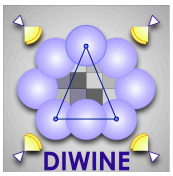
Ref. Ares(2016)6551792 - 22/11/2016



DIWINE

Contract No. FP7-ICT-318177

Cloud network processing for advanced scenarios D3.03

Contractual date:	M36
Actual date:	M36
Authors:	Monica B. NICOLI, Umberto SPAGNOLINI, Gloria SOATTI, Maria Antonieta ALVAREZ VILLANUEVA, Stefano SAVAZZI, Jan SÝKORA, Tomáš HYNEK, Alister G. BURR, Mehdi MORTAZAWI MOLU, Yi WANG, Justin P. COON, Shuping DANG, William H. THOMPSON, Stojan DENIĆ, Pin-Hsun LIN, Eduard A. JORSWIECK
Participants:	PdM, CTU, UoY, UOXF, TREL, TUD
Work package:	WP3
Security:	Public
Nature:	Report
Version:	1.0
Number of pages:	200

Abstract

This deliverable presents a comprehensive overview of the algorithms for distributed channel/system state identification, relay coding and cloud self-organisation, refined after their successful implementation in HW/SLS demonstrators. Advanced techniques for cloud network processing are also discussed as prosiming solutions for cloud network evolution in perspective.

Keywords

Distributed processing, synchronisation, consensus algorithm, cooperative communication, multiuser interference mitigation, resource allocation, relay coding.

Executive summary

Cloud network processing involves the set of processing steps that are carried out within the cloud to guarantee that it can be regarded as an independent and self-contained entity. All partners have been involved to act complementarily with regard to demonstration activities in WP5 and synergically for the node capabilities defined in WP4, using the scenarios defined in WP2 for designing and assessing the performance of cloud network processing techniques.

This report is the final deliverable focused on advanced scenarios of cloud network processing where algorithms have been designed by taking into account the limits, specifications and practical requirements that have been provided as feedback from the experimental activities undertaken in WP5.

The deliverable includes:

- a comprehensive overview of the cloud network processing methods refined after their successful implementations in HW/SLS demonstrators;
- novel algorithms for cloud network processing that are too complex (or not timely for the DIWINE timeframe) for demonstration at present, but are promising candidate solutions for future cloud networks.

Contents

Notations **9**

1 Introduction **13**
1.1 Summary and structure of the report 13

2 Distributed channel and system state identification **19**
2.1 Distributed time-frequency synchronisation 19
2.1.1 General synchronisation system model 19
2.1.2 Distributed phase locked loop (D-PLL) 21
2.1.3 Super-frame structure 21
2.1.4 Synchronisation error estimation 24
2.1.5 Duplexing for distributed synchronisation 26
2.1.6 Numerical results 30
2.1.7 Conclusions 32
2.2 Synchronisation algorithm evaluation 34
2.2.1 Test 1: Estimation and tracking ability, 2 Tx – 1 Rx 36
2.2.2 Test 2: Estimation and tracking ability, 3 Tx – 1 Rx 38
2.2.3 Conclusions 40
2.3 Assessment of distributed algorithms 40
2.3.1 Review of the bivariate channel model 41
2.3.2 Distributed estimation of channel model parameters 42
2.3.3 Network localisation 45
2.3.4 Conclusions 49
2.4 Distributed interference sensing and coordination for cloud scheduling . . 49
2.4.1 Introduction 50
2.4.2 Spectrum sensing model (SSM) 52
2.4.3 MAP testing for PU detection 53
2.4.4 Estimation of SSM parameters 54
2.4.5 Cooperative detection of interference 57
2.4.6 Performance analysis 57
2.4.7 Experimental case study 59
2.4.8 Concluding remarks 61

3 Relay processing and coding **63**
3.1 Adaptive energy efficient scheme for QMF relaying 63
3.1.1 Introduction 63
3.1.2 Model of adaptive asymmetric cooperative relaying 64

3.1.3 Adaptive policies for cooperative relaying and slow fading 65
3.1.4 Case study for finite state and fading channels 69
3.1.5 Conclusion 74
3.2 Relay selection in cloud networks 75
3.2.1 Introduction 75
3.2.2 Combined bulk/per-subcarrier relay selection in two-hop OFDM systems 77
3.2.3 Comparison of multicarrier relay selection schemes in super dense cloud networks 89
3.2.4 Conclusions 103
3.3 Degradedness of fast fading Gaussian multiple-antenna wiretap channels 104
3.3.1 Introduction 104
3.3.2 Preliminaries 106
3.3.3 Multiple antennas without channel enhancement 108
3.3.4 Conclusion 113

4 Cloud self organisation 115
4.1 Cloud resource scheduling by distributed interference coordination 115
4.1.1 Introduction 115
4.1.2 Problem setting 116
4.1.3 Self-adaptive resource allocation algorithm 119
4.1.4 Numerical results 122
4.1.5 Conclusion 125
4.2 Distributed learning process for HNC selection – Advanced scenarios . . 127
4.2.1 Utility function related to instantaneous channel conditions . . . 127
4.3 Non-cooperating relays 129
4.3.1 Intentionally malicious relay 131
4.3.2 Intentionally over-selfish relay 139
4.4 Self-selection of physical layer network coding parameters 145
4.4.1 Introduction 145
4.4.2 System model 146
4.4.3 Relay strategy 148
4.4.4 Numerical results 151
4.4.5 Conclusion 153
4.5 Novel approach for computing network coding function in compute-and-forward 154
4.5.1 Introduction 154
4.5.2 System Model and Preliminaries 156
4.5.3 Finding Optimal $\boldsymbol{a}$ at High SNR 158
4.5.4 Numerical Results 166
4.5.5 Complexity analysis 168
4.5.6 Conclusion 170
4.5.7 Proof of (4.54) and (4.55) 171

4.5.8 Proof of (4.60) 172
4.6 Design for adaptive physical layer network coding over cooperative relaying 173
4.6.1 Design criteria 175
4.6.2 Simulations 184
4.6.3 Conclusions 185

5 Conclusions **187**

Bibliography **189**

Notations

Abbreviations

3GPP	3rd Generation Partnership Project
ADC	Analog-to-Digital Converter
AF	Amplify and Forward
AP	Access Point
APNC	Adaptive Physical Layer Network Coding
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BNE	Bayesian Nash Equilibrium
BPSK	Binary Quadrature Phase Shift Keying
BS	Base Station
CAZAC	Constant Amplitude Zero Auto-Correlation
C&F	Compute-and-Forward
CCDF	Complementary Cumulative Density Function
CDF	Cumulative Density Function
CF	Compress-and-Forward
CFO	Carrier Frequency Offset
CIMC	Critical Industrial Monitoring and Control
C-MLE	Centralised-Maximum-Likelihood Estimation
CQI	Channel Quality Information
CoMP	Coordinated MultiPoint
CRB	Cramer–Rao Bound
CRC	Cyclic Redundancy Check
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
D	Destination
D2D	Device-to-Device
DAC	Digital-to-Analog Converter
DD	Decision-Directed
DEIB	Dipartimento di Elettronica Informazione e Bioingegneria
DF	Decode-and-Forward
DLA	Distributed Learning Algorithm
D-MLE	Distributed-Maximum-Likelihood Estimation
D-MLE-AE	Distributed-Maximum-Likelihood Estimation with Accuracy Exchange

