INFSO-ICT-317941 iJOIN

Deliverable D2.3

Final definition and evaluation of PHY layer approaches for RANaaS and joint backhaul-access layer

Editor:	Dirk Wübben (UoB)
Deliverable nature:	Deliverable
Suggested readers:	iJOIN GA
Due date:	April 30th, 2015
Delivery date:	April 30th, 2015
Version:	1.0
Total number of pages:	96
Reviewed by:	GA members
Keywords:	RAN PHY, Backhaul PHY, RANaaS, Functional Split, Joint RAN-BH, iJOIN Architecture
Resources consumed	24.64 PM

Abstract

This deliverable provides the final definition and evaluation of physical layer approaches for the RANaaS concept and for a joint design of access and backhaul links. It shows how the WP2 approaches are integrated in the iJOIN architecture and also reviews the WP2 functional architecture. For each candidate technology performance evaluations for different communication and network scenarios are provided. Thus, the applicability as well as the performance gains of physical layer approaches for the novel mobile communication architecture are demonstrated.

List of authors

Company	Author
UoB	Dirk Wübben (wuebben@ant.uni-bremen.de)
	Henning Paul (paul@ant.uni-bremen.de)
	Guang Xu (xu@ant.uni-bremen.de)
	Ban-Sok Shin (shin@ant.uni-bremen.de)
SCBB	Massinissa Lalam (massinissa.lalam@sagemcom.com)
CEA	Valentin Savin (valentin.savin@cea.fr)
	Zeina Mheich (zeina.mheich@cea.fr)
UNIS	Lei Zhang (lei.zhang@surrey.ac.uk)
	Atta Quddus (a.quddus@surrey.ac.uk)
	Efstathios Katranaras (e.katranaras@surrey.ac.uk)
IMC	Umer Salim (umer.salim@intel.com)
TI	Boldi Mauro Renato (mauro.boldi@telecomitalia.it)
	Bruno Melis (bruno1.melis@telecomitalia.it)
TUD	Jens Bartelt (jens.bartelt@ifn.et.tu-dresden.de)

History

Modified by	Date	Version	Comments
Dirk Wübben	April 30th, 2015	1.0	Final version of D2.3

Table of Contents

List of Figures

List of Tables

Abbreviations

3GPP	3rd Generation Partnership Program
A/D	Analogue to Digital
ADMM	Alternating Direction Method of Multipliers
ALCE	Augmented Lagrangian based Cooperative Estimation
ASIC	Application Specific Integrated Circuit
AT	Area Throughput
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BH	Backhaul
CEf	Cost Efficiency
CB	Coordinated Beamforming
CDF	Cumulative Distribution Function
CoMP	Coordinated Multi-Point
CP	Cyclic Prefix
CPRI	Common Public Radio Interface
C-RAN	Centralized RAN
CPU	Central Processing Unit
CRC	Cyclic Redundancy Check
CS	iJOIN Common Scenario
CSI	Channel State Information
CT	Candidate Technology
CWDM	Coarse Wavelength Division Multiplexing
D/A	Digital to Analogue
DiCE	Distributed Consensus-Based Estimation
DL	Downlink
DMRS	Demodulation Reference Signal
DSP	Digital Signal Processor
EEf	Energy Efficiency
E-LMMSE-IRC	Enhanced Linear Minimum Mean Square Error with Interference Rejection Combining
eNB	Evolved Node B
ERD	Error Resilient Decoder
EPA	Extended pedestrian A
FD	Frequency Domain
FDD	Frequency Division Duplex
FEC	Forward Error Correction
FER	Frame Error Rate
FFT	Fast Fourier Transform
FLOPS	Floating-point Operations per Second
FPGA	Field Programmable Gate Array
HARQ	Hybrid Automatic Repeat Request
I	In-phase (sample)
I/Q	In-phase / quadrature (sample)
ICI	Inter-Cell Interference
ICIC	Inter-Cell Interference Coordination
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
iJOIN	Interworking and JOINt Design of an Open Access and Backhaul Network Architecture for Small Cells based on Cloud Networks
iNC	iJOIN Network Controller
InH	Indoor / Hotspot (channel model)
INP	In-Network Processing
IR	Incremental Redundancy
iRPU	iJOIN virtual RAN Processing Unit
iSC	iJOIN Small Cell

ISD	Inter Site Distance
iTN	iJOIN Transport Node
ITU	International Telecommunication Union
iveC	iJOIN virtual eNB Controller
J1	logical interface connecting an iSC with a RANaaS instance
J2	logical interface connecting two iSCs
J3	logical interface connecting entities of a veNB to the iNC
J4	logical interface between iNC and core network
JNCC	Joint Network-Channel Coding
JT	Joint Transmission
KPI	Key Performance Indicator
LLR	Log Likelihood Ratio
LOS	Line Of Sight
LTE	Long Term Evolution
LTE-A	LTE Advanced
MAC	Medium Access Control
MARC	Multiple-Access Relay Channel
MCS	Modulation and Coding Scheme
MeNB	Macro eNB
MH	Multi-Hop
MIMO	Multiple-Input Multiple-Output
MMSE	Minimum Mean Square Error
mmWave	Millimetre Wave
MPTD	Multi-Point Turbo Detection
MUD	Multi-User Detection
NGMN	Next Generation Mobile Networks
NLOS	None Line Of Sight
OFDM	Orthogonal Frequency Division Multiplexing
PALCE	Priority-aided Augmented Lagrangian Cooperative Estimation
PAPR	Peak to Average Power Ratio
PC	Power Constraint
PHY	Physical Layer
PRB	Physical Resource Block
PSS	Primary Synchronization Signal
PtP	Point to Point
PUSCH	Physical Uplink Shared Channel
Q	Quadrature (sample)
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RAM	Random Access Memory
RAN	Radio Access Network
RANaaS	RAN-as-a-Service
RAP	Radio Access Point
RB	Resource Block
RC	Relay Channel
RLC	Radio Link Control
RO-DiCE	Reduced Overhead Distributed Consensus-Based Estimation
RoF	Radio over Fibre
RRH	Remote Radio Head
RRM	Radio Resource Management
RS	Reference Signal
RSRP	Reference Signal Received Power
RTT	Round Trip Time
RV	Redundancy Version
SCF	Small Cell Forum
SC-FDMA	Single-Carrier Frequency Division Multiple Access
SDN	Software Defined Network

SINR	Signal to Interference and Noise Ratio
SISODQ	Soft-Input Soft-Output DeQuantizer
SNR	Signal to Noise Ratio
SPTD	Single-Point Turbo Detection
SSS	Secondary Synchronization Signal
SRS	Sounding Reference Signals
TBS	Transport Block Size
TD	Time Domain
TDD	Time Division Duplex
UE	User Equipment
UEf	Utilization Efficiency
UL	Uplink
UMi	Urban Micro (channel model)
vCPU	virtual Central Processing Unit
veNB	virtual Evolved Node B
WP	Work Package
ZF	Zero Forcing

List of Symbols

Parameters and Variables

a	information word
A_{net}	geometrical area
A_{veNB}	area of veNB
B	RAN bandwidth
B_{BH}	bandwidth used by BH for in-band backhauling
B_{RAN}	bandwidth used by RAN for in-band backhauling
c_0	speed of light
C	capacity
C_{BH}	backhaul capacity per unit area
C_{FB}	CSI feedback capacity from UE to iSC
D	data rate
D_{J_1}	backhaul data rate on J1 link
D_{J_2}	backhaul data rate on J2 link
D_P	payload message size
F	receiver noise figure
f_D	Doppler shift
f_s	sampling frequency
G_{AT}	gain in area throughput
F_{FLOPS}	RANaaS total computational capability (FLOPS)
K_{coop}	number of cooperating iSCs for partially centralized precoding
N_{block}	no. of bits in one code block
N_{CP}	length of cyclic prefix
N_{CPU}	No. of CPU cores for precoding matrix calculation
N_{FB_time}	total number of times for channel coefficients feedback
N_{FFT}	no. of FFT points
N_G	no. of PRBs sharing the same precoding matrix
N_{Hop}	no. of backhaul hops
N_{iSC}	no. of iSCs
N_{iSC}^{NET}	no. of iSCs in a network covered by a RANaaS
N_{iSC}^{s}	no. of iSCs per macro sector
N_{It}	no. of iterations
N_L	no. of transmitted layers
N_{PMC}	no. of precoding matrix calculation for the whole bandwidth
N_Q	no. of quantization bits
$n_{PRB,u}^{j}$	no. of physical resource blocks allocated to the *u*-th user served by the *j*-th iSC
N_{PRB}	no. of available physical resource blocks
N_R^{iSC}	no. of receive antennas at iSC
N_R^{UE}	no. of receive antennas at UE
N_{sc}	no. of subcarriers available for transmission
$N_{sc,u}^{j}$	no. of subcarriers allocated to the *u*-th user served by the *j*-th iSC
N_{SYMB}^{SUB}	no of symbols per frame
N_T^{iSC}	no. of transmit antennas at iSC

N_T^{UE}	no. of transmit antennas at UE
N_{UE}	no. of UEs
N_{veNB}	no. of virtual eNBs in the network
OF	oversampling factor
P	transmit power
P_{UE}	UE transmit power
$\bar{P}_x$	desired signal power without precoding
$\bar{P}_I$	interference power without precoding
q_c	RANaaS computational resource division factor
Q_m	modulation order
Q_{99}	99th percentile
R	spatially averaged rate per user
$R_{A_{net}}$	area throughput
$R_{A_{net}}^{Base}$	area throughput of baseline approach
$R_{A_{net}}^{CT}$	area throughput of CT
R_c	code rate
R_c^{BH}	code rate on backhaul
R_c^{RAN}	code rate on access link
R_s	sum rate of all UEs
$r_u(t)$	bits correctly delivered to (from) UE u
SNR_{BH}	SNR on backhaul link
SNR_{ij}	SNR on link between node i and node j
SNR_{RAN}	SNR on access link
$\mathcal{S}_{sc,u}^{j}$	set of subcarriers allocated to the *u*-th user served by the *j*-th iSC
TBS_j	transport block size for the *j*-th codeword
T_s	symbol duration
T_{SUB}	subframe duration
UE_u	user u
Δt	total CSI delay of the precoding operation
Δt_{BH}	backhaul latency per hop
Δt_{FB}	channel feedback delay
$\Delta t_{Precoder\_cal}$	precoding matrix inversion calculation delay
Δt_{Prop_total}	total propagation delay
Δt_{Tx}	total Tx processing time in LTE system
γ	overhead introduced by BH protection coding
ϵ	safety factor
κ	path loss exponent
λ	temporal correlation factor
ρ_{UE}	density of UE
ρ_{iSC}	density of iSC
ζ_1	number of equivalent addition operation times for each division operation
ζ_2	number of equivalent addition operation times for each multiplication operation
σ_n^2	noise variance
θ	Bandwidth sharing parameter
η	backhaul throughput

Indices

i, j	iSC
u	user

1 Executive Summary

Within this work package (WP) physical layer approaches for a dense deployment of small cells which allow for a flexible functional split and a joint design of the radio access and the backhaul network have been investigated. Based on the initial definition of the iJOIN architecture and the introduction of promising physical layer candidate technologies (CT) in deliverable D2.1 [1], the precise definition of these candidate technologies as well as preliminary results were provided in D2.2 [2]. The corresponding discussion of medium access control and radio resource management aspects are provided in the WP3 reports D3.1 [3] and D3.2 [4]. The network layer aspects are considered by WP4 and the reader is referred to D4.1 [6] and D4.2 [7]. The overall iJOIN system was initially defined in D5.1 [9] and the results provided by the individual work packages lead to the definition of reference scenarios and system requirements in D5.2 [10].

This report provides the final definition of the proposed set of novel physical layer technologies and presents detailed performance investigations indicating the benefits of the two main iJOIN concepts, i.e. joint access/backhaul operations and flexible functional split. The achieved performance results are combined with the results of WP3 in D3.3 [5] and WP4 in D4.3 [8] leading to the project wide analysis in D5.3 [11]. To this end, this final report of WP2 is organized as follows:

Section 2 provides a brief introduction to this report underlying the addressed scope and objectives. In addition, the key contributions of this last project period as well as the main contributions of WP2 are summarized.

Section 3 reviews the logical and functional architecture and discusses the approach for virtualized implementation within the virtual eNodeB concept. In particular, the lower layer functional split options are discussed by means of rate requirements and centralization gains, and numerical results based on the parameters of the iJOIN common scenarios as defined in D5.2 [10] are provided. Moreover, relevant physical layer approaches for a joint optimization of the access and the backhaul links are discussed.

Section 4 provides the final description of the WP2 candidate technologies initially introduced in D2.1 [1] and further developed in D2.2 [2]. After a brief description of the final implementation, detailed performance investigations are provided for each CT by means of achievable area throughput and backhaul traffic.

In Section 5 the intra WP2 evaluations are discussed based on the evaluation framework and the common scenarios described in D5.2 [10]. In particular, the interoperability of CTs is addressed in Section 5.1, the relevant metrics and evaluation parameters are defined in Section 5.2, and for all common scenarios numerical results for promising CTs are provided in Section 5.3.

Finally, the report is summarized and concluded in Section 6.

Also enclosed in this report are two appendixes. Appendix I provides for the considered set of backhaul technologies the key parameters as described in D4.2 [7]. Appendix II contains the analytical results for the backhaul load calculations provided in D2.2 [2] as well as the corresponding numerical results for the exemplary calculations in Section 3.2.1.

2 Introduction

2.1 Motivation and Background

The evolution towards 5G mobile networks is characterized by an exponential growth of traffic caused by an increased number of user terminals and a more frequent usage of powerful internet-capable devices. Regionally and temporally fluctuating traffic patterns as well as an increasing diversity of terminal classes and services require a higher scalability of mobile networks. Current mobile networks are not able to support this diversity efficiently but are designed for peak-provisioning and typical internet traffic. The use of very dense, low-power, small-cell networks with very high spatial reuse is a promising way to allow for handling future data rate demands. In small-cell networks, the distance between the radio access points (RAP) and terminals is reduced, and the spatial spectrum reuse is significantly increased. However, due to the density of the network, inter-cell interference increases and interference scenarios become more complex due to multi-tier interference. In such a scenario, centralized processing would allow for efficient interference avoidance and cancellation algorithms across multiple cells as well as joint detection algorithms.

Within iJOIN (Interworking and JOINt Design of an Open Access and Backhaul Network Architecture for Small Cells based on Cloud Networks) the flexible centralization of radio access network (RAN) considering a dense deployment of iJOIN small cells (iSCs) and the joint operation of access and backhaul (BH) networks have been investigated. The novel RAN-as-a-Service (RANaaS) concept was introduced which allows virtualized implementation of the centralized functionality on general purpose IT platforms. Within the project, WP2 develops novel physical (PHY) layer approaches for the dense deployment of small cells that allow for a flexible functional split and a joint design of the radio access and the backhaul network.

In deliverable D2.1 [1] the physical layer in 3GPP LTE networks is studied and relevant PHY layer techniques for the radio access and backhaul have been reviewed to identify the main bottlenecks and challenges. Based on this analysis, a set of promising candidate technologies (CTs) has been derived. These CTs have been further developed in D2.2 [2] and the compliance to the iJOIN objectives was addressed. In addition, the logical and functional architecture from a PHY perspective was introduced and implementation aspects for the virtual eNodeB have been analysed. Furthermore, detailed investigations for principle functional split options as well as implementations on general purpose processors have been provided.

In this report, the previously introduced technologies are finalised and elaborated in detail. In particular, the logical and functional architecture is revised and the lower layer functional split options are discussed by means of rate requirements and potential gains. In addition, numerical results based on the parameters of the iJOIN common scenarios (CS) are provided and relevant physical layer approaches for a joint optimization of the access and the backhaul links are discussed. The final evaluation of the WP2 CTs is provided according to the methodology defined in D2.2 [2]. Finally, the intra WP evaluation for the four iJOIN common scenarios is presented, where the interoperability of CTs is discussed, the common evaluation parameters are defined, and results for the area throughput as the main metric for PHY layer approaches are provided for all common scenarios.

2.2 Key Contributions

This report focuses on the final evaluation of PHY layer technologies and project wide harmonization. The list below highlights the key contributions of this deliverable, as well as the main scientific advances of the technologies described in this deliverable, pointing out the main differences over the state of the art as well as the most significant results achieved.

- Together with WP3, the quantitative analysis of principle functional splits in terms of throughput, latency, and achievable gains for the most promising functional splits was finalized. These results are part of a paper submitted to IEEE Wireless Communications Magazine [17] and detailed discussions with the small cell forum (SCF). Based on the analysis, WP4 has identified the possible functional splits for each iJOIN common scenario as reported in D4.3 [8].
- In collaboration with WP3 and WP4, the final logical and functional architecture is defined in D5.3 and the implementation of all WP2 technologies in this iJOIN architecture is provided.

- For the iJOIN common scenarios the specific functional split options of promising PHY technologies are identified and a harmonized evaluation per scenario is provided. This analysis demonstrates that WP2 technologies are able to increase the area throughput by factor of 40% to 100%. These results are also provided for the project wide analysis in D5.3 [11] of WP5.
- The PHY and MAC architecture for a virtualized implementation of RAN functionality is provided together with WP3. Exemplarily, the most demanding PHY functional block given by the turbo decoder has been implemented in the RANaaS testbed together with WP6 indicating the feasibility as well as the constraints of virtualized implementations. The investigation of the decoder implementation is part of the paper [28] being published at the IWCPM workshop at IEEE ICC and the more general analysis of virtualized implementation is part of the paper [18] submitted to the IEEE Communications Magazine.
- A detailed comparative study investigating the detection performance, the computational complexity as well as the backhaul rate of the previously developed approaches for distributed multi-user detection has been provided in [26]. Furthermore, a novel algorithmic approach achieving superior detection performance with reduced inter-node communication has been presented at ITG SCC [27].
- Based on the analytical results for the Joint Network-Channel Coding framework a backhaul-aware cell selection algorithm was developed in order to improve the overall network area throughput. This joint work with CT3.2 in WP3 has been submitted to IEEE SPAWC [68].
- The Multi-Point Turbo Detection performance assessment at the link and system level has been submitted to IEEE GLOBECOM [69] as a joint WP2/WP3 contribution.
- A framework for cluster size optimization in the presence of precoder processing and channel state information feedback delay was developed. In order to formulate the optimization problem, the various delay factors were modelled and the sum-rate of the network as a function of the cluster size was derived. The centralized precoding algorithm has been implemented on the RANaaS testbed to achieve realistic processing delays. The corresponding theoretical analysis and results have been submitted to IEEE Transactions on Wireless Communications [70].
- With joint access and backhaul coding, a new technology has been derived to improve performance of a joint access and backhaul link, especially when using wireless backhaul technologies. It improves the throughput while keeping the latency low and without adding additional overhead in the form of a dedicated backhaul code. Corresponding results have been published in [37].
- The approaches of In-Network Processing, the hierarchical precoder and the joint access and backhaul coding approach have been investigated on the mmWave testbed as reported in D6.2 [13].

In addition to the above results that have been produced for this deliverable, it is also worth highlighting the previous outcome from WP2 which was reported in deliverables D2.1 and D2.2:

- The benefits of cloud-computing for 5G mobile networks and the implications on signal processing algorithms have been discussed in the paper [16] as part of the special issue “Signal Processing for the 5G Revolution” of the IEEE Signal Processing Magazine.
- With In-Network Processing a technology for distributed multi-user detection and corresponding novel estimation algorithms have been developed achieving the performance of centralized detection if sufficiently strong links between small cells are available as published in [21], [22], and [23].
- Novel algorithmic approaches for designing partially centralized precoders have been developed, which are adaptable to heterogeneous and varying backhaul topologies. For the sake of low communication requirements, a hierarchical approach increasing the area throughput by a factor of 50% has been developed [31].
- The investigations of cloud-based implementations of modern forward error correction decoders for low density check codes and turbo codes have demonstrated that a realization on general purpose hardware of this most demanding task in the physical layer processing chain meets the strict timing requirements if some adaptions are performed [25].

3 Definition of iJOIN Architecture to support PHY layer approaches

Based on the preliminary definitions in D5.1 [9] and D5.2 [10] the final iJOIN architecture is defined in deliverable D5.3 [11]. It introduces the RAN-as-a-Service (RANaaS) concept by defining the virtual evolved node B (veNB) in order to handle several iJOIN small cells (iSCs) and a central processing instance running on a cloud platform. In addition, the key techniques of functional split and joint RAN/BH optimization have been developed for this architecture. In this section, we briefly illuminate the logical and functional architecture as well as both key concepts from a physical layer (PHY) perspective.

3.1 iJOIN Architecture

3.1.1 Logical and RANaaS Architecture

The final iJOIN architecture is defined in Section 5.2 of deliverable D5.3 [11]. Figure 3-1 illustrates the logical architecture introducing the veNB concept which consists of a RANaaS instance running on a cloud platform and one or several iSCs which are physically deployed where traditional small cells would be [14], [66]. In order to allow seamless integration in the existing 3rd Generation Partnership Program (3GPP) Long Term Evolution (LTE) architecture, the veNB appears as a classical evolved node B (eNB) in the network by terminating the S1 and the X2 interfaces.

The iSCs usually implement the lower part of the protocol stack and are connected through the J1 interface to the cloud architecture which hosts the remaining upper part of the protocol stack. In order to allow for cooperative processing among iSCs, the iSCs can exchange information directly with each other using the J2 interface. The actual split execution of the RAN functionalities between the two domains is managed by the iJOIN veNB controller (iveC) running in the RANaaS data centre. For more details on the logical system architecture refer to [16] and deliverable D5.3 [11].

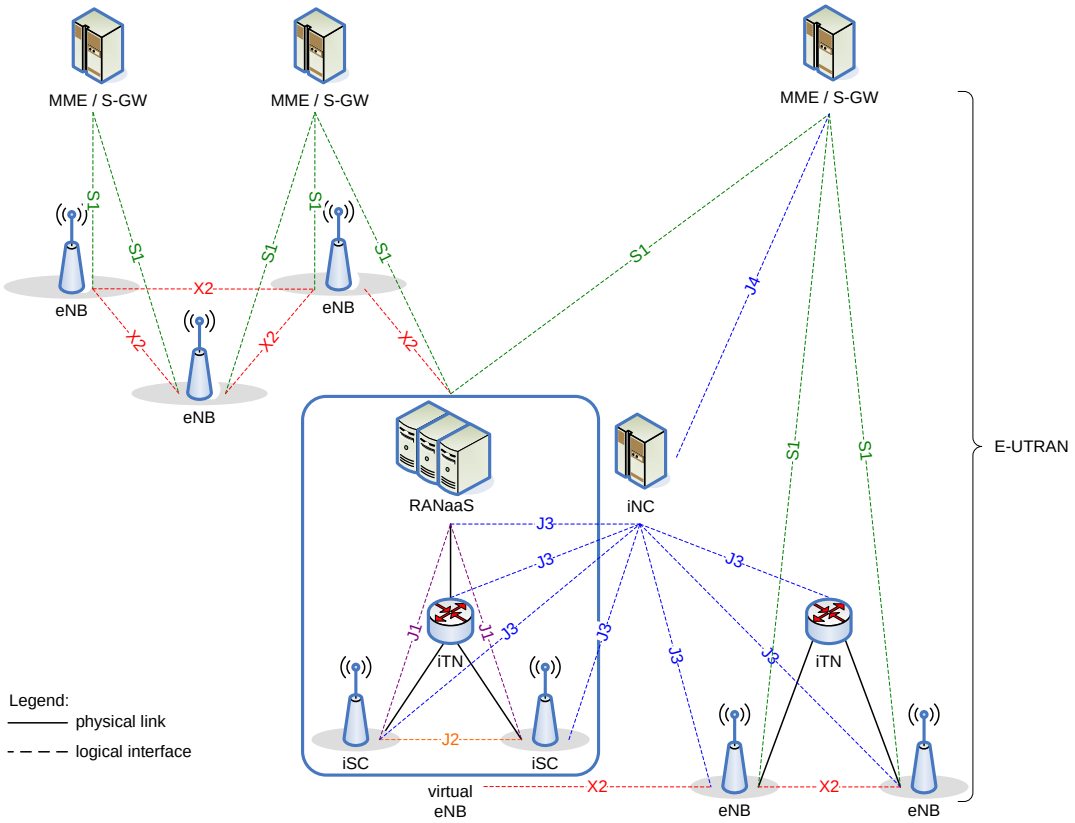


Figure 3-1: iJOIN Logical Architecture

In order to allow for joint RAN and backhaul optimization, the iJOIN Network Controller (iNC) is added to operate the routing among the iJOIN Transport Nodes (iTNs). To enable the software defined network (SDN) based controller solution, the iNC is connected via logical J3 interfaces with the veNB components and by the J4 interface to the core network. Further details are provided in D4.2 [7] and [65], [67].

Figure 3-2 presents the veNB integration within the RANaaS cloud architecture as described in detail in Section 5.2.1 of D5.3 [11]. The upper processing parts of the veNB are executed within iJOIN RAN virtual Processing Units (iRPUs) in order to implement the specific functions according to the selected functional split. This functionality can be cell, user, or radio bearer-based. The iveC is in charge of managing the veNB (interface toward the core network for legacy support) and iRPUs handling.

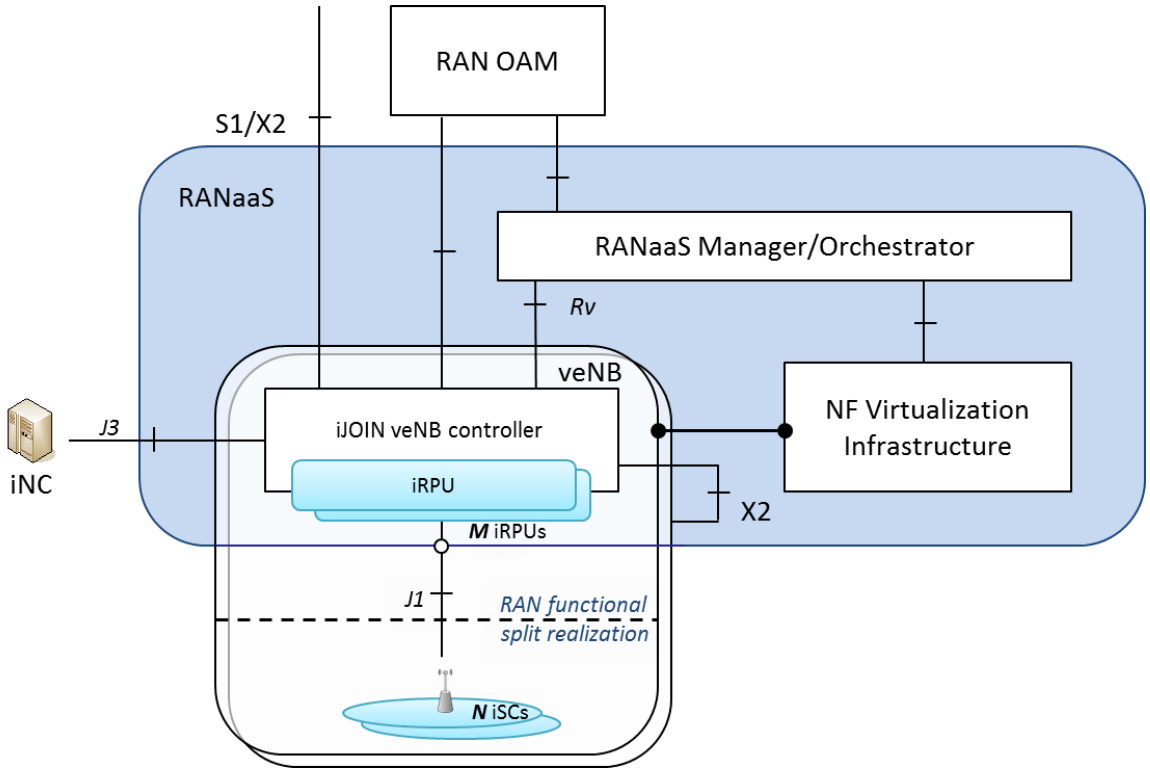


Figure 3-2: RANaaS Cloud Architecture

In terms of WP2 perspective, where functional splits are investigated at the physical layer, the iRPUs will usually implement a given physical processing for a subset of user equipments (UEs), e.g. single user decoding, multi-user joint detection, or single-user multi-point beamforming. Therefore, the iRPUs will usually be more UE-based than cell-based. Messages going to/coming from the iSCs can be forwarded from/to a dedicated iRPU for non-cooperative processes or from/to a common iRPU for cooperative processes, thus covering the various candidate technologies investigated within iJOIN.

3.1.2 Functional Architecture

The dense deployment of small cells naturally leads to strong interference between neighbouring cells. Sophisticated PHY processing approaches exploiting the introduced coupling of neighbouring cells by using joint transmission techniques for the downlink (DL) and joint reception approaches for the uplink (UL) are necessary. These approaches may operate fully centralized within the RANaaS or in a distributed fashion by exploiting distributed processing among iSCs. Based on the available BH resources, BH topology, processing capability and the current communication needs, the iveC may select the appropriate approach for the considered set of candidate technologies.

In deliverable D2.1 [1], we introduced seven promising CTs which expand the common PHY layer processing by introducing additional functionality for the RANaaS system concept. Their final definitions are provided in deliverable D2.2 [2] and performance evaluations are given in Section 4 of this report. The CTs are listed in Table 3-1, where the specific functionality addressed is also indicated. In particular, CT2.1 and CT2.2 implement different distributed as well as centralized approaches for joint multi-user detection in the UL. CT2.3 addresses the joint processing of the RAN and the BH link assuming low rate BH connections between cooperating iSCs. CT2.4 and CT2.5 investigate centralized and decentralized approaches for joint transmission in the DL. The impact of the functional split on the BH data rate is elaborated by CT2.6 and CT2.7 provides a joint forward error correction (FEC) coding scheme for the RAN and the BH link assuming millimetre wave (mmWave) technology as an example.

Table 3-1: iJOIN PHY Layer Candidate Technologies

CT	Topic	Abbreviation	Function
2.1	In-Network Processing	INP	Distributed UL Detection
2.2	Multi-Point Turbo Detection	MPTD	Distributed / Centralized UL Detection
2.3	Joint Network-Channel Coding	JNCC	Joint RAN/BH processing for UL Detection
2.4	Sum-Rate and Energy-Efficiency Metrics of DL COMP with backhaul constraints	CoMP	Centralized DL CoMP
2.5	Partially Centralized Inter-Cell Interference Coordination	ICIC	Distributed DL CoMP
2.6	Data Compression over RoF	RoF	Functional Split and BH processing for UL and DL
2.7	Millimetre wave backhauling	mmWave	mmWave / coding for UL Detection

Figure 3-3 depicts the WP2 specific functional architecture and indicates the interaction of the considered CTs with PHY basic functions on the RAN and the BH as well as medium access control (MAC) functionality in WP3 and network functionality in WP4.

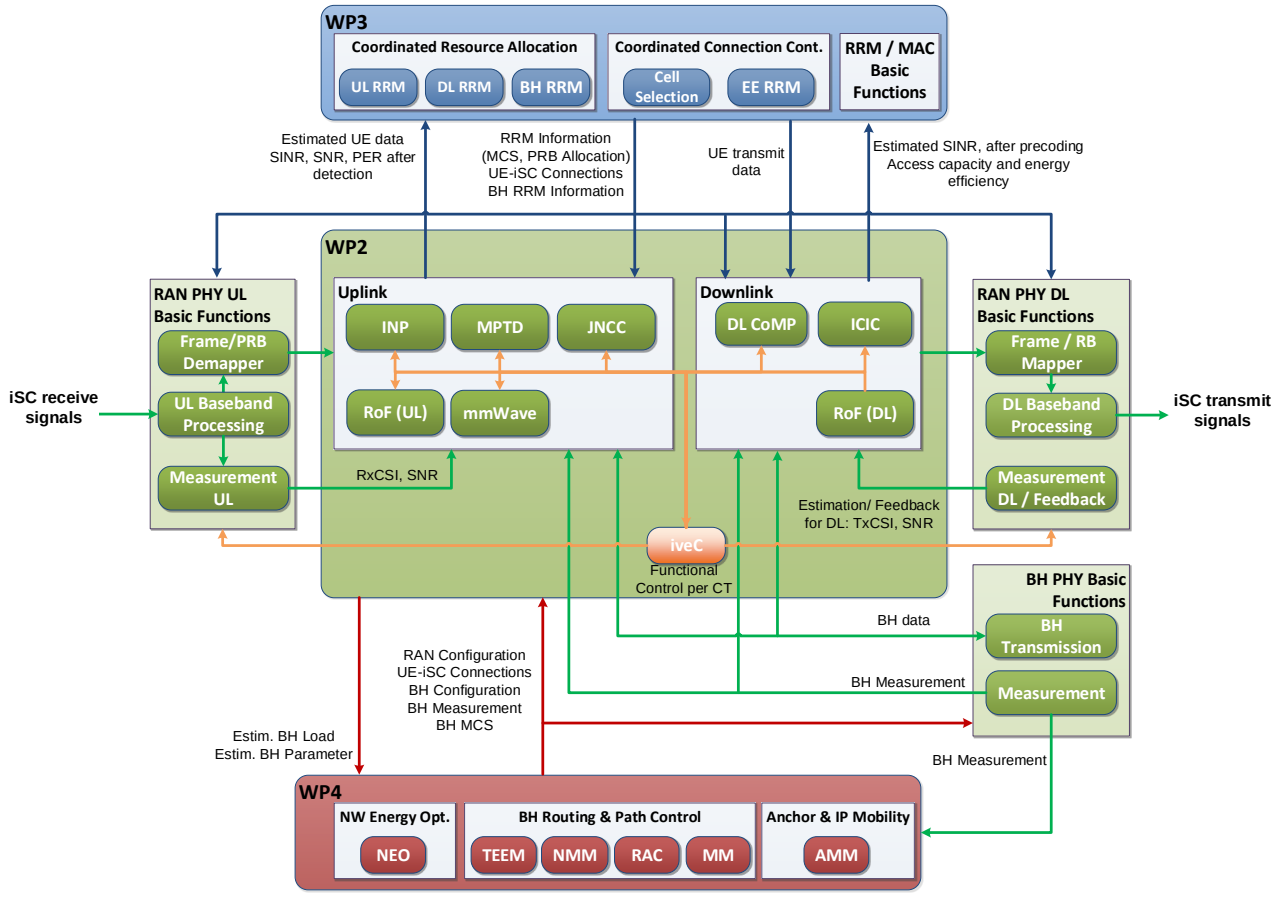


Figure 3-3: Functional Architecture of PHY processing

The "RAN PHY UL Basic Functions" block comprises the first common steps of the LTE UL processing chain. The received signals at the iSCs contain the local observations of the user messages. In the "UL Baseband Processing" block the receive signals at the iSCs are sampled to inphase/quadrature (I/Q) signals and converted to frequency domain (FD). The optional "Frame/PRB Demapper" forwards only the receive signals of allocated physical resource blocks (PRBs) to the uplink CTs based on the available radio resource

management (RRM) information provided by the MAC layer. Additionally, the block "Measurement UL" estimates the receive channel state information (CSI) and the receive signal to noise ratio (SNR) in order to deliver this information to the CTs as well. It should be noted, that the execution of these basic functions really depends on the functional split, which will be discussed in Section 3.2.1. In general within WP2, the CTs have some lower veNB functionality executed in the iSCs and some higher veNB functionality implemented in the iRPUs running on the cloud platform. The required exchange of signals over the backhaul links is organized by the iNC in cooperation with the iveC. This cooperative control also requires BH information provided by the "BH Measurement" block. Based on the available UL signals, the CTs perform an estimation of the transmitted information signals and provide the estimates to higher layer functionality, e.g. WP3 CTs. For the RRM, either common LTE functions or WP3 CT-specific functions are used.

For the DL the principle processing is done accordingly. In case of time division duplex (TDD) transmission mode, the DL oriented CTs receive CSI estimates from the corresponding UL channels. In case of frequency division duplex (FDD) explicit feedback schemes are necessary. Based on this measurement and the RRM information from WP3, the user data is processed for joint transmission. After mapping to resource blocks, the DL baseband processed signal is transmitted over the iSCs. Again, either general LTE RRM functions or WP3 CT-specific RRM functions are used.

It should be noted, that CT2.4 provides also guidelines for the RANaaS manager to decide how many adjacent iSCs should be combined for joint transmission in case of DL coordinated multi-point (CoMP). To decide the optimal veNB size, network deployment information like inter site distance (ISD), iSC density, UE density, and channel properties as well as information about the available computational resources from the RANaaS manager are required as discussed in Section 4.4.

3.2 Functional Split and Virtual eNB

One of the main benefits of the network architecture proposed by iJOIN is the ability to flexibly assign functionality to either the iSCs or the RANaaS platform within a virtual eNodeB. This functional split can be different for the various iRPUs within one veNBs depending on both location and time according to, e.g. the traffic demand, backhaul technology or the deployment scenario. As discussed in D2.2 [2], the actual split has some implications on the processing needs for iSCs and the RANaaS platform, the reliability and latency requirements of the backhaul links (i.e. iSC-iSC and iSC-RANaaS links), and the backhaul load (associated to the payload and the signalling information) that needs to be transmitted on the backhaul connection between the iSCs and the RANaaS platform. These implications result in a number of decision factors that have to be taken into account to decide the actual splits executed in the veNBs as discussed in more detail in D5.2 [10] and D5.3 [11]. In this section, we review the functional split options that are most relevant from WP2 perspective for both downlink and uplink and provide exemplarily some backhaul load calculations using the formulas derived in D2.2 [2] and summarized for completeness in Appendix II.

3.2.1 Functional Split Options

Within iJOIN, four main functional split options A, B, C and D that split between local processing in the iSCs and central processing in the RANaaS platform [15], [17] have been defined. These splits are depicted in Figure 3-4 and are further described in D5.2 [10]. From the PHY-layer perspective, split A, split B, and split C are most interesting, as for all splits above C the PHY layer is terminated in the iSC and thus it forms the transition to the MAC layer being considered in D3.2 [4]. Thus, all splits on MAC layer will be interpreted as split C from PHY layer perspective.

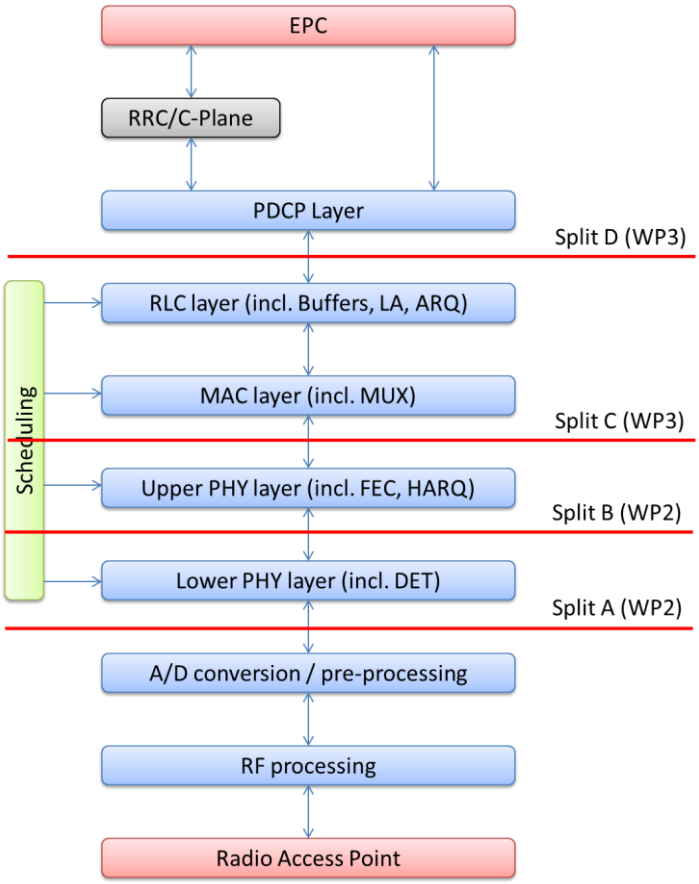


Figure 3-4: Project-wide functional split options

In WP2 we additionally investigate more detailed variants on how to split the PHY layer processing. In particular, for the functional split option A three variants labelled as A.1, A.2 and A.3 are identified. Similarly, for the split option B the two variants B.1 and B.2 are considered, while for split option C the two variants C.1 and C.2 are identified. Figure 3-5 shows these variants and the main functional blocks of the LTE DL/UL baseband signal processing chain, for which a detailed description is provided in D2.2 [2]. In the DL, the information bits processed by the radio link control (RLC) and MAC layers are first subject to FEC encoding, rate matching and scrambling. For simplicity, these last two operations are not represented by separate blocks and thus it can be assumed that they are executed in the FEC encoding block. The encoded bits are then converted to complex symbols in the modulation block and subsequently precoded through the multiplication with a specific precoding vector or matrix. The complex signals after precoding are then mapped on the orthogonal frequency division multiplexing (OFDM) subcarriers of the different transmit antennas in the mapping block. After that, the signals are converted from frequency to time domain by means of an inverse fast Fourier transform (IFFT) operation. The IFFT output after oversampling and low pass filtering, not indicated for simplicity in the figure, is then converted to the analogue domain by the digital to analogue (D/A) conversion block.

