," *IEEE Transactions on Wireless Communications*, vol. 8, issue 10, pp. 5176-5186, Oct. 2009.
- [DJG+09] G. Dartmann, M. Jordan, X. Gong, G. Ascheid, "intercell Interference Mitigation with Long-term Beamforming and Low SINR Feedback Rate in a Multiuser Multicell Unicast Scenario," *IEEE Vehicular Technology Conference-spring*, 2009
- [DY08] H. Dahrouj and W. Yu, "Coordinated beamforming for the multi-cell multiantenna wireless system," Proc. Of Conf. on Information Sciences and Systems(CISS'08) March 2008.

- [FF09] FemtoForum, "OFDMA Interference Study: Evaluation Methodology Document," *Femto Forum Working Group 2 Document*, 2009.
- [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.
- [FLT98] F. Rashid-Farrokhi, K.J. Liu, and L. Tassiulas, "Transmit Beamforming and Power Control for Cellular Wireless Systems," *IEEE Journal on Selected Areas in Communications*, vol.16, no.8, pp.1437-1450, October 1998
- [GA04] D. Gesbert and M. Alouini, "How much feedback is multi-user diversity really worth?," *IEEE International Conference on Communications*, pp. 234-238, 2004.
- [GA10] J. Giese and M. A. Amin, "Performance upper bounds for coordinated beam selection in LTE-Advanced", 2010 International ITG Workshop on Smart Antennas (WSA), Bremen, Germany, 2010
- [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.
- [JFJ+10] V. Jungnickel, A. Forck, S. Jaeckel, F. Bauermeister, S. Schiffermueller, S. Schubert, S. Wahls, L. Thiele, L.; Haustein, T.; Kreher, W.; Mueller, J.; Droste, H.; Kadel, G.; "Field Trials using Coordinated Multi-Point Transmission in the Downlink", in Proc. 3rd International Workshop on Wireless Distributed Networks (WDN), held in conjunction with IEEE PIMRC 2010, Sept. 2010
- [KR03] S. S. Kulkarni and C. Rosenberg, "Opportunistic Scheduling Policies for Wireless Systems with Short Term Fairness Constraints", in Proc. *GLOBECOM*, 2003.
- [LBS+10] T.R. Lakshmana, C. Botella, T. Svensson, X. Xu, J. Li, and X. Chen, "Partial joint processing for frequency selective channels," Proc. *IEEE Vehicular Technology Conference-fall*, 2010.
- [MGF10] P. Marsch, M. Grieger and G. Fettweis "Field Trial Results on Different Uplink Coordinated Multi-Point (CoMP) Concepts in Cellular Systems" *GLOBECOM'10*, Miami, FL, USA, 06. Dec 2010
- [MZ10] W. Mennerich, W. Zirwas, "User Centric Coordinated Multi Point Transmission", Proc. *IEEE VTC-fall*, 2010
- [NGMN08] NGMN, "NGMN Radio Access Performance Evaluation Methodology version 1.0," 2008.
- [PBG+04] G. Piñero, C. Botella, A. González, M. de Diego, N. Cardona, "Downlink power control and beamforming for a cooperative wireless system," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2004.
- [PBG+08] A. Papadogiannis, H.J. Bang, D. Gesbert, and E. Hardouin, "Downlink overhead reduction for Multi-Cell Cooperative Processing enabled wireless networks," *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2008.
- [PGH08] A. Papadogiannis, D. Gesbert, E. Hardouin, "A dynamic clustering approach in wireless networks with multi-cell cooperative processing," in *Proc. IEEE International Conference on Comm.*, 2008
- [PHG09] A. Papadogiannis, E. Hardouin, and D. Gesbert, "A framework for decentralising multi-cell cooperative processing on the downlink," *2008 IEEE GLOBECOM Workshops*, 2008, p. 1–5.