DMT	Diversity-Multiplexing Tradeoff
D-PLL	Distributed Phase Locked Loop
D-SC	Distributed-Simple Consensus
EVD	Eigenvalue Decomposition
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correcting
FER	Frame Error Rate
FG	Fixed Gain
FIFO	First-In First-Out
FIR	Finite Impulse Response
FP7	7th Framework Programme
FSC	Finite State Channels
FTP	File Transfer Protocol
HBS	Hub Base Station
HNC	Hierarchical Network Coding
HTTP	HyperText Transfer Protocol
HW	Hardware
IC	Interference Channel
i.d.	isotropically distributed | Isotropic Distribution
i.i.d.	independent, identically distributed
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
IQ	In-phase and Quadrature
LDPC	Low-Density Parity-Check
LLL	Lenstra–Lenstra–Lovasz
LNCB	Linear Network Coded Bits
LNCF	Linear Network Coded Function
LOS	Line-Of-Sight
LPNC	Linear Physical Layer Network Coding
LQD	LQ-Decomposition
LRT	Likelihood Ratio Test
LTE	Long-Term Evolution
MAC	Multiple Access Channel | Media Access Control
MC	Mutually Coupled
MIMO	Multiple Input Multiple Output
MIMOME	Multiple Input Multiple Output Multiple-antenna Eavesdropper
MLE	Maximum-Likelihood Estimation
MRC	Maximum Radio Combining
MS	Master-Slave
MSE	Mean Square Error
MT	Mobile Terminal
NC	Network Coding
NCS	Network Coded Symbols

NCV	Network Coded Vector
NE	Nash Equilibrium
NLOS	Non-Line-Of-Sight
OFDM	Orthogonal Frequency-Division Multiplexing
PA	Power Amplifier
PBE	Perfect Bayesian Equilibrium
PDF	Probability Density Function
PNC	Physical Layer Network Coding
PRB	Physical Resource Block
PU	Primary User
QAM	Quadrature Amplitude Modulation
QMF	Quantise-Map-and-Forward
QoS	Quality of Service
QP	Quadratic Programming
QPSK	Quadrature Phase Shift Keying
R	Relay | Receiver
RB	Resource Block
RF	Radio Frequency
RMSE	Root Mean Square Error
RSS	Received Signal Strength
RX	Receiver
S	Source
SDR	Software Defined Radios
SER	Symbol Error Rate
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-and-Noise Ratio
SISOSE	Single-Input Single-Output Single-Antenna Eavesdropper
SLS	System Level Simulator
SMN	Smart Meter Network
SNR	Signal-to-Noise Ratio
SSM	Spectrum Sensing Model
SU	Secondary User
SVD	Singular Value Decomposition
SVP	Shortest Vector Problem
SW	Software
TAS	Transmit Antenna Selection
TF	Time-Frequency
TO	Time Offset
TWRC	Two-Way Relay Channel
TX	Transmitter
UE	User Equipments
USRP	Universal Software Radio Peripheral
UT	Utility Target

VG	Variable Gain
VoIP	Voice over IP
WCN	Wireless Cloud Network
WPLNC	Wireless Physical Layer Network Code/Coding
WP	Work Package
WSN	Wireless Sensor Network
XOR	eXclusive OR

Mathematical notations

$\mathbf{R}$	set of real numbers
$\mathbf{R}_+$	set of non-negative real numbers
$\mathbf{R}_{++}$	set of positive real numbers
$\mathbf{x}$	column vector
$\mathbf{X}$	matrix
$\mathbf{X}^T$	transpose of matrix $\mathbf{X}$
$\mathrm{tr}(\mathbf{X})$	trace of matrix $\mathbf{X}$
$\|\Omega\|$	cardinality of set Ω
$\mathrm{diag}\,(\cdots)$	diagonal matrix
$E[\,\cdot\,]$	expected value of
$Cov[\,\cdot\,]$	covariance of
$\sim$	distributed as
$\approx$	approximated by
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
$\mathcal{H}$	Hypothesis
$\mathcal{X}$	Dataset
$\delta(\,\cdot\,)$	Dirac impulse

1 Introduction

WP3 covers the cloud organisation in terms of self-organisation, cooperative processing, coding and interference mitigation. Deliverable D3.03 summarises the activities carried out on advanced cloud network processing during the last two years of the DIWINE project. The report summarises not only some experimenting activity that fed back the guidelines for further theoretical advances, but also all the cloud processing techniques that have been designed and experimented for distributed channel/system state identification, relay coding and cloud self-organisation. The algorithms were designed taking into account the specifications, practical constraints, limits and pitfalls raised during the cooperation in demonstration scenario setup of WP5.

Activities of D3.03 are mainly addressing two objectives: (i) optimisation and performance assessment of the processing algorithms that have been transferred to the HW/SLS demonstrators for validation and tuning (this activity was in close interaction with WP5); (ii) investigation of theoretical aspects and advanced methodologies that are of strategic importance for advancing the know-how in the DIWINE context, though not ready at this stage of the research for implementation into the DIWINE demonstrator platforms.

1.1 Summary and structure of the report

The report is organised as follows. Chapter 2 presents the distributed algorithms designed for the estimation of the key parameters of the network state that are needed to set up the cloud functionalities, namely the timing-frequency offset for network synchronisation, the channel/interference state information and the cloud network topology. Chapter 3 discusses the techniques for physical network coding operations, and particularly algorithms for improving the energy efficiency of Quantise-Map-and-Forward (QMF) relaying, for relay node selection and physical layer security. Advanced methods for cloud self-organisation are in Chapter 4, addressing distributed algorithms for intra-cloud resource management, distributed learning of information needed for physical network coding and self-selection of coding parameters.

The technical contributions from Deliverable D3.02 have been successfully implemented and validated into the HW/SLS demonstrators over the last two years providing important feedbacks for algorithm advances or refinements. Namely, the distributed consensus-based algorithm for time-frequency synchronisation of the cloud nodes (Section 2.2) has been successfully transferred into the USRP-based Smart Meter Network (SMN) demonstrator (see D5.42 [1] for further practical details), with several cross-feedbacks between

theoretical developments and practical setup (Section 2.1). The distributed consensus-based algorithm for interference sensing (Section 2.4) has been implemented and it is part of the functionalities of the Critical Industrial Monitoring and Control (CIMC) demonstrator (see D5.52 [2]). The adaptive energy-efficient scheme for QMF relaying (Section 3.1) has been partially implemented by means of the USRP hardware platforms used for the SMN demonstrator (see D5.42). The advanced distributed learning algorithm (Section 4.2 [1]) was successfully transferred into the SMN demonstrator (see D5.42 [1]). Finally, the proposed approach for self-selection of the physical network coding function has been implemented in the system level simulator (SLS) (see D5.33 [3]).

Technical contents are summarised below for each section, highlighting the contributions that have been transferred into the HW/SLS demonstrators.