The UL signal processing operations are basically complementary to the DL ones. The only difference is the equalization, which is specifically done only in UL to compensate the frequency selectivity of the radio channel. In general the equalization requires also performing channel estimation that, still for simplicity, is considered to be executed in the equalization block.

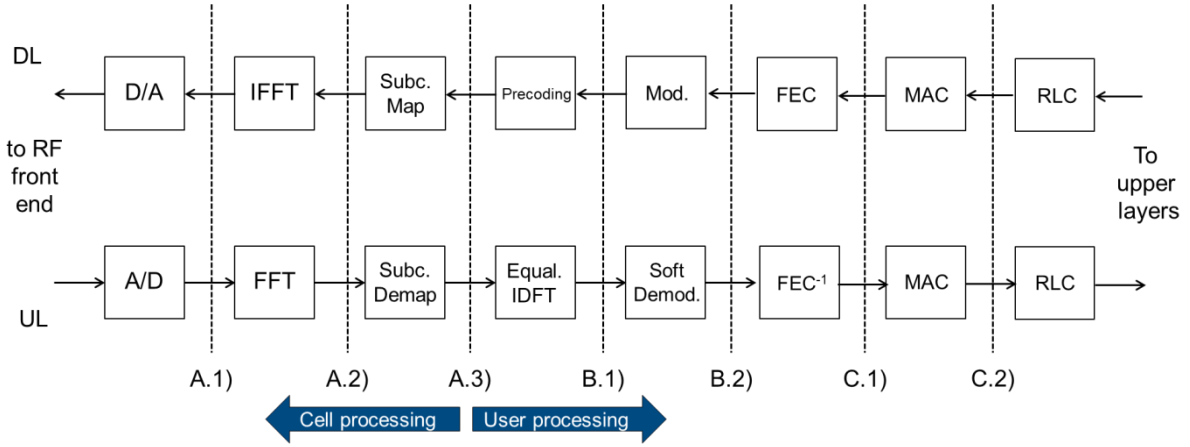


Figure 3-5: PHY-layer detailed functional split options

With the split option A.1 the I/Q signal samples in the time domain (TD) are transmitted between the RANaaS platform and the iSCs. This split option is used in centralized RAN (C-RAN) architectures where the remote radio heads (RRHs) are connected to the central unit by means of common public radio interface (CPRI) [47]. The main drawback of this split option is that the required BH capacity is very high and static, i.e. it does not depend on the actual user traffic. On the other hand, there are no limitations regarding the type of centralized processing that can be performed. By moving the IFFT/FFT operations at the iSC, as in the split option A.2, there is a significant reduction of the required BH capacity by about 3-4 times as demonstrated in the subsequent exemplary calculation, mainly due to the lower peak-to-average power ratio (PAPR) of the signals in frequency domain that reflects in a lower quantization resolution. Furthermore, with the transmission in the frequency domain the guard subcarriers at the band edge and the cyclic prefix (CP) do not need to be transmitted on the BH. However, also for the option A.2 the required BH capacity is static and determined by the system bandwidth, number of antennas per sector, number of carriers, and number of sectors per site. By further moving the resource mapping/demapping at the iSCs side, as in the option A.3, the required BH capacity scales with the actually occupied resource blocks, allowing for exploiting the statistical multiplexing gain. It must be noted that the options A.2 and A.3 enable the same physical layer centralization schemes of the C-RAN architectures and share the same BH latency requirements, in the order of few hundred of microseconds. Besides, as shown in Figure 3-5, the split option A.3 represents also the boundary between cell processing, where all the user signals are processed together, and user processing where instead the various user signals are extracted and processed separately.

In case of split option B.1 equalization and precoding are executed locally, while FEC encoding/decoding and all functionality at higher layers are still performed centrally. It follows that joint detection and centralized precoding techniques cannot be implemented, while distributed detection and/or precoding is possible. The required BH capacity scales now with the number of spatial layers instead of the number of antennas, because multiple-input multiple-output (MIMO) processing is performed locally. This reduces the required BH capacity significantly for users with low rank channels. The BH latency requirement is dominated by the HARQ processing in the uplink, which requires that the overall processing is finalized after 3 ms. For very fast moving UEs, the coherence time of the CSI required for downlink link-adaptation could also become limiting, although in the current implementation of LTE the link adaptation works on a too slow timescale for this to have an impact. With split option B.2 also the modulation/soft-demodulation operations are executed locally so that soft (uplink) or hard (downlink) coded bit are transmitted on the BH. This further reduces the required BH capacity as the rate now depends on the modulation scheme and thus ultimately on the signal to interference and noise ratio (SINR) experienced by the users. Further, with split option C.1 all physical layer processing is performed locally such that only Layer 2 and 3 functionality is centralized. The centralization gain may come in this case from joint scheduling, interference coordination, and path management techniques. The required BH capacity is closely tied to the actual user throughput determined by the user channel quality. Finally, with split option C.2 also the MAC functionality is executed locally. It must be noted that for the option C.1 the hybrid automatic repeat request (HARQ) procedure is managed in the RANaaS platform while for C.2 it is managed at the iSC. Hence, in case of split option C.1, the BH latency requirement is dominated by the HARQ processing, while for the option C.2 it can be relaxed but may not exceed some tens of ms.

The choice of the actual functional split shown in Figure 3-5 depends upon the deployed BH technology, the timing constraints within the protocol stack and the possible centralization gains [17]. Regarding the different technologies that can be employed in the backhaul, a thorough analysis is performed in Section 4 of D4.2 [7], providing a classification in terms the latency and throughput imposed by each technology. For reference, the derived table listing relevant backhaul technologies is provided in Appendix I. Based on this classification and on the backhaul load calculation performed in D2.2 [2] for the different functional splits, in Figure 3-6 it is shown the grouping between the backhaul technologies and the preferred functional splits that have been applied in the iJOIN concept evaluation within the different WPs.

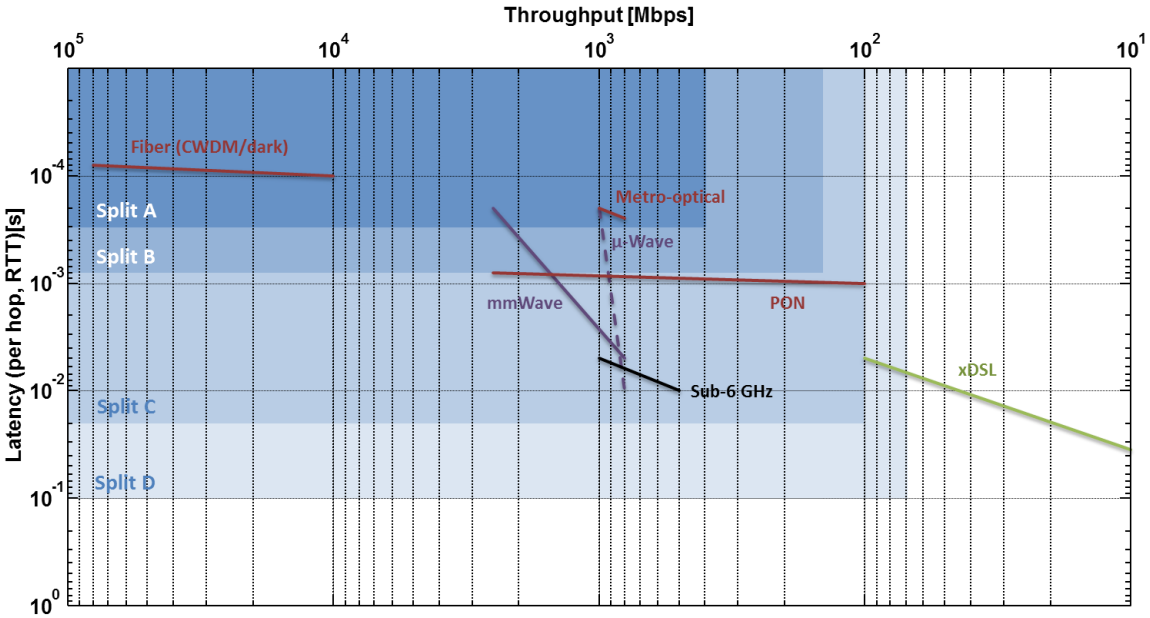


Figure 3-6: Ordering of backhaul technologies and preferred functional splits

For sake of simplicity, the functional splits for one iSC-RANaaS link are analysed in the following. However, this analysis is also important for more general systems with distributed processing among iSCs as applied by some CTs. The split options per CT are described in detail in Section 4, while the corresponding BH rates have been derived in D2.2 [2].

Numerical Example

The BH throughput for the different functional splits is calculated in the following using the equations provided in Appendix II based on the derivations in D2.2 [2]. For this exemplary evaluation of the BH throughput we consider the downlink and the uplink of a 2x2 MIMO LTE cell. The system bandwidth B is assumed equal to 10 MHz in order to be aligned with the iJOIN common scenarios (CS), while the calculation for $B = 20$ MHz has been included in D2.2 [2]. The values of the parameters used in the calculations are summarized in Table 3-2. The exemplary BH throughput has been calculated assuming the usage of the 64-points quadrature amplitude modulation (QAM). In the LTE/LTE-A case the support of the 64-QAM modulation for the uplink is conditioned by the UE capabilities, as specified in [50]. In this analysis, we assume a full utilization of the transmission resources (i.e. PRBs) on the radio interface. In case of partial utilization of the transmission resources the required BH bandwidth for split options A.3 and above scales down proportionally. Finally for the split options A.2, A.3 and B.1 a slightly higher quantization resolution for the uplink (9 bits) compared to the downlink (7 bits) is assumed in order to cope with strong interfering signal that can be received in the uplink, as postulated also in [51].

The different functional splits are compared in terms of message size related to the payload. As the message frequency is the same for all the functional splits (i.e. $1/T_{SUB} = 1$ kHz), the message size is also indicative for the BH throughput without including the BH coding overhead γ.

Table 3-2: System parameters used in the numerical example

Parameter	Exemplary Value	Parameter	Exemplary Value	Parameter	Exemplary Value
B	10 MHz	$N_{sc,1}^{j}$	600	N_T^{iSC}	2
f_s	15.36 MHz	N_{PRB}	50	N_R^{iSC}	2
OF	1	$n_{PRB,1}^{j}$	50	Q_m	6
N_{FFT}	1024	N_{SYMB}^{SUB}	14	N_L	2
N_{CP}	72	T_{SUB}	1 ms	N_Q downlink	15 (A.1) 7 (A.2, A.3, B.1) 1 (B.2, C.1)
N_{sc}	600	T_s	66.6 µs	N_Q uplink	15 (A.1) 9 (A.2, A.3, B.1) 4 (B.2) 1 (C.1)
N_{UE}	1	TBS_1	36696 bit	TBS_2	36696 bit

Figure 3-7 shows the result of the application of the equations provided in Appendix II with the exemplary system parameters of Table 3-2. For completeness, the results in Figure 3-7 are also tabulated in Appendix II. The bar diagram provides the required BH throughput for the payload as a function of the considered functional split options. The red bars show the calculation for the downlink with and without the overhead γ introduced by the BH coding techniques that may be used for error detection or synchronization purposes, while the blue bars show the same calculation for the uplink. For simplicity the BH coding overhead is assumed equally as $\gamma = 4/3$ for all functional split options [2]. As expected, the required BH throughput decreases when shifting the split section from the lower PHY layer towards the upper PHY. The only exception here is the split option B.2 for the uplink, where the transmission of the log likelihood ratios (LLRs) increases a bit the BH throughput compared to option B.1 due to the considered 64-QAM modulation. It must be remarked again that for the split options where the resource mapping/demapping is performed at the iSC side (i.e. A.3, B.1, B.2 and C.1) there is an inherent and further throughput reduction in case of partial PRB utilization as discussed in Section 3.3. Note, that functional split C.2 is not explicitly considered, as the difference to C.1 is negligible in terms of BH data rate.

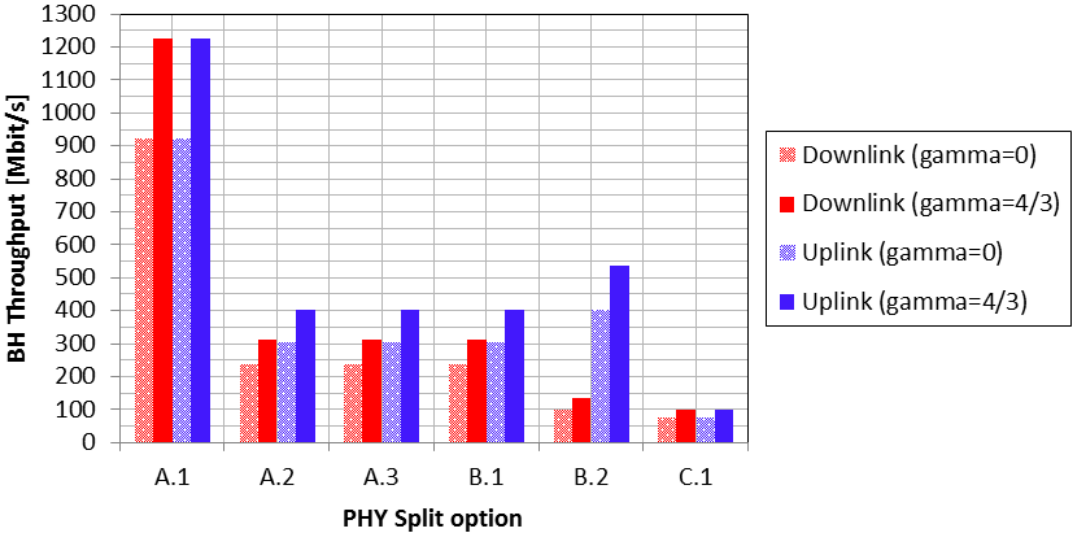


Figure 3-7: Exemplarily required BH throughput for different functional split options (LTE cell with 2x2 MIMO, B=10 MHz bandwidth and full PRB utilization)

It should be noted that the actual BH throughput depends on specific values assumed for the various system parameters and, in particular, on the system bandwidth B and the number of quantization bits N_Q of the signals transmitted on the BH. Correspondingly, the rates derived in the literature differ. Thus, we provide subsequently a brief comparison of the results provided by NGMN (next generation mobile networks) [51], Alcatel-Lucent [55] and SCF [63] for the analysis of the different functional split options at the physical layer. In these documents also evaluations of the BH throughput are provided for some of the functional splits analysed in the following. In particular the BH throughput values provided in [51] are referred to a $B = 20$ MHz LTE radio site with 3 sectors and 4 antennas per sector. The overhead is taken into account and includes control, signalling and protocol overhead. The number of quantization bits for split option A.2 is assumed equal to 7 bits for DL and 10 bits for UL. The BH throughput values provided in [55] are referred to a 20 MHz LTE cell with 2x2 MIMO and assuming a 33% overhead for both split option A.1 and A.2. The number of quantization bits for split option A.2 is assumed equal to 8 bits for both DL and UL. As reported in both [51] and [55], a 33% percent overhead (i.e. 16/15 for CPRI overhead and 10/8 for 8B/10B line coding) relative to the base rate is a typical assumption for CPRI (i.e. split option A.1). For the other split options, the overhead to be added depends on the networking protocol and may be lower than the one considered for CPRI. Also in [63] a 20 MHz LTE cell with 2x2 MIMO is assumed. In this case a fixed quantization of 16 bits for each I/Q signal component is assumed for both split options A.1 and A.2.

Table 3-3: System parameters and BH throughput evaluation from literature

Parameter	NGMN [51]	Alcatel-Lucent [55]	SCF [63]	iJOIN
B	20 MHz	20 MHz	20 MHz	10 MHz
N_Q downlink	15 (A.1) 7 (A.2, A.3)	15 (A.1) 6 – 8 (A.2)	16 (A.1 – B.1) 1 (B.2)	15 (A.1) 7 (A.2, A.3)
N_Q uplink	15 (A.1) 9 – 10 (A.2, A.3)	15 (A.1) 6 – 8 (A.2)	16 (A.1 – B.1) 8 (B.2)	15 (A.1) 9 (A.2, A.3)
Configuration	3 sector LTE site with 4x4 MIMO	Single LTE cell with 2x2 MIMO	Single LTE cell with 2x2 MIMO	Single LTE cell with 2x2 MIMO
BH throughput (DL)	14.7 Gbit/s (A.1) 3.02 Gbit/s (A.2)	2.46 Gbit/s (A.1) 716 Mbit/s (A.2)	2.46 Gbit/s (A.1) 1.07 Gbit/s (A.2)	1.23 Gbit/s (A.1) 314 Mbit/s (A.2)
BH throughput (UL)	14.7 Gbit/s (A.1) 4.78 Gbit/s (A.2)	2.46 Gbit/s (A.1) 716 Mbit/s (A.2)	2.46 Gbit/s (A.1) 922 Mbit/s (A.2)	1.23 Gbit/s (A.1) 403 Mbit/s (A.2)

Summary

In Table 3-4 a summary of the functional split options and the corresponding properties is provided. It must be noted that the round trip time (RTT) requirements given in Table 3-4 accounts for the sum of the round trip BH delay and the processing delay. It should be noted, that for split option A.1 only the CPRI requirement for BH latency is considered here.

Table 3-4: Summary of functional split options

Split Option	Lowest functionality centralized	Impact on LTE	RTT requirements	Centralization schemes	Advantage	Disadvantage
A.1	FFT/IFFT	small	5 µs (BH latency for CPRI)	Fully centralized PHY processing	- Very simple iSCs - All centralized processing options	- Very high BH bandwidth - Strict latency requirements
A.2	Subcarrier mapping	small	CSI constrained	Fully centralized PHY processing	- Reduced BH load - All centralized processing options	High BH bandwidth
A.3	Precoding/ Equalization	- Sync. signalling between iSC/ RANaaS req. - Subc. mapping signalling req.	CSI constrained	Fully centralized PHY processing of payload	- Reduced BH load - All centralized processing options - BH load proportional to radio interface load	Additional signalling required
B.1	Modulation/ demodulation	Precoding signalling req.	3 ms (HARQ)	Centralized modulation/ detection	- Reduced latency requirements - Joint detection possible - BH load proportional to the radio interface load	- No centralized precoding - Turbo Equalization unfeasible - Additional signalling required
B.2	FEC	small	3 ms (HARQ)	Centralized coding/decoding	- Joint decoding/ NC possible - BH load proportional to the radio interface load	- No centralized precoding - Turbo Equalization unfeasible
C.1	MAC	small	3 ms (HARQ)	Centralized MAC	- Greatly reduced BH load - BH load proportional to the radio interface load	- No joint PHY processing
C.2	RLC + long-term scheduling	Split scheduling, dedicated signalling for resource allocation	Several frames (10ms each)	Interference mitigation, cooperative schemes	- Reduced BH latency requirements	- No coordinated fast scheduling

3.2.2 Centralized Processing in Cloud

Depending on the functional split, some lower part of the protocol stack is implemented in the iSCs consisting of field programmable gate arrays (FPGAs), application specific integrated circuits (ASICs) or digital signal processors (DSPs). The remaining upper part of the protocol stack is implemented in the iRPUs of the corresponding RANaaS data centre. These iRPUs shall be implemented in virtual machines running on a cloud platform. Thus, the software implementation of the functionality is not executed on "bare metal" (i.e. directly on a processor and other physical resources), but it is follows a virtualized implementation. Hence, an additional abstraction mechanism called hypervisor is laid between the application software implementing the iRPU and the physical hardware. This additional layer has an effect on the temporal behaviour of the application program. For example, the response times of the software running virtualised are increased in a non-deterministic fashion, leading to significant jitter. This effect is discussed in the deliverable D6.2 [13]. Compared to the "bare metal" case, the measured processing times showed a significantly higher variance. So far, it is not clear whether this effect stems from the measurement, or the processing time is actually increased. However, the strong indication is obtained that with centralized processing on a cloud platform, additional latency margins have to be provided for the processing in order to ensure that the critical timeouts are not exceeded.

For functional splits A and B, the FEC decoding is performed in the RANaaS and needs to be implemented in iRPUs. Since this functionality is the computationally most expensive functional block in uplink PHY processing chain, the computational load of the iRPU is dominated by the turbo decoder [62]. In order to investigate its computational complexity and assess the influence of virtualization on its performance, simulations and theoretical analyses have been performed. Initial results were presented in deliverables D2.2 [2] and D6.1 [12].

Extending these works further, in [28] and in the accompanying D6.2 [13] the actual performance of a 3GPP LTE compliant turbo decoder running on a practical commercial cloud platform has been measured. It was confirmed that due to the software implementation and the underlying virtualization the variance of the decoding time is increased compared to the "bare-metal" case, in addition to the non-deterministic behaviour of the turbo decoder requiring a random number of iterations for successful decoding.

In order to meet the strict timing constraints for FEC decoding, the computational aware scheduler [11] is proposed. As demonstrated in deliverable D6.2 [13], it is able to reduce the decoding time and thus the computational load by proper selection of modulation and coding schemes (MCSs), effectively avoiding computational outage.

3.3 Joint RAN/BH Optimization

A traditional eNB is dedicated to – and therefore optimized for – a good RAN link performance. Howevere, as the veNB introduced by iJOIN contains distributed nodes connected by BH links (J1 and J2), it is consequent to consider also a joint optimization of both RAN and BH links. Within iJOIN, several approaches have been investigated that are summarized here.

BH Multiplexing Gain

In Section 3.2.1, we analysed how the flexible functional splits impact the data rate requirements on the J1 BH link. Additionally, varying traffic on the access links can also reduce the required aggregation capacity on the BH by means of statistical multiplexing. The existence of this multiplexing gain is well known; however we will present a quantification of the effect in the following and show the influence of an optimal functional split. A more detailed investigation can be found in [17].

For this investigation, we combine the expressions for the required rates from Section 3.2.1 with statistics of parameters obtained from a calibrated 3GPP LTE system level simulator. The main time-variant parameters that influence the BH capacity are the load and the MCS used, i.e. the utilization of transmission resources and the channel quality of the users. Figure 3-8 shows the cumulative distribution function (CDF) of the BH data rates of the three exemplary splits A.1, B.1 and C.1 each for a fully loaded system (i.e. the case that all PRBs are utilized), and for a variable loaded system. The system parameters are aligned with those in Section 3.2.1 and the considered scenario is described in more detail in [17].

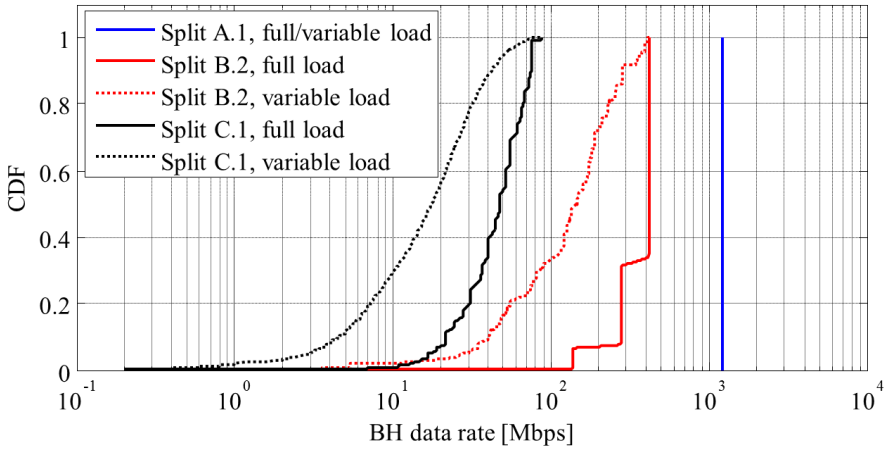


Figure 3-8: CDF of BH data rates for three different split options with full and variable load

The BH data rate of split A.1 is independent of the load leading to a constant BH rate for both PRB allocations. Split B.1 varies only with the load while split C.1 also depends on the MCS and thus also varies for a full load. The variance of the CDFs can be exploited for a multiplexing gain. A widely accepted strategy for dimensioning the BH network capacity is to guarantee a certain outage percentile for the BH data

rate (per base station), e.g. the Q_{99} which implies a packet loss rate of less than 1%, and to divide the outage percentile by a safety factor ϵ, e.g. 0.9 [53]. Hence, it is guaranteed that the actual load does not exceed 90% of the transport network capacity with 99% probability. However, the BH traffic of multiple iSCs is aggregated at, e.g. iTNs before it is processed at the RANaaS. At each aggregation point, we can exploit a statistical multiplexing gain which implies that the peak-demand of the aggregated traffic leaving the aggregation point is less than the sum of the peak-demand of the individual traffic arriving at the aggregation point. The multiplexing gain can be exploited in any scenario with non-static traffic and can be well determined using the central limit theorem. Figure 3-9 illustrates the scaling of the required BH data rates within the number of aggregated base stations. The multiplexing gain is the gap between the solid curves (without multiplexing) and the dashed curves (with multiplexing). As can be seen, there is no gain to be exploited for Split A.1 as the data rate does not vary. For split B.1, which only varies with the load, the multiplexing gain is in the order of a factor 2. For split C.1, which has a very high variance as it depends on the load as well as the MCS, the multiplexing gain can be larger than factor 3. Additionally, the throughput of the different BH technologies as discussed in Section 3.2.1 are illustrated above the figure.

Of course these results can slightly differ based on the considered scenario. However, the fact remains that the multiplexing gain increases for a higher variance in BH data rates. To exploit the multiplexing gain as far as possible, it would be beneficial to aggregate cells with a great variance of traffic, potentially combining areas with different daily traffic profiles (like residential and commercial areas) and MCS distributions with sufficient variance. The flexible functional split can also increase the variance of BH traffic if it is adapted to the current network load.

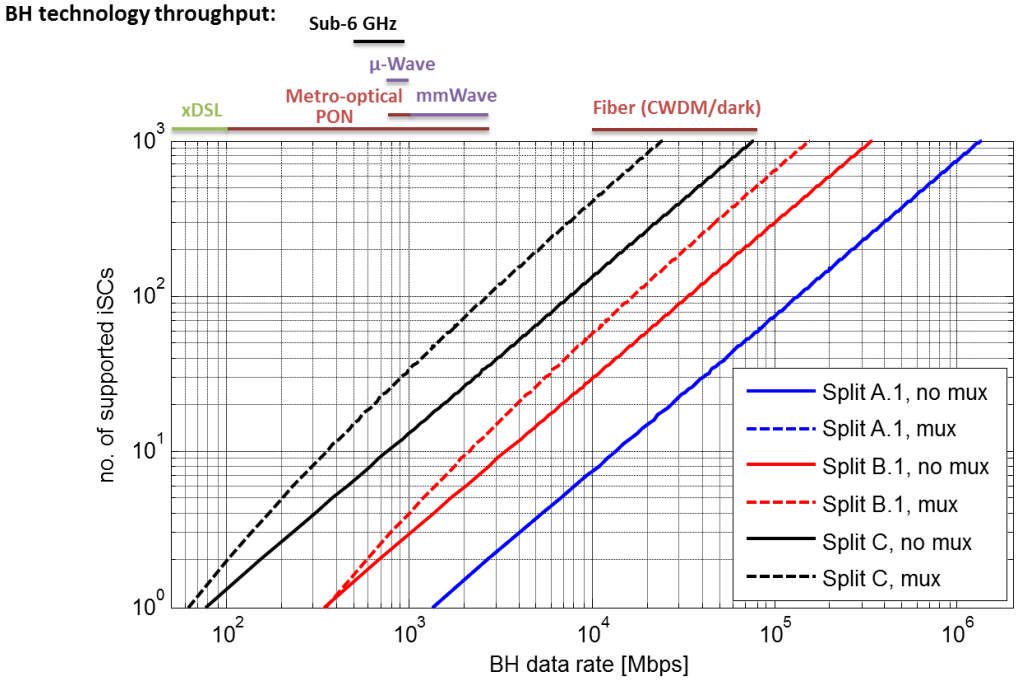


Figure 3-9: Number of supported iSCs per deployed aggregation capacity

Joint RAN/BH Coding

Different types of coding play an important role in any digital communication network, e.g. to increase reliability by means of forward error correction (FEC) or to use the deployed links as efficiently as possible by applying network coding. In current networks, the RAN and BH networks are viewed as separate entities. With the introduction of the veNB concept it is beneficial to facilitate coding across both network domains.

Forward error correction coding is used to increase the reliability of communication over noisy channels. In a conventional LTE network, a turbo code is employed on the RAN links while BH vendors offer separate proprietary solutions to ensure reliability on BH links, which is especially important for wireless BH links due to their higher attenuation. However, in the functional split options A and B, the RAN FEC code of the uplink is only decoded in the RANaaS and thus after the signal has been forwarded over the BH link. This allows combining the coding for RAN and BH in order to protect both transmission links simultaneously as investigated in detail in CT2.7 in Section 0. Instead of using separate FEC codes for RAN and BH as shown in Figure 3-10 a), it is also possible to use only one FEC encoder at the UE to protect the transmission of the

user data over the access and the backhaul link simultaneously as illustrated in Figure 3-10 b). This requires adapting the code rate of the RAN FEC to account for the quality of the RAN channel and the BH channel jointly. While removing the BH code entirely has benefits regarding latency and required hardware, a joint code can only protect the user data but not uncoded reference or synchronization signals. These uncoded signals would have to be encoded separately. Therefore, iJOIN developed advanced decoding schemes that allow for a forwarding of channel quality information or soft-bits between the RAN and BH link receivers that can improve the end-to-end performance both when a BH code is used or when it is skipped entirely.

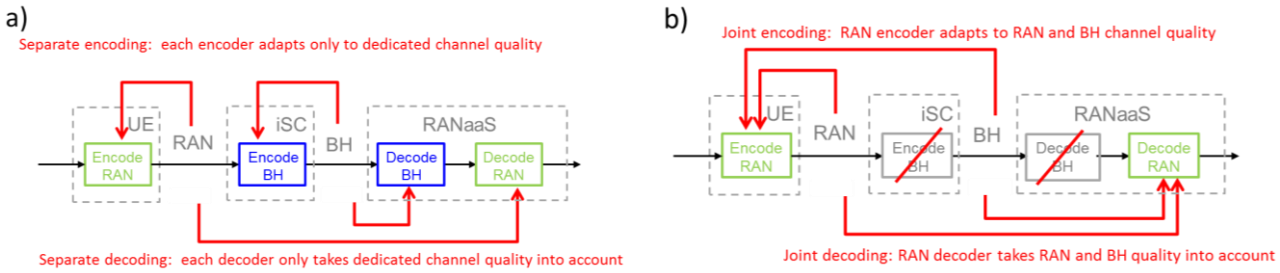


Figure 3-10: Separate coding for RAN and BH (left) compared to joint coding (right)

As a different aspect of coding, network coding can efficiently combine multiple messages (e.g. from different UEs or iSCs) at an intermediate node. Figure 3-11 shows a system composed of one iSC connected to the RANaaS platform via a wireless backhaul link. The corresponding iRPU of the RANaaS instance attempts to decode the messages sent by the active UEs. In a pure relaying system, the iSC plays the role of a relay to send all UE messages to the RANaaS instance via the backhaul link. Consequently, pure relaying is not efficient when the backhaul link quality is poor because it constraints the overall system capacity. In order to increase the system performance when the backhaul data rate is limited, a well-known coding strategy called Joint Network-Channel Coding (JNCC) was proposed in the literature [58]. The principal idea of JNCC is that the final destination attempting to decode all UE messages uses the direct links from UEs as well as the backhaul link which relaxes the backhaul rate requirements. In other terms, the RANaaS in Figure 3-11 extracts the information of all UEs using jointly the signals obtained from the intermediate node (iSC) via a backhaul link and directly from UEs due to the broadcast nature of wireless channels. JNCC is very important for systems with limited backhaul data rate because it allows decreasing the traffic on the backhaul link by combining RAN and BH at the decoding node.

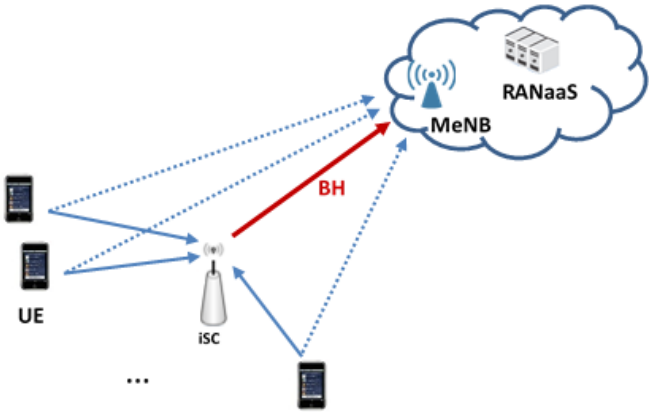


Figure 3-11: JNCC-based multiple access relay MeNB

Figure 3-12 shows the uplink area throughput (AT) as a function of the backhaul rate for a system composed of a RANaaS located in a three-sectoral Macro eNB (MeNB). In each sector, we consider the square scenario with one iSC connected to the RANaaS via a backhaul with rate D_{J1}. Moreover, 30 UEs are uniformly dropped inside the iSCs coverage area and are all associated to iSCs. The iSCs are dropped in a cluster with a radius equal to 50 m whereas the UEs are dropped with a radius equal to 70 m according to iSC deployment in the square scenario [10]. Simulation parameters concerning channel model, deployment parameters, transmit power, and antenna gains for the radio access points (RAPs) are based on 3GPP specifications [59]. Thus, in each sector, the system is similar to that in Figure 3-11. It can be observed that when the backhaul data rate D_{J1} is limited (less than 120 Mbit/s), JNCC brings noticeable gains compared to pure relaying strategy. This motivates the adoption of JNCC in systems characterized by a limited backhaul data rate.

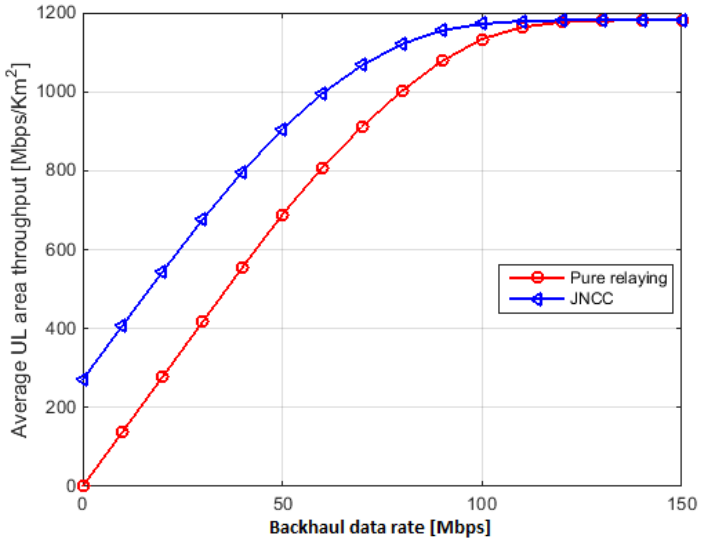


Figure 3-12: Average area throughput versus backhaul rate D_{J1} with and without JNCC

In CT2.3, the performance of JNCC is investigated as a function of the BH link quality and compared to that of pure relaying systems for different user association methods. Therefore, this CT provides a benchmark to quantify the gains using JNCC as a function of the BH link quality in order to identify the cases where JNCC is beneficial with respect to pure relaying.

Joint Precoder and Backhaul Design

In a network with a dense distribution of iSCs, managing interference between iSCs is particularly critical to achieve the benefits of network densification. With sufficient backhaul, it is possible to implement a centralized joint processing and fully suppress the interference by forming a so-called "virtual MIMO" or "network MIMO" system [29].

This comes however at a high cost in terms of backhaul rate and/or backhaul latency. In particular, the multi-user CSI has to be available at the RANaaS within a sufficiently short delay to allow for an efficient precoding. This leads to the need of designing robust precoding schemes taking into account the CSI feedback delay [32] and precoding matrix calculation delay. This robust precoder design is very well known and is the most common way to take into account the imperfections of the backhaul. Extensive work in literature exists on that topic due to the similarities with the conventional (single transmitter) MIMO broadcast channel, which has been heavily studied in the past decade [61].

Optimizing the number of cooperative cells as a function of the delay introduced in the network also allows to design a transmission scheme which adapts to the quality of the backhaul network. This corresponds to the approach followed by CT 2.4 as discussed in detail in Subsection 4.4 of deliverable D2.2 [2]. Theoretically, a larger number of jointly processing iSCs leads to better interference cancelation and higher spectrum efficiency can be achieved. However, a large number of cooperation cells can also cause a significant delay due to the CSI feedback and the precoding matrix calculation, which creates a mismatch between the actual channel coefficients and the outdated ones used for precoding. Consequently, performance degradation arises. In order to trade-off the interference and the delay introduced mismatch, a whole network that shares the same RANaaS platform with limited computational resources can be divided into several veNBs consisting of one or more iRPUs and a number of cooperative iSCs as investigated by CT2.4. The number of cooperating iSCs can be adapted in order to optimize the area throughput as discussed in D2.2 [2].

Another constraint for centralized processing is the restriction to transmit only quantized versions of the precoded signals from the RANaaS to the iSCs over the backhaul. With imperfect backhaul, the level of accuracy of the quantization impacts the design of the precoder. In particular, it is possible to take into account the deformations introduced by the quantization in the optimization of the precoder, thus jointly designing both aspects [30]. Interestingly, a weak backhaul to one iSC impacts the design of the precoded signal transmitted to the other iSCs to "compensate" the inaccurate transmission of this iSC.

Going beyond the limitations of centralized precoding, another opportunity is given by the utilization of J2 links between the iSCs and the precoding capabilities at the iSCs. Indeed, with J2 links of high quality, it is

then possible to directly exchange the global multi-user CSI between the iSCs and design accurate joint precoders in a distributed manner as investigated by CT2.5. Clearly, this requires to choose the appropriate functional split for the given backhaul topology. When the backhaul links are of heterogeneous qualities, a partially centralized precoder design is proposed in deliverable D2.2 [2] and its performance is evaluated in a realistic scenario in Section 4.5. The precoding algorithm consists in exploiting the high quality CSI available at some iSCs while adapting to the precoder design at the other iSCs having less accurate CSI available. This precoding scheme allows a smooth transition from non-cooperative distributed transmission to fully cooperative one as the number of iSCs having a high quality CSI increases. Another application for this transmission algorithm is to adapt to the computation capabilities of the iSCs in case iSCs having different computation capabilities.

Altogether, the strong potential of designing jointly the precoder and the backhaul network has been highlighted and state of the art solutions have been discussed. Interestingly, even though improvements could be realized over state of the art solutions, only preliminary solutions have been found due to the very high complexity of the optimization problem. This confirms the potential of this approach and that these topics need to be further investigated, both by academics in a theoretical level and by industrial to develop practical algorithms, in order to fully achieve the high gains promised.

Joint RAN / BH Energy Optimization

Another important aspect that can be improved through RAN/BH optimisation regards the energy efficiency of the system. While most effort in literature has been thrown on throughput maximisation, energy consumption at the backhaul network is equally important in order to realise the next generation green communication networks. Energy allocation between the backhaul and access link can heavily affect the network throughput [56], [57].

In general, to analyse system energy efficiency one should consider the total achieved desired utility (e.g. area throughput, etc.) and the sum of all power contributions in the network, which is directly related to the power usage of each network element over a given period of time. Specifically in the iJOIN architecture, the instantaneous total power consumption is a sum of contributions from the RANaaS instance, backhaul network and iSCs (see D5.2 [10] and [64] for a detailed energy model). In addition, the total power consumption at RAN and backhaul will depend on different functional split options, transmission schemes (e.g. joint transmission (JT) and coordinated beamforming (CB) CoMP), backhaul technologies (e.g. microwave, fibre) and topologies (e.g. star, point to point (PtP)), complexity of access/backhaul link algorithms, etc. Considering any specific functional split implementation (in which case power consumption at RANaaS instance is constant and relatively independent of the backhaul and access network) approaches for joint RAN/backhaul energy optimization can be devised.