- [PGB07] A.S. Prabhu, H. Gupta, M.M. Buddhikot, Fast spectrum allocation in coordinated dynamic spectrum access based cellular networks, Proceedings of IEEE DySPAN 2007
- [PGD07] A. Prabhu, H. Gupta, S. Das: Minimum interference channel assignment in multi-radio wireless mesh networks, IEEE SECON, 2007.
- [PAG+08] A.S. Prabhu, M. al Ayyoub, H. Gupta, S. Das, M.M. Buddhikot: Near-optimal dynamic spectrum allocation in cellular networks, Proc. IEEE Dynamic Spectrum Access Networks (DySPAN), 2008.
- [Ros60] J. B. Rosen: The Gradient Projection Method for Nonlinear Programming. Part I: Linear constraints", SIAM Journal, vol. 8, N° 4, march 1960, pp. 181-217.
- [SB04] M. Schubert and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem With Individual SINR Constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no.1, pp. 18-28, January 2004.
- [SCWY] Y. Song, L. Cai, K. Wu, and H. Yang, "Collaborative MIMO Based on Multiple Base Station Coordination," *Contribution to IEEE*, vol. 802.
- [STS+07] M. Sadek, A. Tarighat, and A. H. Sayed, "A Leakage-Based Precoding Scheme for Downlink Multi-User MIMO Channels". *IEEE Trans. on Wireless Comm.*, vol. 6, no. 5, pp. 1711–1721, May 2007.
- [TCJ08] A. Tolli, 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, pp. 1428-1440, 2008.
- [TSH+2009] L. Thiele, T. Wirth, M. Schellmann, Y. Hadisusanto, V. Jungnickel, V., "MU-MIMO with Localized Downlink Base Station Cooperation and Downtilted Antennas," 2009 IEEE International Conference on Communications workshops, , pp. 1-5 2009.
- [VM99] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," *Proc. IEEE Vehicular Technology Conference-spring*,, 1999.
- [WIN2D112] IST-4-027756 WINNER II, Deliverable D1.1.2, "WINNER II channel models: Part I channel models," September 2007.
- [WIND14] D1.4 – "Initial Report on Advanced Multiple Antenna Systems" vol. 1, 2009, pp. 1-110. WINNER+
- [WIND18] D1.8- "Intermediate Report on CoMP (Coordinated Multi-Point ) and Relaying in the Framework off CoMP. Wireless World Initiative" - WINNER+, 2009.
- [WWK+09] X. Wei, T. Weber, A. Kuhne, and A. Klein, "Joint transmission with imperfect partial channel state information," *Proc. IEEE Vehicular Technology Conference-spring*, 2009.
- [XWT+08] W. Xiaoting, W. Wenbo, Z. Tianlin, and Z. Zongyin, "Spatial multiuser pairing scheduling strategies for virtual MIMO systems," 11<sup>th</sup> IEEE Singapore International Conference on Communication Systems, IEEE, 2008, pp. 823-827.

## List of acronyms and abbreviations

3GPP	3rd Generation Partnership Project
ANR	Automatic Neighbour Relation
ARTIST4G	Advanced Radio Interface Technologies for 4G SysTems
BBU	Baseband Unit
BS	Base Station
BSCBS	Base Station Coordinated Beam Selection
CA	Cooperation Area
CCN	Central Coordination Node
CDF	Cumulative Distribution Function
CoMP	Coordinated Multi Point
CQI	Channel Quality Indicator
CS	Channel State
CS/CB	Coordinated Scheduling / Coordinated Beamforming
CSG	Closed Subscriber Group
CSI	Channel State Information
CSIT	Channel State Information at the Transmitter
CU	Central Unit
DL	DownLink
DSA	Dynamic Spectrum Allocation
eNB	enhanced Node B
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
FA	Frequency Adaptive
FDD	Frequency Division Duplex
FDPS	Frequency Domain Packet Scheduling
HARQ	Hybrid Automatic Repeat Request
HeNB	Home enhanced Node B
Het Net	Heterogeneous Network
HUE	HeNB User Equipment
HZ	Hotzones
ICIC	Inter-Cell Interference Coordination
JP	Joint Processing
KPI	Key Performance Indicator
L1	Layer 1
L2	Layer 2
LTE	Long Term Evolution
MAC	Media Access Control
MCS	Modulation and Coding Scheme
MIESM	Mutual Information Effective SNR Mapping
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MUE	Macro User Equipment
MU-MIMO	Multi User Multiple-Input Multiple-Output

NA	Non Adaptive
OD	Orthogonal Deficiency
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
ORR	Opportunistic Round Robin
OTA	Over The Air
PAPR	Peak to Average Power Ratio
PDU	Protocol Data Unit
PF	Proportional Fair
PJP	Partial Joint Processing
PMI	Precoding Matrix Index
PRB	Physical Resource Block
QoS	Quality of Service
RAN	Radio Access Network
RB	Resource Block
RI	Rank Indicator
RoF	Radio over Fiber
RR	Restriction Request
RRC	Radio Resource Control
RRH	Radio Remote Heads
RRM	Radio Resource Management
RSRP	Reference Symbol Received Power
RSS	Received Signal Strength
SC	Single Carrier
SC FDMA	Single Carrier Frequency Division Multiple Access
SC SFBC	Single Carrier Space-Frequency Block Code
SCM	Spatial Channel Model
SDMA	Space Division Multiple Access
SDP	Semi Definite Programming
SFBC	Space Frequency Block Codes
sIF	Strongest Interfering
SIMO	Single-Input Multiple-Output
SINR	Signal to Interference and Noise Ratio
SISO	Single-Input Single-Output
SLNR	Signal to Leakage plus Noise Ratio
SLR	Signal to Leakage Ratio
SM	Spatial Multiplexing
SON	Self-Organizing Network
STBC	Space Time Block Codes
SU-MIMO	Single User Multiple-Input Multiple-Output
SVD	Singular Value Decomposition
TDD	Time Division Duplex
TDPS	Time Domain Packet Scheduling
TM	Transmission Mode
TMS	Transmission Mode Switching
TS	Tabu Search

TTI	Transmission Time Interval
TxD	Transmit Diversity
UE	User Equipment
UL	UpLink
WCI	Worst Companion Indicator
ZF	Zero Forcing