Chapter 2: Distributed channel and system state identification

Section 2.1 Distributed time-frequency synchronisation

The distributed nature of cooperative networks results in both Carrier Frequency Offset (CFO) and timing offset (TO) that degrade the performance of collaborative communication. Thus, a global synchronisation is mandatory to guarantee the proper communication and coding among multiple uncoordinated nodes in dense networks. Within DIWINE a novel method has been proposed a novel method to perform joint time and frequency synchronisation in presence of large TOs and CFOs by extending the principle of consensus method to synchronisation; the method has been finalised to the implementation in SMN demonstrator (WP5) for practical setting validations. The estimation is based on assigning to all nodes the *same training sequence* based on Constant Amplitude Zero AutoCorrelation (CAZAC) sequences, that is properly arranged to decouple CFO from TO in symbol and frame synchronisation. The synchronisation is based on weighted consensus algorithm to reach synchronisation in a connected network by exploiting the uncoordinated superposition (or collision) of the waveforms from all transmitting nodes used as an ensemble of reference timing for synchronisation. In the realistic scenario of the SMN demonstrator, half duplex constraint represents a limit in synchronisation as each device can only transmit its synchronisation frames, or receive (and update) the synchronisation from the other nodes. It is part of the research to investigate a duplex scheduling as the trade-off between the time allocated to transmit or receive in dense cooperative networks. The analytical model shows that to maximise the convergence of timing and frequency synchronisation the duplex scheduling can be random and independent on each node, with probability of transmission $\rho < 0.5$. Thus, time and frequency synchronisation is guaranteed in multi-node interference scenario without the need to assign to every node an independent CAZAC sequence, and thus without any coordination in a random duplexing with $\rho < 0.5$.

Section 2.2 Synchronisation algorithm evaluation

This section presents hardware validation of the proposed distributed synchronisation algorithm. All transmitting nodes send the same synchronisation frames, and the synchronisation configuration is estimated from these superimposed waveforms. This method greatly simplifies the CFO and TO estimation procedure when implemented in the SMN demonstrator (WP5). A real-time algorithm is implemented on Software Defined Radios (SDR), and is tested in various configurations to investigate the tracking and accuracy abilities compared to the theoretical developments (Section 2.1). The hardware tests verified the algorithm as a suitable candidate for synchronisation in dense networks, with accurate estimation of the timing and frequency offsets of the transmitting nodes. Its ability to decouple the estimations was also demonstrated with experiments.

Section 2.3 Assessment of distributed algorithms for channel estimation and network localisation

In this section distributed consensus for self-learning of the cloud topology has been extended to joint node-to-node measurements of the average power gain and the degree of temporal fading, namely the mean and the variance of the received signal strength (RSS). These two RSS moments were proved to be (on average) linearly related to the log-distance between the cloud nodes, see Deliverable D3.02 [4]. The RSS statistics, modelled according to the stochastic propagation model developed in WP2, are characterised by some unknown environment-dependent features, common to all the cloud nodes, which are cooperatively estimated during a cloud network calibration procedure before proceeding with the self-learning of the cloud network geographical topology. The extended method has been tested on real channel data collected by IEEE 802.15.4 devices. Performance analysis was carried out in terms of convergence speed, error at convergence and communication overhead for network calibration and localisation using both experimental and simulated data.

Section 2.4 Distributed interference sensing and coordination for cloud scheduling

The problem of distributed spectrum sensing is analysed in a scenario where multiple cloud networks, deployed in close proximity, need to self-coordinate their access to a set of shared time-frequency resources so as to minimise the mutual interference. In a simplified setting with two coexisting networks, the devices of the sensing cloud (secondary users, SUs) exchange local spectrum estimates to cooperatively recognise and track the overall time-varying interference patterns caused by the devices of the other pre-existing cloud network (primary users, PUs). PUs are assumed to perform a periodic transmission over predefined (but unknown to the SUs) time-frequency hopped resources. Cooperation between the SUs is crucial to detect the complete interference pattern of the primary network. The weighted consensus algorithm developed in D3.02 [4] is here combined with an

iterative decision-directed procedure for distributed detection of the PU spectrum occupancy. The distributed approach is shown to provide the estimate of the complete interference pattern to each SU regardless of the incomplete visibility at each node. The method has been validated by experimental tests on unlicensed 2.4 GHz spectrum sharing, within the CIMC demonstrator, as detailed in D5.52 [2].

Chapter 3: Relay processing and coding

Section 3.1 Adaptive energy efficient scheme for QMF relaying

In point-to-point communication systems, physical layer abstraction techniques can be readily used to quantify the performance and adapt the parameters. Since such abstraction has not been clearly defined for cooperative schemes such as QMF, we propose a heuristic method based on a combination of adaptive blind policy and cyclic redundancy check coding. Following the proposed method, we demonstrate how the QMF relay node can adapt its communication parameters. It is shown that the proposed technique can considerably improve the energy efficiency of the relay node and at the same time provide good performance gains. The technique has been partially implemented by means of USRP hardware platform.

Section 3.2 Relay selection in cloud networks

In this section, we present our current research progress on relay selection in cloud networks. Relay selection is regarded as one of the most promising technologies for use in future cellular networks, especially super dense cloud networks. There are two commonly used selection schemes, i.e. bulk selection and per-subcarrier selection. However, the performance of the former is too poor to satisfy the increasingly high requirements of cloud networks, while the latter strategy normally requires too many relays in one transmission interval. Therefore, we propose a joint selection scheme combining the characteristics of both and prove its priority over both in small-scale networks. Meanwhile, we compare the performance of these three selection schemes in super dense cloud networks and it is shown that bulk selection is preferable to the other two, because of the selection contention phenomenon. Three types of relays, i.e. Decode-and-Forward (DF), Fixed-Gain (FG) Amplify-and-Forward (AF) and Variable-Gain (VG) AF are employed when comparing the performance. Also, a variable termed bulk gain factor is defined and used as a powerful tool to analyse different relay selection schemes in super dense cloud networks.

Section 3.3 Degradedness of fast fading Gaussian multiple-antenna wiretap channels with statistical CSIT

The characterisation of secrecy capacities of Gaussian wiretap channels depends on the knowledge of Channel State Information at the Transmitter (CSIT). When there is only statistical CSIT, the secrecy capacity is unknown in general. In this section we investigate the relation between the usual multivariate stochastic order

and the degradedness among the legitimate receiver's and eavesdropper's channels under fast fading with only statistical CSIT. The considered transmitter has multiple antennas and the legitimate receiver and the eavesdropper both have multiple antennas. Based on the technique of coupling we derive criteria to check the degradedness of the fast fading Gaussian wiretap channel including some commonly used channel models even there is only statistical CSIT. We illustrate one example of a Rician $2 \times 2 \times 2$ channel where the ergodic secrecy capacity can be derived under different K-factors.