Regarding the RAN, power consumption is a composition of the power used for signal transmission and power consumption at sites, e.g. for active site cooling, signal processing, etc. For the backhaul network, however, power consumption highly depends on the backhaul technology used. Taking the microwave backhaul technology as an example, the total energy consumption consists of consumption on the switches and power used for transmission. In addition, implementation of cooperative schemes among iSCs will require more backhaul links, leading essentially to higher total backhaul power consumption.

The key point of the joint energy efficiency optimisation relies on the dependency between the power used for data transmission at the backhaul and access links and the achieved utility. If insufficient power is available for information transmission through backhaul links, they will become a bottleneck for the RAN. In that case, even if more power is used at the access links in order to achieve better throughput for example, performance will not improve. As an example approach for this concept, the idea of adaptive CoMP scheme selection has been investigated in D2.1 [1]. More specifically, energy efficiency for two of the most representative CoMP schemes JT and CB was analysed in the presence of backhaul capacity constraints. The results indicated that the different CoMP schemes should be selected in different scenarios in order to optimise the system energy efficiency.

4 Final Description and Evaluation of iJOIN PHY Candidate Technologies

In deliverable D2.1 [1] the basic ideas of promising physical layer approaches for the radio access and the backhaul of densely deployed small cell networks were introduced. These PHY CTs were defined in detail and preliminary results per CT were provided in deliverable D2.2 [2]. In this section, we will provide final updates for the CTs listed in Table 3-1 (also presented below as Table 4-1 for convenience) and present extensive performance evaluations.

Table 4-1: List of iJOIN PHY Candidate Technologies

CT	Topic	Abbreviation
2.1	In-Network Processing	INP
2.2	Multi-Point Turbo Detection	MPTD
2.3	Joint Network-Channel Coding	JNCC
2.4	Sum-Rate and Energy-Efficiency Metrics of DL COMP with backhaul constraints	CoMP
2.5	Partially Centralized Inter-Cell Interference Coordination	ICIC
2.6	Data Compression over RoF	RoF
2.7	Millimetre wave backhauling	mmWave

In the discussion of the iJOIN Architecture in Section 3.1 we already pointed out the key topics of joint detection for the uplink, joint transmission for the downlink and joint optimization of access and backhaul considered by WP2 CTs. Joint multi-user detection techniques for the uplink are investigated by CT2.1 “In-Network Processing” and CT2.2 “Multi-Point Turbo Detection”. The joint optimization of the FEC codes for the access links and the backhaul link applying network coding is considered by CT2.3 “Joint Network-Channel Coding”. Regarding joint transmission, CT2.4 “Sum-Rate and Energy-Efficiency metrics of DL CoMP with backhaul constraints” and CT2.5 “Partially Centralized Inter-Cell Interference Coordination” analyse different precoding approaches and interference coordination techniques with respect to complexity and backhaul constraints. Techniques to reduce the backhaul throughput for radio over fibre (RoF) links is the topic of CT2.6 “Data Compression over RoF”, whereas CT2.7 “Millimetre Wave Backhauling” addresses wireless backhaul links and joint access and backhaul FEC.

4.1 CT2.1: In-Network Processing

4.1.1 Final implementation of CT

This CT investigates distributed multiuser detection (MUD) in a dense deployment of iSCs using In-Network Processing (INP). Multiple iSCs observe several users in their coverage area and exchange some information over their J2 BH connections in order to jointly detect those user messages by an iterative algorithm. This can be accomplished by the so-called distributed consensus-based estimation (DiCE) algorithm [19] which was described in detail in D2.1 [1]. Variants of the DiCE algorithm like the reduced overhead DiCE (RO-DiCE) [22] or the Fast-DiCE [23] algorithm were developed in order to reduce the required communication overhead among the iSCs and to improve the convergence speed [24], [26], respectively. A further reduction of the communication overhead can be achieved by the augmented Lagrangian based cooperative estimation (ALCE) approach that avoids the exchange of some variables compared to DiCE [27]. To be more robust in case of erroneous BH links and to keep low communication overhead the variant priority-aided ALCE (PALCE) was developed as well [27]. All these algorithms were discussed in detail in deliverable D2.2 [2]. The received UE signals are processed using these algorithms in the iSCs and are then forwarded over the J1 connection to the RANaaS platform where the final processing of the UE data is performed in the associated iRPUs.

In principle, functional split B and C can be applied with this CT. In case of functional split B.1 the local estimates are forwarded by means of quantized soft-symbols, whereas for split B.2 quantized log likelihood ratios (LLRs) after demodulation are transmitted to the RANaaS platform where the turbo decoding is

performed. To limit the J1 rate, for each UE only one iSC needs to forward its estimates and different possibilities to choose the forwarding iSCs exist. For example the iSC with the best J1 connection or the iSC with the best local estimate for a specific user (e.g. determined by the lowest estimated BER) forwards the corresponding values. In the sequel, we will concentrate on split B.2 as it provides an easy way to estimate the BER using LLRs and it leads to less traffic compared to split B.1.

In case of functional split C the decoding is performed within the iSCs and only the estimated user messages are forwarded of the J1 links to the RANaaS platform. Here, split C.2 provides the advantage of allocating the HARQ process in the iSC which leads to relaxed requirements for the J1 latency. Additionally, a successful cyclic redundancy check (CRC) as part of the HARQ process would allow for an early termination of the iterative estimation of UE messages within the INP algorithm. For the sake of simplicity this possibility has not been considered subsequently, although it would reduce the computational complexity per iSC, the computational latency per user, and the total J2 BH rate. In principle, both functional splits B.2 and C.2 differ only in the J1 BH rate and the required J1 latency. As the estimation performance is the same, we will always discuss the results for the functional split B.2 implementation in the sequel.

In this section we provide performance results for the developed approaches considering an LTE uplink following the iJOIN scenarios. Centralized detection, where all receive signals are forwarded to a central processing node, acts as an upper bound benchmark for distributed procedures as all observations are incorporated into the detection process. In contrast, local detection acts as a baseline performance measure since each iSC detects the users on its own without any cooperation. For the overall investigation, the same MCS is assumed for all served users and either a complete overlapping of allocated resources or a simple partly orthogonal allocation is assumed. The specific allocation of resources for the UEs in combination with the presented INP algorithms is investigated in CT3.8 and described in D3.3 [5].

4.1.2 Evaluation of the CT

Evaluation Methodology

In order to investigate this CT in the overall iJOIN architecture, the INP algorithms are implemented in an LTE PHY uplink chain based on release 10 and link-level simulations are performed. Within the uplink processing chain the commonly used equalizer is substituted by the various INP algorithms which provide a distributed equalization of the received UE signals among cooperating iSCs. For the evaluations, we assume perfect channel knowledge at the iSCs and apply appropriate channel models for the different considered scenarios. The J2 backhaul links connecting cooperating iSCs are assumed to be perfect such that exchanged signals among the iSCs are received error-free. Investigations regarding erroneous BH links among iSCs can be found in [20] and in D2.2 [2]. As performance measures we analyse the area throughput (AT) based on the frame error rate (FER) and the corresponding required J1 and J2 backhaul rates.

In order to classify the performance of this CT, centralized detection incorporating all received signals is used as an upper performance bound. On the other hand, local detection based only on locally received information serves as the baseline. Due to its limited performance, the UEs located within one small cell are equally assigned on orthogonal PRBs. In contrast, for central and distributed detection, each UE occupies either the whole bandwidth or shares the bandwidth orthogonally with users in the same small cell depending on the investigated scenario. The gains are shown with respect to the specified baseline system in each scenario. For both centralized and local detection, linear minimum mean square error (MMSE) equalization is used which considers interference as an additional noise part. Correspondingly, the MMSE variants of the INP algorithms are applied for distributed detection as described in D2.2 [2].

Performance Results

Wide-area Continuous Coverage without External Inter-Cell Interference

For the wide-area continuous coverage scenario we assume the parameters defined in Section 4.3 of deliverable D5.2 [10] with 2 UEs per iSC. In this scenario, $N_{iSC} = 7$ iSCs are placed within 2 tiers covering $N_{UE} = 14$ UEs. The distance between two iSCs is 50 m and the total coverage area of the 7 small cells is $A_{veNB} = 0.011\ \text{km}^2$. For INP and central detection, all UEs are considered to be detected jointly among the iSCs such that no additional interference from outer cells is present. The access channel between UEs and iSCs is modelled by the urban micro (UMi) channel model using the "Pedestrian A" power profile and non-line of sight (NLOS) path loss [49]. Each iSC is equipped with $N_R^{iSC} = 2$ receive antennas while each UE has

$N_T^{UE} = 1$ transmit antenna. For central and distributed detection by INP algorithms, each UE uses the same MCS configuration and allocates the whole available bandwidth of $B = 10$ MHz. Consequently, all 14 UEs are completely interfering representing a worst-case scenario for MUD. In practise, the RRM would appropriately adapt the resource allocation to improve the overall performance. In case of local detection, each iSC detects only the messages of the two UEs located in its cell and the UEs of other cells lead to inter-cell interference (ICI) which is considered as noise within the local MMSE detector. Again, all UEs use the same MCS configuration. However, within each cell the UEs are scheduled on orthogonal PRBs, i.e. each UE occupies only 5 MHz of bandwidth per cell.

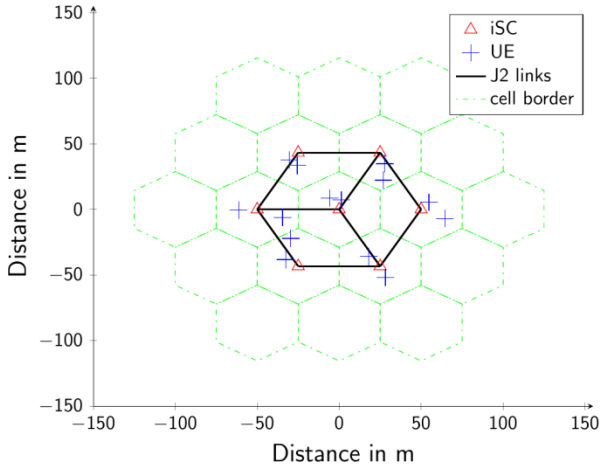


Figure 4-1: Considered wide-area continuous coverage deployment

Figure 4-1 depicts the considered dropping of UEs and iSCs where the black lines represent the logical J2 BH connections among iSCs. The different INP algorithms are then applied to the whole network of all connected iSCs. The 14 UEs are randomly placed in the seven inner hexagonal cells once and the deployment is then fixed for all simulations in order to avoid influences of arbitrary deployments on the performance investigations.

a) Local Detection

b) Central Detection

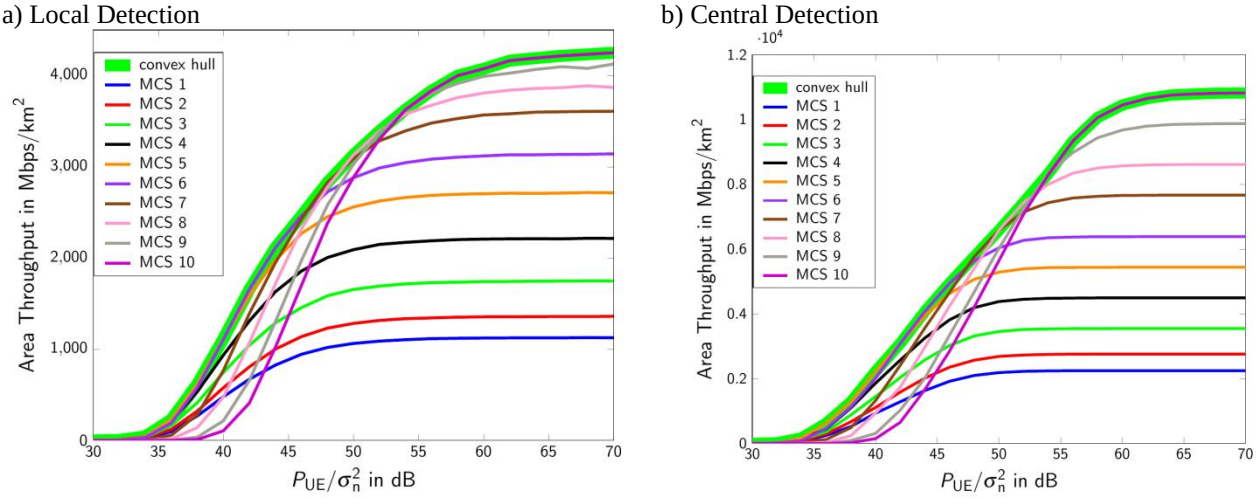


Figure 4-2: Area throughput for wide-area scenario with local and central detection for MCS=1-10

Figure 4-2 depicts the area throughput for the considered wide-area deployment. The performance of local and centralized linear MMSE detection is shown for the MCS 1 to 10 specified by LTE [43] over a varying SNR (defined as the ratio of UE transmit power P_{UE} to per-subcarrier noise level σ_n^2). These MCSs use quadrature phase shift keying (QPSK) symbols with effective code rates ranging from 0.138 to 0.666. Note that the y-axis uses a different scale these plots. For central detection all UEs are considered by the iSCs and larger area throughput can be obtained by choosing a higher MCS following the increased SNR. The convex hull indicates the maximum throughput if the best MCS per SNR is selected. In contrast, for local detection it can be observed that LTE MCS 5 to 10 lead to the highest achievable throughput over the considered SNR range. The interference generated by UEs from other cells is not considered, such that the performance of local detection is interference limited. This leads to significant performance degradations compared to central detection.

a) Area throughput per MCS for DiCE in wide-area

b) Area throughput for detection modes in wide-area

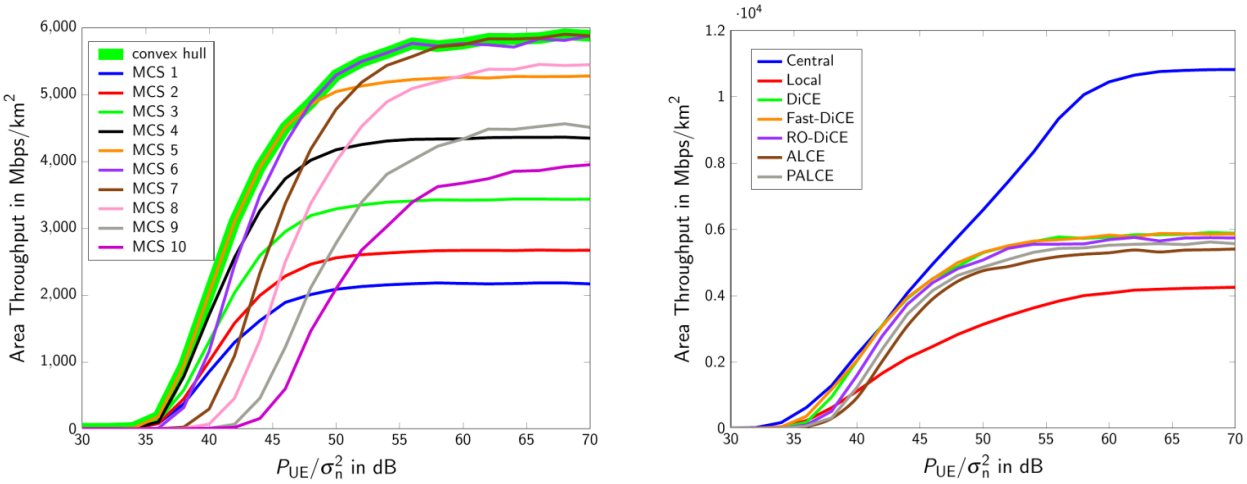


Figure 4-3: Area throughput per MCS for DiCE in wide-area without external ICI and maximum achievable area throughput over MCSs for central, local detection and distributed algorithms after 2 iterations

Similar effects can be observed for the corresponding throughput curves for the INP algorithms, e.g. the DiCE algorithm as shown in Figure 4-3 a). Here, the DiCE algorithm is stopped after 2 iterations and it can be seen that MCS 5, 6 and 7 lead to the highest throughput over the considered SNR range. Figure 4-3 b) shows the convex hulls for local detection, central detection, and all INP algorithms. In principle, the INP algorithms achieve a performance close to the upper bound for an SNR in the range of $35-45$dB, but a degraded performance for higher SNR due to the low number of iterations can be observed. For a higher number of iterations the performance gap can be closed, e.g. with 5 iterations no degradation occurs below $SNR = 50$ dB. In comparison to the baseline system (local detection), the INP algorithms achieves a significant gain over the whole SNR range. It can be observed that a gain of roughly 60% in area throughput can be achieved by all INP algorithms.

Interestingly, the RO-DiCE shows a similar performance compared to the other INP algorithms, although it uses an approximation on its estimates in order to reduce the J2 rate. Furthermore, the faster convergence of the Fast-DiCE algorithm does not lead to improved performance compared to the DiCE in the LTE simulation environment. The application of the turbo decoder basically compensates the difference in equalization performance such that the same detection performance is achieved.

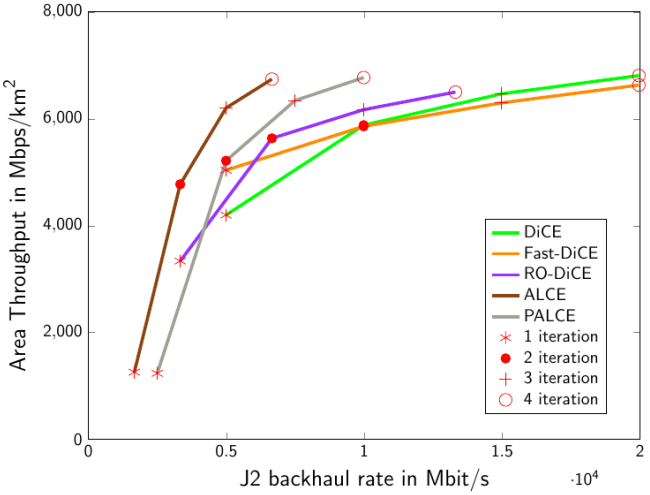


Figure 4-4: Area throughput over J2 backhaul rate for INP algorithms (MCS=7) in wide-area scenario

Figure 4-4 depicts the achievable area throughput over the J2 backhaul rate per iSC-iSC connection for the INP algorithms with 1 to 4 iterations. The SNR is fixed to 70dB and all UEs use MCS 7 as it achieves the highest area throughput for the considered SNR. For the calculation of the J2 BH rate $N_Q = 9$ bits for quantizing the J2 messages are used. The corresponding J2 BH rates of one iSC-iSC connection for the INP algorithms are given in Table 4-2 (see D2.2 [2] for details).

Table 4-2: J2 BH rate calculations for INP algorithms

Algorithm	J2 BH rate	
DiCE	$D_{J2} = 3 \cdot 2 \cdot N_L \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB} \cdot N_{It} \cdot T_{SUB}^{-1}$	(4.1)
Fast-DiCE	$D_{J2} = 3 \cdot 2 \cdot N_L \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB} \cdot N_{It} \cdot T_{SUB}^{-1}$	(4.2)
RO-DiCE	$D_{J2} = 2 \cdot 2 \cdot N_L \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB} \cdot N_{It} \cdot T_{SUB}^{-1}$	(4.3)
ALCE	$D_{J2} = 2 \cdot N_L \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB} \cdot N_{It} \cdot T_{SUB}^{-1}$	(4.4)
PALCE	$D_{J2} = 1.5 \cdot 2 \cdot N_L \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB} \cdot D_{J2} \cdot T_{SUB}^{-1}$	(4.5)

As discussed in D2.2 [2], the J2 BH rates are independent of the functional split, as in both cases the same kind of quantized estimates and auxiliary variables are exchanged among the iSCs. The J2 BH rate is then proportional to the number of iterations N_{It} for the INP algorithms. Note that the calculations in Table 4-2 are for a single iSC-iSC connection only. Therefore, the aggregated rates per J2 link will be higher depending on the chosen physical BH topology as discussed in D2.2 [2]. Here, a physical star topology with an iTN is assumed, where signals between iSCs are exchanged over the iTN.

Figure 4-4 shows the area throughput of the different INP algorithms for a varying number of iterations. Obviously, with more iterations a higher area throughput can be achieved, but causing also a higher J2 BH traffic. The Fast-DiCE shows a higher throughput after the first iteration compared to DiCE and RO-DiCE. However, for more iterations its performance is similar to DiCE. The RO-DiCE algorithm achieves the same throughput performance as DiCE and Fast-DiCE, but with a lower BH rate when 2, 3 and 4 iterations are used. Compared to DiCE and its variants, the ALCE and PALCE algorithms achieve a significant reduction in J2 BH rate, which is roughly 2 times lower than the rate of the DiCE and Fast-DiCE for the same area throughput. Thus, the ALCE and PALCE algorithms should be considered as the main approaches of this CT.

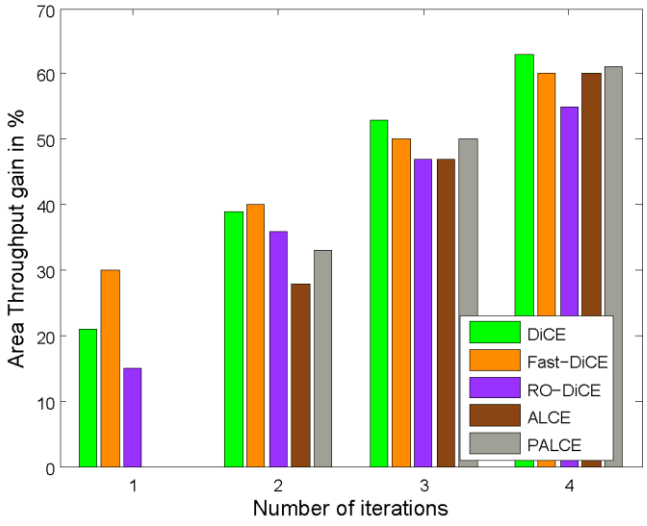


Figure 4-5: Area throughput gain over number of iterations for INP algorithms

Figure 4-5 shows the gain G_{AT} in area throughput over the number of iterations for all INP algorithms compared to the baseline system. As depicted in Figure 4-4 the ALCE and PALCE algorithms lead to worse performance if only one iteration is used. Thus, at least two iterations are required to provide a gain compared to the baseline.

As discussed before, the J1 BH rates are the same for all INP algorithms as they all forward the same kind of local estimates for the user messages. Thus, the J1 rate depends only on the applied functional split. Furthermore, depending on the chosen forwarding approach only the iSC with the best local estimate for a specific user may transmit its message to the RANaaS platform leading to a varying J1 load per iSC. However, as the total J1 rate, i.e. the aggregated rate over all iSC-RANaaS links, is constant and mainly given by the number of resources allocated to the served users, we derive subsequently this total rate. In addition, we also indicate the rate per iSC-RANaaS link assuming an equal load of all links by dividing the total rate with the number of iSCs $N_{iSC} = 7$.

The J1 BH rate for functional split B.1 is given according to the BH rate formula in Appendix II as

$$D_{J1}^{B.1} = 2 \cdot \sum_{u=1}^{N_{UE}} N_{sc,u} \cdot N_Q \cdot N_L \cdot N_{SYMB}^{SUB} \cdot T_{SUB}^{-1} = 2 \cdot 14 \cdot 600 \cdot 9 \cdot 1 \cdot 11 \frac{\text{bit}}{\text{ms}} = 1663 \text{ Mbit/s} \tag{4.6}$$

where 11 symbols are used for physical uplink shared channel (PUSCH) data. The remaining 3 symbols are reserved for demodulation reference signals (DMRSs) occupying 2 pilot symbols and the SRS (sounding reference signal). The average J1 rate per iSC equals $D_{J1}^{B.1}/N_{iSC} = 237.6$ Mbit/s.

In case of functional split B.2, for each UE one iSC forwards the LLRs to the RANaaS platform. Thus, the total J1 BH rate is given by

$$D_{J1}^{B.2} = \sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \cdot Q_m \cdot N_Q \cdot N_L \cdot N_{SYMB}^{SUB} \cdot T_{SUB}^{-1} = 14 \cdot 600 \cdot 2 \cdot 4 \cdot 1 \cdot 11 \frac{\text{bit}}{\text{ms}} = 736 \text{ Mbit/s} \tag{4.7}$$

and the average J1 rate per iSC reduces to 105 Mbit/s.

For functional split C.2 the estimated information bits are delivered to the RANaaS platform by one iSC. Thus, the total BH rate on the J1 link is determined by the transport block sizes of the users. Assuming MCS 7 for all users, the J1 BH rate corresponds to

$$D_{J1}^{C.2} = N_{UE} \cdot TBS_1 \cdot T_{SUB}^{-1} = 14 \cdot 6200 \frac{\text{bit}}{\text{ms}} = 86.8 \text{ Mbit/s}. \tag{4.8}$$

and the average J1 rate per iSC is 12.4 Mbit/s.

Table 4-3: BH rates and gains of INP algorithms after 2 iterations for wide-area scenario without ICI

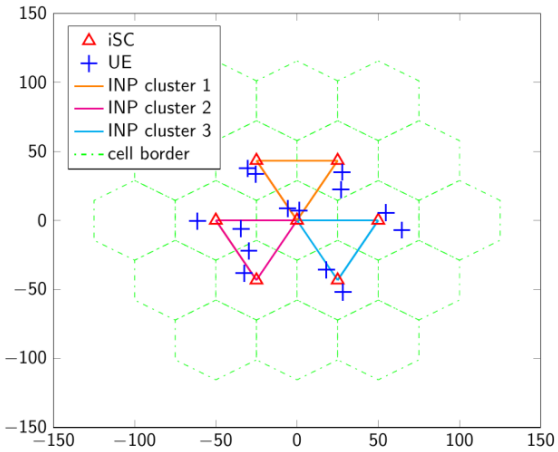
Algorithm	Total J1 BH rate in Mbit/s			J2 BH rate in Gbit/s	Gain AT
	Split B.1	Split B.2	Split C.2		
DiCE	1663	736	86.8	10	39%
Fast-DiCE				10	40%
RO-DiCE				6.7	35%
ALCE				3.3	27%
PALCE				5	33%

Table 4-3 summarizes the achieved gains in area throughput by the various algorithms of this CT and the required BH rates for J1 and J2 assuming 2 iterations for all algorithms. Obviously, for this case the required J2 BH rates of all algorithms are too high to be served by currently available technologies. However, it should be emphasized that in this scenario a worst-case is investigated since all 14 UEs are occupying the whole bandwidth of 10 MHz, i.e. 50 PRBs. By an appropriate RRM the bandwidth per UE is usually reduced such that also the BH requirement is lowered. Furthermore, in this investigation all 7 iSCs are considered for the detection of all 14 UEs. Therefore, each iSC needs to exchange signals regarding the messages of all 14

UEs creating a huge amount of data to be exchanged over the J2 links. In order to reduce this amount, several INP instances may run separately among smaller groups of iSCs, e.g. 3 or 4 iSCs.

As an example, Figure 4-6 a) depicts the implementation of 3 INP instances each covering 3 iSCs. The cooperation among iSCs is indicated by the coloured lines. Within each INP instance, the served 6 UEs are detected and all other UEs generate non-treated interference. Thus, by combining the results of the three INP instances, all 14 UEs are covered. Such clustering reduces the J2 BH load per iSC-iSC link by a factor of 2.3. Of course, incorporating these modifications will also have an impact on the overall detection performance of the INP algorithms such that the gains in area throughput will be reduced by a certain amount.

a) 3 INP clusters covering all 14 UEs



b) Area throughput of detection modes

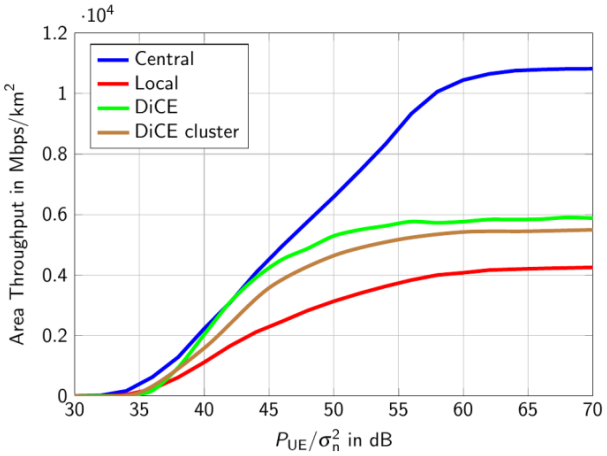


Figure 4-6: Considered deployment of INP clusters and corresponding area throughput curves

The resulting maximum achievable area throughput of the clustered DiCE over MCS 1 to 10 is shown in Figure 4-6 b) together with the results for central detection, local detection and the DiCE without clustering. Slight performance degradation for the clustered DiCE algorithm compared to common implementation can be observed. This is due to a higher interference level per INP cluster compared to the case when one INP instance is used for all 14 UEs. However, still a gain of roughly 30% at an SNR of 70dB compared to the baseline can be seen. At the same time, the required J2 backhaul rate is reduced by a factor of 2.3 since each DiCE algorithm needs to detect 6 UEs per cluster only, compared to 14 UEs if one DiCE algorithm is used among all iSCs.

Wide-area Continuous Coverage with External Inter-Cell Interference

This deployment corresponds to the scenario for the project wide analysis described in D5.3 [11] where $N_{iSC} = 19$ iSCs are placed within 3 tiers covering in total $N_{UE} = 38$ UEs. However, only the inner 7 iSCs are used for joint detection of the corresponding 14 UEs within the area. The remaining 24 UEs of the outer tier produce a constant interference level at the detecting iSCs as external ICI.

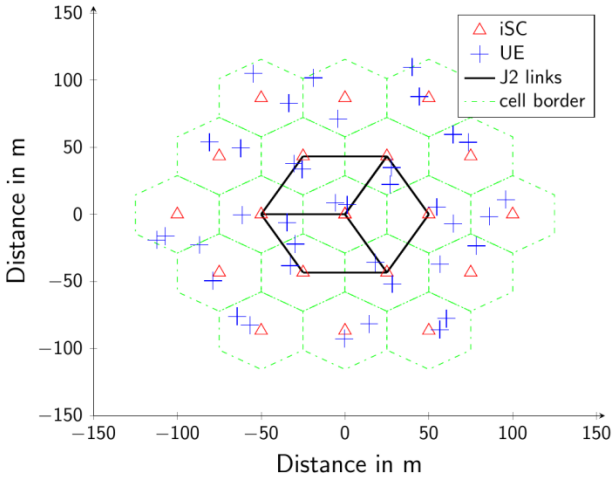


Figure 4-7: Considered wide-area continuous coverage deployment with external ICI

Figure 4-7 shows the same deployment of UEs and iSCs within the inner tiers as in the previous scenario, but external interfering UEs are randomly dropped in the outer tier as well. For central and distributed detection, again each UE uses the same MCS configuration and allocates the whole available bandwidth of $B = 10$ MHz. As baseline system, local MMSE detection of two UEs per iSC is again considered where in each cell both UEs are scheduled orthogonally on the PRBs.

Figure 4-8 a) shows the area throughput for the DiCE algorithm after 2 iterations in the wide-area scenario with external ICI. It can be seen that MCS 5 is leading to the highest area throughput over the considered SNR. Due to a constant interference level caused by the remaining 24 UEs in the outer tier all other MCSs lead to degraded detection performance. Obviously, the non-considered ICI leads to a significant performance degradation compared to the previous scenario without external ICI. To avoid such interference limitation of the performance, an appropriate RRM including also uplink power control should be applied. Corresponding RRM approaches with orthogonal resource allocation are considered by CT3.8 and are discussed in Section 4.8 of D3.3 [5].

a) Area throughput per MCS for DiCE in wide-area

b) Area throughput for detection modes in wide-area

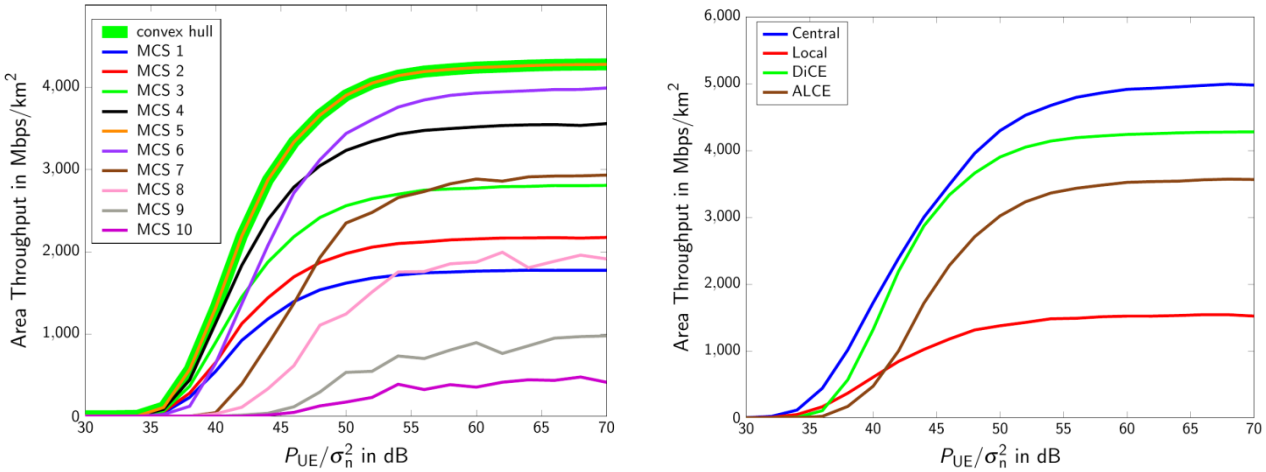


Figure 4-8: Area throughput per MCS for DiCE in wide-area with external ICI and maximum achievable area throughput over MCSs for central, local detection and distributed algorithms after 2 iterations

Figure 4-8 b) depicts the maximum achievable area throughput of the different detection modes. It can be seen that the DiCE algorithm achieves a throughput performance close to the centralized detection which serves as the upper bound performance. The DiCE algorithm outperforms the baseline system and a maximum gain of approximately 180% in area throughput can be observed. The ALCE achieves a lower maximum area throughput, but still realises a maximum gain of approximately 130% compared to the baseline with decreased J2 throughput.

In the following, we investigate the J1 and J2 BH rates for the DiCE and ALCE algorithms. As mentioned before, detailed calculations for the J2 backhaul traffic can be found in Section 4.1.3 of D2.2 [2]. For functional splits B.1 and B.2 the J1 BH rates do not differ from the wide-area scenario without external ICI. For functional split C.2 the BH rate on the J1 link is determined by the transport block size for MCS 5 which leads to the maximum achievable area throughput. The total J1 BH rate is given by

$$D_{J1}^{C.2} = N_{UE} \cdot TBS_1 \cdot T_{SUB}^{-1} = 14 \cdot 4392 \frac{\text{bit}}{\text{ms}} = 61.5 \text{ Mbit/s}. \quad (4.9)$$

and the average J1 rate per iSC is $D_{J1}^{C.2}/N_{iSC} = 8.8$ Mbit/s.

Table 4-4: BH rates and gains of INP algorithms after 2 iterations for wide-area scenario with ICI

Algorithm	Total J1 BH rate in Mbit/s			J2 BH rate in Gbit/s	Gain AT
	Split B.1	Split B.2	Split C.2		
DiCE	1663	736	61.5	10	180%
ALCE				3.3	130%

Table 4-4 gives an overview of the achievable gains in area throughput and the required BH rates for J1 and J2 for the DiCE and the ALCE algorithms. The ALCE algorithm achieves a slightly reduced gain as the DiCE algorithm, but requires a significantly lower J2 BH rate. As mentioned above, the J2 BH rate can be further reduced by applying several INP instances for a smaller number of UEs and using an appropriate RRM for the UE scheduling.

Indoor (Shopping Mall / Airport)

For the indoor scenario, we assume the parameters defined in deliverable D5.2 [10] with 5 UEs per iSC leading to total $N_{UE} = 20$ UEs detected by $N_{iSC} = 4$ iSCs. Each iSC is located in the centre of one area with 5 UEs randomly dropped within this area consisting of a hall and 16 rooms as shown in Figure 4-9.

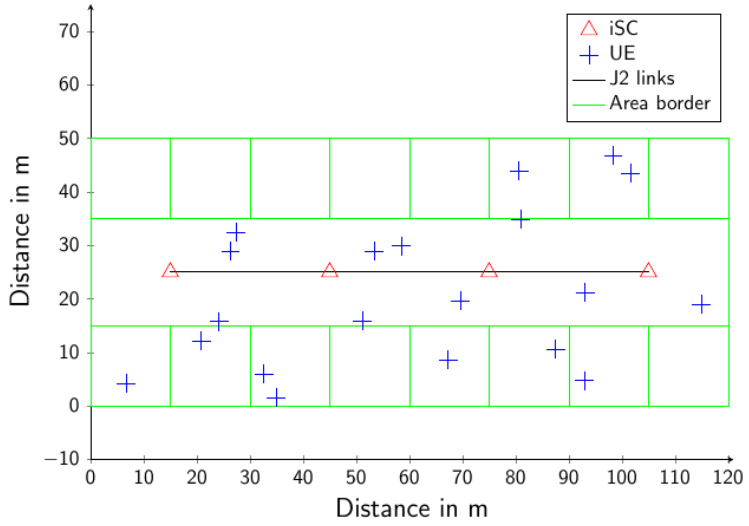


Figure 4-9: Considered indoor deployment

The access channel between UEs and iSCs uses the "Pedestrian A" power profile [49]. For UEs dropped in the hall the channels to the iSCs adopt a line of sight (LOS) path loss model. For the remaining UEs placed in the rooms an NLOS path loss model is used to model the channel between UEs and iSCs [44]. Each UE is equipped with $N_T^{UE} = 1$ transmit antenna and each iSC has $N_R^{iSC} = 2$ receive antennas. Due to the dense deployment of UEs, we consider reducing the intra-cell interference by allocating orthogonal resources to UEs of one iSC. Thus, for all detection modes the 5 UEs corresponding to one iSC use the same MCS configuration and allocate orthogonally the bandwidth of $B = 10$ MHz. Therefore, each UE has equal resources with bandwidth of 1.8 MHz, i.e. each UE occupies 10 PRBs.

a) Local detection b) Central detection

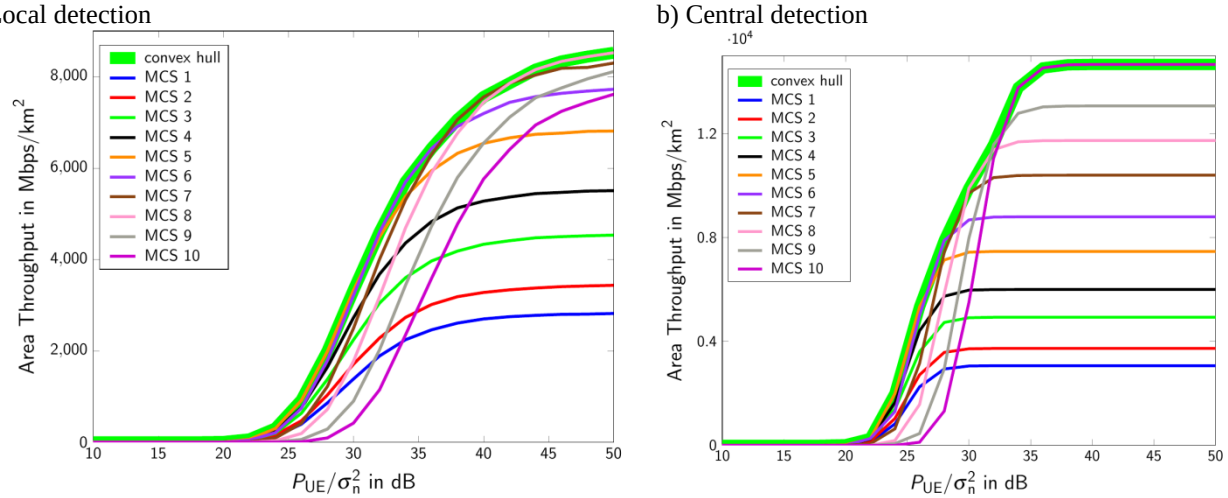
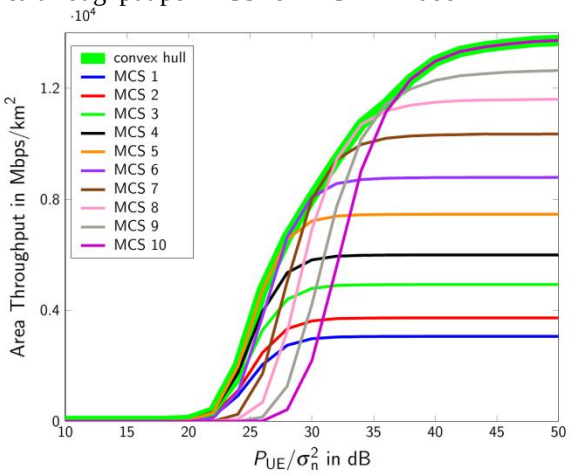


Figure 4-10: Area throughput for local and central detection for MCS=1-10

Figure 4-10 shows the area throughput curves for local and central detection for MCS 1 to 10 over varying SNR. For local detection only the five UEs corresponding to the iSC in the specific area are considered for detection. Since the five UEs per iSC are allocated on orthogonal PRBs, no intra-cell interference exists. All remaining UEs in the other areas produce inter-cell interference, which is considered as equivalent noise

variance in the local MMSE detection. It can be observed that only MCS 5-7 should be used to maximize the area throughput and the other MCSs lead to a lower throughput. For central detection, higher throughput can be achieved by selecting a higher MCS following the increased SNR. Correspondingly, depending on the SNR different MCSs are chosen to achieve the maximum throughput for the central detection.

a) Area throughput per MCS for DiCE in indoor



b) Area throughput of considered detection modes

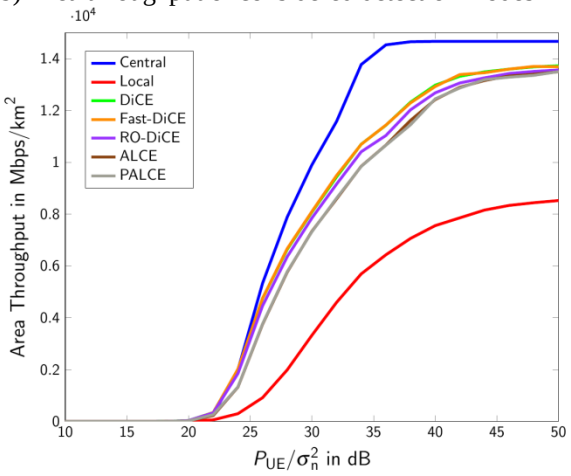
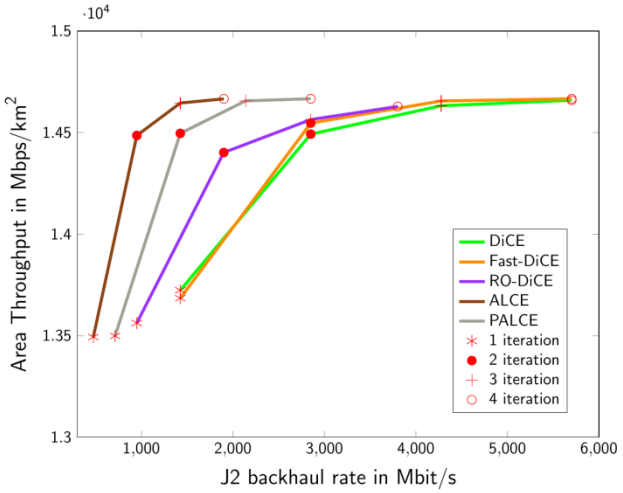


Figure 4-11: Area throughput per MCS for DiCE applied to indoor and maximum achievable area throughput for central, local detection and distributed algorithms after 2 iterations

The area throughput of the DiCE as one example for INP is shown in Figure 4-11 a). Again, the DiCE algorithm is stopped after 2 iterations. As can be seen, similar to central detection a higher throughput is obtained by choosing a higher MCS following an increased SNR.