Chapter 4: Cloud self organisation

Section 4.1 Cloud resource scheduling by distributed interference coordination

We consider a cloud scenario where multiple node-to-node links coexist on the same spectrum and have to coordinate to self-adapt their Time-Frequency (TF) allocation based on distributed interference detection. This context addresses the IoT or any Device-to-Device (D2D) communication, and it is crucial to design a radio resource management system that allocates the TF resources to the links in a distributed way, i.e. without any centralised coordination, so as to guarantee the Quality of Service (QoS), particularly in heterogeneous traffic conditions. Interference-aware resource allocation has the capability to adapt the resource management to a context where multiple links coexist in the same spectrum. In this section, a distributed scheduling approach is proposed where each node-to-node link reacts to the locally sensed interference by self-adapting its own TF allocation. Each node autonomously trades the QoS request in term of packet service with the resource availability by inflating/deflating the spectrum allocation based on the sensed interference level. The change of the interference pattern perceived in turn by other links serves as inter-link signalling of the need/release of TF resources. Each node optimises the allocation by iterated local adjustments, till an equilibrium with the other links is reached. The section shows that, in perspective, the proposed scheduling algorithm is able to maximise the total throughput in a fully distributed way, arranging efficiently the radio resources over the TF domain so as to satisfy the QoS requirements for each node.

Section 4.2 Distributed learning process for HNC selection – Advanced scenarios

An extension of distributed learning algorithm (DLA) towards more complex networks and scenarios is shown. It includes optional additional channel related utility function – based on the symbol error rate – which is used for local decision making. The advanced DLA was successfully transferred to WP5 for demonstration. In depth implementation details are provided in D5.42 [1].

Section 4.3 Non-cooperating relays

This section focuses on relays that do not fully follow the idea of the cloud concept. A scenario with intentionally malicious relay is analysed first. Second, we assume a scenario where a relay node starts to behave over-selfish for example because of emptying battery. For both of the scenarios we provide a game theoretic analysis, show equilibria existence and suitable utility function. Both scenarios are far advanced to be shown by the demonstrator now.

Section 4.4 Self-selection of physical layer network coding parameters

Compute-and-Forward (C&F) is an important relaying technique with good potential for implementing physical layer network coding in dense relay networks such as the DIWINE cloud. However it requires the choice of coefficients at relays. In this section, C&F relaying in a multi-source multi-relay network is studied. Two novel algorithms for coefficient selection are proposed, addressing the complexity of cloud organisation. The algorithms assume (i) no coordination and (ii) partial coordination between the relay nodes and prove to be near optimal based on extensive simulations with large number of relay nodes in a cloud.

Section 4.5 Novel approach for computing network coding function in compute-and-forward

Coefficient selection may require very complex integer optimisation, especially in dense cloud networks with many nodes. In this section, a novel approach is proposed to determine the integer variables for C&F relaying, making it more practically feasible to increase efficiency of dense multi-hop wireless networks for the DIWINE cloud.

Section 4.6 Design for adaptive physical layer network coding over cooperative relaying

In this section, an approach is described to ensure that each relay can choose a non-singular Physical layer Network Coding (PNC) function to overcome all the singular fade states, and that the destination can unambiguously recover the source messages. The proposed design is divided into an off-line search algorithm which could find a small number of best efficient coefficient matrices, and an on-line search algorithm which determines the optimal full rank mapping matrix for each fading state. The off-line search greatly reduces the computation complexity of the on-line search algorithm with a small performance loss. The whole system is suited to conventional modulation schemes whose cardinality is a power of 2, and in which binary constellation labels can be used, which is preferable from the point of view of engineering implementation. This approach has been implemented in the SLS, and is described in D5.33 [3].

2 Distributed channel and system state identification

2.1 Distributed time-frequency synchronisation

The wireless cloud network (WCN) accounted in DIWINE comprises nodes scattered in a bounded area. Moreover, nodes are coupled with neighbours within a connectivity radius that depends on nodes' transmission power and receiver sensitivity, as well as propagation settings. To enable the global synchronisation of the network to guarantee the proper coding and communication with self-coordination, a distributed algorithm for timing/frequency synchronisation is necessary. Timing offset (TO) is caused by propagation delay (here neglected) and timing misalignment due to the initial arbitrary phase of each node, whilst carrier frequency offset (CFO) is caused by the oscillator frequency mismatch, and possibly Doppler effect (here neglected).

A preliminary study about the usage of constant amplitude zero autocorrelation (CAZAC) sequence as synchronisation symbol was developed in D3.02. Here, it is proposed a new synchronisation frame structure based on two consecutive CAZAC sequences that allows to decouple the coupling effect between TO and CFO [5]. Furthermore, a modified distributed phase locked loop (D-PLL) has been studied to achieve the convergence of the network by exchanging between nodes the *same* signature. The proposed distributed synchronisation algorithm allows scalability of the network and reduces computational complexity in comparison with conventional synchronisation methods. A realistic half-duplex constrain scenario has been studied during the project and an optimum solution that maximise the convergence of the network synchronisation has been developed.

2.1.1 General synchronisation system model

The uncoordinated network considered here is composed of K fully connected nodes without any time and frequency master clock reference, where each node has local clock based on carrier oscillator that runs independently before synchronisation. We consider that the network employs OFDM modulation scheme for communication with N sub-carriers, sub-carrier spacing f_s and sampling rate $F_s = Nf_s$. To enable the global synchronisation of the network, nodes exchange the same training sequence that, once synchronised, superimpose and collide one another without any impairment for the synchronisation. For the k-th node, the local (clock) time $t_k[n]$ and carrier (oscillator) instantaneous phase $\psi_k[n]$ are modelled by a discrete reference time n as

$$\begin{aligned} t_k[n] &= nT_k + \tau_k, \\ \psi_k[n] &= \phi_k + 2\pi f_k n, \end{aligned} \tag{2.1}$$

where T_k is the local period, τ_k is the timing in samples, f_k is the carrier (oscillator) frequency and ϕ_k is its phase. We assume (to simplify) that nodes are frame synchronous, that is all clocks share the same period $T_1 = \cdots = T_k = \cdots = T_K$. Time is discretised into synchronisation steps, e.g. super-frame period, within the n-th time each node can either listen or transmit a packet, but not both (half-duplex constraint).

Let denote $\mathcal{T}[n]$ as a set of transmitting nodes and $\mathcal{R}[n]$ the set of receiving nodes (also can be defined a set of idle nodes) at time n with $\mathcal{T}[n] \cap \mathcal{R}[n] = \varnothing$ and $K = |\mathcal{T}[n] \cup \mathcal{R}[n]|$. The set of transmitting nodes $\mathcal{T}[n]$ broadcast the same preamble sequence distributing their synchronisation state to nodes $\mathcal{R}[n]$, these nodes $\mathcal{R}[n]$ receive a superimposition of transmitted signals colliding together and update locally the TO and CFO. Elements of set $\mathcal{T}[n]$ and $\mathcal{R}[n]$ change dynamically over synchronisation step n in order to let each node broadcast its synchronisation state or be able to correct its clock/oscillator. The parameters regarding TO and CFO associated to local clock/oscillator of i-th transmitting node with respect to the reference clock/oscillator can be expressed as $\boldsymbol{\theta}_i = [\tau_i, f_i]^T$. The transmitted waveform $x[m]$ by every node contains a preamble for synchronisation that is followed by data payload.

Each node i belonging to the set $\mathcal{T}[n]$ transmits the *same* pilot impaired with its own TO and CFO that collide with the others at each receiver nodes that are in the neighbourhood. The superimposition of the pilots at the receiver is also impaired with the TO and CFO of the receiver node k as depicted in Figure 2.1. Therefore, the relative TO and CFO normalised by sub-carrier spacing f_s between the i-th transmitter and k-th receiver nodes are $\Delta\tau_{ik} = \tau_i - \tau_k$ and $\Delta f_{ik} = f_i - f_k$, respectively. The m-th sample of the basedband received signal by k-th node, $y_k[m]$, can be represented

$$y_k[m] = \sum_{i \in \mathcal{T}[n]} \sum_{p=1}^{P} h_{i,k}[p] x[m|\boldsymbol{\theta}_{ik}] + w_k[m], \tag{2.2}$$

where $x[m|\boldsymbol{\theta}_{ik}]$ represents the impaired pilot waveform

$$x[m|\boldsymbol{\theta}_{ik}] = \exp\left(j\frac{2\pi}{N}(\Delta f_{ik})m\right) x[m - \Delta\tau_{ik}], \tag{2.3}$$

$h_{i,k}[p]$ is the amplitude of multi-path fading Rayleigh channel between nodes i and k and is distributed as $h_{i,k}[p] \sim \mathcal{CN}(0, \sigma_h^2[p])$. AWGN samples are distributed as $w[m] \sim \mathcal{CN}(0, \sigma_w^2),\ \forall m$.

Cross-correlation of the received signal $y_k[m]$ with local copy of the training sequence $x[m]$ at receiver is a statistic sufficient for the estimation of $\Delta\tau_{ik}$ and Δf_{ik} that is a linear combination of all the relative TOs and CFOs errors between receiver node k and transmitter nodes of $\mathcal{T}[n]$. Distributed synchronisation is the iterative method that locally corrects the synchronisation state $\boldsymbol{\theta}_k = [\tau_k, f_k]^T$ based on the relative error $\Delta\boldsymbol{\theta}_k = [\Delta\tau_k, \Delta f_k]^T$ to reach asymptotically a synchronisation status of the network $\boldsymbol{\theta}_k[n] \to \boldsymbol{\theta}_\infty$ for $n \to \infty$.

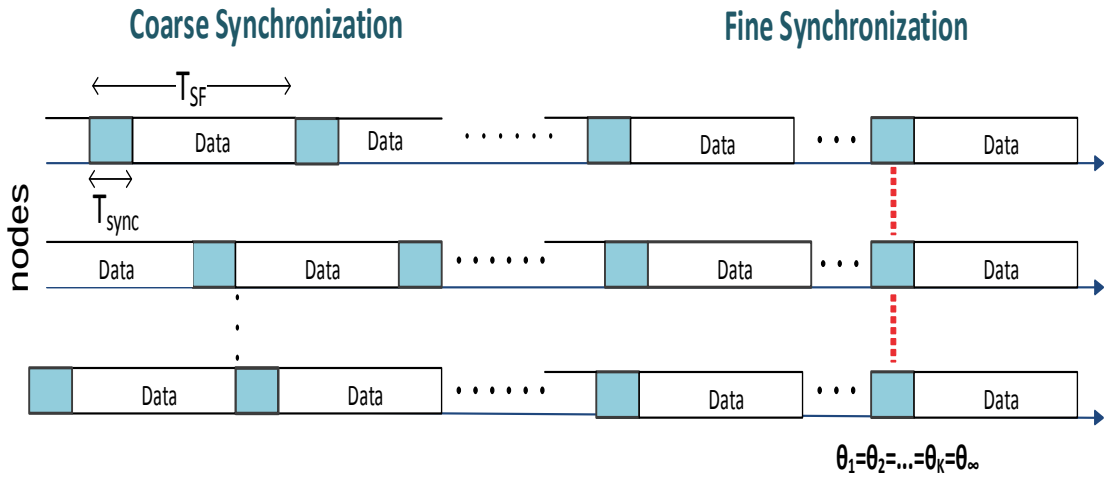


Figure 2.1: Evolution of the frame structure during synchronisation.

2.1.2 Distributed phase locked loop (D-PLL)

All nodes in the network aim at reaching a consensus via local communication with other nodes based on the estimation of the synchronisation parameters $\Delta\boldsymbol{\theta}_k$ at n-th synchronisation step. These parameters are the average difference of timing and carrier frequency offsets between transmitting nodes, acting as a reference, and the k-th receiver node. Distributed phase locked loop (D-PLL) is based on extraction of synchronisation parameters from ensemble of signal $y_k[m]$ from the transmitting nodes and generation of common synchronisation state where each node in the network is correcting itself through the distributed consensus algorithm. The updating of the distributed consensus algorithm for the synchronisation problem is [6]

$$\begin{aligned} f_k[n+1] &= f_k[n] + \varepsilon_\tau \Delta f_k[n], \\ \tau_k[n+1] &= \tau_k[n] + \varepsilon_f \Delta\tau_k[n], \end{aligned} \tag{2.4}$$

where ε_τ and ε_f are the step sizes that depend of the network topology and convergence needs. The connectivity of the network changes dynamically over time according to the duplexing schedule that guarantee the convergence of the network (a proposed random duplex scheduling is detailed in Subsection 2.1.2). To estimate the synchronisation parameters $\Delta\boldsymbol{\theta}_k = [\Delta\tau_k, \Delta f_k]^T$, it is designed a training sequence that allows to decouple the coupling effect between TO and CFO. The analysis of preamble design is detailed below.

2.1.3 Super-frame structure

The selection of CAZAC sequence as synchronisation symbol for HW/SW demonstrators was investigated in D3.02. The CAZAC sequences have special properties that make

it possible to perform jointly timing and frequency synchronisation using only one preamble signature for both TO and CFO estimation. To investigate the coupling effect between TO and CFO, we consider the preamble waveform to be made of one CAZAC sequence. The m-th complex sample of preamble waveform is given by

$$x[m] = e^{j\frac{\pi}{N}um^2} \text{ for } 0 \leq m \leq N-1, \tag{2.5}$$

where N is even and root index u is relatively prime to N. At the receiver node k, a filter for detection of preamble and estimation of synchronisation parameters is matched to $x[m]$. Therefore, the estimation is based on the cross-correlation of received signal with $x[m]$:

$$r_k[\ell] = \sum_{m=0}^{N-1} y_k[m+\ell]x^*[m]. \tag{2.6}$$

The cross-correlation in (2.6) provides a coarse estimation of TO and the coupling effect of TO and single integer CFO [7], so they cannot be separately used for synchronisation. Namely, the CFO normalised to the sub-carrier spacing can be represented as the sum of an integer and a fractional part: $f = f^I + f^F$. The lag that maximises $|r_k[\ell]|$ depends of both fractional and integer part of CFO. In presence of integer CFO f^I, given a specific u, the lag of cross-correlation peak shifts linearly by s samples with a unit shift in f^I, this term adds up to the TO and causes ambiguity. Since shift s depends on the root index u and the length of sequence N, as N is fixed, the appropriate choice of u is required to maximise the f^I to be estimated and corrected by synchronisation algorithm. Figure 2.2 shows an example of the coupling effect between TO and CFO estimation for $u = 3$. The lag of cross-correlation peak shifts $s = 21$ samples for a relative CFO error of $\Delta f_{ik} = 1$. Thus, the choice $u = 1$ yields the smallest shift ($s = 1$) and this is adopted for the design of the synchronisation frame. The use of the *same* CAZAC sequence for all nodes requires only one matched filter. Therefore, cross-correlation (2.6) includes the effect of multiple TOs and CFOs.