The maximum achievable area throughput of local, central and distributed detection over the considered SNR is shown in Figure 4-11 b). In this scenario all INP algorithms executing only 2 iterations lead to a small performance degradation compared to the optimum central detection. This gap can be closed further by allowing for more iterations, but also leading to a higher J2 traffic. Compared to the baseline system, the INP algorithms achieve a significantly higher area throughput with a gain of approximately 70%.

a) Area throughput over J2 backhaul rate for INP algorithms in indoor



b) Area throughput gain over number of iterations for INP algorithms

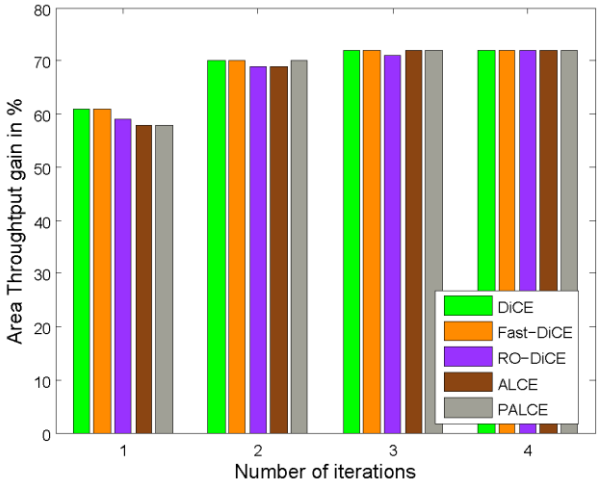


Figure 4-12: Area throughput over J2 backhaul rate for INP algorithms (MCS=10) in indoor and area throughput gain over number of iterations for INP algorithms

Figure 4-12 a) depicts the achievable area throughput versus the J2 backhaul rate per iSC-iSC connection for the INP algorithms with 1 to 4 iterations at a fixed SNR of 50dB using MCS 10. As topology, we assume that all iSCs are logically connected in a line as shown in Figure 4-6, since it requires lower J2 BH overhead compared to other topologies. Obviously, the area throughput increases with the J2 backhaul rate due to the higher number of iterations. However, the throughput gains do not increase significantly after the third iteration. Again, the ALCE algorithm requires the lowest J2 BH rate for the same area throughput.

Figure 4-12 b) shows for the different INP approaches the area throughput gain versus the number of iterations. The ALCE algorithm achieves a gain of 69% wrt. the baseline system after 2 iterations. The corresponding J2 BH rate is 950 Mbit/s, which can be realized by the light licenced mmWave BH technology as specified in Appendix I. The other approaches achieve similar gains and the corresponding J2 BH rates are specified in Table 4-5.

In the following we investigate the J1 BH rates for the INP algorithms. For this indoor scenario, the total J1 BH rate for functional split B.1 is given as

$$D_{J1}^{B.1} = 2 \cdot \sum_{u=1}^{N_{UE}} N_{sc,u} \cdot N_Q \cdot N_L \cdot N_{SYMB}^{SUB} \cdot T_{SUB}^{-1} = 2 \cdot 20 \cdot 120 \cdot 9 \cdot 1 \cdot 11 \frac{\text{bit}}{\text{ms}} = 475.2 \text{ Mbit/s} \quad (4.10)$$

leading to an average J1 rate per iSC of $D_{J1}^{B.1}/N_{iSC} = 118.8\,\text{Mbit/s}$. For functional split B.2, the corresponding total BH rate is given as

$$D_{J1}^{B.2} = \sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \cdot Q_m \cdot N_Q \cdot N_L \cdot N_{SYMB}^{SUB} \cdot T_{SUB}^{-1} = 20 \cdot 120 \cdot 4 \cdot 2 \cdot 1 \cdot 11 \frac{\text{bit}}{\text{ms}} = 211.2 \text{ Mbit/s} \quad (4.11)$$

resulting in the lower average J1 rate per iSC of $52.8\,\text{Mbit/s}$. Thus, functional split B.2 should be chosen if the turbo-decoder is implemented centrally in the RANaaS platform. For functional split C.2, the turbo decoding is done at each iSC. Thus, the iSCs can terminate the iterative process of INP as soon as a successful CRC is achieved and the corresponding iSC forwards the decoded bits to the RANaaS. The total BH rate for this split on the J1 link is determined as

$$D_{J1}^{C.2} = N_{UE} \cdot TBS_1 \cdot T_{SUB}^{-1} = 20 \cdot 1736 \frac{\text{bit}}{\text{ms}} = 34.72 \text{ Mbit/s}, \quad (4.12)$$

which can be realized by using mmWave BH. The average J1 rate per iSC is then $8.68\,\text{Mbit/s}$.

Table 4-5: BH rates and gains of INP algorithms after 2 iterations for indoor scenario

Algorithm	Total J1 BH rate in Mbit/s			J2 BH rate in Mbit/s	Gain AT
	Split B.1	Split B.2	Split C.2		
DiCE	475.2	211.2	34.72	2850	70%
Fast-DiCE				2850	70%
RO-DiCE				1900	69%
ALCE				950	69%
PALCE				1245	70%

Table 4-5 summarizes the achievable gains in area throughput and the required BH rates for the J1 and J2 connections for different functional splits. For the corresponding J2 BH rates, all INP algorithms can achieve almost 70% gain in area throughput after 2 iterations.

Conclusions

In this section we presented performance evaluations for the INP approach considered by CT2.1 for joint MUD. For the wide-area scenario a complete overlapping of allocated PRBs by all UEs was considered in order to achieve a high utilization of the available spectrum. Without external ICI the proposed INP algorithms show gains of up to 40% compared to the baseline, whereas with external ICI gains of up to 180% can be observed. For the indoor scenario, gains of 70% can be achieved by the INP algorithms. The investigations show that the INP approach realizes significant gains in area throughput similar to central detection. In contrast to central detection, the traffic on the J1 link from the iSCs to the RANaaS platform is significantly reduced. However, this comes with a high requirement for the J2 backhaul for the considered scenarios where the UEs completely allocate the physical resources leading to severe multi-user interference. In practise the application of an appropriate RRM will avoid complete overlapping of PRBs resulting in

smaller BH traffic. In addition, the proposed INP clustering is a promising approach to control the BH rate. Among the INP algorithms the best performance-rate trade-off is achieved by the ALCE algorithm. Consequently, this approach should be the preferred INP algorithm for distributed MUD in the considered dense deployment of iSCs. In general, the operation of the INP can be adapted in the number of iterations, the used iSC-iSC links, the execution of the turbo decoder in the iSCs or in the RANaaS platform, etc. Thus, the available BH and the processing resources can be utilized in an optimum way improving the utilization efficiency of the overall network.

4.2 CT2.2: Multi-Point Turbo Detection

4.2.1 Final implementation of CT

This CT investigates the benefit of using the turbo detection principle to increase the aggregated user throughput in the uplink direction. In a dense small cell deployment, one user can more easily see other small cells in addition to its serving one, especially if he is at the edge of the cell. By scheduling the (edge) users on the same resources and exploiting the created interference as a source of information through the turbo detection principle in each concerned small cell, the aggregated throughput of the system should be improved, as "more" spectrum and diversity are made available.

We use turbo detection processing to perform the multi-user detection, which can be done in two variants (associated equations are explicitly given in deliverable D2.1 [1]):

- Centrally at the RANaaS data centre: we call this approach **Multi-Point Turbo Detection** (MPTD). This approach relies on a high quality J1 interface to be LTE-compliant and uses functional split A. One iRPU is associated to the paired of UEs selected for MPTD and implements the joint turbo detection by gathering messages for the paired iSCs. There are as many WP2-based iRPUs dedicated to the turbo detection as MPTD UE pairs.
- Locally at each iSC: we call this approach **Single-Point Turbo Detection** (SPTD). The J1 interface is used to coordinate long term scheduling between the iSCs and for forwarding the estimated UE messages. Any functional split above the MAC HARQ can be used. The J2 interface may be used for iSC cooperation but we did not investigate this option in this project. Since no centralised physical processing is performed in the SPTD case, no WP2-based iRPU runs in the RANaaS.

In both cases, there is at least one iRPU per iSC to handle non-MPTD UEs in the MPTD case and all UEs in the SPTD case. These iRPUs do not implement any functionality of the physical layer (only the upper part of the RAN protocol stack) and are outside the scope of WP2. There is also an additional iRPU which communicates with all iRPUs associated to iSCs to perform the centralised RRM algorithm in charge of selecting and pairing the UEs that will benefit from the advanced physical processing (locally or centrally). This additional iRPU implements the RRM algorithm defined inWP3 (see CT3.7 in deliverable D3.3 [5]). We assume that all iSCs deployed to cover a dense local area belong to the same veNB, so only one iveC is needed.

Compared to D2.2 [2], we evaluate our CT on different channels. In the previous deliverable, we assumed that the channel on one subcarrier was the same for all the subcarriers. Since we only simulated one PRB, this assumption was still valid. For the final evaluation, we will consider 5 PRBs (number in relation with the evaluation performed in CT3.7) and the ITU-R Urban Micro channel model in addition to the Indoor Hotspot one. The SPTD solution will also be simulated.

4.2.2 Evaluation of the CT

Evaluation Methodology

To evaluate the benefit of MPTD/SPTD solutions compared to the previous baseline, link-level simulations have been carried out. The same setup given in Figure 4-13 is used. In this setup, two iSCs are connected to the RANaaS platform through the J1 interface. Both UEs are scheduled on the same resources such that they interfere with each other at their respective serving cell. We used LTE-compliant MCSs based on turbo-code [41], [42]. All equations and deep technical descriptions associated to the MPTD and the SPTD without cooperation strategies were already provided in deliverable D2.1 [1].

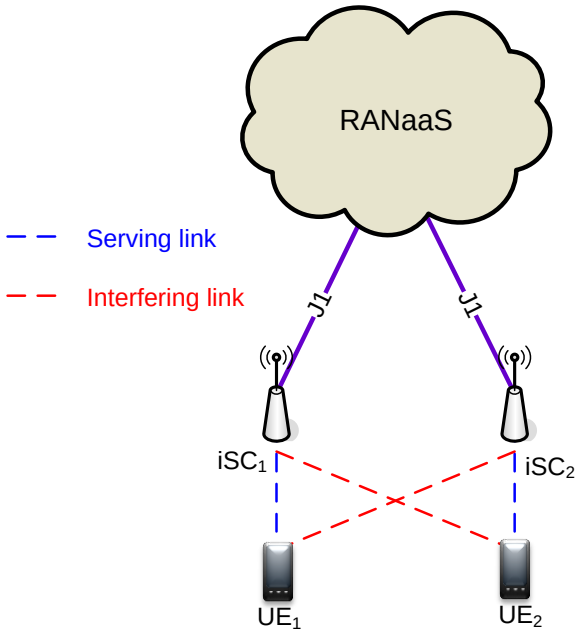


Figure 4-13: MPTD/SPTD Simulation Set-up

We assume that both UEs will send their frames with the same MCS and that they are received by each iSC with the same power. This clearly represents an edge user worst-case situation. Through Monte-Carlo simulations, the FER of both UEs can be captured when the noise is changing. This noise represents the interference coming from other UEs which are not represented here. In an LTE system, the FER of 10^{-1} is usually targeted counting on the HARQ process to recover from the errors thanks to incremental redundancy (IR).

Figure 4-14 shows how the LTE-compliant HARQ procedure was captured in our Monte-Carlo framework when enabled. At a given SNR, the two users send their frame toward their respective serving iSC. If one frame cannot be correctly decoded (check with the CRC code), then it is retransmitted 8 ms later with a different redundancy version (RV) which sequence follows the LTE standard [43]. The statistics are collected only on the two first frames and the number of retransmissions needed to convey them is stored. We allow up to four transmissions just like in the LTE standard. If one user has its frame successfully decoded and the other does not, the first user will still send dummy frames which will not be accounted in the metrics computation. Figure 4-14 shows such process where the frame of user 2 has needed 4 transmissions to pass.

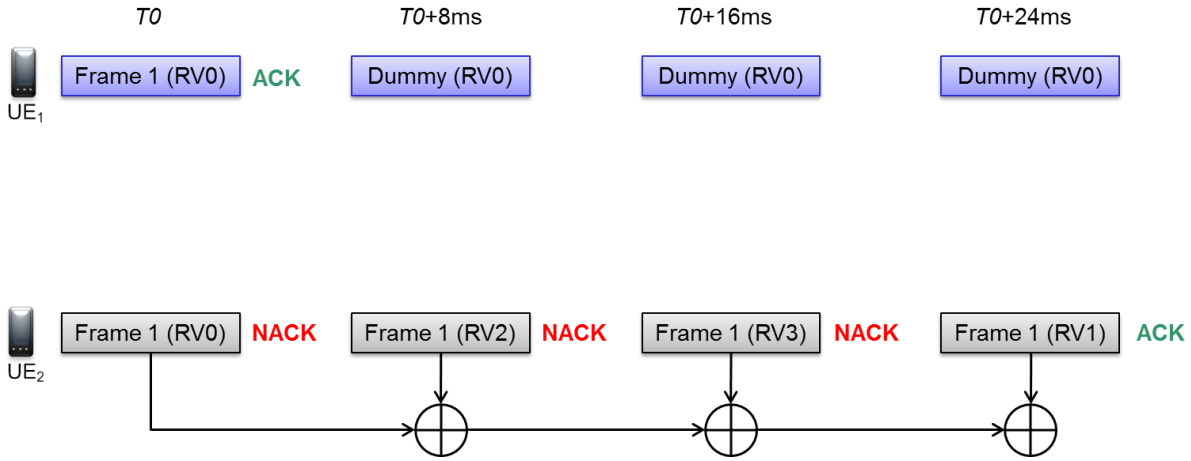


Figure 4-14: HARQ modelling in the Monte-Carlo framework

The average FER and the average number of iterations will be collected for various SNR values and different receiver strategies. We use a realistic fast fading model from the ITU-R [44] associated for a common scenario investigated.

When MPTD is used, only I/Q messages (functional split A.3) associated to the subcarriers allocated for the UEs are sent from the iSCs to the RANaaS through the J1 interface in the uplink direction. The backhaul

load required by each iSC is then directly proportional to the number of PRBs allocated to a UE benefiting from MPTD operation:

$$D_P^{A.3} = 2 \cdot N_R^{iSC} \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB} \tag{4.13}$$

If we assume that the iSC j:

- is equipped with $N_R^{iSC} = 2$ receive antennas,
- uses $N_Q = 9$ bits for I or Q quantization,
- allocates 5 PRBs to UE u ($N_{sc,u}^{j} = 60$), and
- is configured with a normal cyclic prefix ($N_{SYMB}^{SUB} = 14$),

then the throughput required on the J1 interface is equal to 30.24 Mbit/s for the payload only. In the downlink direction, only the acknowledgement information should be forwarded, which obviously requires less throughput. Also to be LTE-compliant, the MPTD processing within RANaaS and the round-trip time should be done in less than 3 ms to respect the synchronous acknowledgment procedure.

When SPTD is used, there is no message exchange over the J2 interface for the physical processing. Only higher layer messages are exchanged over the J1 interface (including the user data after decoding), which are beyond the scope of WP2.

Performance Results

We compare the MPTD/SPTD processing to the baseline enhanced linear minimum mean square error filter with interference rejection combining (E-LMMSE-IRC) [45] in different scenarios. We add the GENIE curves for the iterative processing, which represent a perfect user cancellation (lower bound for FER, upper bound for performance).

Airport/Shopping mall

For the Airport/Shopping mall scenario (CS 4) defined in deliverable D5.2 [10], we use the ITU-R Indoor Hotspot (InH) [44] channel model for the fast fading generation, assuming NLOS. We assumed a 2.6 GHz carrier, a 3 km/h UE speed and a transmission over 5 PRBs. The (contiguous) PRBs are randomly selected in a 10 MHz bandwidth for each transmission and they stay the same during an HARQ procedure. With a symmetric setup, the performance obtained for both users is equivalent. Therefore, we only display the results of the first UE.

Figure 4-15 shows the FER when HARQ procedure is not engaged, with a QPSK-based MCS targeting coding rate of 0.5. Due to the use of an LTE-compliant MCS tool chain, the number of information bits (transport block size) can only be chosen in a finite set of values, which is defined per PRB. Therefore, we choose the transport block size (TBS) which will give us the closest coding rate to our target: 680 information bits and 1440 coded bits.

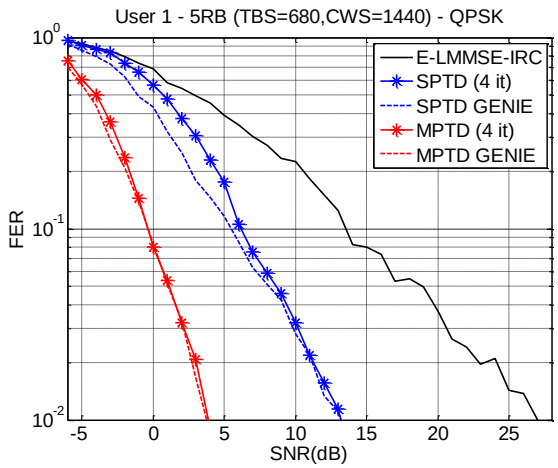


Figure 4-15: ITU-R InH - FER vs SNR for the QPSK MCS without HARQ

We produce the same curves when HARQ is engaged and capture the average number of transmissions needed in Figure 4-16.

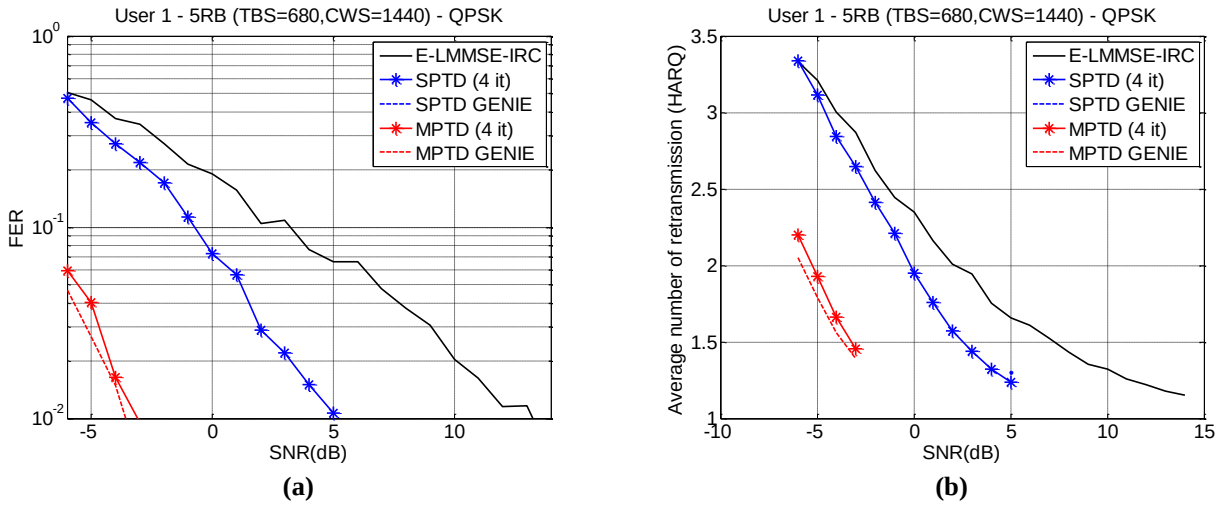


Figure 4-16: ITU-R InH - FER vs SNR (a) and Average number of transmissions vs SNR (b) for the QPSK MCS

The SPTD case provides also better result than the baseline but to an expected lesser extent than the MPTD. For the both FER curves, we can notice that the MPTD and the SPTD tend to perfectly cancel the contribution of the "interfering" user: MPTD and SPTD curves are quite close to their respective GENIE curves in terms of FER for all the SNR range simulated. Regarding, the number of HARQ retransmissions, the GENIE approach usually needs fewer repetitions to achieve the same FER (more visible for higher modulation order which is not displayed here).

In system-level simulations, we usually target FER of 10^{-1} for the MCS. Therefore, if we want to model the turbo detection processing with a perfect user cancellation (greatly facilitating the link to system abstraction), then we can without too much approximation assumes a perfect cancellation for both MPTD and SPTD procedure for the ITU-R Indoor Hotspot channel.

Wide-area Continuous Coverage

For the wide-area continuous coverage scenario (CS 3) defined in deliverable D5.2 [10], we use the same QPSK-based MCS as previously, but we implement the ITU-R Urban Micro (UMi) NLOS [44] channel model to capture the fast fading of such environment. We consider a 2 GHz carrier frequency following the assumptions defined in D5.2 [10].

Figure 4-17 shows the FER without HARQ for the MPTD, SPTD and the baseline receive strategies. The same curves are displayed in Figure 4-18 when HARQ is engaged, as well as the average number of transmissions needed.

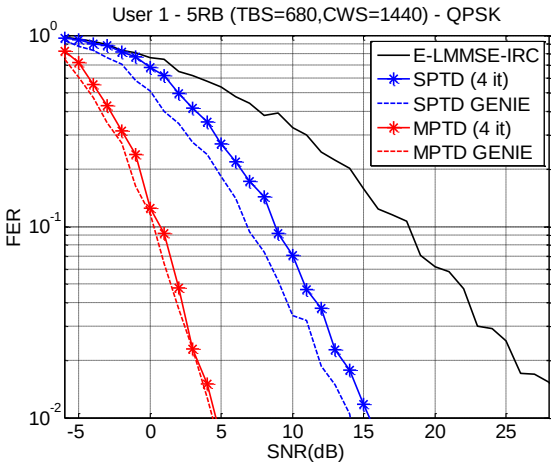


Figure 4-17: ITU-R UMi NLOS - FER vs SNR for the QPSK MCS without HARQ

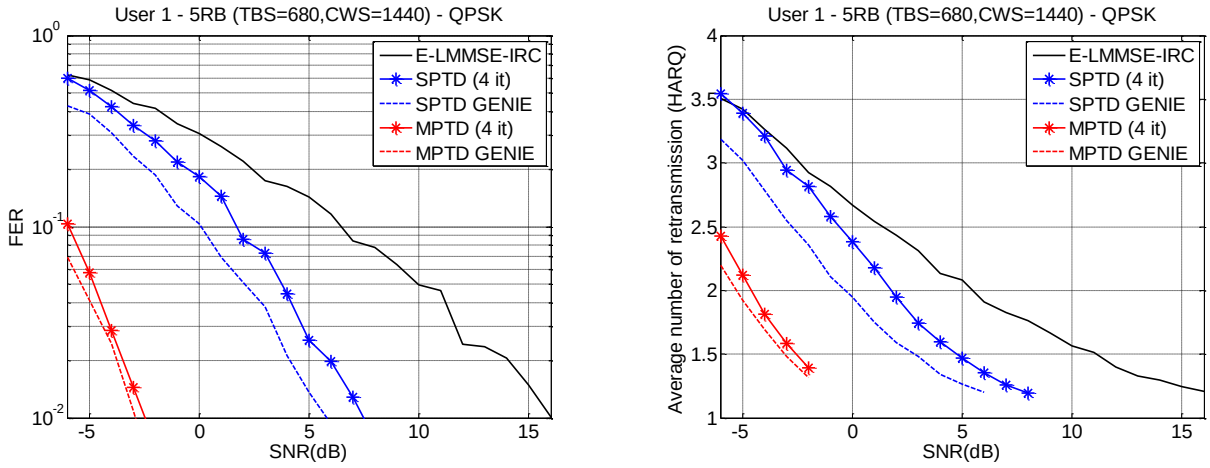


Figure 4-18: ITU-R UMi NLOS - FER vs SNR (a) and Average number of transmissions vs SNR (b) for the QPSK MCS

For this new channel model, we observe the same ranking in terms of performance without and with HARQ. The MPTD outperforms the SPTD which is better than the baseline. Without HARQ the gain observed in terms of SNR at a FER of 10^{-1} is around 10 dB, while it decreases to 5 dB with HARQ (with always more repetitions needed for the worst receivers). With just four iterations, the MPTD process is really close to its GENIE counterpart in terms of FER and average number of transmissions. Just like the InH case, we can safely approximate the MPTD approach by a perfect user cancellation. For the SPTD case, four iterations are not sufficient to achieve the same level of convergence for this worst-case scenario (both UEs are received with the same average power at both iSCs). However, we are around the one dB worse than the perfect user cancellation, so we may still use this assumption for the system-level assumptions, keeping in mind that the SPTD results we will obtain may be a bit over-estimated.

Stadium

To capture the Stadium environment, we still use the ITU-R UMi channel on a 2 GHz carrier frequency but with LOS assumed in the modelling [44]. Figure 4-19 displays the FER vs SNR without HARQ just like previously with the same receiver strategies. Figure 4-20 shows the FER and the average number of transmissions needed when HARQ is engaged.

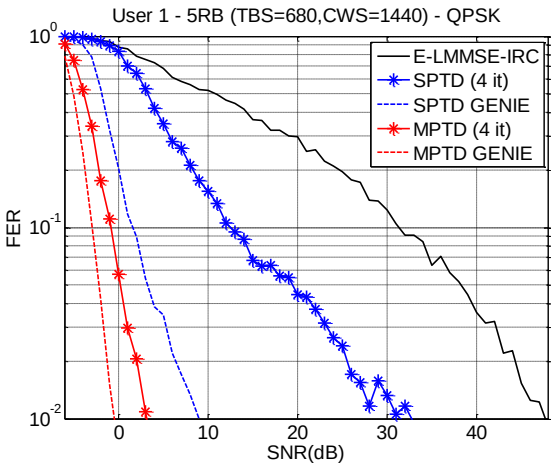


Figure 4-19: ITU-R UMi LOS - FER vs SNR for the QPSK MCS without HARQ

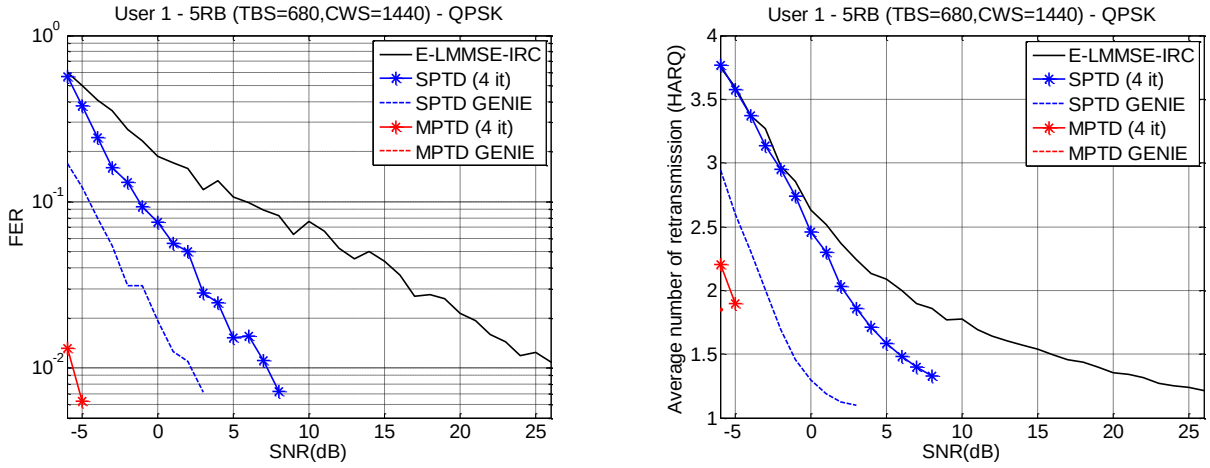


Figure 4-20: ITU-R UMi LOS - FER vs SNR (a) and Average number of transmissions vs SNR (b) for the QPSK MCS

First of all, we notice that the baseline receiver performance is worse than the one obtained in the previous scenario. Without HARQ, an SNR of 30dB is needed to get a FER of 10^{-1} for the baseline, while only 17 dB is required in the NLOS case. The HARQ procedure improves the performance, but for lower FER a gap is still present between the LOS and the NLOS case. This is due to the symmetric setup we used combined with the LOS assumption. Under such assumptions, the E-LMMSE-IRC has difficulties to cope with an interferer which exhibits a similar power than the desired signal on average with a dominant LOS path. With MPTD, this effect is well mitigated with FER curves slightly better than previously (NLOS case) as we benefit from better channel condition (LOS vs. NLOS) and receiver diversity.

On the contrary, the SPTD is far from its GENIE bound without HARQ (cf. Figure 4-19) and manages to get within 5 dB of it with HARQ (cf. Figure 4-20). With only two receive antennas (twice less than the MPTD), the SPTD has difficulties to deal with the symmetric setup with LOS. It still outperforms the baseline approach, though. In the ITU-R LOS modelling, the power of the LOS path is randomly chosen, so there are cases where one UE has a stronger path than the second UE. In our observations, the power imbalance that comes from the LOS model has led to weird cases where, for instance, the iteration devoted to the decoding of UE_1 produces the codeword of UE_2 instead as it was really the most dominant one at this iSC. Since the procedure halts on the CRC check and CRC of UE_2 was valid, this leads to false convergence.

If we unbalance our symmetric setup by reducing the power of the interfering user by a 6dB factor at each iSC, iSC1 receives UE_1 with $P(\text{dB})$ and UE_2 with $P(\text{dB}) - 6\text{dB}$, while iSC_2 receives UE_1 with $P(\text{dB}) - 6\text{dB}$ and UE_2 with $P(\text{dB})$. The FER vs SNR curve is given in Figure 4-21. We plot the GENIE curve as well, which assuming a perfect user cancellation will be the same as the symmetric setup: the interfering user contribution being perfectly removed, its receive power has no influence in the result.

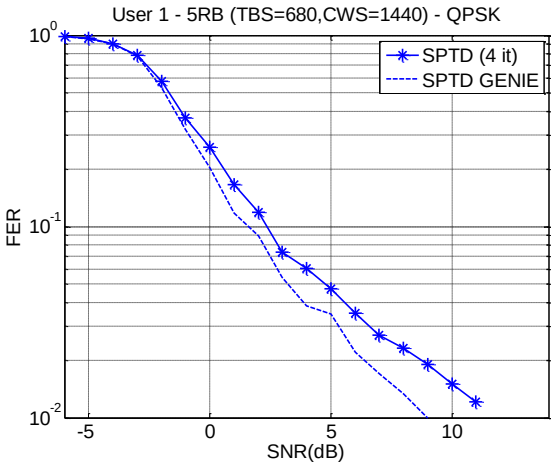


Figure 4-21: ITU-R UMi LOS - FER vs SNR for the QPSK MCS without HARQ, asymmetric setup

As we can see, this time the SPTD with four iterations is quite close to the GENIE curve (less than 1 dB at a FER of 10^{-1}). It means that assuming perfect cancellation is a valid approach for SPTD abstraction as well for ITU-R UMi LOS (and NLOS) channel, with only few exceptions where we may be overestimating the performance. Indeed, our worst-case scenario with the same power received at each iSC from two users is less likely to happen than a classical setup where the interferer will bring a smaller contribution to the received signal toward an iSC which is not its serving one, especially with power control engaged.

Conclusions

As a general conclusion, MPTD significantly outperforms the baseline MMSE receiver as expected. For a given SNR, better FER is achieved with fewer HARQ retransmissions as the quality of the detection is significantly increased. While the link-level simulations were performed in the worst-case scenario (edge users with equivalent receive power), a better scenario will most likely reduce the gap between the baseline and the turbo detection as the interferer should be less powerful than the served user (otherwise a handover would solve this power imbalance). The comparison with the GENIE curves on the various fast fading channels let us model in a simplified way the turbo detection for the link-to-system abstraction layer by assuming a perfect user cancellation.

Table 4-6: CT2.2 Area Throughput Gain

Common Scenario	# UEs	#iSCs	Pairing Ratio	MPTD AT Gain	SPTD AT Gain
Airport / Shopping mall	40	4	74%	31%	24%
Wide-area	38	19	56%	51%	20%
	190	19	72%	65%	26%
Stadium	15	15	90%	56%	Not Simulated
	75	15	85%	60%	Not Simulated

Despite being targeted to edge users essentially, using MPTD or SPTD approach should have benefits on the AT overall. To assess them in a large scale deployment, system-level simulations have been performed in WP3 relying on the previous link-to-system abstraction (see CT3.7 in deliverable D3.3 [5] for more details). Table 4-6 gives the AT gain for each of our approach (when simulated) compared to the baseline receiver. Simulations were run for three common scenarios. An aggressive pairing factor has been used in the RRM algorithm in CT3.7 ($\Delta_{\text{threshold}} = 15$ dB), increasing the number of UEs selected to be paired together and for which either MPTD or SPTD will be applied, while open-loop uplink power control defined in [54] was used ($P0 = -86$ dB, $\alpha = 1$).

Through a calibrated system-level simulator, we can see that MPTD and SPTD (which over-estimates a bit the real performance) can improve the AT by a significant amount in dense deployment scenarios. The gain in AT is less pronounced in the indoor scenario, because the cell throughput for the baseline is already quite high, around 27.5 Mbit/s per iSC on average. Considering only one antenna and 5 PRBs per UE, the peak uplink throughput on 10 MHz is then equal to 37.5 Mbit/s with the best supported MCS. Therefore, we only have room for a 10 Mbit/s improvement, if we do not change the UEs and the number of deployed iSCs. With MPTD, we have around 36.1 Mbit/s on average per iSC, which means that we are almost maximising the air interface base on the UE capabilities.

The Indoor results have been submitted to IEEE GLOBECOM as a joint WP2-WP3 contribution [69].

4.3 CT2.3: Joint Network-Channel Coding

4.3.1 Final implementation of CT

This study aims at investigating the gain that can be obtained using joint network channel coding (i.e. Figure 4-22 b)) instead of pure relaying (i.e. Figure 4-22 a)) in cooperative systems. In deliverable D2.2 [2], we proposed and evaluated several bit-interleaved network coding designs which were combined with different channel codes applied by the UEs in order to evaluate the overall performance of the joint network coding and channel (JNCC) coding scheme. It was shown that significant load reduction can be achieved on the

backhaul traffic using the JNCC technique depending on the quality of user-to-destination links. Motivated by these results, we investigate also in this deliverable the performance of JNCC. While in deliverable D2.2 [2], the performance of JNCC was evaluated using practical codes, we concentrate in this deliverable on the performance in terms of area throughput using random coding with long codewords which allows finding the performance limit for any practical scheme. Since the design and analysis of practical JNCC codes become more complex when the number of users (and channel parameters) increases, the use of the area throughput as a metric in this deliverable, along with capacity achieving codes, offers a simple tool to analyse the performance of JNCC for a system with arbitrarily number of users. Thus in this deliverable, we focus on the performance limit of JNCC, while the construction of practical codes achieving this limit was done in the deliverable D2.2 [2].

This CT focuses on the uplink communication and assumes that iSCs are connected to the RANaaS platform via wireless BH links with limited capacity (Figure 4-22). The RANaaS platform is collocated within a Macro eNB (MeNB). In pure relaying systems, the corresponding iRPU of the RANaaS instance decodes the messages of users being connected to an iSC using the backhaul only, as there is no direct connection between users and the RANaaS with the consequence of possible congestion over the iSC-to-RANaaS connections. To avoid traffic congestion, we allow the iRPU of the RANaaS instance to decode jointly user messages using both the information obtained from iSC, which acts as a relay via backhaul link, and directly from users due to the broadcast nature of wireless systems. The investigated architecture can be modelled as a multiple-access relay channel (MARC) where multiple users communicate with a final destination (RANaaS) with the help of a relay (iSC).

In order to evaluate the gain brought by the joint network channel coding we compare the performance of our system with that of a baseline system, consisting of multi-hop (MH) transmission which models the pure relaying system. In fact, using the information obtained via direct links by the destination in the decoding process distinguishes the MARC from the MH channel, as illustrated in Figure 4-22.

Throughout this study, we assume that the iRPUs of the RANaaS instance is the final destination attempting to decode all user messages and that iSCs play the role of a relay for some users to the RANaaS instance. We assume also, that an iRPU among the multiple iRPUs of the RANaaS instance performs decoding of users which are connected to the MeNB, only using one access link. Another iRPU performs decoding of users connected to the iSC, only using relayed signal. Finally, a third iRPU performs joint decoding of users connected to the iSC, using both relayed and direct signals. Users can transmit all the time. Two types of links are considered: BH and RAN. The bandwidth of the RAN is shared equally between all users associated to the same iSC or to the MeNB. The iSC has an orthogonal access to the RANaaS w.r.t. the RAN via the BH link. The main goal is to see how the system performance can be improved using MARC model and compare it with the baseline.

4.3.2 Evaluation of the CT

Evaluation Methodology

The evaluation of the CT is done for two cases of backhaul links. The first case is the in-band BH where the system has a total bandwidth B which has to be shared between BH and RAN. In this case, we are interested in finding the optimal sharing of B between RAN and BH such that the area throughput is maximized. The second case is the out-band BH in which the backhaul is characterized by a fixed data rate (the BH can be millimetre wave backhauling, xDSL, etc.).