The α-power barycenter of the cross-correlations is the metric for determining the TO and CFO errors:

$$d_k = \frac{\sum_\ell \ell \times |r_k[\ell]|^\alpha}{\sum_\ell |r_k[\ell]|^\alpha} = \Delta\tau_k + s \times \Delta f_k^I. \tag{2.7}$$

The metric is the averaging over lags weighted by the α-power of the absolute value of the cross-correlation as in (2.7), this metric yields an estimation of average TO and CFO that are coupled together. For a root index equal to $-u$, the shifts caused by TOs are the same while the shift induced by integer CFOs are in opposite direction. So the new metric based on this root index is represented as $\Delta\tau_k - s \times \Delta f_k^I$. We can take advantage of this property and introduce a preamble waveform containing two ZC sequences with root index u and $-u$. The additional part of preamble waveform requires the receiver to have one matched filter for the first part, and another for the second part. A block diagram of the proposed distributed synchronisation method is depicted in Figure 2.3.

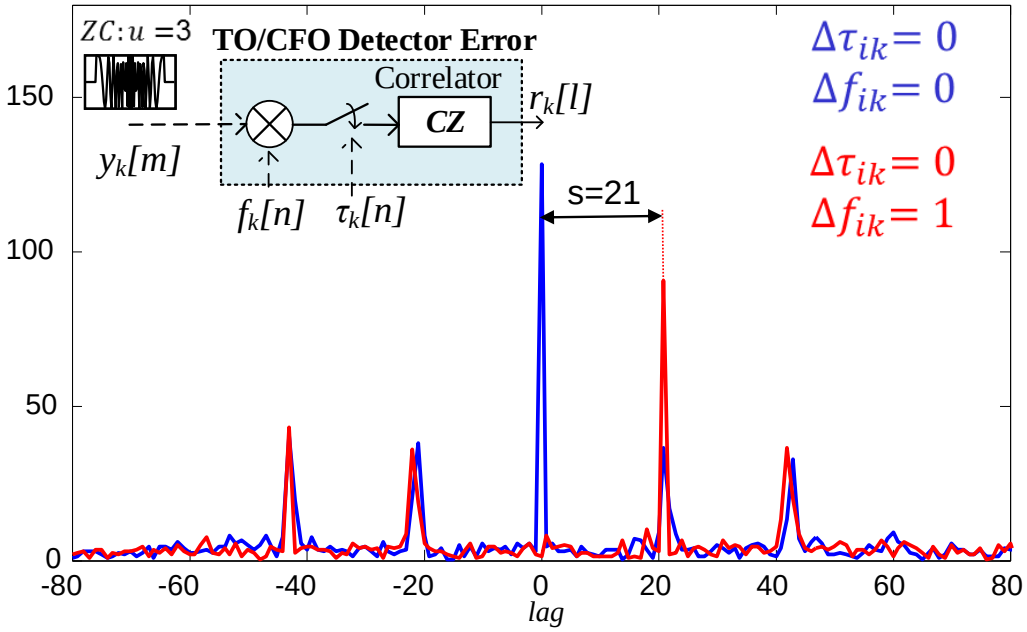


Figure 2.2: Coupling effect between TO and CFO estimation for $u = 3$, integer CFO $\Delta f_{ik} = 0$ (blue line) and $\Delta f_{ik} = 1$ (red line).

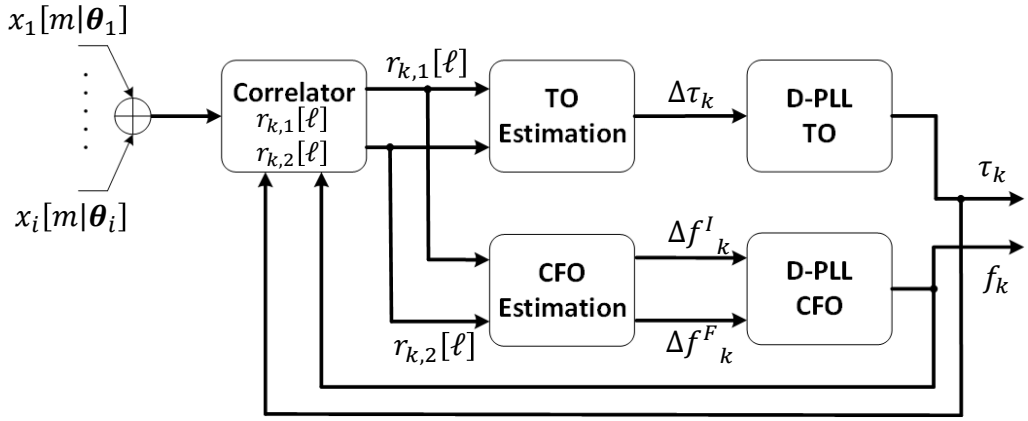


Figure 2.3: Block diagram of the distributed synchronisation method.

Preamble design

The synchronisation frame structure is composed by two consecutive CAZAC sequences with the same root index but opposite chirp sweep as this structure guarantees joint TO and CFO synchronisation. Figure 2.4 depicts the super-frame (considered in DIWINE) based on the synchronisation frame as preamble, followed by payload. The m-th sample

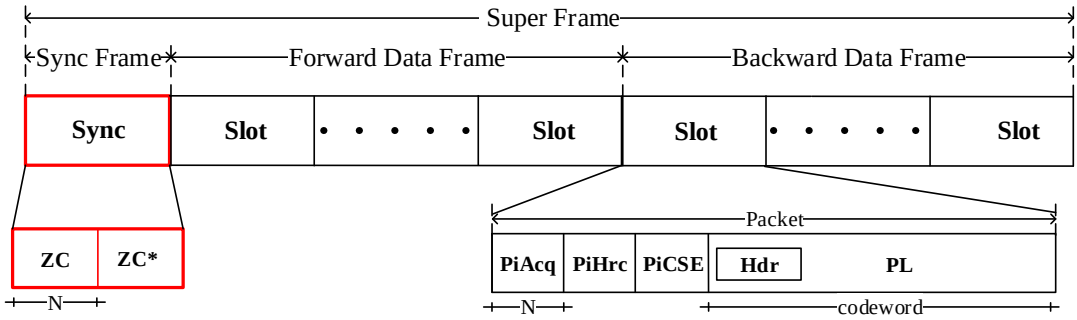


Figure 2.4: DIWNE super-frame structure.

of this synchronisation frame is

$$x[m] = \begin{cases} e^{j\frac{\pi}{N}um^2} & 0 \leq m \leq N-1, \\ e^{-j\frac{\pi}{N}u(m-N)^2} & N \leq m \leq 2N-1, \end{cases} \tag{2.8}$$

Considering the received waveform in (2.2), two cross-correlation with the local copy of the CAZAC sequence should be carried out at each receiver node k. The cross-correlation $r_{k,1}[\ell]$ and $r_{k,2}[\ell]$ can be represented as

$$r_{k,1}[\ell] = \sum_{m=0}^{N-1} y_k[m+\ell] e^{-j\frac{\pi}{N}um^2}, \tag{2.9a}$$

$$r_{k,2}[\ell] = \sum_{m=0}^{N-1} y_k[m+\ell] e^{j\frac{\pi}{N}um^2}. \tag{2.9b}$$

The barycenter of both cross-correlation are used as a metric for TO and CFO estimation:

$$\begin{aligned} \hat{d}_{k,1} &= \frac{\sum_\ell \ell \times |r_{k,1}[\ell]|^\alpha}{\sum_\ell |r_{k,1}[\ell]|^\alpha}, \\ \hat{d}_{k,2} &= \frac{\sum_\ell \ell \times |r_{k,2}[\ell]|^\alpha}{\sum_\ell |r_{k,2}[\ell]|^\alpha}. \end{aligned} \tag{2.10}$$

The metrics are associated to each cross-correlation in (2.9a) and (2.9b), Figure 2.5 shows the histograms of $\hat{d}_{k,1}$ and $\hat{d}_{k,1}$. The value of the metric change due to the TO and CFO simultaneously. The estimation of TO and integer CFO relative errors are based on (2.10).

2.1.4 Synchronisation error estimation

Both timing and frequency synchronisation are usually performed in two phases: coarse and fine estimation. Coarse TO estimation generates an initial estimate of the starting point of the frame while fine timing improves the estimate by eliminating the effect of

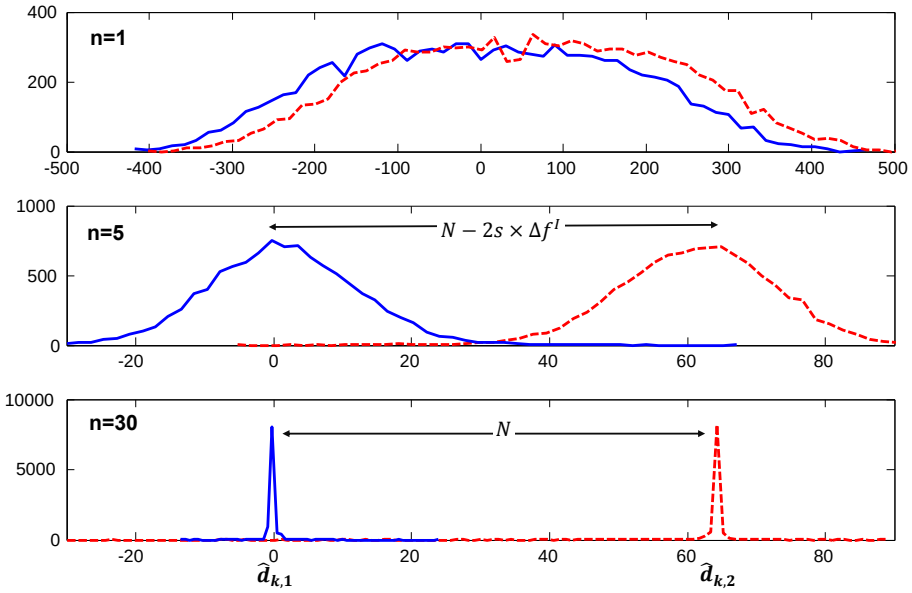


Figure 2.5: Histograms of metrics $\hat{d}_{k,1}$ (solid line) and $\hat{d}_{k,1}$ (dashed line) vs. synchronisation-steps: $n = 1$, $n = 5$ and $n = 30$ for $K = 10$ nodes and $N = 64$ with uniform initialisation ($n = 0$).

coupling f^I and it gives the true starting point. Similarly, coarse frequency synchronisation provides a rough CFO estimate while fine frequency synchronisation estimates the fractional CFO f^F.

Joint TO and integer CFO estimation

The TO can be estimated by using (2.9a) and (2.9b). According to the fact that the received signal used to extract the synchronisation information is the superimposition of multiple waveforms, and it has the TO and CFO contributions of all transmitting nodes, TO estimation in typical receiver node k is a combination of all the TOs and CFOs contributions. Hence, the two metrics in (2.10) can be paired with synchronisation mismatches as

$$\hat{d}_{k,1} = \Delta\tau_k + s \times \Delta f_k^I. \tag{2.11a}$$

$$\hat{d}_{k,2} = N + \Delta\tau_k - s \times \Delta f_k^I. \tag{2.11b}$$

Thus, both metrics (2.11a) and (2.11b) yield the TO and CFO estimation error. For $\Delta f^I = 0$, the difference between two metric is N. On the other hand for Δf^I different from zero the distance between them is $\hat{d}_{k,2} - \hat{d}_{k,1} = N - 2s \times \Delta f^I$ (see Figure 2.5), where shift s associated with specific root index u. The joint CFO and TO estimation at k-th node

with respect to all the transmitter are

$$\Delta f_k^I = \frac{\hat{d}_{k,2} - \hat{d}_{k,1} - N}{2s}. \tag{2.12a}$$

$$\Delta \tau_k = \frac{\hat{d}_{k,2} + \hat{d}_{k,1} - N}{2}. \tag{2.12b}$$

Both (2.12a) and (2.12b) are individually used for timing and carrier frequency synchronisation.

Fractional CFO estimation

The fractional CFO estimate Δf_k^F and correction takes place after TO and integer CFO synchronisation has reached a convergence, or a steady-state behaviour such that $\Delta \tau_k \sim 0$ and $\Delta f_k^I \sim 0$. In this situation, multiple peaks have narrowed down to a single peak in each of the cross-correlation $r_{k,1}[\ell]$ and $r_{k,2}[\ell]$ and Δf_k^F estimation can be performed by linear regression between the residual phase components of two peaks spaced apart by N samples. Since the two estimates $\hat{d}_{k,1}$ and $\hat{d}_{k,2}$ from (2.10) are not integer, the cross-correlations $r_{k,1}[\ell]$ and $r_{k,2}[\ell]$ should be interpolated. A simple and quite accurate method is the linear interpolation:

$$\begin{aligned} \tilde{r}_{k,1} &= (1-\lambda_1) \cdot r_{k,1}[\lfloor \hat{d}_{k,1} \rfloor] + \lambda_1 \cdot r_{k,1}[\lceil \hat{d}_{k,1} \rceil], \\ \tilde{r}_{k,2} &= (1-\lambda_2) \cdot r_{k,1}[\lfloor \hat{d}_{k,2} \rfloor] + \lambda_2 \cdot r_{k,2}[\lceil \hat{d}_{k,2} \rceil], \end{aligned} \tag{2.13}$$

where $\lambda_1 = \hat{d}_{k,1} - \lfloor \hat{d}_{k,1} \rfloor$ and $\lambda_2 = \hat{d}_{k,2} - \lfloor \hat{d}_{k,2} \rfloor$. The fractional CFO is thus estimated from these interpolated values:

$$\Delta f_k^F = \frac{1}{2\pi} \left(\angle(\tilde{r}_{k,2}) - \angle(\tilde{r}_{k,1}) \right), \tag{2.14}$$

where $\angle(\tilde{r}_{k,2})$ and $\angle(\tilde{r}_{k,1})$ are the accumulated residual phase.