Case 1: In-band BH

We consider a system of two users (UE_1, UE_2), one relay node R (i.e., iSC) and the final destination D (i.e., RANaaS) as illustrated in Figure 4-22. The notation SNR_{ij} where $i \neq j$, $i \in \{UE_1, UE_2, R\}$ and $j \in \{D, R\}$ will be used to denote the signal to noise ratio for the link between i and j. We assume also that $SNR_{UE_1R} > SNR_{UE_1D}$ and $SNR_{UE_2R} > SNR_{UE_2D}$. Gaussian transmitted signals are assumed and a capacity-achieving code is used for each link. We consider the capacity function defined by $C(x) = \log_2(1 + x)$. Note that this study can be easily extended to the case with arbitrary number of users.

a) Baseline: Multi-hop channel b) Multiple access relay channel

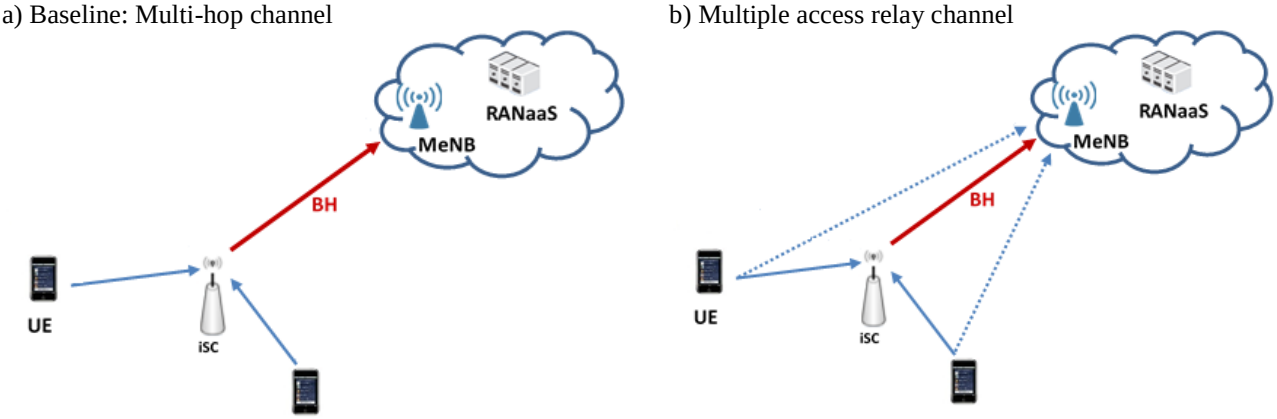


Figure 4-22 System model: a) pure relaying scheme (baseline model) b) JNCC-based MARC scheme

Subsequently, the sum rate derivation is studied for the following channel models:

- ***Point to Point (PtP) channel:*** The destination D decodes the information of users UE_1 and UE_2 using direct links only. We recall that both UEs share equally together the complete RAN bandwidth B. The sum rate is given by:

$$R_{s-PtP} = \frac{1}{2} \cdot B \cdot \left(C\left(SNR_{UE_1D}\right) + C\left(SNR_{UE_2D}\right)\right) \tag{4.14}$$

- ***Multi-hop (MH) channel:*** The destination D decodes the information of users UE_1 and UE_2 using the information obtained from the iSC R only (Figure 4-22 a)). The iSC should be able also to decode the information of users UE_1 and UE_2. Then, the sum rate is given by [60]

$$R_{s-MH} = \min\left(\frac{1}{2} \cdot B_{RAN} \cdot \left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right],\ B_{BH} \cdot C\left(SNR_{RD}\right)\right) \tag{4.15}$$

where $B_{RAN} = \theta \cdot B$ denotes the RAN bandwidth, $B_{BH} = (1-\theta) \cdot B$ is the bandwidth for the BH and $\theta \in [0,1]$ is the bandwidth sharing parameter. It is easy to demonstrate that the maximum value of R_{s-MH} is obtained when $\frac{1}{2} \cdot B_{RAN} \cdot \left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right] = B_{BH} \cdot C\left(SNR_{RD}\right)$. Equivalently, the optimum value of θ is given by

$$\theta^* = \frac{C\left(SNR_{RD}\right)}{C\left(SNR_{RD}\right) + \frac{1}{2} \cdot \left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right]} \tag{4.16}$$

and the maximum sum rate is $R_{s-MH} = \frac{1}{2} \cdot \theta^* \cdot B \cdot \left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right]$.

- ***Multiple access relay channel (MARC):*** The destination D decodes the information of users UE_1 and UE_2 using the information obtained from the iSC R and from direct links (Figure 4-22 b)). The relay should be able to decode the information of users UE_1 and UE_2 as well. The sum rate is given by [60]

$$R_{s-MARC} = \min\left(\frac{1}{2} \cdot B_{RAN} \cdot \left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right],\ B_{BH} \cdot C\left(SNR_{RD}\right) + \frac{1}{2} \cdot B_{RAN} \cdot \left[C\left(SNR_{UE_1D}\right) + C\left(SNR_{UE_2D}\right)\right]\right) \tag{4.17}$$

It is easy to verify that when $C\left(SNR_{RD}\right) > \frac{1}{2} \cdot \left[C\left(SNR_{UE_1D}\right) + C\left(SNR_{UE_2D}\right)\right]$, the value of θ which maximizes the sum rate is such that

$$\frac{1}{2} \cdot B_{RAN} \cdot \left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right] = B_{BH} \cdot C\left(SNR_{RD}\right) + \frac{1}{2} \cdot B_{RAN} \cdot \left[C\left(SNR_{UE_1D}\right) + C\left(SNR_{UE_2D}\right)\right] \tag{4.18}$$

In this case, the optimal value of θ (i.e. θ^*) is given by

$$\theta^* = \frac{C(SNR_{RD})}{C(SNR_{RD}) + \frac{1}{2}\cdot\left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right] - \frac{1}{2}\cdot\left[C\left(SNR_{UE_1D}\right) + C\left(SNR_{UE_2D}\right)\right]} \quad (4.19)$$

We can see also from the expression of R_{s-MARC} that if $\left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right] < \left[C\left(SNR_{UE_1D}\right) + C\left(SNR_{UE_2D}\right)\right]$ then the Point-to-point transmission gives better sum rate. This is because in this case, the maximum sum rate R_{s-MARC} is obtained for $\theta^* = 1$, and is equal to $R_{s-MARC} = \frac{1}{2}\cdot B\cdot\left[C\left(SNR_{UE_1R}\right) + C\left(SNR_{UE_2R}\right)\right]$ which is strictly less than R_{s-PtP}. Otherwise, the optimal θ^* is obtained using numerical simulations.

- ***Point-to-point/Relay channel (PtP/RC):*** Since the direct transmission could give better performance than MARC if the SNR of the direct link of one user is better than the SNR to the relay, we consider also a mix of point-to-point channel and relay channel (i.e. a MARC with one user only). Let's suppose that the destination decodes the information of user UE_i from direct link only (where $SNR_{UE_iD} > SNR_{UE_iR}$), while it decodes the information of user UE_j (where $SNR_{UE_jD} < SNR_{UE_jR}$) jointly based on the direct link and the $R-D$ link. The sum rate is given by [60]

$$R_{s-PtP/RC} = \frac{1}{2}\cdot B_{RAN}\cdot C\left(SNR_{UE_iD}\right) + \min\left\{\frac{1}{2}\cdot B_{RAN}\cdot C\left(SNR_{UE_jR}\right), \frac{1}{2}\cdot B_{RAN}\cdot C\left(SNR_{UE_jD}\right) + B_{BH}\cdot C(SNR_{RD})\right\} \quad (4.20)$$

where $B_{RAN} = \theta\cdot B$, $B_{BH} = (1-\theta)\cdot B$ and $\theta \in [0,1]$. As in the MARC case, the optimal value of θ can be obtained using numerical simulations. The area throughput can be easily obtained by dividing the sum rate by the coverage area of D which is equal to 0.216 Km^2 [59] .

Case 2: Out-band BH

In case 2, the backhaul is characterized by a limited data rate. We consider as a case study, a system with a RANaaS platform located within a three sectorial MeNB representing the final destination. The iRPUs of the RANaaS instance attempt to decode the messages sent by the active UEs. In each sector, a square scenario is considered consisting of a cluster of N_{iSC}^{S} iSCs connected to the associated RANaaS instance via a wireless backhaul link shared amongst the N_{iSC}^{S} iSCs (Figure 4-23).

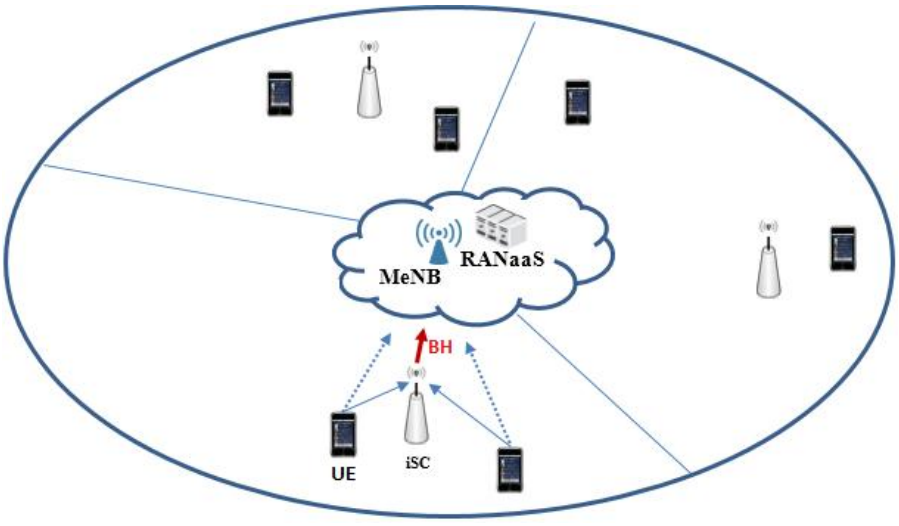


Figure 4-23: System model for the out-band BH case: a three sectorial MeNB with 1 iSC/MeNB sector

The UEs associated to the MeNB transmit directly to it without relaying. However, when the UE is associated to the iSC, two decoding schemes will be considered. In the first scheme, the corresponding iRPU of the RANaaS instance decodes the user messages using the information obtained via the wireless backhaul link only (MH scheme). In the second scheme, the corresponding iRPU of the RANaaS instance decodes jointly the information obtained via the backhaul link and the direct link (MARC scheme).

We evaluate the area throughput using MH and MARC schemes when the following methods are used to associate the users to the MeNB or the iSC:

- ***The association based on the strongest reference signal received power (RSRP):*** The received RSRP by each UE depends on the transmit power (of the MeNB or the iSC) and the channel gain including antenna gains, path loss, and shadowing. The RSRP association is defined as the association in which each UE is associated to the RAP (the iSC or the MeNB) from which it receives the strongest RSRP. The disadvantage of the RSRP association is that it does not take into account the UL performance of the system which depends on the inter-cell interference, the number of users associated to the iSC and the available backhaul capacity.
- ***The optimized association:*** The algorithm used to optimize the assignment of users to the MeNB or the iSCs in order to maximize the area throughput, is a modified version of the algorithm "Evolve" [38] for the uplink transmission and which was proposed in CT3.2 [4]. The uplink transmit power for each UE depends on the association, thus RSRP and optimized association could not be fairly comparable if the sum transmit power of all UEs in the optimized association is greater compared to RSRP association. Hence, we have optimized also the area throughput in UL with a sum power constraint (PC) on all UEs.

Performance Results

Case 1: In-band BH

We consider two scenarios A and B. In scenario A, all UEs have better channel SNR to the iSC than to the destination. The parameter settings are as follows: $SNR_{UE_1R} = 9$ dB, $SNR_{UE_2R} = 7$ dB, $SNR_{UE_1D} = 1$ dB and $SNR_{UE_2D} = 3$ dB. We assume also a total bandwidth $B = 10$ MHz Figure 4-24 a) shows the maximum area throughput using different strategies as a function of the backhaul SNR, SNR_{RD}. The upper bound, obtained when $SNR_{RD} \to \infty$, is given by

$$R_{s-max} = max\left\{\frac{1}{2}\cdot B\cdot\left[C\left(SNR_{UE_1D}\right)+C\left(SNR_{UE_2D}\right)\right],\frac{1}{2}\cdot B\cdot\left[C\left(SNR_{UE_1D}\right)+C\left(SNR_{UE_2R}\right)\right],\frac{1}{2}\cdot B\cdot\left[C\left(SNR_{UE_1R}\right)+C\left(SNR_{UE_2D}\right)\right],\frac{1}{2}\cdot B\cdot\left[C\left(SNR_{UE_1R}\right)+C\left(SNR_{UE_2R}\right)\right]\right\} \quad (4.21)$$

and can be achieved logarithmically when $SNR_{RD} \to \infty$. It is obvious that in this case, the MARC provides better performance than the case of a mixed channel ($UE_1 - D$ is PtP and $UE_2 - R - D$ is a RC) since the direct link has worse SNR than the link from UE_1 to R.

In scenario B, one of the two UEs has better channel SNR to the destination than to the iSC. The parameter settings are as follows: $SNR_{UE_1R} = 1$ dB, $SNR_{UE_2R} = 7$ dB, $SNR_{UE_1D} = 9$ dB and $SNR_{UE_2D} = 3$ dB. Figure 4-24 b) shows the maximum area throughput using different strategies as a function of SNR_{RD}. In this case, it is better to use a direct transmission for UE_1 and a relay channel for UE_2 since the direct $UE_1 - D$ link has better SNR than the link from UE_1 to R. As explained before, using only direct transmission from two users to destination (PtP case) is even better in this case than doing a MARC. This is because the relay needs also to decode the information of user UE_1 with a bad SNR which decreases the total throughput. Thus, optimizing the association of UEs to the MeNB or the iSC is crucial to achieve the best performance in terms of AT.

a) Scenario A b) Scenario B

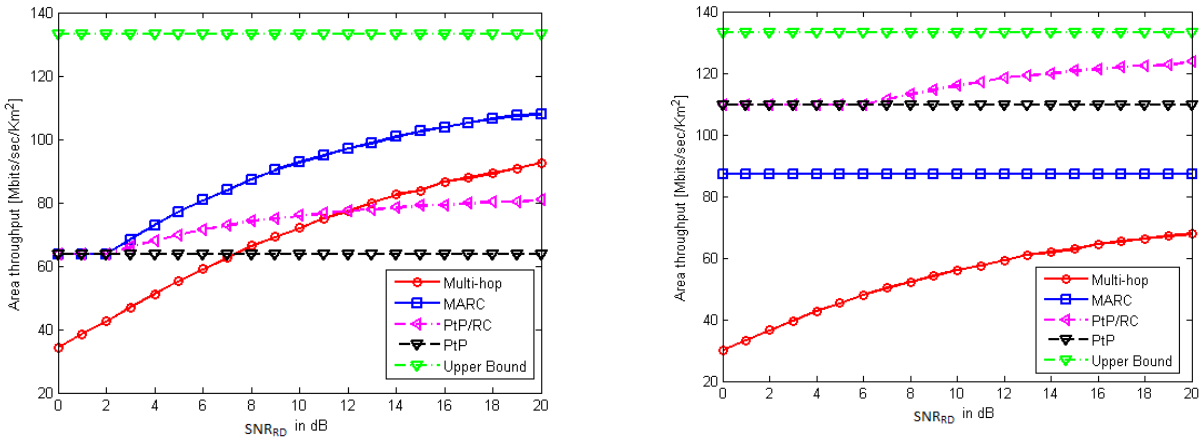


Figure 4-24: Area throughput versus backhaul SNR for In-band BH case: (a) Scenario A, (b) Scenario B

Case 2: Out-band BH

We consider the square scenario defined in D5.2 [10] with the following parameter settings: 10 UEs are deployed per sector and a total bandwidth of 10 MHz is used. The square scenario is considered in each sector of the MeNB. We consider that 2/3 of UEs are randomly and uniformly dropped inside a cluster of N_{iSC}^{s} iSCs and the remaining UEs are randomly and uniformly located in the MeNB area [59]. Simulation results are averaged on 1000 to 4000 independent runs, where a network deployment is generated randomly in each run. We assume that users associated to the MeNB and those associated to iSCs share the total bandwidth so they don't interfere each other. Figure 4-25 shows the average UL area throughput as a function of the backhaul data rate D_{J1} per sector when N_{iSC}^{s} = 1 and N_{iSC}^{s} = 4 iSC(s)/MeNB sector, respectively. In these figures, we observe that the optimal association provides significant gains in terms of average UL area throughput with respect to the RSRP association. For the sake of a fairly comparison, Figure 4-25 shows also the PC optimal association in which the sum transmit power of all UEs is constrained to be less or equal to that of RSRP association. The PC optimal association can achieve up to 34% and 84.5% of gain, for N_{iSC}^{s} = 1 and N_{iSC}^{s} = 4 respectively, w.r.t RSRP association.

Table 4-7: Area Throughput Gain using JNCC wr.t. pure relaying

D_{J1} per cluster of N_{iSC}^{s} iSCs in Mbit/s	20	40	60	80	100
N_{iSC}^{s} = 1	20.53 %	13.48 %	9.67 %	3.9 %	0.61 %
N_{iSC}^{s} = 4	32.04 %	18.82 %	12.94 %	8.81 %	5.49 %

The gain obtained using MARC instead of pure relaying MH is given in Table 4-7 for the PC optimal association. It can be observed that the gain is significant only when the backhaul data rate is the bottleneck of the system performance (up to 20.5% and 32% for N_{iSC}^{s} = 1, and N_{iSC}^{s} = 4 respectively). However, when the backhaul data rate is sufficiently large, MARC and MH achieve the same performance. This is also the reason why the average network area throughput versus backhaul data rate converges to a fixed value when the backhaul data rate increases. Thus, using joint decoding for the direct and the relayed signals can be useful only for systems with limited backhaul data rate (e.g., xDSL and microwave backhaul). In the in-band BH case, we have observed that there is a gap in area throughput between the MARC and MH channel case even when SNR_{RD} increase, because in that case we have optimized the portion of bandwidth for the RAN and BH. However in the out-band case, the total bandwidth is used for RAN only.

a) N_{iSC}^{S} = 1 iSC/MeNB sector b) N_{iSC}^{S} = 4 iSCs/MeNB sector

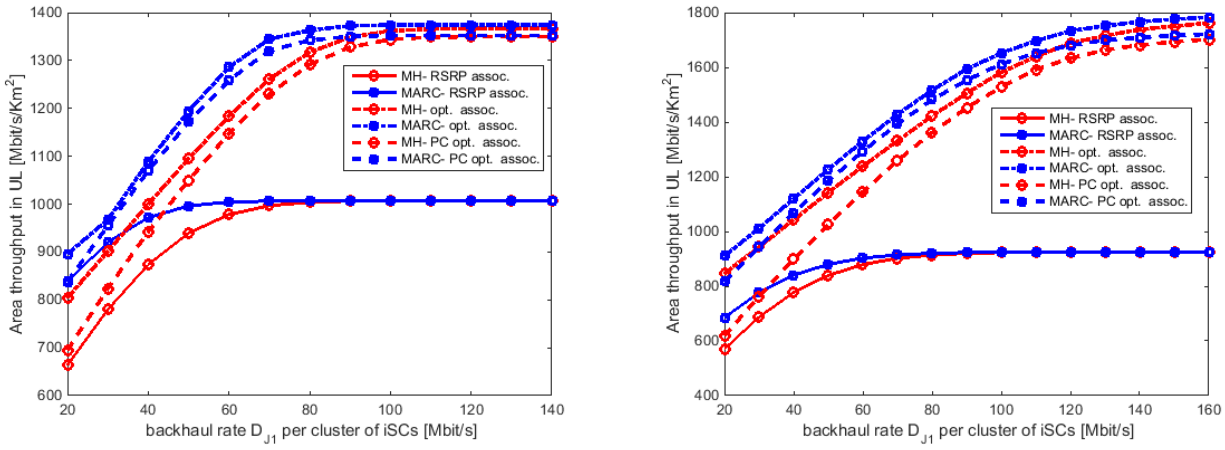


Figure 4-25: Area throughput in UL versus backhaul rate D_{J1} per cluster of iSCs for out-band BH case

Conclusions

We have demonstrated in this CT that JNCC can achieve significant gain in UL area throughput w.r.t. pure relaying scheme when the BH data rate is limited. However, both lead to the same performance when the BH data rate is large. Furthermore, the optimization of user association results to noticeable gains in UL area throughput w.r.t. the RSRP association for the same sum power constraint, especially when the BH data rate is large. Finally, we mentioned that both systems considering pure relaying and JNCC are using the same total power and BH resources. However, the available resources are used more efficiently in JNCC, by combining messages by means of network coding, since the area throughput is better than the case of pure relaying in some cases. This shows also the effectiveness of the JNCC-based systems in terms of utilization efficiency.

4.4 CT2.4: Sum-Rate and Energy-Efficiency metrics of DL CoMP with backhaul constraints

4.4.1 Final implementation of CT

In this CT, we consider centralized multi-user joint transmission (JT) to benefit from the strong computational capability of the centralized RANaaS platform and from the fact, that global CSI can be collected from each individual iSC through J1 links. Theoretically, larger joint processing size (i.e. number of cooperating iSCs) gives the potential for better interference cancelation in order to achieve higher spectral efficiency. However, this is generally not true in practice when factors such as latency in the network are taken into account. As discussed in Section 3.2.2.1 of deliverable D2.2 [2], the processing delay for calculating the precoder matrices depends on the available computational resources, the total number of transmit antennas of cooperating iSCs, and total number of receive antennas of served users. In addition, the amount of channel coefficients required for joint processing increases with the number of cooperating antennas and introduces a significant CSI feedback delay. All of this additional latency can cause a severe mismatch between the actual transmit channel and the channel coefficients used for the precoder calculation, consequently leading to performance and throughput degradation.

In order to process a large number of iSCs by a RANaaS platform with restricted computational resources, one feasible solution is to divide the network into a set of veNBs with each one comprising a reasonable number of cooperating iSCs. In that case, each veNB processes independently the DL transmission and requires only CSI between the iSCs and the UEs within the veNB. Therefore, the computational complexity per veNB is reduced leading also to a lower overall computational effort. Meanwhile, the feedback amount will drop since the number of considered UE-iSC links per veNB is reduced. Consequently, the introduced delay could be mitigated. However, extra inter-veNB interference due to veNB of small sizes may reduce the system performance and throughput. Therefore, there must be an optimal veNB size that is not so small (to mitigate inter-veNB interference into a reasonable level) and not so large either (to save the performance loss due to the delay caused channel mismatch), which optimises system performance.

Within each veNB the calculation of the precoder matrices is the most computational demanding module of DL processing. Thus, we analyse subsequently the computational effort of the iRPU which implements this precoder calculation for joint DL transmission per veNB. To this end, it is assumed that multiple veNBs share the same and fixed computational resource. However, one should also notice that the assumption that only one iRPUs calculates all precoder of one veNB is only one possible implementation. Similarly, multiple iRPUs per veNB could for example be implemented to process groups of UEs and iSCs within the veNB separately or to parallelize some computations. However, instead of optimizing the number of iSCs per veNB and the number of iSC cooperating within the same iRPU, we restrict our self to one iRPU processing all iSCs connected to the veNB of interest.

The optimal veNB size in terms of area throughput provides a useful guideline for the RANaaS manager to decide how many adjacent iSCs should be combined for joint transmission. In order to obtain the optimal veNB size, some basic network deployment information such as inter site distance (ISD), iSC density, UE density, etc. are required. In addition, also the available computational resources, the precoding granularity (i.e. how many subcarriers use the same precoding matrix), the CSI feedback capacity, and the channel statistics (especially the temporal correlation and path loss) need to be considered. Note that some of the required information is changing with time, for example, the precoding granularity could be changed subframe by subframe depending on the channel condition and the applied modulation and coding scheme. However, re-clustering the cooperative iSCs is a complex processing that requires some high layer information exchange, which is relatively long-term operation compared with the subframe duration. Therefore, the adaptation for optimal number of cooperative iSC should be based on averaged long term observation of the required information rather than the instantaneous one.

In this report, for the random deployment of iSCs (i.e. specified by an iSC density) we extended the SINR relation to consider also multiple transmit antennas at the iSCs. Correspondingly, the impact of the delay on the area throughput is investigated taking the available computational resources per RANaaS instance into account. Furthermore, the precoder has been implemented on the RANaaS testbed at University of Bremen in order to gain realistic measurement results for the precoding processing delay. Together with the investigation of the precoding granularity, the area throughput is evaluated for different scenarios.

Note that the main target of the CT is the maximisation of the area throughput with given computational resources and iSC transmission power. However, energy efficiency is another important criterion of joint transmission algorithm design. Maximizing the area throughput with given power is equivalently to achieve the maximizing energy efficiency with given area throughput constraint.

4.4.2 Evaluation of the CT

Evaluation Methodology

The evaluation of this CT considers two cases: iSCs with fixed location and randomly distributed iSCs. For the evaluations with fixed iSC locations, parameters of the common scenarios are used together with the results of the testbed implementation in order to compare the DL JT CoMP approach with the baseline, where no cooperation between iSCs is assumed. On the other hand, for the evaluation with randomly distributed iSCs the densities of UEs and iSCs are considered for solving the sum-rate optimisation problem in terms of the cluster size. For both cases, the area throughput is investigated.

iSCs with fixed location

For the case with fixed iSC locations, parameters from the wide-area scenario and the stadium scenario as defined in deliverable D5.2 [10] are used. For the wide-area we assume that the iSCs are located on a hexagonal grid with ISD = 50 meters, whereas a rectangular placement of iSCs with ISD = 20 meters is assumed for the stadium scenario. In both cases, the UEs are randomly distributed in the considered area and the minimum distance between iSCs and UEs is constrained to 5 metres. One UE per iSC on average is assumed in this evaluation. However, the extension to the multiple UEs per iSC case is straightforward. In addition, one iSC can always serve multiple UEs at orthogonal subcarriers/sub-band with adapted radio resource management algorithms.

For both scenarios, the path loss is modelled using the ITU channel model with urban micro NLOS path loss [44]. To investigate the influence of channel frequency selectivity by means of varying precoder granularity, the fast fading factor of the LTE EPA (extended pedestrian A) model has been adopted. Finally, a Doppler spread of $f_D = 6$ Hz for all iSC to UE links is assumed, which corresponds to 3 km/h moving speed at carrier

frequency 2 GHz. In addition, the processing delay that is mainly caused by the precoding matrix calculations is obtained from the iJOIN testbed at University of Bremen, which has been described in detail in Annex A of [13].

In case of an OFDM system, it would be optimal to calculate a specific precoder for each subcarrier. However, to reduce computational complexity, one can also apply the same precoding matrix to a set of neighbouring subcarriers. The new parameter precoding granularity N_G denotes the number PRBs in frequency domain that use the same precoding matrix. Generally speaking, smaller granularity leads to an accurate precoding, but requires more computational resources. On the other hand, larger precoding granularity means more subcarriers will use the same precoding matrix, which may cause throughput reduction due to channel mismatch. Therefore, in order to trade-off the delay and the precoding matrix mismatch in the frequency domain, a reasonable precoding granularity is one key of achieving optimal area throughput.

For the considered scenario with $B = 10$ MHz we assume that to each UE the available N_{PRB} =50 PRBs are assigned for DL transmission. Thus, per veNB

$$N_{PMC} = \frac{N_{PRB}}{N_G} \tag{4-22}$$

precoder matrix calculations are required. By considering the precoding granularity, the total delay per veNB defined in equation (3-46) of D2.2 [2] can be approximated as follows

$$\Delta t \approx \frac{N_{veNB} \cdot N_{PMC}}{N_{CPU}} \Delta t_{Precoder\_cal} + N_{FB\_time}\left(\Delta t_{FB} + \Delta t_{Prop\_total}\right) + \Delta t_{Tx} + N_{Hop}\Delta t_{BH} \ . \tag{4-23}$$

The second term represents the channel coefficient feedback delay per veNB, which consists of the onetime channel feedback delay Δt_{FB} due to channel waiting and scheduling, the two-way propagation delay per feedback t_{Prop_total} , and the total number of times N_{FB_time} for feedback. Note that per veNB the UEs feedback their channel coefficients indecently and in parallel. Therefore, the total delay is not directly related to the number of veNBs N_{veNB}. The third item Δt_{Tx} denotes the transmission chain processing delay as discussed in Section 3.2.2.1 of D2.2 [2]. The last term contains the backhaul latency given by the backhaul latency Δt_{BH} per hop multiplied with the number of backhaul hops N_{Hop} from the RANaaS platform to the iSCs. The values for the backhaul latency Δt_{BH} of various backhaul technologies are given in Appendix I. Note that the channel estimation delay is not considered in (4-23) since it is negligible comparing with the precoding matrix calculation delay or CSI feedback delay as discussed in D2.2 [2].

The dominating factor in (4-23) is the first term if the available computational resource is limited. It represents the total delay for calculating the precoding matrices, where $\Delta t_{Precoder\_cal}$ denotes the delay for calculating one precoding matrix per veNB by one CPU. This delay depends on the total number of transmit antennas $N_{iSC} \cdot N_T^{iSC}$ and the total number of receive antennas $N_{UE} \cdot N_R^{UE}$ per veNB. The factor $N_{veNB} \cdot N_{PMC}$ denotes the total number of precoder calculations for the considered number of veNBs N_{veNB}. N_{CPU} gives the available number of central processing units (CPUs) for recoding matrix calculations within the RANaaS platform. With a given total number N_{iSC}^{NET} of iSCs connected to the RANaaS platform of interest and fixed computational resources (i.e. N_{CPU}), a larger veNB size (i.e. larger number of N_{iSC} of cooperating iSCs) leads to a larger single time precoding matrix calculation delay $\Delta t_{Precoder\_cal}$. However, that does not imply that the total calculation delay will be increased since also the number of veNBs N_{veNB} depends on veNB size. A larger veNB size results in less veNB $N_{veNB} = N_{iSC}^{NET}/N_{iSC}$, which in turn reduces the total delay of the first term in (4-22). However, one should note that the required number of CPUs for precoding matrix calculation should be designed carefully to optimizing the area throughput with reasonable cost, i.e, a RANaaS platform serving a large number of iSCs deserve to be allocated more CPUs.

iSC with random location

The CT-specific evaluations will be based on the veNB size optimisation that has been derived in Section 4.4 of D2.2 [2]. Specifically, both iSC and UE are randomly distributed in a planar deployed network with given densities. Instead of using measurement results for the precoding matrix calculation delay from the RANaaS testbed, the analytical derivation based on the theoretical capability of a processor from Section 3.2.2.1 in D2.2 [2] is used. By considering also multiple transmit antennas per iSC, the SINR equation (4-42) derived in D2.2 [2] is generalised as follows:

$$SINR = \frac{\frac{\left(N_T^{iSC} N_{iSC} - N_{UE}\right)\lambda^2 + 1}{N_T^{iSC} N_{iSC}} \bar{P}_x}{\frac{(N_{UE} - 1)(1 - \lambda^2)}{N_T^{iSC} N_{iSC}} \bar{P}_x + \frac{N_{UE}}{N_T^{iSC} N_{iSC}} \bar{P}_I + \sigma_n^2} \tag{4-24}$$

Here $\bar{P}_x$ and $\bar{P}_I$ denote the desired signal power and the interference power without considering precoding and fast fading channel factors, respectively. Both terms have been derived in (4-37) and (4-40) in D2.2 [2]. Correspondingly, the temporal correlation factor λ has been defined in (4-31) in D2.2 [2] and σ_n^2 represents the noise variance. By replacing the new definition of the *SINR* in (4-24) into equation (4-51) of D2.2 [2], we can obtain the sum-rate and the area throughput accordingly.

Performance Results

Wide-area Continuous Coverage

We assume in total $N_{iSC}^{NET} = 61$ (4 tiers) iSCs is connected to the RANaaS platform with fixed computational resources. Let's consider the delay factor in equation (4-23) one-by-one: 3GPP defines a worst-case for the transmit processing time as Δt_{Tx} = 2.3 ms [41]. The propagation delay Δt_{Prop_total} depends on the distance between UEs and iSCs plus the distance between iSCs and RANaaS platform, which is assumed 1000 meters on average resulting in $\Delta t_{Prop\_total} = 1000\ \text{m}\ / c_0$, where c_0 is the speed of light. The feedback rate of the backhaul link from the iSC to the RANaaS platform is $C_{FB} = 24000\ bit/s$, which is the 1/20 of the peak rate of 50.4 Mbit/s for the 10 MHz LTE system. We assume that the channel coefficients are quantized to 8 bits per channel coefficient and the UEs in each veNB feedback their channel coefficients in parallel and independently.

Figure 4-26 shows the average processing delay $\Delta t_{Precoder\_cal}$ per matrix calculation versus the number of cooperating iSCs as measured on the RANaaS testbed at University of Bremen [13]. For this measurement, the virtual machine is equipped with one virtual central processing unit (vCPU) and 1 Gbyte random access memory (RAM). Note that with the number of cooperating iSCs also the number of served UEs increases and that $N_T^{iSC} = 2$ transmit antennas per iSC are assumed. Furthermore, the figure shows the delay for only one matrix calculation per veNB per CPU. In the case, the whole bandwidth divides into several sub-bands use different precoding matrices, the delay at each veNB multiplies by the factor of N_{PMC} as shown in the first term of equation (4-23). However, when more CPUs are allocated to the veNB, the delay can be reduced linearly with the increase of N_{CPU} as indicated by the first item in (4-23). To show the relationship of area throughput in terms of veNB size with fixed computational resource, we assume that in the simulations 8 CPUs are assigned to serve in total 61 iSCs (4 tiers) in the next simulation.

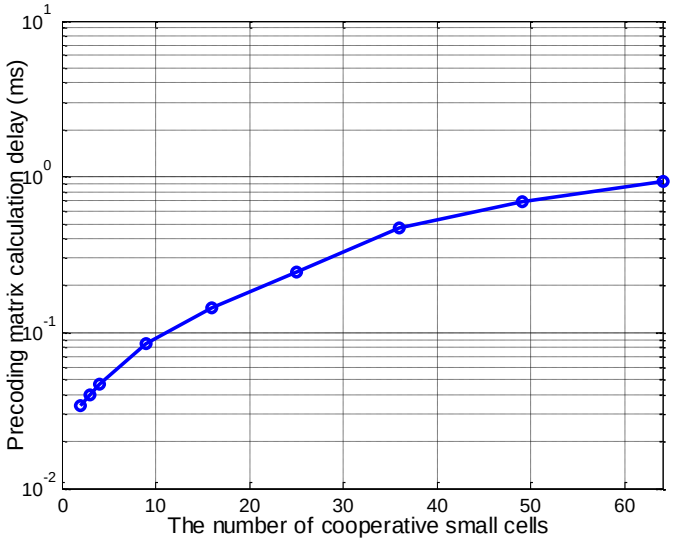
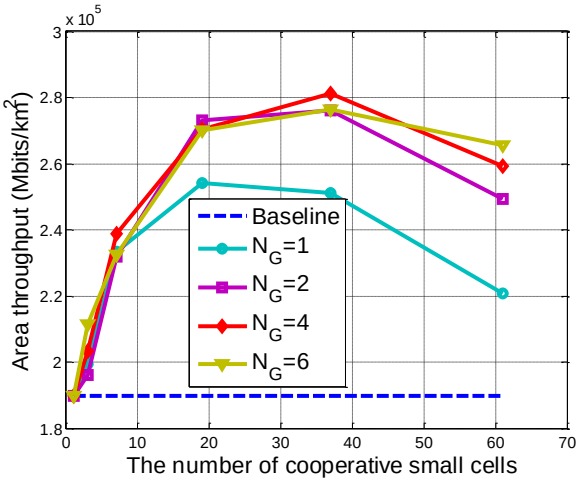


Figure 4-26: Processing delay versus the number of cooperative small cells (for a single veNB only)

Based on the delay curves provided in Figure 4-26, the output SINR of each UE versus the number of cooperating iSCs for a varying precoding granularity $N_G = \{1, 2, 4, 6\}$ is achieved. The corresponding area

throughput and the gain over the baseline are calculated by using Shannon equation and the results are shown in Figure 4-27.

a) Area Throughput



b) Area Throughput Gain

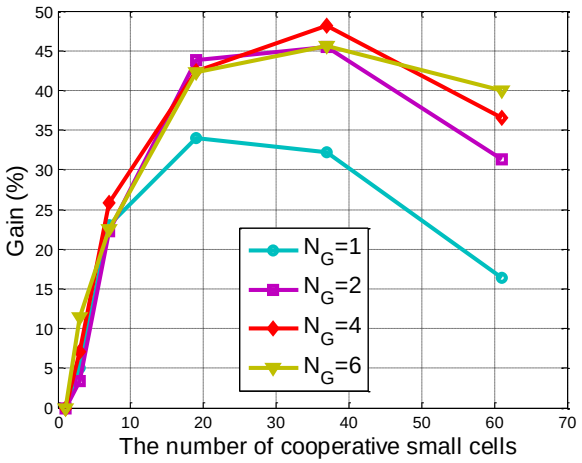


Figure 4-27: Area throughput and gain versus number of cooperative iSCs for wide-area scenario

According to the simulation results, all curves with different granularity N_G show similar trends in Figure 4-27 a). Specifically, the area throughput is increasing first, reaching a maximum value when the number of cooperative iSCs reaches 37 (3 tiers), and then decreasing for higher number of cooperative iSCs. This trend can be explained as follows: with a relative small number of cooperative iSCs (≤37), the complexity of the precoding matrix calculation is affordable for the virtual machine, while larger cooperative small cells leads to a better interference cancellation. However, the curves drop for a larger number (>37) of cooperating iSCs as the increased processing delay results in a severe precoding mismatch. The advantage of inter-veNB interference cancellation is minor compared to the delay caused channel mismatch such that the area throughput reduces. Note that the baseline area throughput is achieved by using maximum ratio transmission precoding using the 2 transmit antennas per iSC, which has a negligible complexity compared to zero-forcing (ZF) precoding. Therefore, the precoder calculation caused delay in the baseline is set zero for simplicity. In addition, we set the baseline precoding granularity to $N_G = 1$.

From the performance results we can also recognize the performance trade-off when selecting the precoding granularity N_G for the given 8 CPUs to calculate the precoding matrices for 61 iSCs. With smaller granularity (i.e. $N_G = 1$), the area throughput is relatively small due to the increased delay caused by the higher number of precoder calculations. However, the largest precoding granularity ($N_G = 6$) does not lead to the maximum throughput neither, because large granularity leads to stronger mismatch between the precoding matrices and the transmit channel. From the figure we can also see that for all N_G, the maxim throughput is achieved for 37 cooperating iSCs, i.e. 3 tiers.

Figure 4-27 b) shows the area throughput gain in terms of the number of cooperative iSCs for different N_G. Obviously, when using 8 CPUs to serving in total 61 iSCs, a gain of 48% can be achieved in comparison comparing with the baseline by taking 37 iSCs for cooperation and $N_G = 4$ PRB.

Figure 4-28 depicts the three main factors of the total delay Δt in equation (4-23): the transmission chain delay Δt_{Tx}, the CSI feedback delay and precoding matrix calculation delay. The total delay for the case with no cooperation (i.e. when number of cooperative iSC equals to 1) is only composed of the Tx chain delay as the complexity of the assumed maximum ratio transmission precoding and the CSI feedback delay are negligible. Furthermore, we see that the CSI feedback delay increases significantly in the number of cooperating iSC and it even exceeds the precoding matrix calculation delay. This effect is caused by the relative large number of CPUS used in the RANaaS platform to calculate the precoder matrices and the relative small CSI feedback capacity. With less available CPUs, or more CSI feedback capacity, the precoding calculation delay would be the dominating factor. Note that different number of CPUs allocated to the RANaaS platform leads to different optimization results; it also affects the optimal precoding granularity. However, the computational resource allocated to precoding matrix calculation should be designed carefully to trade off the performance and the cost.

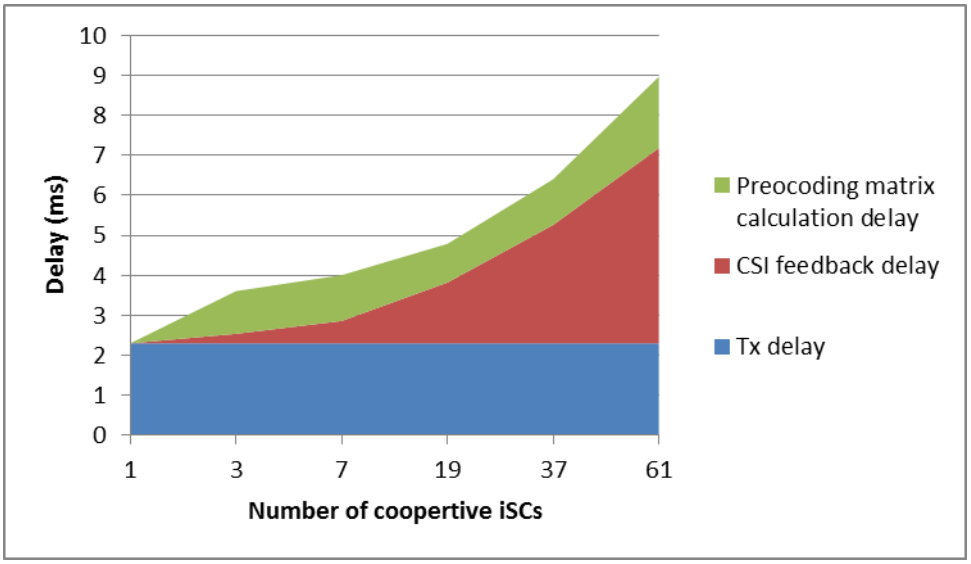


Figure 4-28: Delay versus the number of cooperative small cells for wide-area scenario (N_G=4 PRB)

According to the proposed method in Section 4.4.2 in [2], the J1 backhaul load with functional split option A.1 can be obtained as

$$D_{J1}^{A.1} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot OF \cdot (N_{FFT} + N_{CP}) \cdot N_{SYMB}^{SUB} \cdot T_{SYMB}^{-1} = 2 \cdot 2 \cdot 15 \cdot 1 \cdot (2048 + 144) \cdot 14 \cdot 1 \frac{\text{bit}}{\text{ms}} = 1841.28 \text{ Mbit/s} \quad (4.25)$$

Similarly, the J1 backhaul load with functional split option A.2 equals

$$D_{J1}^{A.2} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot N_{sc} \cdot N_{SYMB}^{SUB} \cdot T_{SYMB}^{-1} = 2 \cdot 2 \cdot 7 \cdot 1200 \cdot 14 \cdot 1 \frac{\text{bit}}{\text{ms}} = 470.4 \text{ Mbit/s} \quad (4.26)$$

and for split A.3 it reduces to

$$D_{J1}^{A.3} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot N_{sc,u}^{i} \cdot N_{SYMB}^{SUB} \cdot T_{SYMB}^{-1} = 2 \cdot 2 \cdot 7 \cdot 600 \cdot 14 \cdot 1 \frac{\text{bit}}{\text{ms}} = 235.2 \text{ Mbit/s}. \quad (4.27)$$

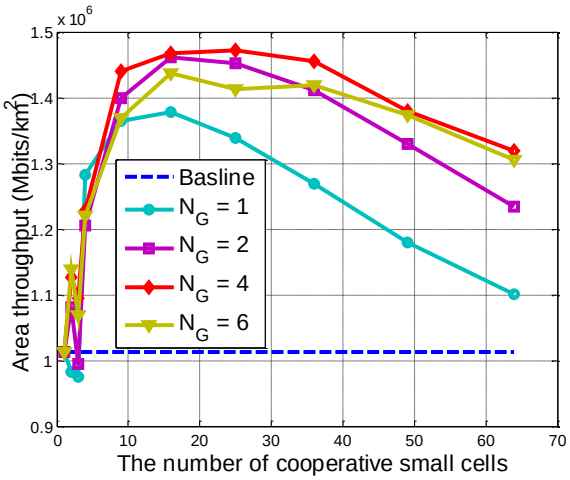
According to Appendix I, the dark fibre backhaul (No. 4a) is required to support the information exchange between iSC and RANaaS.

Stadium Scenario

For the stadium scenario, a ring backhaul topology with up to 10 Gbit/s capacity per ring is assumed. For the evaluation we further assume that cooperation among iSC connected by different rings is possible to achieve the maximum area throughput. Consequently, we suppose that the RANaaS platform transmits the precoded data to all of the rings that contain the cooperative iSCs in order to operate the cross-ring cooperation. In this evaluation, we assume that in total 8 CPUs are assigned to 64 iSCs for precoding matrix calculation.

For the delays, all of the factors are set exactly as in the wide-area scenario. The precoding granularity also follows the setup in the wide-area scenario, e.g. taking N_G = 1, 2, 4, 6 PRBs, respectively. However, different from the wide-area scenario, the small cell deployment is rectangular-grid based instead of hexagon-grid based, therefore, the cooperation strategy and the available number of cooperation iSCs are different from the wide-area scenario, which can be chosen from 1 (i.e. no cooperation, baseline), 2, 3, 4 (i.e. 2^2), 9 (3^2), 16 (4^2) and so on.

a) Area Throughput



b) Area Throughput Gain

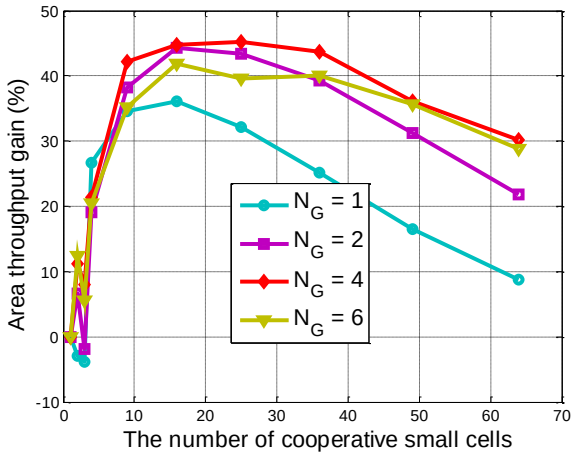


Figure 4-29: Area throughput and gain versus the number of cooperative small cells for stadium scenario

Based on different precoding granularity, Figure 4-29 a) depicts the area throughput of ZF-based precoding in terms of different number of cooperative small cells. The maximum area throughput is achieved for a precoding granularity of $N_G = 4$ PRB and 25 cooperating iSC. Similar to the wide-area scenario, all curves for different N_G show the similar shape. Figure 4-29 b) shows the area throughput gain versus the number of cooperative iSCs for different N_G. A maxim gain of 46% can be achieved compared to the baseline.

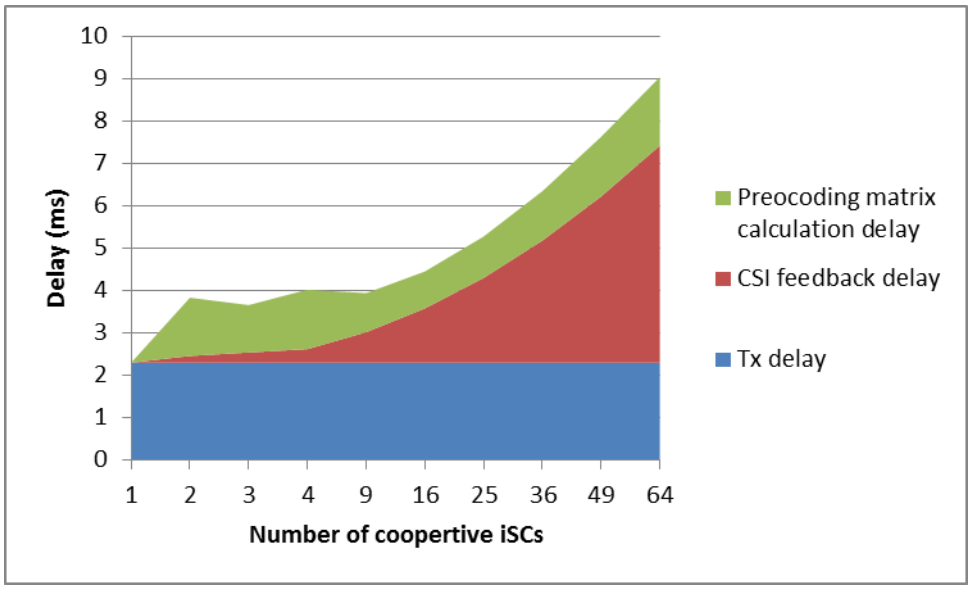


Figure 4-30: Delay versus the number of cooperative small cells for stadium scenario (N_G=4 PRB)

The total delay comprising of the transmission chain delay, the CSI feedback delay and the precoding matrix calculation delay are shown in Figure 4-30. Concerning the J1 backhaul load, the same backhaul rates as determined for the wide-area scenario are required. Thus, the J1 backhaul load with functional split options A.1, A.2 and A.3 are 1.84128 Gbit/s/iSC, 470.4 Mbit/s/iSC and 235.2 Mbit/s/iSC, respectively. According to Appendix I, the dark fibre backhaul (No. 4a) is needed to support the information exchange between the iSCs and RANaaS platform.

Both iSCs and UE are randomly distributed

For this scenario, the simulation setup parameters are the same as in D2.2 [2], except that each iSC is now equipped with 2 antennas instead of single antenna. For easy understanding, we will briefly introduce the parameters in the following. 2000 iSCs and 1000 active UEs are uniformly distributed in a circle with a network radius of 500 metres. All of the cooperative veNBs in the network share the same resource elements. We assume that all UEs experience the same noise power and the input SNR is set to 20dB. The path loss exponent equals $\kappa = 2.2$. A Rayleigh fast fading factor is considered for the channel between iSC and UE by

equation (4-29) in [2]. The temporal correlation of the channel is modelled by (4-31) in [2] with the Doppler spread $f_D = 10$Hz for all iSCs to UE links.

For the delay factors illustrated in equation (4-23): the channel estimation is set to zero since it is negligible compared to the processing and feedback delay. We use the worst-case transmit processing time defined in 3GPP, i.e. Δt_{Tx}=2.3 ms. The propagation delay Δt_{Prop_total} depends on the distance between UEs and iSCs plus the distance between iSCs and RANaaS platform, which is assumed 1000 meters on average, then $\Delta t_{Prop\_total} = 1000\ m/c_0$, where c_0 is the speed of light. The feedback rate of the backhaul link from iSC to RANaaS is $C_{FB} = 10^7$ bit/s.

Similar to the evaluations in D2.2 [2] a server consisting of an Intel Xeon Processor E5-2680 with computational capability of F_{FLOPS}=344 GFLOPS is assumed to determine the delay for calculating the precoding matrix. To this end, the measured processing delay in equation (4-23) is replaced by the analytical expressions as provided in equations (4-46), (4-47) and (4-48) of D2.2 [2] with parameters $\zeta_1 = 15$ and $\zeta_2 = 1$ [46]. Note here we assume the addition operation is double-precision floating point (64-bit). In addition, the computational resources are shared by multi-tasks and multiple sub-carriers of the networks. We introduce a parameter q_c as the computational resource division factor. Thus, $1/q_c$ of total computational resources can be used for precoding matrix calculation as discussed in Section 4.4.3 of D2.2 [2]. Therefore the available computational capability for precoding calculation is $1.72\text{E}+11/q_c$ addition operations per second as derived in D2.2 [2].

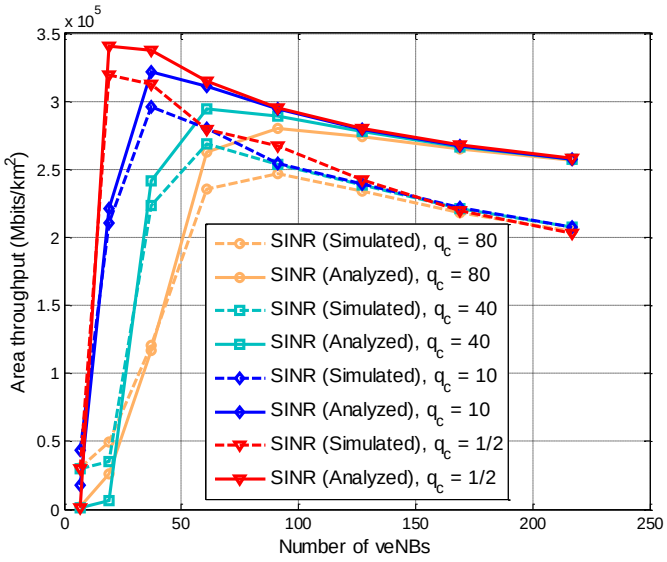


Figure 4-31: Area throughput for ZF-based planar deployment versus available computational resource factor q_c

Based on different cloud processing capabilities (q_c changing from 1/2, 10, 40, to 80), Figure 4-31 depicts the area throughput of the ZF-based precoding in terms of different veNB sizes. From the figures we can see that each curve calculated by the proposed theoretic method roughly matches the corresponding simulation results and the peak point (optimal veNB size) are strictly overlapped to each other. One should notice that the peak points tend to move towards to the right as the latency reduces (computational capacity factor changing from 1/2 to 80). This could be explained by the fact that when q_c goes up, the system is affected more by the latency and smaller veNB size is needed to improve the delay caused channel mismatch and achieve the optimal solution.

Again, the J1 backhaul load with functional split options A.1, A.2 and A.3 are 1.84128 Gbit/s/iSC, 470.4 Mbit/s/iSC and 235.2 Mbit/s/iSC, respectively. According to Table A-6-1, the dirk fibre backhaul (No. 4a) is sufficient to support the information exchange between iSC and RANaaS.

Conclusions

To verify the proposed veNB size optimisation problem in the presence of precoding processing delay and CSI feedback delay, we have considered two cases: i.e. iSCs with fixed location and iSCs with random location. The first case refers to the iSCs being placed at fixed grid and two scenarios are considered: for the wide-area scenario, the evaluation results show that the optimal setup is taking 37 (3 tiers) small cells for

cooperation and the precoding granularity is 4 PRB, given 8 CPUs for 61 iSCs. 48% gain can be achieved comparing with the baseline where no cooperation between iSCs is assumed. On the other hand, for the stadium scenario with 8 CUPs assigned to 64 iSCs, 46% gain has been achieved over the baseline when a veNB consists of 25 cooperating iSCs with precoding granularity $N_G = 4$ PRBs. For the second case that the iSCs are with randomly distributed locations, the evaluation results show that the proposed analytic method is perfectly consistent with the simulation results.

4.5 CT2.5: Partially Centralized Inter-Cell Interference Coordination

4.5.1 Final implementation of CT

This CT studies the partially centralized design of a joint multi-user precoder for JT CoMP which is described in detail in iJOIN deliverable D2.2 [2]. In this CT, the iSCs are divided in two groups as a function of their available CSI and/or signal processing capabilities. In the first group, only locally available CSI is exploited and simple matched precoding is used for the transmission. In contrast, iSCs in the second group exploit a global multi-user CSI estimate. Furthermore, they adapt to the iSCs in the first group by taking into account their transmitted signal in the design of their own precoder, This is made possible by the fact that the iSCs in the second group have received more CSI than the iSCs in the first group and are hence able to reconstruct the signals that are going to be transmitted by the iSCs in the first group.

The design of this second-part of the precoder is also more demanding in terms of computational complexity than the simple matched precoder such that this algorithm also allows adapting to the heterogeneity between the iSCs. This partially centralized design of the precoder can be seen as a "hierarchical" or "2-level" precoder design. This contribution has led to the publication [31] and a generalization of the approach to multiple levels of precoding along with the details of the mathematical derivations can be found in [31]. This precoding approach is illustrated in Figure 4-32.

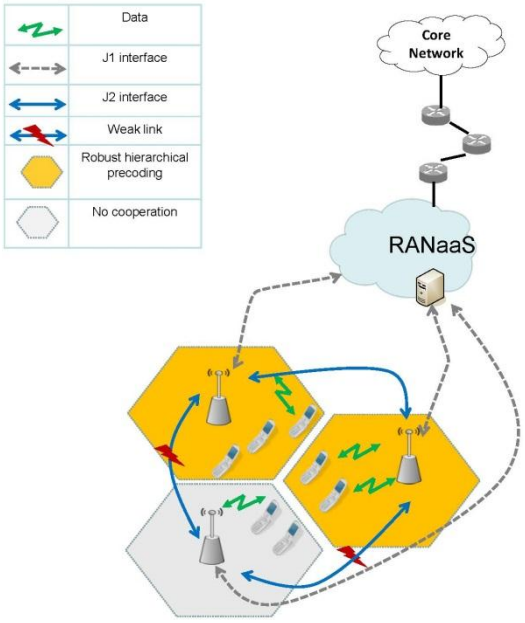


Figure 4-32: General description of the robust hierarchical precoding.

The main interest of this partially centralized design of the precoder is to obtain a precoding scheme which can adapt to different backhaul topologies and/or different iSC's signal processing capabilities. Indeed, by varying the number of iSCs belonging to the second group of iSCs and apply the advanced precoding algorithm, it is possible to go from a non-cooperative precoding to a fully centralized precoding. This precoding scheme appears then as one possible approach to bridge the gap between fully distributed and fully centralized precoding.

This precoding algorithm assumes that the precoding is done at the iSCs and is hence adapted to the functional splits B and C as described in Section 3.2. An efficient multi-user precoder design relies critically on the availability of an accurate channel estimate at the iSCs in a timely manner as shown in the performance evaluation of the technical contribution. Hence, the key parameter for the applicability of the CT is the delay introduced by the backhaul when sharing the CSI between the iSCs. The sensitivity clearly

depends on the fading scenario but a rough rule of thumb is that the total delay should not exceed the tens of milliseconds.

4.5.2 Evaluation of the CT

Evaluation Methodology

This CT is evaluated via link level simulations using Monte-Carlo simulations in a Rayleigh block fading environment. The path loss is modelled using the ITU UMI channel model with urban micro NLOS path loss formula. Our primary interest is on the precoder design with imperfect CSI at the iSCs and in particular on the impact of delay in the CSI acquisition which has been recognized in the past years as one of the main obstacles for JT CoMP [34]. As a consequence, we assume that the UEs have perfect CSI. This is also in agreement with the fact that obtaining the CSI on the receiver side requires only the transmission of pilots and not feedback, and is hence less delay constrained. We study how the proposed CT allows improving the area throughput which is one of the main iJOIN key performance indicators (KPIs).

Performance Results

Wide-Area Coverage scenario

We consider the wide-area continuous coverage scenario as described in deliverable D5.2 [10]. The network configuration studied consists of a two-tier network with 19 iSCs. The iSCs are positioned according to a hexagonal layout while one UE is dropped uniformly at random in each cell. We study the performance of the central 7 cells while the other iSCs simply generate uncoordinated interference. We investigate in the following the implementation of the partially centralized precoding algorithm across the 7 central iSCs. In particular, we study how the number of iSCs belonging to the second group of iSCs (i.e., the number of cells which apply the robust precoding algorithm) varies. For the sake of completeness, we recall the simulation parameters in Table 4-8.

Table 4-8: Simulations parameters for the performance evaluation of CT2.5

Parameter	Value	Parameter	Value
Number of small cells N_{iSC}	19	ISD	50 m
Number of antennas at each iSC N_T^{iSC}	2	UE dropping	Random dropping
Number of antennas at each UE N_R^{UE}	1	UE mobility	10km/h
Power constraint at each iSC	30 dBm	Minimum distance	UE-iSC 5 m
Number of UEs	1 UE per small cell (avg.)	Path loss model	ITUR NLOS urban micro
Small cell dropping	Regular on Hexagonal Grid	Backhaul Capacity / Latency	50 - 100 Mbit/s, 1-10 ms 10 Gbit/s, 5µs

We show in Figure 4-33 the area throughput achieved by the proposed robust hierarchical precoding as a function of K_{coop} the number of iSCs running the partially centralized precoding algorithm. To evaluate the gains of this precoding approach, we also show the performance obtained with the reference transmission scheme, i.e., in the absence of cooperation. It is also interesting to show the area throughput achieved with full centralization of the precoding as it shows the potential gain of coordination. Finally, we also present a "naive" partially centralized approach consisting of simply forming small cooperation clusters but without taking into account the iSCs outside the cooperation cluster. We can also observe that the hierarchical precoding approach allows exploiting more efficiently these capabilities in comparison with the simple distributed ZF precoding approach. We can see that the proposed approach allows to move smoothly from the non-cooperative approach to the fully cooperative one as the number of cooperative iSCs increases, i.e.,

as the acquisition to an accurate multi-user channel estimate becomes possible for more iSCs. This is possible if a low delay J2 link is available to these iSCs. The exact meaning of "low delay" will be discussed in the following when we will investigate the effect of the delay in the CSI.

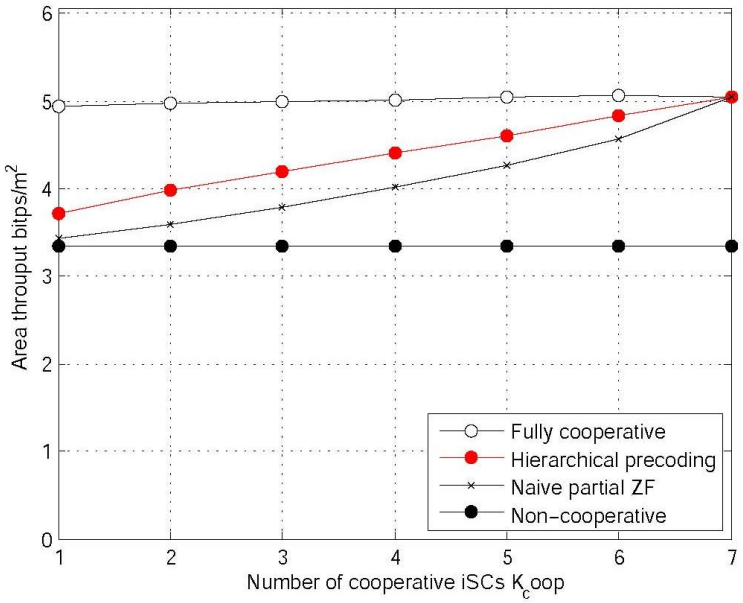


Figure 4-33: Area throughput as a function of the number of cooperative iSCs

It should be noted that the improvement in the AT can be traded-off against a reduction of the transmitted power. Indeed, the proposed precoding scheme allows to transmit in a more efficient manner, thus requiring less power to achieve a given target rate. The translation from the AT improvement to an energy efficiency (EEf) improvement is straightforward and is not investigated in further details.

The impact of the functional split and the backhaul required for this precoding algorithm are discussed in detail in Section 5.3.3 along with the performance improvement.

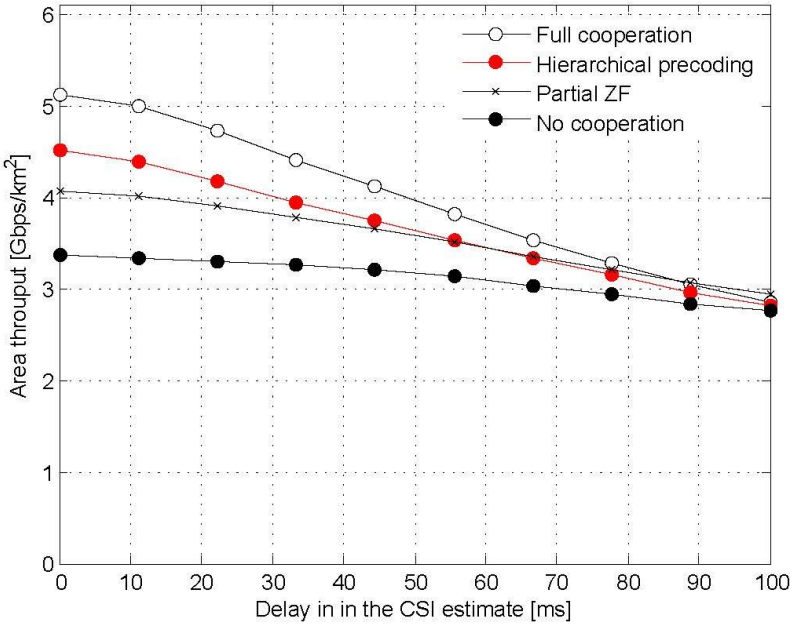


Figure 4-34: Area throughput as a function of the delay in the CSI at the iSCs.

In the following, we discuss another important behaviour which is the sensitivity of the transmission with respect to delay in the CSI at the iSCs. We consider in the following the well-known Jake's model with a speed of 10 km/h. In that scenario, the area throughput is shown as a function of the delay in Figure 4-34. The sensitivity to the delay in the CSI estimate dictates the requirements for the J2 backhaul links in order to apply our precoding algorithm. Indeed, if the delay is larger than 50 ms, the improvement provided by our algorithm practically vanishes. This delay constraint can be realistically fulfilled with existing backhaul

technologies as described in Appendix I. The most critical aspect consists then in developing cooperation protocols and transmit strategies being efficient in terms of delay. Another interesting consequence of this analysis is that it potentially allows reducing the cost of sharing the precoded signal by reducing the frequency with which the precoder is updated. The analysis of the impact of the delay is therefore very helpful for the design of the partially centralized JP CoMP transmission.

Conclusions

We have provided a novel precoding scheme and shown how its use in a realistic iJOIN scenario could lead to significant performance improvements. It can hence be seen that 30 % of improvement can be achieve over a wide range of delay with only few cooperative iSCs in the simulated iJOIN setting. Beyond this interesting precoding scheme, the problem of the design of partially centralized cooperation schemes remains mostly open and is expected to be key to further interesting innovations. Hence, this contribution is a first step in a very important and promising research direction and further works should lead to further insights and interesting results.

4.6 CT2.6: Data Compression over RoF

4.6.1 Final implementation of CT

In this CT, the possible functional splits at the physical layer between RANaaS and iSCs are investigated. The backhaul load of the different split options has been evaluated analytically in D2.2 [2] together with the related advantages and constraints, which may also impact the DL and UL coordination approaches discussed in the other CTs. A summary of this analysis is also provided in this document in Section 3.2.1.

Downstream of this analysis the split option A.3, shown in Figure 4-35, is selected and analysed in detail. In particular the impact on the area throughput and on the BH capacity requirement is investigated and compared with a fully centralized physical layer. The split option A.3 combines several advantages: it provides a significant backhaul load reduction compared to the fully centralized solution but, at the same time, it gives the possibility to perform the statistical multiplexing of the backhaul traffic generated by different iSCs (characteristic not applicable in the fully centralized case with time domain I/Q transmission) and gives also the possibility to implement centralized coordination algorithms.

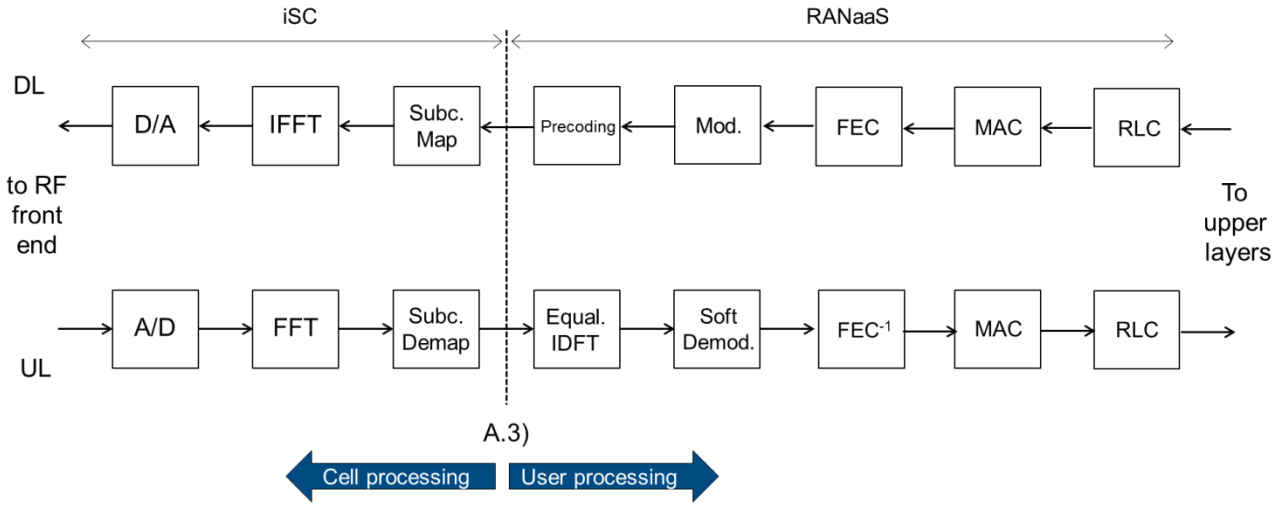


Figure 4-35: Functional split A.3

This functional split places the RANaaS/iSC boundary at the input of the resource mapping block in the downlink. Symmetrically, the boundary is placed at the resource demapping output in the uplink. The resource mapping/demapping and FFT/IFFT operations are therefore executed at the iSC side so that only the allocated transmission resources need to be transmitted on the backhaul. This split provides a significant reduction of the backhaul load compared to the functional split A.1 with I/Q time domain transmission [2]. Besides, it allows the centralization in the cloud of some downlink and uplink physical layer functions that are computationally intensive, like multi-cell precoding, channel estimation, multi-user detection and FEC decoding.

A further aspect related to this functional split is the generation and multiplexing of downlink channels or signals that carry content with no (or infrequent) variation in time directly at the iSC side, thus providing a further reduction of the backhaul load. In the specific case of the LTE system, this would apply to the cell specific reference signal (RS) and to the primary and secondary synchronization channels (PSS, SSS) that could be generated directly at the iSC and multiplexed in the frequency domain with the other data and control channels received from the RANaaS.

Besides, as the backhaul load is proportional to the radio interface load, it opens the possibility to design the backhaul capacity without satisfying the full load condition (i.e. all iSCs with full resource utilization at the same time). This undersized design brings advantages in terms of scalability and backhaul cost reduction but entails a certain probability of blocking as a function of the expected traffic distribution in space and time.

4.6.2 Evaluation of the CT

Evaluation Methodology

The reduction of the backhaul load of this CT compared to the baseline is evaluated analytically by means of the mathematical relations derived in [2], which express the backhaul load as a function of the signal characteristics transmitted between RANaaS and iSC. The improvement of the considered functional split on the iJOIN area throughput metric [10] is evaluated using the theoretical model described in [35].

Performance Results

The backhaul load of the functional split A.3 is calculated and compared with the fully centralized solution (split option A.1) using the formulas derived in the analytical evaluation and summarized in Appendix II. The exemplary case of the downlink for an LTE cell equipped with two transmit antennas $N_T^{iSC} = 2$ and bandwidth $B = 10$ MHz is considered. The message size for the payload, not including the overhead introduced by the BH protection coding, is calculated with the equation given in Appendix II for the split option A.3

$$D_P^{A.3} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j}\right) \cdot N_{SYMB}^{SUB} \qquad \text{[bit/iSC]} \tag{4-28}$$

while for the split option A.1 the following equation is used

$$D_P^{A.1} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot OF \cdot (N_{FFT} + N_{CP}) \cdot N_{SYMB}^{SUB} \qquad \text{[bit/iSC]} \tag{4-29}$$

The quantization resolution N_Q required by the two functional split options is different with a significant lower resolution required for the split option A.3 due to the operation in the frequency domain [2]. The lower quantization resolution together with the avoidance in the transmission of the guard subcarriers and cyclic prefix are the factors that determine of the reduction of the backhaul load for the split option A.3. The application of the equations above with $N_Q = 7$ for the split option A.3 and $N_Q = 15$ for the option A.1 allows quantifying the BH load reduction, as shown in Table 4-9. The BH throughput η without protection overhead is calculated from the message size considering that the message frequency is equal the reciprocal of the subframe period (i.e. $T_{SUB} = 1$ ms) and assuming a full load condition (i.e. $\sum N_{sc,u}^{j} = 600$). The other parameters, namely $N_{FFT} = 1024$, $N_{CP} = 72$, OF=1 and $N_{SYMB}^{SUB} = 14$ are derived from the numerology of the LTE system in case of bandwidth $B = 10$ MHz.

Table 4-9: BH throughput per iSC in full load condition

Split option	BH throughput η without protection overhead in Mbit/s/iSC	BH throughput η·γ with protection overhead (γ=4/3) in Mbit/s/iSC
A.1	920.6	1227.5
A.3	235.2	313.6

The downlink BH throughput reduction in full load condition of the functional split A.3 is therefore equal to about 3.9 times compared to a fully centralized physical layer and increases as the radio interface load

decreases, due to the fact that the BH throughput of the split option A.1 is constant and does not depend on the radio interface load.

A further reduction of the BH throughput is achievable by taking advantage of the statistical multiplexing of the traffic generated by multiple iSCs. This characteristic can be exploited in case of a transport network topology where the BH capacity is shared among multiple iSCs. An estimate of the further BH throughput reduction is performed considering a regular hexagonal iSC grid (like for example in the wide-area scenario). It is assumed that each iSC and the surrounding first tier of iSCs (i.e. in total 7 iSCs) are connected through a fibre ring to the RANaaS. The BH throughput per iSC in such condition is calculated analytically in Table 4-10, where it is further assumed a BH blocking probability of 1% and a PRB utilization in each iSC with Gaussian distribution. The average PRB utilization in each iSC is set to 50% (medium load) or 90% (high load) with a variance equal to 5 PRBs.

Table 4-10: BH throughput per iSC when exploiting the statistical multiplexing (1% BH blocking probability, 7 iSCs per group)

Split option	Average PRB usage in each iSC [%]	BH throughput η without protection overhead in Mbit/s/iSC	BH throughput η·γ with protection overhead (γ=4/3) in Mbit/s/iSC
A.1	Not relevant	920.6	1227.5
A.3	50	128.6	171.4
	90	224.5	299.3

The BH throughput reduction of the functional split A.3 compared to A.1 when exploiting the statistical multiplexing, increases with respect to the full load case to about 7.1 times in case of 50% PRB utilization and 4.1 times for 90% resource utilization.

Wide-Area Continuous Coverage scenario

The impact of the functional split A.3 on the area throughput metric [10] is evaluated using the analytical model provided in [35]. Given the required backhaul capacity per iSC $\eta \cdot \gamma$ for the considered functional split and given the iSC density ρ_{iSC} expressed in iSCs/km^2 it is possible to determine the required BH capacity per area C_{BH} expressed in Gbit/s/km^2

$$C_{BH} = \rho_{iSC} \cdot \eta \cdot \gamma \quad [\text{Gbit/s/km}^2] \tag{4-30}$$

The spatially averaged rate per user R in bit/s/Hz is determined using the analytical model given in [35], which is derived on the base of the Shannon capacity, assuming iSCs equipped with omnidirectional antennas that transmit a power P and using a bandwidth B

$$R = \frac{\pi^{5/2}}{2}\sqrt{\frac{\rho_{UE}\,\rho_{iSC}P}{\sigma_n^2}} \cdot \operatorname{erfc}\left(\frac{\pi^2\rho_{UE}}{4}\sqrt{\frac{P}{\sigma_n^2}}\right) \cdot \exp\left(\frac{\pi^4\rho_{UE}^2 P}{16\sigma_n^2}\right) \quad [\text{bit/s/Hz}] \tag{4-31}$$

where ρ_{UE} is the assumed user density expressed in UE/km^2 and σ_n^2 is the thermal noise power. The exponential term in equation (4-31) usually dominates the other terms. If, however

$$\rho_{UE} \cdot \sqrt{P/\sigma_n^2} > 4/\pi^2 \tag{4-32}$$

as it occurs in an interference limited system, the spatially averaged rate R per user can be approximated by

$$R \approx 2\sqrt{\frac{\rho_{iSC}}{\rho_{UE}}} \quad [\text{bit/s/Hz}] \tag{4-33}$$

Finally, the area throughput $R_{A_{net}}$ is calculated using the equation below

$$R_{A_{net}} = \rho_{UE} \cdot R \cdot B \quad [\text{bit/s/km}^2] \tag{4-34}$$

Considering the wide-area evaluation scenario the inter-iSC distance ISD is equal to 50 meters [10]. For a hexagonal cellular layout this corresponds to an iSC density of about

$$\rho_{iSC} = \frac{2}{\sqrt{3} \cdot ISD^2} = \frac{2}{\sqrt{3} \cdot 0.05^2} \approx 462 \quad [\text{iSC/km}^2] \tag{4-35}$$

Given that for this scenario there are in average $N_{UE} = 2$ users served per iSC [10] the user density is equal to

$$\rho_{UE} = N_{UE} \cdot \rho_{iSC} = 2 \cdot 462 = 924 \quad [\text{UE/km}^2] \tag{4-36}$$

Applying the equations (4-30) and (4-34) it is then possible to calculate the required downlink backhaul capacity per area and the achievable area throughput respectively. It must be pointed out that the used model does not assume any form of coordination among the iSCs. The results are summarized in Table 4-11. In this calculation, the transmit power P of each iSC is set to 30 dBm and the equivalent noise bandwidth is 9 MHz. The thermal noise power σ_n^2 at the UE side is calculated assuming a receiver noise figure F equal to 9 dB [39].

Table 4-11: Required BH capacity and Area Throughput for wide-area scenario (fixed iSC density)

Split option	Required BH capacity C_{BH} in Gbit/s/km^2	Area Throughput $R_{A_{net}}$ in Gbit/s/km^2
A.1	576.1	11.8
A.3	144.9	11.8

The obtained area throughput given in Table 4-11 corresponds to an average throughput per iSC of about $R_{A_{net}}/\rho_{iSC} \approx 25.5$ Mbit/s that for a bandwidth $B = 10$ MHz corresponds also to a spectrum efficiency of 2.55 bit/s/Hz. The BH load reduction provided by the split option A.3 can be also exploited in a different way, namely to increase the iSC density when the available BH capacity is fixed. The assumption of a limited BH capacity per area holds for example in case of wireless backhaul with specific network topologies (e.g. tree, ring, chain) or also for other BH technologies, like fibre, when is it difficult or even not possible to add further capacity. For example assuming an available BH capacity of about 145 Gbit/s/km^2, as in the previous calculation, the corresponding area throughput for the two split options shows a gain of about 100% for the option A.3, as shown in Table 4-12.

Table 4-12: Required BH capacity and Area Throughput for wide-area scenario (fixed BH capacity)

Split option	Required BH capacity C_{BH} in Gbit/s/km^2	Area Throughput $R_{A_{net}}$ in Gbit/s/km^2	Gain
A.1	144,9	5.94	–
A.3	144.9	11.8	98.6%

Stadium scenario

Concerning the stadium evaluation scenario, we consider $N_{iSC} = 15$ iSCs placed over a uniform grid for a total covered area of 80×40 m [10]. The aggregated BH capacity required for this scenario in order to satisfy the full load condition is calculated from the BH throughput per iSC provided in Table 4-9. In case of the split option A.1 the aggregated BH throughput is equal to about $N_{iSC}\eta\gamma = 15 \cdot 1.23 = 18.4$ Gbit/s while for the split option A.3 it reduces to about $N_{iSC}\eta\gamma = 15 \cdot 0.3136 = 4.7$ Gbit/s. Besides, considering that the iSCs are closely deployed, there is the possibility that a unique BH physical link, shared among the iSCs, is used to connect them to the RANaaS, as shown in the physical deployment example of Figure 4-36 [10].

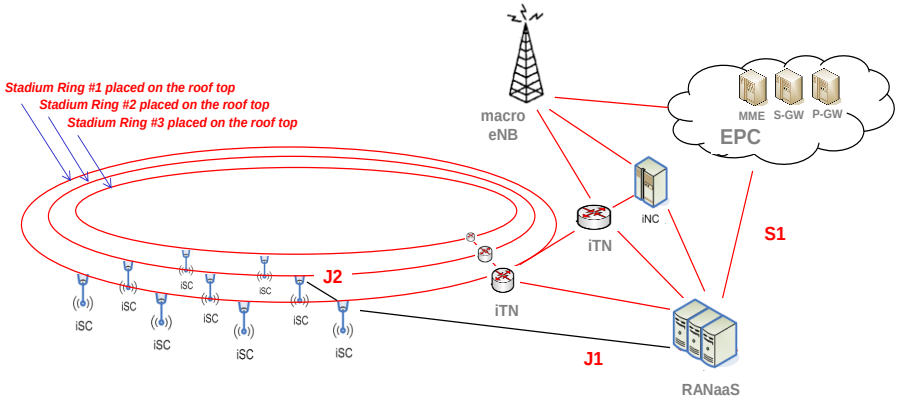


Figure 4-36: Stadium – Physical deployment example

In such kind of deployment, using the split option A.3, it becomes possible to further reduce the required BH capacity as the BH load generated by each iSC is proportional to the radio interface load. As the full load condition (i.e. all iSCs with full PRB utilization at the same time) occurs with a certain probability, it becomes then possible to reduce the BH capacity requirement, if a certain level of blocking in the BH can be tolerated. The Figure 4-38 shows a numerical example where the BH blocking probability is calculated assuming that the PRB utilization in one downlink subframe for a given iSC is Gaussian distributed with mean value of {5, 15, 25, 35, 45} PRBs and variance of 5 PRBs.. This situation corresponds to an average PRB usage in each iSC of about {10%, 30%, 50%, 70%, 90%} respectively, as the total number of PRBs for a 10 MHz bandwidth is 50. The PRB utilization is supposed statistically independent among the 15 iSCs of the stadium scenario, which corresponds to a favourable condition in terms of reduction of the blocking probability for a given backhaul capacity. However, in a real system design, this condition has to be verified with the actual traffic and, more in general, it should be favoured by possibly aggregating sources that are statistically independent, The Figure 4-37 shows the probability distribution of the number of allocated PRBs in each subframe for a given iSC.

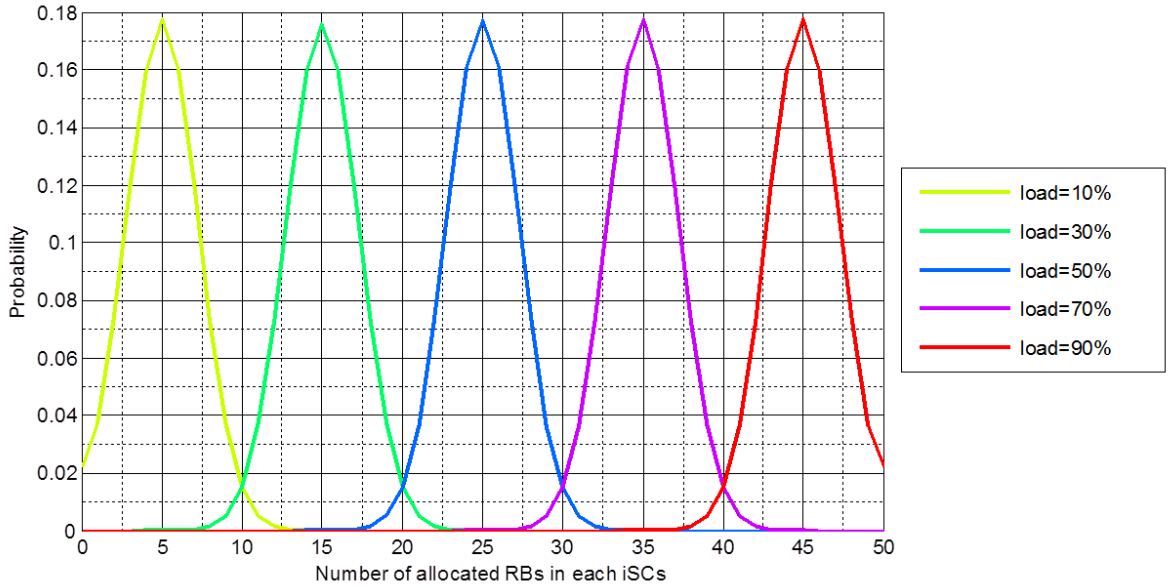


Figure 4-37: Probability density function of the number of allocated PRB per subframe in each iSC

Figure 4-38 shows the BH blocking probability calculated as a function of the available BH capacity. As shown before, the blocking probability becomes zero for an aggregated BH capacity of about 4.7 Gbit/s.

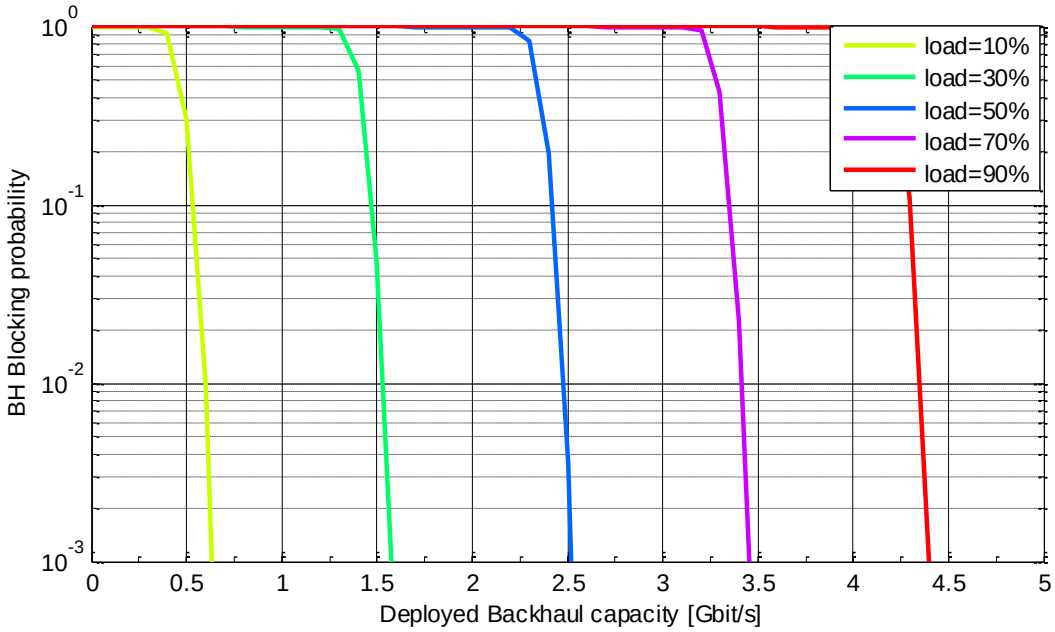


Figure 4-38: BH blocking probability vs. BH available capacity

However, assuming that in a real design, a 1% blocking probability in the BH can be accepted, the deployed BH capacity can be reduced to the values listed in Table 4-13.

Table 4-13: Required BH capacity for the stadium scenario vs. average PRB usage (split option A.3)

Average PRB usage in each iSC [%]	Required BH capacity in Gbit/s	Blocking probability
10	0.6	1%
30	1.58	1%
50	2.49	1%
70	3.46	1%
90	4.39	1%
100	4.7	0%

Conclusions

Based on the evaluation results it can be asserted that the data compression over RoF provides a reduction of the BH capacity requirement of a factor ranging from 3 to 4 times, compared to a fully centralized architecture (i.e. option A.1), for the same iSC density and designing the BH network to support the full load condition (i.e. design for zero BH blocking probability). A further reduction of the BH load is achievable by designing the BH network for a tolerable level of blocking (e.g. 1% plus eventually a margin ε [40]), assuming that it is known or estimated in some way the traffic statistics. It also follows that the split option A.3 makes the BH costs/investments proportional to the actual traffic growth and may also improve the BH utilization efficiency through the exploitation of the statistical multiplexing. The BH throughput reduction can be also exploited in a different way, namely to densify the network with respect to a fully centralized architecture, when the BH deployed capacity is fixed (e.g. already deployed). In such a case the estimated gain on the AT metric is about a factor of 2 (i.e. 100%) for the wide-area scenario. Last point to be considered is that split option A.3 retains the same potential centralization gain of the fully centralized architecture (i.e. multi-user detection/decoding, coordinated transmission).

4.7 CT2.7: Millimetre Wave Backhauling

4.7.1 Final implementation of CT

This CT deals with PHY layer aspects of introducing wireless BH into the RANaaS architecture. The unique challenge arising in mmWave links is that, in contrast to fibre links, the channel conditions can vary, as mmWave links are attenuated by, e.g. rain. The usual way to deal with this would be adaptive coding and modulation on the BH links. However, this introduces latency, so this CT aims to not use an additional BH en/decoder but to utilize the RAN code for this. The two main concepts, which are both applicable to the uplink only, are depicted in Figure 4-39:

- Joint encoding: the encoder at the UE can take the BH channel into account, lowering the code rate R_c^{RAN} in case of unfavourable BH conditions
- Joint decoding: the decoder of the RAN-code can be adapted to take an erroneous BH channel into account

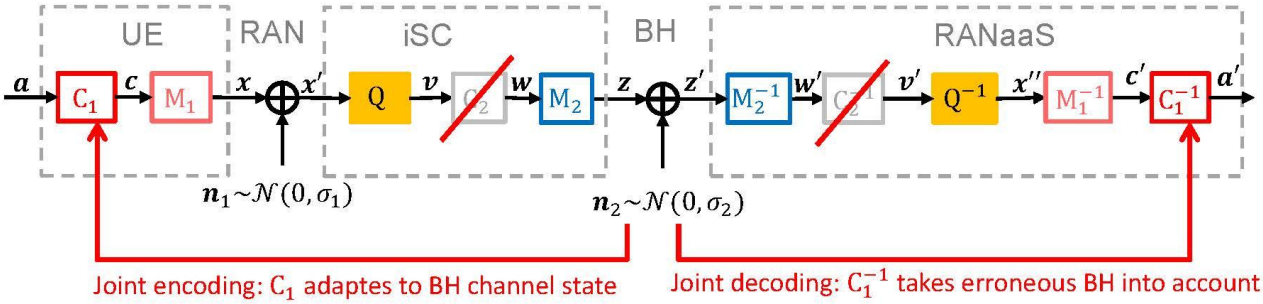


Figure 4-39: Overview of the approach for joint design of access and backhaul

Detailed descriptions of these two approaches can be found in the previous deliverable [2] as well as in publications [36], [37]. Here, we only want to recapture the most important aspects.

The concept of joint encoding is very simple and easily integrates into the iJOIN architecture. An erroneous BH transmission will decrease the overall performance of the end-to-end link. Since the access code protects the data from end to end, BH errors can be mitigated by adapting the RAN's MCS.

For joint decoding we have developed two techniques: the error resilient decoder (ERD) and the soft-input/soft-output dequantizer (SISODQ). Both use the same concept: channel state information of the BH channel is forwarded and utilized in calculating the LLR values that serve as input to the RAN decoder. While the ERD only uses the (average) BER as channel state information of the BH, the SISODQ takes the symbol-by-symbol LLR values of the BH symbols and combines them into LLR values of the access symbols.

4.7.2 Evaluation of the CT

The new approaches developed within CT2.7 are single-link technologies, meaning that they improve the performance of a single UE-iSC-RANaaS link. To make it comparable to other system wide approaches the evaluation has been performed in 2 stages: first, the performance for a single link is investigated and then these results are mapped to a larger scenario by the methodology described in the following. The main KPI investigated is area throughput, however, there are also short discussions on energy and cost efficiency.

Evaluation Methodology

The algorithms introduced in CT2.7, namely joint en/decoding, aim to improve the performance of a single link in terms of BER, which is the basic KPI for this CT:

$$BER = \|\boldsymbol{a} - \boldsymbol{a}'\|_0 \tag{4-37}$$

The notation $\|.\|_0$ denotes the 0-norm, i.e. the number of bit errors between the binary vectors $\boldsymbol{a}$ and $\boldsymbol{a}'$. This BER can be mapped to a normalized throughput via

$$\begin{aligned} \eta &= (1 - BER)^{N_{block}} \cdot S \\ &= (1 - BER)^{N_{block}} \cdot Q_m^{RAN} \cdot R_c^{RAN} \cdot R_c^{BH} \end{aligned} \tag{4-38}$$

The BER, and subsequently the throughput, depend on a number of parameters that will be varied in the evaluations These parameters include the access SNR SNR_{RAN}, the BH SNR SNR_{BH} and the used MCSs which are identified by the code rates and the modulation schemes $Q_m^{RAN}, R_c^{RAN}, Q_m^{BH}, R_c^{BH}$. Of these, Q_m^{BH} is usually set as constant to 2 for sake of simplicity, as the BH performance only depends on the relation SNR_{BH} and Q_m^{BH}. Therefore, the extension to other Q_m^{BH} is straightforward.

The performance also depends on the applied algorithm ERD or SISODQ, both using an *uncoded* BH. As baseline, a conventional detection/decoding scheme with *uncoded* BH without forwarding channel quality information of the BH is used. For comparison, we also evaluate a conventional scheme which incorporates a *coded* BH. It is obvious that a coded BH is advantageous to uncoded BH in terms of BER. However it introduces latency, which can be partially avoided by the proposed algorithms, and additional BH overhead, which is not the case for uncoded BH. To make a 'fair' comparison possible we introduced the notion of a 'total code rate' $R_c^{tot} = R_c^{RAN} \cdot R_c^{BH}$ in previous deliverables. In this deliverable, to be able to use the standardized MCS, we instead include the R_c^{BH} into the throughput in Equation (4-38). The notion behind is this following: we assume that the BH data rate is limited. If using coded BH, this means that the access rate has to be reduced, because the available BH rate is used up by the coding overhead.

The single link evaluations also require a number of different parameters, which are kept constant and are summarized in Table 4-14. The parameters are as far as possible aligned with the common assumptions in Section 5.

Table 4-14: Parameter values for link evaluation

Parameter	Value
B	10 MHz
N_Q	4 for $Q_m^{RAN} = 2$ 6 for $Q_m^{RAN} = 4$ 8 for $Q_m^{RAN} = 6$
N_R^{iSC}	2
N_{block}	256
Code RAN/BH	LTE turbo code with rate matching
Access channel	Rayleigh block fading with variable SNR
Equalizer	MMSE with perfect CSI
BH channel	AWGN with variable SNR

For system level evaluation, the convex hull of the throughput curves of multiple MCS is required. To also account for features like HARQ that were not simulated, we calculate the system level throughout by fitting a curve of the shape

$$R = \min(\alpha_1 \cdot \log_2(1 + \alpha_2 SNR), \alpha_3) \tag{4-39}$$

to the simulated MCS curves (see also Figure 4-40 b)).

To map the single link results to system level, we require the distribution of the SNR_{RAN} in a certain scenario. By weighting the single-link throughput for each SNR_{RAN} with its probability, we can calculate the average system throughput of the given scenario. The final area throughput is then calculated by scaling the normalized throughput with the number of subcarriers, symbols per time frame and frame duration and by normalizing it to the cell area:

$$R_{A_{net}} = \mathrm{E}[R(SNR_{RAN})] \cdot N_{SC} \cdot N_{SYMB}^{SUB} \cdot T_S^{-1} \cdot A_{veNB}^{-1} \tag{4-40}$$

The SNR distributions are acquired by a simple system level simulation that calculates the uplink SNR of iSCs and aggregates this information for a large number of user drops. Interference from users in other cells is treated as noise.

It should be noted that for each resource a user is scheduled on, a maximum of one user in other cells can cause interference. Therefore, the number of users in a cell does not influence the SNR distribution. The number of users only has an impact on the throughput per user, yet not on the area throughput, as the total throughput is simply split between the users in the area. Therefore, the single-user TP can also be seen as the area TP normalized to the cell area.

Link level results

Figure 4-40 a) shows exemplary the BER performance of the ERD for varying SNR_{RAN} with $SNR_{BH} = 5$ dB and using the LTE MCS 2-28. Figure 4-40 b) shows the corresponding normalized throughput as well as the fitted system level throughput. Similar results have been produced for all four described algorithms and two different SNR_{BH}.

a) BER b) normalized throughput

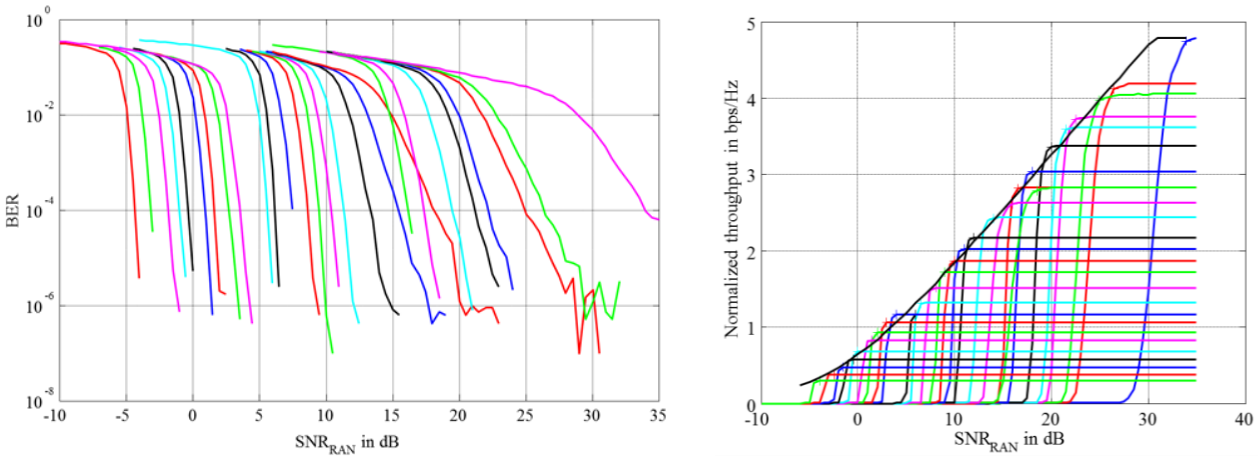


Figure 4-40: Exemplary BER (a) and normalized throughput for different MCS and fitted system throughput (b) for the ERD

Figure 4-41 shows the direct comparison of BER of all four techniques for the exemplary MCS 11. We can see that the proposed schemes are clearly superior to the conventional scheme with uncoded BH, with the SISODQ having the greatest improvement. The coded BH is advantageous in terms of BER performance, which is to be expected as explained earlier, because it sends additional redundant information over the BH. However, this will reduce the user throughput as can be seen below.

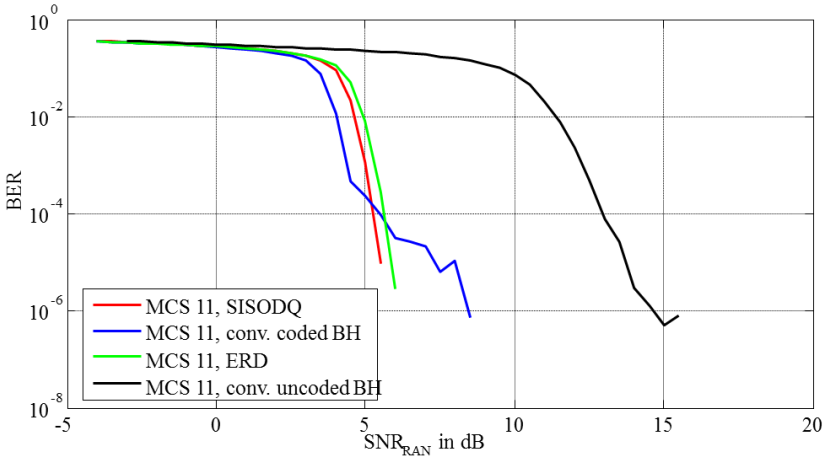


Figure 4-41: BER comparison of proposed (SISODQ, ERD) and baseline (conv) algorithms for exemplary MCS

Figure 4-42 shows the system throughput for all four schemes for two different SNR_{BH}. As can be seen, for a low SNR_{BH}, the SISODQ performs best, especially for higher MCS. The ERD also performs better than the baseline algorithms.

For a higher SNR_{BH} it is apparent that ERD and SISODQ perform equally well. The difference of the SISODQ as compared to the ERD is that symbol-by-symbol information on the BH is forwarded while the

ERD only forwards general statistical information. For a sufficiently high SNR_{BH} the symbol-by-symbol information seems to not provide an additional gain.

From Figure 4-42 it is can also be seen that the gain in AT could be traded for a lower SNR, i.e. less transmit power, which results in a higher energy efficiency. However, this would only reduce the UEs transmit power, which commonly have a maximum of 200 mW. Compared to the total UE power consumption including baseband, CPU, etc (about 2 W), the iSC power (up to a few kW) and the RANaaS platform power (MW), even a large reduction in UE TX power would only marginally increase the overall energy efficiency of the network.

a) $SNR_{BH} = 5$ dB b) $SNR_{BH} = 10$ dB

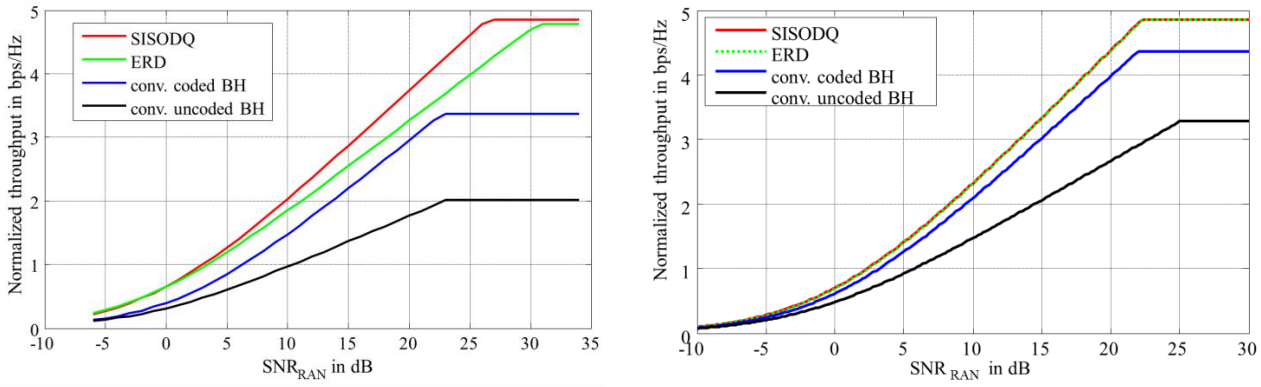


Figure 4-42: Comparison of normalized system throughput of proposed (SISODQ, ERD) and baseline (conv) algorithms for different SNR_{BH}

System level results

Three of the common scenarios that are described in deliverable D5.2 [10] have been investigated: CS 1 (Stadium), CS 2 (Square) and CS 3 (Wide-area). As mmWave links require a LoS link, they are not an option for indoor deployment and thus CS 4 is not investigated. Figure 4-43 show the CDFs of the SINR obtained for the three scenarios. The stadium exhibits the highest SINR due to the small ISD. For wide-area, the ISD is larger and consequently the SINRs are smaller. In comparison to the other two, the square scenario uses a random deployment and subsequently interference between users is higher and thus it exhibits the lowest SINRs of the three scenarios.

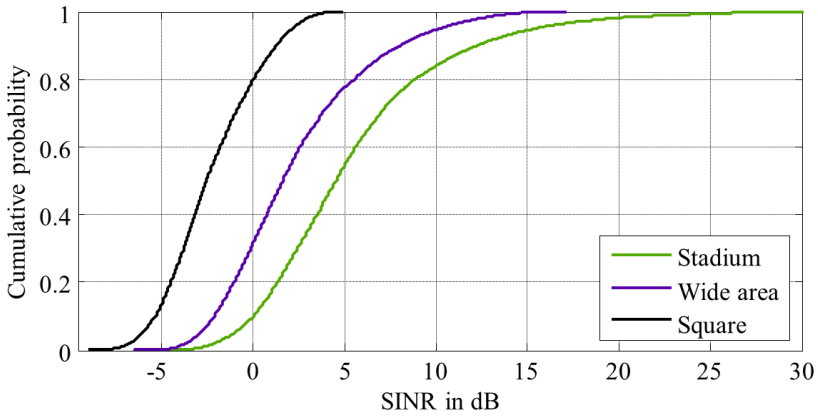


Figure 4-43: CDF of SINRs for the three investigated scenarios

Figure 4-44 shows the corresponding area throughput evaluation, again for two different SNR_{BH}. The results are very similar in all three scenarios and match the link level results discussed above. The two proposed schemes, ERD and SISODQ, perform better than the two baseline approaches, with the SISODQ having a slight advantage over the ERD. As these results now also incorporate the limitations of the BH code, it becomes apparent that the conventional coded BH scheme is inferior to the proposed schemes.

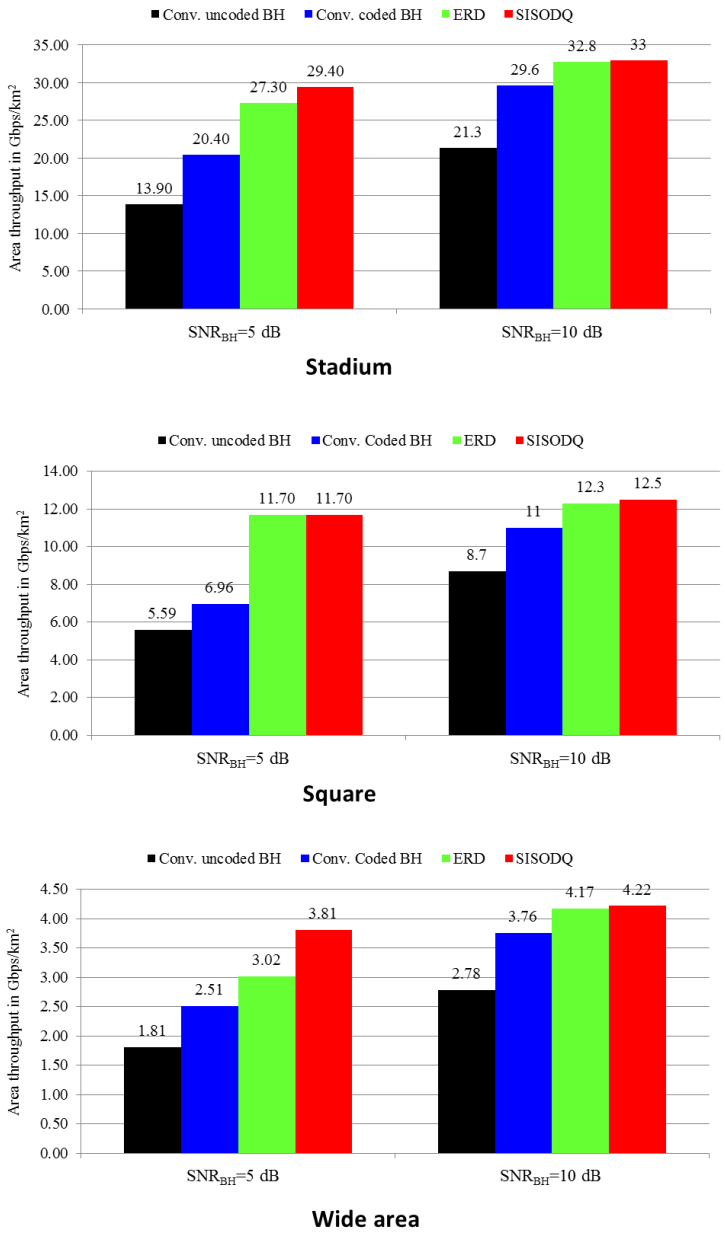


Figure 4-44: Area throughput comparison of CT2.7 schemes

The square scenario achieves the highest total throughput, due to its dense and systematic deployment. As the square scenario is more random, the SINR is general more unfavourable and thus the AT is reduced. As the SISODQ only performs better than the ERD for high SNRs, both algorithms perform equally in terms of AT in this scenario. The wide-area scenario uses a less dense deployment with larger inter-site distances and thus has a larger cell area and subsequently a lower total AT. As discussed for link level results above, there is only a very small difference between SISODQ and ERD for a higher SNR_{BH}, and the gain compared to the baseline schemes is lower.

Conclusions

In summary, we showed that the proposed schemes for a joint coding of access and backhaul can lead to significant improvements in area throughput. While they slightly increase the complexity of decoding in the RANaaS platform, they offer the possibility to completely remove a dedicated BH de/encoder, thus reducing the hardware required and lowering the backhaul latency. Combined with the fact that mmWave backhaul deployments are generally cheaper than fibre, CT2.7 with the proposed schemes can thereby increase cost efficiency, although a detailed cost analysis could not be performed within the scope of this CT.

5 Performance Evaluation

In Section 4 the considered PHY layer CTs were presented and the performance evaluations were discussed in detail for each CT separately. In this section we discuss the interaction of these CTs by elaborating on their interoperability. Furthermore, for the four iJOIN common scenarios (CS) defined in D5.2 [10] the performance results of a selection of suitable CTs and their functional split options are presented. Thus, we aim to demonstrate the flexibility of the considered CTs for the varying test and deployment scenarios.

In Subsection 5.1, we provide a summary of the primary function, the affected area and the processing scope for each WP2 CT. Subsection 5.2 provides a brief summary of the key performance indicators (KPIs) and collects for each CT the main KPIs and the evaluation approach. Furthermore, an overview of the mapping between WP2 CTs and the iJOIN CSs is given. In Subsection 5.3 the performance evaluations per CS are provided. First, for each CS the backhaul technologies for the functional split options are defined, before the gains achieved by considered CTs are provided. Subsection 5.4 ends this section with a summary.

5.1 Interoperability

As a summary, Table 5-1 lists for each WP2 CT the primary function, the optimization area and the processing scope. CT2.1 and CT2.2 are alternative approaches for joint multi-user detection in the UL, where CT2.1 estimates the UL signals distributively among iSCs and CT2.2 considers mainly centralized processing in the RANaaS platform or purely local detection within the iSCs. Thus, depending on the BH deployment, the appropriate approach can be chosen. CT2.3 considers also the estimation of UE signals, but focusses more on a joint design of RAN and BH in case of low rate BH links. Centralized as well as decentralized approaches for joint transmission in the DL are addressed by CT2.4 and CT2.5. CT2.4 focusses on the centralized precoding design taking the constrained complexity of the RANaaS platform and the different latency impacts of the overall transmission link into account. For centralized processing, high rate J1 links are assumed. Conversely, CT2.5 considers rate limited J1 and J2 BH links for distributed precoding. A combination of both approaches is possible in principle, but has not been considered as the preferred functional splits of these CTs are different. In addition to these CT which consider a set of UEs and iSCs, CT2.6 and CT2.7 concentrate on the backhaul and the joint RAN/BH optimization for point to point (PtP) links. Due to its basic nature, CT2.6 provides general results and can be combined with all UL and DL CTs in principle. As CT2.7 considers a joint coding for RAN and BH link with focus on mmWave backhauls, it can be combined with all UL oriented CTs.

Table 5-1: Categorization and compatibility of WP2 CTs

CT	Function	Optimization area	CT scope	Interoperation
2.1	UL Detection	RAN	Distributed	Different alternatives for UL Detection
2.2	UL Detection	RAN	Distributed / Centralized	
2.3	UL Detection	Joint RAN/BH	Centralized	
2.4	DL CoMP	RAN	Centralized	Interoperation is potentially possible, however due to the choice of functional splits not exploited
2.5	DL CoMP	Joint RAN/BH	Distributed	
2.6	Functional Split/BH for UL and DL	BH	PtP	Support for other UL & DL approaches
2.7	mmWave / coding for UL	Joint RAN/BH	PtP	

For each of the four iJOIN CS defined in D5.2 [10] we will discuss the principle functional split options given a specific BH deployment and the achieved performance gains. Based on the actual scenario, the iveC may select the appropriate transmission approach. As the UL and the DL communication flows are basically independent of each other, the corresponding CTs are interoperable and it is even possible to apply different functional splits for UL and DL at the same time.

5.2 Evaluation Methodology

The final description of the technical contributions has been provided in Section 4. The main approaches are described and for each CT the evaluation set up is described and quantitative simulation results are provided. In the following, we provide a global vision on the evaluations by first presenting the main figures of merit discussed in this project in Subsection 5.2.1 before summarizing the key evaluation parameters for each CT in Subsection 5.2.2.

5.2.1 Evaluation Metrics

The iJOIN project has defined in deliverable D5.2 [10] the following four metrics to be improved by using a joint design of RAN and BH for the envisioned dense deployment of small cells with cloud processing.

- **Area Throughput (AT)** is the main metric targeted by CTs in WP2 by applying sophisticated PHY layer approaches. The area throughput is given by the data rate successfully delivered to or received form the users in a given geometrical area A_{net} over time. With $r_u(t)$ denoting the bits correctly delivered to (from) user u at the time slot t, the area throughput in a network comprising N_{UE} users is given by

$$R_{A_{net}} = \frac{R_S}{A_{net}} = \frac{1}{A_{net} \cdot T} \sum_{u=1}^{N_{UE}} \int_{t=0}^{T} r_u(t) dt \quad [\text{bit/s/km}^2] \tag{5.1}$$

  where R_S is the average sum rate measured in the time interval T.

- **Energy Efficiency (EEf)** considers the power consumption of the different entities in the iJOIN network and depends for example on the complexity of algorithms, the placement of functions due to the applied split, the RF transmit power, the BH technology, etc.

- **Utilization Efficiency (UEf)** measures how well the utilized resources are used to realize a given system performance. CTs which allow for flexible functional split depending on the current needs (e.g. BH load, required user data rate, etc.) are appropriate to utilize the different resources well-balanced and prevent over-provisioned system design.

- **Cost Efficiency (CEf)** depends on the number and kind of nodes involved in the network and recognizes in addition the cost of the backhaul infrastructure determined by factors like distances, BH technology, etc.

In the following, we will present the performance gains with respect to the corresponding baseline solution specified for each CT in Section 4. With $R_{A_{net}}^{CT}$ denoting the area throughput of the considered CT and $R_{A_{net}}^{Base}$ specifying the AT of the baseline, the gain G_{AT} is defined as:

$$G_{AT} = \frac{R_{A_{net}}^{CT} - R_{A_{net}}^{Base}}{R_{A_{net}}^{Base}} \tag{5.2}$$

5.2.2 Evaluation Configuration Sets

As shown in Table 5-2, all WP2 CTs address the improvement of the area throughput. This is meaningful as this KPI is directly influenced by the physical layer processing. It allows considers both the densification of the network and the advanced physical layer schemes. It should be noted that this AT is an average figure of merit such that the distribution of the rates per user is not taken into account. In particular, the allocation of the rate could be unfair and leads to a bad quality of service to weak users. This is, however, a problem which is usually dealt with in higher layers, in particular when allocating the resources to the users. For each CT, the procedure of how the AT performance is evaluated is briefly described. In addition, the secondary KPI addressed by a CT is defined where qualitative explanations of the achieved results have been provided for each CT in Section 4. In order to demonstrate the properties of a CT, it is meaningful to investigate the KPIs with respect to varying system parameters like number of users N_{UE}, number of iSCs N_{iSC}, the BH rate on the J1 or J2 interface, and the BH latency.

Table 5-2: Evaluation of main KPIs per CT

CT	Main KPI	Evaluation of main KPI	Sec. KPI	Ev. parameters	Comments
2.1	AT	FER is evaluated to determine AT together with CT3.8	UEf	N_{UE}, N_{iSC}, BH Rate	RRM by CT3.8
2.2	AT	FER is evaluated, AT investigated together with CT3.7	UEf	fixed N_{UE}, N_{iSC} in WP2, varying N_{UE}, N_{iSC} in WP3	RRM by CT3.7
2.3	AT	SINR is evaluated analytically, the AT is derived usding network capacity for the corresponding SINR values	UEf	J2 rate	Algorithm for user association from CT3.2
2.4	AT	Output SINR is evaluated, the AT is derived based on SINR	EEf	N_{iSC}, computational resource	
2.5	AT	AT is evaluated through link level simulations.	EEf	J1 delay, Number of cooperating iSCs	
2.6	AT	Analytical evaluation	UEf	J1 BH rate	
2.7	AT	Single link throughput	EEf CEf	one UE-iSC pair, RAN SNR, BH SNR	

Furthermore, the different PHY layer CTs discussed in Section 4 allow for different functional splits. Table 5-3 lists the possible functional splits for each CT with a brief explanation to indicate the chosen variant as defined in Section 3.2.1. Based on this, promising CTs are identified for the common scenarios in Section 5.3.

Table 5-3: Functional splits per CT

CT	A	B	C
2.1	–	INP forwarding soft bits	INP with decoding per iSC
2.2	MPTD with FD IQ forwarding of PRBs	–	SPTD
2.3	–	JNCC in RANaaS	JNCC in iSC_2
2.4	Centralized precoding forwarding PRBs	–	–
2.5	–	Distributed precoding forwarding IQ signals	Distributed precoding forwarding uncoded UE messages
2.6	RoF	–	–
2.7	mmWave supports RRH	mmWave forwarding of EQ output → decoding in RANaaS	–

5.3 Performance Results

For the numerical performance investigations the different CTs are evaluated in the four iJOIN common scenarios

- CS 1: Stadium
- CS 2: Square
- CS 3: Wide-area continuous coverage
- CS 4: Shopping Mall / Airport

defined in D5.2 [10]. Due to the inherent properties of a CT, not every approach is applied to every CS. In contrast, the CTs are applied where they will show their benefits. Table 5-4 indicates the mapping of PHY CTs to the investigated common scenarios and specifies in addition the evaluated principle functional split. In the subsequent investigations, for each functional split per CS the corresponding backhaul technologies as listed in Appendix I are specified following the guidelines in D4.3 [8] for the scenarios. Based on the actually deployed BH and the current communication needs, it is possible to identify from the different alternative solutions the most promising CT. This allows obtaining a general, adaptive, and robust framework for communications.

Table 5-4: Mapping of PHY CTs and functional splits to iJOIN Common Scenarios

Scenario	2.1 INP	2.2 MPTD	2.3 JNCC	2.4 CoMP	2.5 ICIC	2.6 RoF	2.7 mmWave
CS 1 Stadium		A		A			B
CS 2 Square			C				B
CS 3 Wide-area continuous coverage	B C	A C		A	B C	A	B
CS 4 Indoor (Airport / Shopping Mall)	B C	A C					

5.3.1 CS 1: Stadium

As the deployment of high data rate backhaul technology like fibre or mmWave is quite likely for the stadium scenario, the physical layer functional split options A and B are investigated. In Table 5-5 the suitable backhaul technologies according to Appendix I as well as the investigated CTs are specified. In case of functional split option A, joint uplink detection by CT2.2 and joint downlink transmission by CT2.4 are considered. Furthermore, joint RAN/BH coding for mmWave links is investigated by CT 2.7 with functional split B.

Table 5-5: Backhaul technologies for functional split options in CS 1

Split Option	BH Number	BH Technology	#Hops	Latency	Bandwidth	CT
A	4b	CWDM	J1 Ring topology	5µs	10 Gbit/s	2.2 2.4
B	1c	Millimetre wave	1	< 200 µs	2.5 Gbit/s	2.7

For uplink transmission, CT2.2 realizes joint multi-user detection by implementing the multi-point turbo detection (MPTD) approach centrally in the RANaaS platform in order to increase the aggregated uplink user throughput. To this end, the I/Q receive signals of the cooperating iSCs on allocated RPBs (i.e. functional split option A.3) are forwarded over high rate J1 links to the central processing unit, which is feasibly to the deployed coarse wavelength division multiplexing (CWDM) ring topology. As specified in Table 5-6, gains in area throughput of up to 60% compared to the baseline can be achieved.

For the downlink transmission, the centralized joint transmission CoMP being investigated in CT 2.4 with functional split options A.1, A.2, or A.3 can be applied. CT2.4 investigates the delay impact on the area throughput in terms of parameters such as available computational resources at the RANaaS platform and precoding granularity in order to find the optimal number of cooperative iSCs. Among the delay factors, the precoding matrix calculation delay dominates when the number of cooperative iSCs is large and the available computational resources are small. The CT provides a comparison of area throughput in terms of number of cooperative iSCs for different precoding granularity, aiming to obtain the optimal number of iSCs to cooperate. The investigations show that with the optimal granularity and number of cooperative iSCs, 46 % gain in terms of AT can be achieved compared to the baseline. The summarized gain and backhaul load for the different approaches are shown in Table 5-6., where obviously A.3 requires the lowest BH traffic yielding the same AT gain as the other split options.

CT 2.7 investigates joint RAN/BH encoding over mmWave links by applying functional split option B. To this end, the concepts joint encoding and joint decoding are proposed to adapt to the varying channel conditions. For joint encoding, the encoder at the UE takes also the BH channel into account, lowering the code rate in case of unfavourable BH conditions. Joint decoding of the RAN FEC code can be adapted to take an erroneous BH channel into account. Two techniques are developed for joint decoding: the ERD and the SISODQ as discussed in Section 0. The performance evaluations indicate gains in area throughput ranging from 44% up to 110% with respect to a baseline, i.e. uncoded BH.

Table 5-6: Overview of gains achieved by CTs in CS 1

Split Option	CT	UEs per iSC	J1 BH load in Mbit/s	J2 BH load in Mbit/s	Gain AT
A	2.2	1 5	302.4 per iSC (max)	0	56% 60%
	2.4	1	A.1: 1840 per iSC A.2: 470.4 per iSC A.3: 235.2 per iSC	0	46%
B	2.7	Not relevant	134.4 per iSC	0	<110%

The investigations for the stadium scenario show that gains of up to 60% and 46% over the corresponding baseline schemes are achievable for uplink and downlink transmission, respectively. However, both functional split A approaches relay on a backhaul technology with high capacity and low latency such as dark fibre. With joint RAN/BH coding for mmWave, a gain of 110% over the baseline with functional split B was demonstrated.

5.3.2 CS 2: Square

The square scenario is considered by CT 2.3 and CT 2.7 where joint coding or decoding schemes are investigated for the UL. The considered backhaul technologies for both functional split options are listed in Table 5-7 using the parameters specified in Appendix I. In CT2.3, the joint network channel coding (JNCC) is investigated considering functional split C. The decoding is performed at the RANaaS platform which is supposed to be physically located in the MeNB. In the corresponding iRPU the user messages are decoded using only direct links if the user is associated directly to the MeNB. However, if the user is associated to the iSC, two schemes were investigated: (i) the pure relaying scheme in which the corresponding iRPU decodes its messages using only the signal received via the backhaul link and (ii) the JNCC scheme in which the corresponding iRPU decodes using the backhaul link as well as the direct links. The maximum number of hops is considered to be two. The backhaul is supposed to be with limited capacity (xDSL), since the JNCC achieves gains with respect to pure relaying only when the backhaul capacity is limited. However, due to the large latency of xDSL only non-time critical applications may be served.

CT2.7 considers a joint coding of access and backhaul following functional split B. It deals with mmWave BH in which mmWave links can be attenuated due to bad channel conditions. The usual way to address this problem would be adaptive coding and modulation on the BH links, however this introduces latency. Hence, CT2.7 proposes joint encoding and decoding which takes into account the BH channel as well.

Table 5-7: Backhaul technologies for functional split options in CS 2

Split Option	BH Number	BH Technology	#Hops	Latency	Bandwidth	CT
B	1c	Millimetre wave	1	<200 µs	< 2.5 Gbit/s	2.7
C	5	xDSL	≤ 2	5-35 ms	10-100 Mbit/s	2.3

An overview of gains achieved by CT2.3 and CT2.7 is given in Table 5-8. With CT2.3 a gain of 32 % in AT can be achieved using JNCC instead of pure relaying when the number of UEs per iSC is limited to 10 and

the BH data rate is 20 Mbit/s for 4 iSCs. It was shown in Section 4.3 that this gain decreases when the backhaul capacity increases and is negligible when the backhaul capacity is sufficiently large. In CT2.7, the proposed schemes for a joint coding of access and backhaul can lead to significant improvements in area throughput up to 109% depending on the BH channel quality. While the proposed schemes increase the complexity of decoding in the iRPU of the RANaaS instance, they offer the possibility to completely remove a dedicated BH de-/encoder, thus reducing the hardware required and lowering the backhaul latency.

Table 5-8: Overview of gains achieved by CTs in CS 2

Split Option	CT	UEs per iSC	J1 BH load in Mbit/s	J2 BH load in Mbit/s	Gain AT
B	2.7	Not relevant	134.4 per iSC	0	<109%
C	2.3	≤ 10	20 for 4 iSCs	0	32 %

To summarize, with CT2.3 and CT2.7 two approaches have been developed for improving the AT for the square scenario which is characterized by backhaul links with reduced throughput.

5.3.3 CS 3: Wide-area continuous coverage

For the wide-area continuous coverage scenario the backhaul technologies for the considered functional split options as well as the investigated CTs are specified in Table 5-9. In order to connect the iSCs with RANaaS dark fibre deployment is considered to support functional split A, whereas wireless links for the different J1 and J2 connections are used in case of functional split B and C.

Table 5-9: Backhaul technologies for functional split options in CS 3

Split Option	BH Number	BH Technology	#Hops	Latency	Bandwidth	CT
A	4a	Dark fibre	1 hop	5µs	10 Gbit/s	2.2 2.4 2.6
B	1c	Millimetre wave	J1: 3 hops J2: 2 hops	J1: 600 µs J2: 400 µs	2.5 Gbit/s	2.1 2.5 2.7
C	J1: 3a J2: 1c	J1: Sub-6 GHz PtP NLOS TDD J2: Millimetre wave	J1: 2 hops J2: 2 hops	J1: 10 ms J2: 400 µs	J1: 500 Mbit/s J2: 2.5 Gbit/s	2.1 2.2 2.5

For the uplink, two alternative approaches for joint detection are investigated by means of CT 2.1 and CT 2.2. In CT 2.1, joint MUD via In-Network Processing is performed, considering functional splits B.2 and C.2 as possible alternatives for the actual implementation. In both cases, the iSCs perform joint detection of the user signals by exchanging local estimates over the J2 interfaces based on the applied INP algorithms. Depending on the split option, either LLR values or decoded messages are forwarded to the RANaaS platform. The performance evaluations of CT 2.1 show that the ALCE algorithm achieves a throughput performance close to the centralized detection which serves as the upper bound performance. A gain of approximately 130% in area throughput by the ALCE algorithm can be observed compared to the baseline (local detection) for split options B.2 and C.2. In case of split option B.2, the combination with the joint RAN/BH coding approach for mmWave transmissions investigated by CT2.7 can be realized. This combination has been implemented on the joint access and backhaul testbed as discussed in Section 4.2.2 of D6.2 [13]. This brings together the advantages of MUD and the joint RAN/BH coding.

In CT 2.2 turbo detection among multiple iSCs is applied to increase the aggregated uplink user throughput. The turbo processing can either be done centrally at the RANaaS platform (MPTD) or locally in each iSC (SPTD). In case of MPTD, the functional split A is applied and quantized I/Q samples are sent to the

RANaaS platform requiring a high backhaul rate and, more importantly, low latency for the J1 interface. In comparison to split option A.2, the forwarding of samples after PRB demapping according to split A.3 as investigated by CT2.6 can exploit the advantage of the statistical multiplexing gain on the BH. In case of SPTD, functional split C.2 is applied and the decoded bits are transmitted by the iSCs to the RANaaS platform. The performance evaluation of CT 2.2 shows gains of up to 51% for 2 UEs/iSC and 65% for 10 UEs/iSC, in terms of AT compared to the baseline as specified in Table 5-10.

Table 5-10: Overview of gains achieved by CTs in CS 3

Split Option	CT	UEs per iSC	J1 BH load in Mbit/s	J2 BH load in Mbit/s	Gain AT
A	2.2	2 10	302.4 per iSC (max)	0	51% 65%
	2.4	1	A.1: 1840 per iSC A.2: 470.4 per iSC A.3: 235.2 per iSC	0	48%
	2.6	Not relevant	235.2 per iSC	0	98%
B	2.1	2	105 per iSC	3300	130%
	2.5	1	700 per iSC	Hundreds of bit/iSC	25%
	2.7	Not relevant	134.4 per iSC	0	<110%
C	2.1	2	8.8 per iSC	3300	130%
	2.2	2 10	40 per iSC (max)	0	20% 26%
	2.5	1	350 per iSC	Hundreds of bit/iSC	30%

For the downlink two approaches for coordinated transmission are investigated in CT 2.4 and CT 2.5 respectively. In CT 2.4 the problem of the optimal number of cooperating iSCs for centralized joint transmission schemes when considering the processing delay is investigated. The functional split options A.1, A.2 or A.3 are applicable in this case with a significant reduction of the backhaul load when moving from the first to the latter. The performance evaluations show an optimal number of cooperative small cells equal to 37 (3 tiers) by taking into consideration the processing and CSI reporting delays (precoding calculation delay, CSI feedback delay and transmit chain delay, etc.). The gain in terms of AT for this cluster size is about 48% for a precoding granularity $N_G = 4$ PRBs with respect to a baseline system without cooperation.

In CT 2.5 the design of a hierarchical joint precoding scheme is investigated. The functional split options B.1, B.2 or C are applicable in this case, which are complementary to the options applicable for CT 2.4. The performance evaluations show that the proposed approach allows going smoothly from the non-cooperative approach to the fully cooperative one as the number of cooperative iSCs that apply the second step of "centralized" precoding increases, i.e. as the backhaul networks offers more possibility of cooperation, with a minimum AT gain of 25% compared to the baseline.

Concerning the BH, two alternative options applicable on fibre or mmWave radio are considered in CT 2.6 and CT 2.7 respectively. In CT 2.6 the split option A.3 is investigated and compared to a fully PHY centralized architecture (i.e. option A.1). This split provides a reduction of the BH rate requirement for the same iSC density and designing the BH network to support the full load condition (i.e. design for zero BH blocking probability). The BH throughput reduction can be also exploited to densify the network which brings a gain on the AT metric of about 98% compared to a fully centralized PHY.

In CT 2.7 the joint RAN/BH encoding over mmWave BH links is investigated. The split option B.2 is applied by forwarding the soft demodulator output to the RANaaS. For joint decoding, two techniques are developed namely ERD and SISODQ. The performance evaluation for CS 3 shows area throughput gains

ranging from 44% up to 110% with respect to a baseline represented by an uncoded BH, which does not forward the channel quality information of the BH or also a conventional scheme with coded BH.

Table 5-10 summarizes the various results of the analysis. Concerning split option A the combination of CT 2.2 and CT 2.6 for the uplink is a possible solution that may also bring to a combination of the AT gains if CT 2.6 is used to leverage on the iSC density. For the downlink, the coordinated transmission based on CT 2.4 can be applied with the same functional split of the uplink (e.g. A.2 or A.3). In general these CTs are suitable for scenarios where the J2 connections among the iSCs are not deployed. For split option B the combination of CT 2.1 and CT 2.7 is possible when using the functional split B.2, thus bringing together the MUD and the joint RAN/BH coding advantages. In this case high throughput J2 connections are required mainly to support CT 2.1, while the CTs have rather similar requirements for J1. For the downlink the hierarchical precoding technique of CT 2.5 is a promising solution that can flexibly adapt to the availability of J2 connections. For the split option C.1 the CT 2.5 is still a possible solution for the downlink, while for the uplink CT 2.1 and CT 2.2 can be seen as possible alternatives also in relation to the availability of J2 connection among the iSCs.

5.3.4 CS 4: Indoor (Shopping Mall / Airport)

In CS 4 uplink CTs 2.1 and 2.2 are applied for joint MUD and investigated. For CT2.1 functional split options B.2 and C.2 are considered whereas for CT2.2 functional split options A.3 and C.3 are used. Table 5-11 specifies for the functional split options the backhaul technologies as defined in Appendix I. Regarding the iSC deployment, 4 iSCs connected in a line topology are considered in a given area, where each iSC considers 5 to 10 UEs in its range.

Table 5-11: Backhaul technologies for functional split options in CS 4

Split Option	BH Number	BH Technology	#Hops	Latency	Bandwidth	CT
A	4a	Dark fibre	1 hop	5 µs	10 Gbit/s	2.2
B	J1: 1c J2: 1c	Millimetre wave	J1: 3 hops J2: 2 hops	J1: 600 µs J2: 400 µs	2.5 Gbit/s	2.1
C	J1: 3a J2: 1c	J1: Sub-6 GHz PtP NLOS TDD J2: Millimetre wave	J1: 2 hops J2: 2 hops	J1: 10 ms J2: 400 µs	J1: 500 Mbit/s J2: 2.5 Gbit/s	2.1 2.2

Corresponding to the scenario definition in D5.2 [10] $N_T^{UE} = 1$ transmit antenna per UE and $N_R^{iSC} = 2$ receive antennas per iSC are assumed. In CT2.1, each iSC has 5 UEs connected to it such that in total 20 UEs are present. In contrast, CT2.2 assumes that each iSC is serving 10 UEs such that in total 40 UEs are present. In both CTs an orthogonal resource allocation for the UEs per iSC is used. Consequently, in CT2.1 each UE allocates 10 PRBs or 1.8 MHz of bandwidth whereas in CT2.2 each UE allocates 5 PRBs or 1 MHz of bandwidth. Regarding the channel model, the ITU InH channel model is used with LOS and NLOS path loss formulas.

Split Option A - CT2.2

For functional split option A, CT2.2 is investigated using MPTD. Here, raw quantized I/Q samples after the first FFT stage are sent from the iSCs to the RANaaS. The turbo detection is then done centrally at the RANaaS. Consequently, a high bandwidth and low latency backhaul technology is required for the J1 link between the iSCs and the RANaaS. For this split option dark fibre backhaul is used on the J1 link which delivers a bandwidth up to 10 Gbit/s. The J2 backhaul is not used by CT2.2 since no exchange among iSCs is exploited by MPTD.

As baseline system a local MMSE detection per iSC is considered. Performance results of CT2.2 show a gain in area throughput of 32% for MPTD compared to the baseline system (see deliverable D3.3 [5]).

Split Option B.2 – CT2.1

For functional split option B.2, CT2.1 is investigated. Here, frequency domain symbols are exchanged among the iSCs over the J2 backhaul connections. The detection is done locally at each iSC and the estimated soft symbols are then forwarded to the RANaaS by one iSC. Since a high data traffic is present on the J2 links high bandwidth and low latency is required on these connections. Thus, for the J2 backhaul technology mmWave is assumed which delivers a BH rate of up to 2.5 Gbit/s.

Performance results of CT2.1 show a gain in area throughput of 69% for the ALCE algorithm after 2 iterations in comparison to the baseline system. The baseline system is a local detection by the iSCs regarding the 5 UEs belonging to the specific iSC. Details can be taken from Section 4.1. The average required BH rate for the J1 per iSC connection is 52.8 Mbit/s while a rate of 950 Mbit/s is required for the J2 connections. Both rates can be provided by using mmWave as BH technology.

Split Option C.2 – CT2.1 and CT2.2

For functional split option C.2, both CT2.1 and CT2.2 are considered while CT2.1 does not explicitly investigate retransmissions by HARQ. In CT2.1 decoding is performed at each iSC and the iSC with best decoding performance transmits its decoded bits to the RANaaS. Similar to functional split B.2, frequency domain symbols are exchanged among the iSCs over the J2 backhaul connections. This requires a high bandwidth and low latency backhaul technology for the J2 which is delivered again by the mmWave backhaul technology. In CT2.2 SPTD is used where each iSC performs turbo detection locally and the decoded bits are transmitted to the RANaaS by all iSCs.

For CT2.1 a gain of 69% in area throughput after 2 iterations can be achieved using the ALCE algorithm. This gain requires an average J1 BH rate per iSC of 8.68 Mbit/s and a J2 BH rate of 950 Mbit/s. CT2.2 provides a gain of 23% in area throughput by using SPTD compared to the baseline system.

Table 5-12: Overview of gains achieved by CTs in CS 4

Split Option	CT	UEs per iSC	J1 BH rate in Mbit/s	J2 BH rate in Mbit/s	Gain AT
A	2.2	10	302.4 per iSC (max)	0	32%
B	2.1	5	52.8 per iSC	950	69%
C	2.1	5	8.68 per iSC	950	69%
C	2.2	10	40 per iSC (max)	0	23%

Table 5-12 summarizes the overall gains achieved by CT2.1 and CT2.2 for CS 4 for various functional splits. In case of restrictions on the J2 BH connections, e.g. if no J2 connections between iSCs shall be deployed, CT2.2 is a promising candidate for the indoor scenario. Both MPTD and SPTD do not require any cooperation over the J2 connections among the iSCs to achieve the presented gains in area throughput. However, high BH rates on the J1 connections are required to serve MPTD. In contrast to that, CT2.1 requires rather low BH rates on the J1 connection since only one iSC needs to deliver information to the RANaaS. Nevertheless, this comes with a high requirement on the J2 BH rate since the J2 links need to cope with the exchange of signals among the iSCs in order make the presented gains in area throughput possible. In cases where the J1 connection to the RANaaS might be costly, but short J2 connections among iSCs are feasible, CT2.1 is a promising candidate for the considered CS.

5.4 Summary

We have provided in this section general performance investigations for the CTs considered in WP2 based on the area throughput evaluations for the different scenarios. Each CT enables one or more functional split. The corresponding CTs for the UL/DL approaches are interoperable and it is even possible to apply different functional splits for UL and DL. Based on each scenario, the iveC may select the appropriate transmission approach. Given a specific BH and deployment scenario, the discussions in the previous subsections allow to identify the most suitable and efficient CT which can be used adaptively for communications. For example, the distributed approaches in CT 2.1 and CT 2.5 as well as CT 2.7 are shown to be promising candidates for direct millimetre wave backhauling. The CT2.3 is shown to be gainful for the square scenario in the case of a

limited backhaul. Centralized approaches in CT2.4 and CT2.2 as well as CT 2.6 can be applied in the case of fibre backhauling. Some of these CTs can be applied with different functional splits and in several deployment scenarios. This shows the effectiveness of these approaches in the iJOIN system concept and indicates that our holistic RAN/BH design with flexible functional split realizes the performance gains and the flexibility with respect to deployment scenarios.

6 Summary and Conclusions

This deliverable provides the final definitions and evaluations of physical layer concepts and WP2 candidate technologies for a dense deployment of small cells. The two key concepts of flexible functional split and joint access and backhaul design have been investigated in detail. The utilization of reconfigurable commodity hardware enables a system architecture where the execution of radio access network (RAN) functionality can be changed in time and space. This allows switching between local, cooperative and centralized processing depending on the traffic demand, availability of backhaul resources, and the applied deployment. The most relevant lower layer functional split options have been identified, numerical results for the rate requirements have been provided and potential gains have been discussed. Due to its distributed nature, the backhaul connecting the iJOIN small cells with the RAN-as-a-Service (RANaaS) platform becomes an ingredient of the overall processing chain. In order to avoid principle bottlenecks and to utilize the available resources most efficiently, relevant approaches for the joint optimization of the access and the backhaul have been discussed and the provided numerical results indicate the benefits of such joint operation.

Furthermore, the set of candidate technologies (CTs) introduced in deliverable D2.1 [1] and further developed in D2.2 [2] have been finalized by discussing the implementation in the iJOIN architecture. For each CT detailed evaluations have been provided in order to demonstrate the potential gains as well as the requirements. Furthermore, the overall evaluation of WP2 approaches for the four iJOIN common scenarios is presented by means of the achieved gain in area throughput compared to a corresponding baseline. This evaluation demonstrates the contribution of the physical layer approaches to the overall iJOIN system.

Acknowledgements and Disclaimer

This work was partially funded by the European Commission within the 7th Framework Program in the context of the ICT project iJOIN (Grant Agreement No. 317941).

Appendix I Backhaul Technologies

In this appendix we summarize the parameters of the considered set of backhaul technologies as discussed in D4.2 [7].

Table A-6-1: Backhaul Technologies

Number	BH technology		Latency (per hop, RTT)	Throughput	Topology	Duplexing	Multiplexing Technology
1a	Millimetre wave	60GHz Unlicensed	≤5 ms	≤800 Mbit/s	PtP (LOS)	TDD	–
1b			≤200 µsec	≤1 Gbit/s	PtP (LOS)	FDD	–
1c		70-80GHz Light licensed	≤200 µsec	≤2.5 Gbit/s	PtP (LOS)	FDD	–
2a	Microwave (28-42 GHz) Licensed		≤200 µsec	≤1 Gbit/s	PtP (LOS)	FDD	–
2b			≤10 ms	≤1 Gbit/s	PtmP (LOS)	TDD	TDMA
3a	Sub-6 GHz Unlicensed or licensed		≤5 ms	≤500 Mbit/s	PtP (NLoS)	TDD	–
3b			≤10 ms	≤500 Mbit/s (shared among clients)	PtmP (NLoS)	TDD	TDMA
3c			≤5 ms	≤1 Gbit/s (per client)	PtmP (NLoS)	TDD	SDMA
4a	Dark Fibre		5 µs/km × 2	≤10 Gbit/s	PtP		–
4b	CWDM		5 µs/km × 2	≤10·N Gbit/s (with N≤8)	Ring		WDM
4c	Metro Optical Network		250 µs	≤1 Gbit/s	Mesh/Ring		Statistical Packet Multiplexing
4d	PON		≤1 ms	100 Mbit/s – 2.5 Gbit/s	PtmP		TDM (DL)/ TDMA (UL)
5	xDSL		5-35 ms	10 – 100 Mbit/s	PtP		–
6	1 Gigabit Ethernet		≤200 µsec	≤1 Gbit/s	PtP		–

Appendix II Analytical BH load calculation

The formulas for the analytical calculation of the downlink payload message size D_P transmitted on the backhaul for the different functional splits are summarized in Table A-6-2. The analysis considers the load between the RANaaS and one iSC with N_T^{iSC} antennas serving N_{UE} users. The transmission frequency of the messages is equal to $1/T_{SUB}$ (i.e. the reciprocal of the subframe period that coincides also with the scheduling interval).

Table A-6-2: Downlink payload message size for the different functional splits

Split Option	Name	Payload message size [bit/iSC]
A.1	Time-Domain I/Q	$D_P^{A.1} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot OF \cdot (N_{FFT} + N_{CP}) \cdot N_{SYMB}^{SUB}$
A.2	Frequency-Domain I/Q before IFFT	$D_P^{A.2} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot N_{sc} \cdot N_{SYMB}^{SUB}$
A.3	Frequency-Domain I/Q before resource mapping	$D_P^{A.3} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB}$
B.1	Frequency-Domain I/Q before precoding (performed at iSC)	$D_P^{B.1} = 2 \cdot N_L \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB}$
	Frequency-Domain I/Q before precoding (performed in the RANaaS)	$D_P^{B.1} = 2 \cdot N_T^{iSC} \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB}$
B.2	Frequency Domain after FEC encoding	$D_P^{B.2} = \sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \cdot Q_m \cdot N_L \cdot N_{SYMB}^{SUB}$
C.1	Frequency Domain before FEC encoding	$D_P^{C.1} = \sum_{u=1}^{N_{UE}} \left(TBS_{1,u} + TBS_{2,u} \right)$

In the equations the size of the messages transmitted on the BH are calculated without including the overhead of the coding techniques that may be used for error detection or synchronization purposes. For example, for the split option A.1 based on CPRI [47] a typical 8b/10b line code [48] is used, which corresponds to a rate increase of 1.25 over the backhaul.

The value N_Q represents the number of quantization bits used for each signal (e.g. I/Q component, soft bits) transmitted on the BH, e.g. for each real and each imaginary part of a complex signal sample. For simplicity of notation, the same symbol N_Q is used in all the considered cases, even though the actual value depends on the considered functional split. Some indication about the number of quantization bits is provided in [55]. For example in case of option A.1, where OFDM or single-carrier frequency division multiple access (SC-FDMA) signals in the time domain are transmitted on the BH, a rather high resolution (15 bits per dimension per sample) is necessary to keep quantization noise at a tolerable level [55]. Conversely, for split option A.2 where the OFDM/SC-FDMA signal in the frequency domain is transmitted on the BH, a lower resolution can be used compared to split option A.1, due to the smaller amplitude fluctuations of the signals in the frequency domain. In particular, the authors in [55] postulate a minimum of 5 to 6 bit depending on the actual SINR. A further margin of 1 or 2 bits may be necessary in the presence of strong interfering signals and thus a total number of about 7 to 8 quantization bits for each signal component can be considered reasonable for the split options A.2 to B.1. Concerning the split option B.2, in the uplink case "soft bits" (e.g. log likelihood ratios (LLRs) after symbol-to-bit demapping or their expected values) are transmitted to the channel decoder on the BH. Conversely, in the downlink the channel coded bits after rate matching are transmitted to the subcarrier mapping function located in the iSC. In the downlink case a 1 bit resolution is clearly used, while for the uplink a LLR resolution over 3 to 4 bits leads to a negligible quantization loss

[52]. Finally, concerning the split option C.1, where the MAC bits are transmitted on the BH, a 1 bit resolution is applied on both DL and UL.

Table A-6-3: Uplink payload message size for the different functional splits

Split Option	Name	Payload message size [bit/iSC]
A.1	Time-Domain I/Q	$D_P^{A.1} = 2 \cdot N_R^{iSC} \cdot N_Q \cdot OF \cdot (N_{FFT} + N_{CP}) \cdot N_{SYMB}^{SUB}$
A.2	Frequency-Domain I/Q after FFT	$D_P^{A.2} = 2 \cdot N_R^{iSC} \cdot N_Q \cdot N_{sc} \cdot N_{SYMB}^{SUB}$
A.3	Frequency-Domain I/Q after resource demapping (orthogonal access)	$D_P^{A.3} = 2 \cdot N_R^{iSC} \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB}$
	Frequency-Domain I/Q after resource demapping (non-orthogonal access)	$D_P^{A.3} = 2 \cdot N_R^{iSC} \cdot N_Q \cdot \left(\bigcup_{u=1}^{N_{UE}} \mathcal{S}_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB}$
B.1	I/Q after equalization and IDFT	$D_P^{B.1} = 2 \cdot N_L \cdot N_Q \cdot \left(\sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \right) \cdot N_{SYMB}^{SUB}$
B.2	Quantized soft demodulator output	$D_P^{B.2} = \sum_{u=1}^{N_{UE}} N_{sc,u}^{j} \cdot Q_m \cdot N_Q \cdot N_L \cdot N_{SYMB}^{SUB}$
C.1	FEC decoder output	$D_P^{C.1} = \sum_{u=1}^{N_{UE}} (TBS_{1,u} + TBS_{2,u})$

The formulas for evaluating the uplink message size related to the data payload on the backhaul connection between the iSCs and the RANaaS are summarized in Table A-6-3. As for the downlink, the size of the messages transmitted on the BH are calculated without including the overhead of the coding techniques that may be used for error detection or synchronization purposes. Due to the multi-carrier nature of the SC-FDMA signal, the equations derived for the downlink can be easily extended to the uplink by changing the related parameters. The analysis considers the load between one iSCs and the RANaaS with N_R^{iSC} receive antennas serving N_{UE} users. The transmission frequency of the messages is equal to $1/T_{SUB}$ (i.e. the reciprocal of the subframe period that coincides also with the scheduling interval).

The Table A-6-4 provides the tabulation of the results calculated with the equations given above for the exemplary case described in Section 3.2.1.

Table A-6-4: Exemplarily required BH throughput for different functional split options (LTE cell with 2x2 MIMO, B=10 MHz bandwidth and full PRB utilization)

	Throughput in Mbit/s			
Split	DL ($\gamma = 1$)	DL ($\gamma = 4/3$)	UL ($\gamma = 1$)	UL ($\gamma = 4/3$)
A.1	920.6	1227.5	920.64	1227.5
A.2	235.2	313.6	302.4	403.2
A.3	235.2	313.6	302.4	403.2
B.1	235.2	313.6	302.4	403.2
B.2	100.8	134.4	403.2	537.6
C.1	73.3	97.9	73.3	97.9

References

[1] iJOIN Project, "D2.1- State-of-the-art of and promising candidates for PHY layer approaches on access and backhaul network," November 2013. Available: http://www.ict-ijoin.eu/wp-content/uploads/2014/01/D2.1.pdf

[2] iJOIN Project, "D2.2- Definition of PHY layer approaches that are applicable to RANaaS and a holistic design of backhaul and access network," November 2014. Available: http://www.ict-ijoin.eu/wp-content/uploads/2012/10/D2.2.pdf

[3] iJOIN Project, "D3.1- Final report on MAC/RRM state-of-the-art, Requirements, scenarios and interfaces in the iJOIN architecture," November 2013. Available: http://www.ict-ijoin.eu/wp-content/uploads/2014/01/D3.1.pdf

[4] iJOIN Project. "D3.2- Definition of MAC and RRM approaches for RANaaS and a joint backhaul/access design," November 2014. Available: http://www.ict-ijoin.eu/wp-content/ uploads/2012/10/D3.2.pdf

[5] iJOIN Project. "D3.3- Final definition and evaluation of MAC and RRM approaches for RANaaS and a joint backhaul/access design," May 2015.

[6] iJOIN Project. "D4.1- Report on SotA and requirements for network-layer algorithms and network operation and management," November 2013. Available: http://www.ict-ijoin.eu/wp-content/ uploads/2014/01/D4.1.pdf

[7] iJOIN Project. "D4.2- Network-layer algorithms and network operation and management: candidate technologies specification," November 2014. Available: http://www.ict-ijoin.eu/wp-content/ uploads/2012/10/D4.2.pdf

[8] iJOIN Project, "D4.3- Final definition and evaluation of network-layer algorithms and network operation and management," May 2015.

[9] iJOIN Project, "D5.1- Revised definition of requirements and preliminary definition of the iJOIN architecture," November 2013. Available: http://www.ict-ijoin.eu/wp-content/uploads/2014/01/ D5.1.pdf

[10] iJOIN Project, "D5.2- Final definition of Requirements and Scenarios," November 2014. Available: http://www.ict-ijoin.eu/wp-content/uploads/2012/10/D5.2.pdf

[11] iJOIN Project, "D5.3- Final definition of iJOIN architecture," May 2014.

[12] iJOIN Project. "D6.1- Preliminary proof-of-concept results for selected candidate algorithms," November 2014. Available: http://www.ict-ijoin.eu/wp-content/uploads/2012/10/D6-1.pdf

[13] iJOIN Project, "D6.2- Final proof-of-concept results for selected candidate algorithms," May 2014.

[14] P. Rost, C.J. Bernardos, A. De Domenico, M. Di Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wübben, "Cloud Technologies for Flexible 5G Radio Access Networks," IEEE Communications Magazine, Vol. 52, No. 5, pp. 68-76, Nov. 2014.

[15] A. Maeder, M. Lalam, A. De Domenico, E. Pateromichelakis, D. Wübben, J. Bartelt, R. Fritzsche, and P. Rost, "Towards a Flexible Functional Split for Cloud-RAN Networks," European Conference on Networks and Communications (EuCNC 2014), Bologna, Italy, June 2014.

[16] D. Wübben, P. Rost, J. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, "Benefits and Impact of Cloud Computing on 5G Signal Processing," IEEE Signal Processing Magazine, vol. 31, no. 6, pp. 35-44, Nov. 2014.

[17] J. Bartelt, P. Rost, D. Wübben, J. Lessmann, B. Melis, G. Fettweis, "Fronthaul and Backhaul Requirements of Flexible Centralization in Cloud Radio Access Networks," IEEE Wireless Communications Magazine, submitted January 2015.

[18] P. Rost, I. Berberana, A. Maeder, H. Paul, V. Suryaprakash, M. Valenti, D. Wübben, A. Dekorsy, G. Fettweis, "Benefits and Challenges of Virtualization in 5G Radio Access Networks", IEEE Communications Magazine, submitted December 2014.

[19] H. Paul, J. Fliege, and A. Dekorsy, "In-network-processing: Distributed consensus-based linear estimation," IEEE Commun. Lett., vol. 17, no.1, pp. 59-62, Jan.2013.

[20] H. Paul, B.-S. Shin, and A. Dekorsy, "Distributed Consensus-Based Linear Estimation with Erroneous Links," in Proc. VDE/ITG International Workshop in Smart Antennas (WSA), 2013.

[21] H. Paul, B.-S. Shin, D. Wübben, and A. Dekorsy, "In-Network-Processing for Small Cell Cooperation in Dense Networks," in Proc. IEEE Vehicular Technology Conference (VTC) workshops, First International Workshop on Cloud Technologies and Energy Efficient in Mobile Communications Networks (CLEEN), 2013.

[22] B. S. Shin, H. Paul, D. Wübben, and A. Dekorsy, "Reduced Overhead Distributed Consensus-Based Estimation Algorithm," Proc. IEEE Global Communications Conference (GLOBECOM) workshops, The first International Workshop on Cloud-Processing in Heterogeneous Mobile Communication Networks (IWCPM), 2013.

[23] G. Xu, H. Paul, D. Wübben, and A. Dekorsy, "Fast Distributed Consensus-based Estimation for Cooperative Wireless Sensor Networks," 18th International ITG Workshop on Smart Antennas (WSA 2014), Erlangen, Germany, Mar. 2014.

[24] D. Wübben, H. Paul, B.-S. Shin, G. Xu, and A. Dekorsy, "Comparative Study of Distributed Consensus-based Estimation Schemes for Small-Cell Networks," European Conference on Networks and Communications (EuCNC 2014), Bologna, Italy, June 2014.

[25] D. Wübben, H. Paul, P. Balleydier, V. Savin, and P. Rost, "Decoder Implementation for Cloud Based Architectures," European Conference on Networks and Communications (EuCNC), Bologna, Italy, June 2014.

[26] D. Wübben, H. Paul, B.-S. Shin, G. Xu, and A. Dekorsy, "Distributed Consensus-Based Estimation for Small Cell Cooperative Networks," Globecom 2014 Workshop – Broadband Wireless Access (GC14 WS-BWA), Austin, USA, Dec. 2014.

[27] G. Xu, H. Paul, D. Wübben, and A. Dekorsy, "Distributed Augmented Lagrangian Method for Cooperative Estimation in Small Cell Networks," 10th International ITG Conference on Systems, Communications and Coding (SCC 2015), Hamburg, Germany, Feb. 2015.

[28] H. Paul, D. Wübben, and P. Rost, "Implementation and Analysis of Forward Error Correction Decoding for Cloud-RAN Systems", Proc. IEEE International Conference on Communications (ICC) workshops, The second International Workshop on Cloud-Processing in Heterogeneous Mobile Communication Networks (IWCPM), London, June 2015.

[29] D. Gesbert, S. Hanly, H. Huang, S. Shitz Shamai, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: a new look at interference," IEEE Trans. on Inf. Theory, vol. 28, no. 9, pp. 1380–1408, 2010.

[30] S. Park, O. Simeone, O. Sahin, and S. Shamai, "Joint precoding and multivariate backhaul compression for the downlink of cloud radio access networks", IEEE Trans. Signal Process., vol. 61, no. 22, pp. 5646-5658, Nov. 2013.

[31] P. de Kerret, R. Fritzsche, D. Gesbert, and U. Salim, "Robust precoding for network MIMO with Hierarchical CSIT", 11th International Symposium on Wireless Communication Systems (ISWCS 2014), Barcelona, Spain, Aug. 2014.

[32] R. Fritzsche and G. P. Fettweis, "Robust precoding with general power constraint considering unbounded channel uncertainty," 9th International Symposium on Wireless Communication Systems (ISWCS 2012), Paris, France, Aug. 2012.

[33] Y. Qi, M. Ali Imran, A. Quddus, and R. Tafazolli, "Achievable Rate Optimization for Coordinated Multi-Point Transmission (CoMP) in Cloud-Based RAN Architecture," IEEE International Conference on Communications (ICC), London, UK, June 2014.

[34] Advanced Radio Interface Technologies for 4G Systems (ARTIST-4G), FP7 project, Webpage: https://ict-artist4g.eu/.

[35] V. Suryaprakash, A. Fehske, A. F. dos Santos, and G. P. Fettweis, “On the impact of sleep modes and BW variation on the energy consumption of radio access networks,” IEEE 75th. Vehicular Technology Conference (VTC Spring), Yokohama, Japan, May 2012.

[36] J. Bartelt and G. Fettweis, “A Soft-Input/Soft-Output Dequantizer for Cloud-Based Mobile Networks,” 15th IEEE International Workshop on Signal Processing Advances in Wireless Communications (IEEE SPAWC 2014), Toronto, Canada, 2014.

[37] J. Bartelt and G. Fettweis, “An Improved Decoder for Cloud-Based Mobile Networks under Imperfect Fronthaul,” Globecom 2014 Workshop – Wireless optical network conversion in support of cloud architectures (GC14 WS- WONC), Austin, USA, Dec. 2014.

[38] A. De Domenico, V. Savin, and D. Ktenas, “A backhaul-aware cell selection algorithm for heterogeneous cellular networks,” IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), 2013.

[39] 3GPP TR 36.872, “Small cell enhancements for E-UTRA and E-UTRAN - Physical layer aspects (Release12),” V12.1.0, Dec. 2012.

[40] Y. D’Halluin, P. A. Forsyth, and K. R. Vetzal. "Wireless network capacity management: A real options approach," European Journal of Operational Research. Vol. 167, issue 1, p. 584-609, 2007.

[41] 3GPP TS 36.211, “Evolved Universal Terrestrial Radio Access; (E-UTRA); Physical channels and modulation,” v10.6.0, Dec. 2012.

[42] 3GPP TS 36.212, “Evolved Universal Terrestrial Radio Access; (E-UTRA); Multiplexing and channel coding,” v10.7.0, Dec. 2012.

[43] 3GPP TS 36.213, “Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Release 10),” v10.8.0, Dec. 2012

[44] ITU-R, “REPORT M2135, Guidelines for evaluation of radio interface technologies for IMT-Advanced,” 2008.

[45] 3GPP TR 36.866, “Study on Network-Assisted Interference Cancellation and Suppression (NAIC) for LTE (Release 12),” v12.0.1, Mar. 2014

[46] M. R. Rurter, “An Introduction to Computing,” Lecture notes University of Cambridge, Available: http://www.tcm.phy.cam.ac.uk/~mjr/lect1.pdf, 2001.

[47] Common Public Radio Interface, [online: http://www.cpri.info/]

[48] N. K. Babu and P. S. S. Babu, “Design of Physical Coding sublayer using 8B/10B algorithm,” International Journal of Recent Technolgy and Engineering (IJRTE), ISSN: 2277-3878, Volume 2, Issue 2, May 2013

[49] 3GPP TR 25.996, “Spatial channel model for Multiple Input Multiple Output (MIMO) simulations”, Release 12 (2014-09)

[50] 3GPP TS 36.306, “E-UTRA - User Equipment (UE) radio access capabilities” Release 10, (2011 - 12).

[51] NGMN, RAN Evolution Project, “Further Study on Critical C-RAN Technologies,” Version 1.0 (approved), 20 May 2014.

[52] G. Bosco, G. Montorsi, and S. Benedetto, “Soft Decoding in Optical Systems,” IEEE Transactions on Communications, Vol. 51, NO. 8, August 2003.

[53] Y. d’Halluin, P. A. Forsyth, and K. R. Vetzal. "Wireless network capacity management: A real options approach," European Journal of Operational Research. Vol. 167, issue 1, p. 584-609, 2007.

[54] 3GPP TR 36.814, “Further advancements for E-UTRA physical layer aspects (Release9),” v9.0.0, Mar. 2010.

[55] U. Dötsch, M. Doll, H. P. Mayer, F. Schaich, J. Segel, and P. Sehier, “Quantitative Analysis of Split Base Station Processing and Determination of Advantageous Architectures for LTE,” Bell Labs Technical Journal, vol. 18, no. 1, pp. 105–128, May 2013.

[56] M. Shariat, E. Pateromichelakis, A. Quddus, and R. Tafazolli, ”Joint TDD Backhaul and Access Optimization in Dense Small Cell Networks,” IEEE Transactions on Vehicular Technology, to appear, doi: 10.1109/TVT.2014.2379013

[57] T. Chen, Y. Yang, H. Zhang, H. Kim, and K. Horneman, ”Network Energy Saving Technologies for Green Wireless Access Networks,” IEEE Transactions on Wireless Communications, vol.18, no.5, pp.30-38, October 2011

[58] C. Hausl and P. Dupraz, “Joint Network Channel Coding for the Multiple-Access Relay Channel”, 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON), Reston, VA, Sept. 2006.

[59] 3GPP TSG RAN, “TR 36.932, Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN (Release 12)”, V12.1.0, March 2013.

[60] A. El Gamal and Y. H. Kim, “Network Information Theory”, Cambridge University Press, New York, USA, 2011.

[61] J. Zhang, M. Kountouris, J.G. Andrews, and R. W. Heath, “Multi-Mode Transmission for the MIMO Broadcast Channel with Imperfect Channel State Information,” IEEE Transactions on Communications, vol.59, no.3, pp.803-814, March 2011.

[62] S. Bhaumik, S.P. Chandrabose, M. K. Jataprolu, G. Kumar, A. Muralidhar, P. Polakos, V. Srinivasan, and T. Woo, “CloudIQ: A framework for processing base stations in a data center,” IEEE MobiCom, Istanbul, Turkey, August 2012.

[63] Small Cell Forum, “Small Cell Virtualization Use Cases,“ RPH Group, v0.16, April 2015.

[64] D. Sabella, A. De Domenico; E. Katranaras, M. A. Imran, M. di Girolamo, U. Salim, M. Lalam, K. Samdanis, and A. Maeder, "Energy Efficiency Benefits of RAN-as-a-Service Concept for a Cloud-Based 5G Mobile Network Infrastructure," IEEE Access, Vol.2, pp. 1586-1597, Dec. 2014.

[65] J.C. Zuniga, C.J. Bernardos, A. de la Oliva, T. Melia, R. Costa, and A. Reznik, “Distributed Mobility Management: A Standards Landscape,” IEEE Communications Magazine, Vol. 51, No. 3, pp. 80-87, 2013

[66] D. Sabella, P. Rost, Y. Sheng, E. Pateromichelakis, U. Salim, P. Guitton-Ouhamou, M. Di Girolamo, and G. Giuliani, "RAN as a Service: Challenges of designing a flexible RAN architecture in a cloud-based heterogeneous mobile network," Future Network and Mobile Summit (FuNeMS), Lisabon, Portugal, 2013

[67] C.J. Bernardos, A. De La Oliva, P. Serrano, A. Banchs, L.M. Contreras, H. Jin, and J.C. Zúniga, “An architecture for software defined wireless networking,” IEEE Wireless Communications Magazine, Vol. 21, No. 3, pp. 52-61, June 2014.

[68] Z. Mheich, A. De Domenico, and V. Savin, “Uplink Capacity and User Association for Cooperative Heterogeneous Cellular Networks,” submitted to 16th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2015), Stockholm, Sweden, June 2015.

[69] M.Lalam and T. Lestable, “Centralised Radio Resource Management for Scalable Multi-Point Turbo-Detection in Dense Small Cells Deployment,” submitted to IEEE Globecom, December 2015.

[70] L. Zhang, A. Quddus, E. Katranaras, D. Wübben, Y. Qi, and R. Tafazolli, “Performance Analysis and Optimal Cooperative Cluster Size for Randomly Distributed Small Cells under Cloud RAN,” submitted to IEEE Transactions on Wireless Communications.