The next subsection studies a duplex scheduling strategy based on distributed synchronisation algorithm to minimise the convergence time of the synchronisation network.

2.1.5 Duplexing for distributed synchronisation

The aim of duplex scheduling is to guarantee a fast network synchronisation by selecting the fraction of time each node devotes to transmit or to receive. In a mutually coupled (MC) scenario, all nodes that are transmitting are acting as an aggregated TO/CFO reference for the benefit of the receiving nodes to correct their TO and CFO. We prove below that there is an optimal duplexing equilibrium that maximises the speed of network synchronisation [8].

Random duplex

In a cooperative network, the distributed synchronisation algorithm can exchange the synchronisation status in each node by randomly and independently selecting the duplexing state as this avoids any coordination. The probabilistic model for the k-th node is:

$$\begin{aligned} \Pr[k \in \mathcal{T}[n]] &= p \\ \Pr[k \in \mathcal{R}[n]] &= 1-p. \end{aligned} \tag{2.15}$$

Convergence analysis needs to consider the ensemble of nodes as in any distributed method, and it can be evaluated from the trajectory of the vector containing the timing and frequency offsets of all nodes defined as $\boldsymbol{\tau}[n] = [\tau_1[n], \cdots, \tau_K[n]^T$ and $\boldsymbol{f}[n] = [f_1[n], \cdots, f_K[n]^T$. At n-th synchronisation step all the receiving nodes in $\mathcal{R}[n]$ update jointly their local synchronisation state according to

$$\begin{aligned} \boldsymbol{f}[n+1] &= (\mathbf{I} - \varepsilon_f \mathbf{L}[n])\, \boldsymbol{f}[n], \\ \boldsymbol{\tau}[n+1] &= (\mathbf{I} - \varepsilon_\tau \mathbf{L}[n])\, \boldsymbol{\tau}[n], \end{aligned} \tag{2.16}$$

where $\mathbf{L}[n]$ is the Laplacian matrix which encompasses the topological aspects of network and it is dependent on the instantaneous partitioning sets $\mathcal{T}[n]$ and $\mathcal{R}[n]$, thus, the Laplacian matrix of the graph at time n is $\mathbf{L}[n] = \text{diag}\{|\mathbf{A}_1[n]|, \cdots, |\mathbf{A}_K[n]|\} - \mathbf{A}[n]$, the elements of the adjacency matrix are $[\mathbf{A}[n]]_{i,k} = 1$ if the link $(i,k) \in \mathcal{E}[n]$, $[\mathbf{A}[n]]_{i,k} = 0$ otherwise, and $|\mathbf{A}_i[n]| = \sum_{k=1}^{K} [\mathbf{A}[n]]_{i,k}$ is the degree of the i-th node at time n, i.e. the number of transmitting neighbours from which it receives the superimposed synchronisation frames. The connectivity changes over time according to the randomly duplexing independently from one iteration to the other.

Convergence analysis

The TO dominates the time of convergence of the synchronisation, and consequently the CFO error is affected by the variations on the TO. Therefore, the analysis of the conditions that minimise the time of convergence for a random duplexing is based on the conditions that guarantee the TO synchronisation. Derivation in this subsection is under the approximation that the probability density function (pdf) of TO evolution $\tau_k[n]$ is Gaussian (for central limit) regardless that at initialisation TO $\tau_k[0]$ is uniform, as this reflects the condition of asynchrony when communications are organised in frames (Figure 2.1). This assumption is validated numerically by the inspections of the histograms of $\tau_k[n]$ vs. iterations n (Figure 2.6). At the n-th iteration of TO synchronisation, the pdf of the ensemble of K nodes is Gaussian:

$$\tau_k[n] \sim N(\bar{\tau}[n], \sigma^2[n]) \quad \text{for } k = 1, 2, \ldots, K, \tag{2.17}$$

purpose is to evaluate analytically the evolution of $\sigma^2[n]$ vs. iteration to infer the convergence property of TO for varying probability p. To simplify, the graph is assumed fully

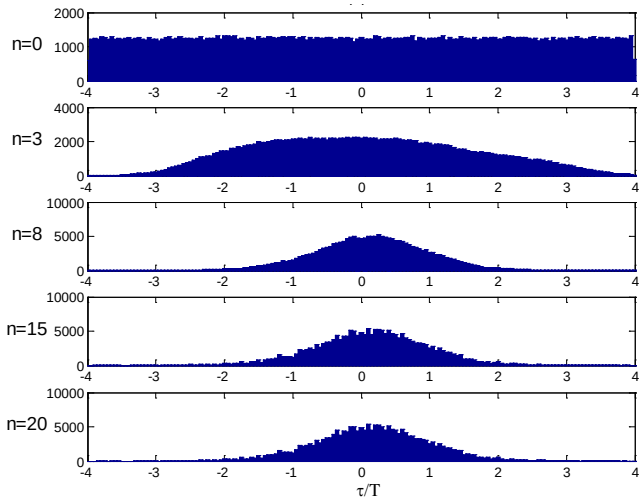


Figure 2.6: Histograms of $\tau_k[n]$ vs. iterations $n = 0, 3, 8, 15, 20$ (for $K = 20$).

connected (all-to-all connectivity) with degree $d = K - 1$, while orientations depend on $\mathcal{R}[n]$ and $\mathcal{T}[n]$ duplex partitioning.

In random duplexing there is a subset of $\ell = |\mathcal{T}[n]|$ nodes labelled as $1, 2, \ldots, \ell$ (node ordering is arbitrary as duplex choice is statistically independent among nodes) and the remaining $K - \ell = |\mathcal{R}[n]|$ nodes are receiving according to the partition:

$$\begin{aligned} \tau^{\mathcal{T}}[n|\ell] &= \{\tau_1[n], \tau_2[n], \ldots, \tau_\ell[n]\}, \\ \tau^{\mathcal{R}}[n|\ell] &= \{\tau_{\ell+1}[n], \tau_{\ell+2}[n], \ldots, \tau_K[n]\}, \end{aligned} \tag{2.18}$$

thus, the probability that ℓ nodes act as transmitters is binomial

$$\Pr[\ell = |\mathcal{T}[n]|] = \binom{K}{\ell} p^\ell (1-p)^{K-\ell}. \tag{2.19}$$

The synchronisation update formula applies only to $\mathcal{R}[n]$ nodes ($k > \ell$), these nodes update their synchronisation status based on the $\mathcal{T}[n]$ nodes with $\ell = |\mathcal{T}[n]|$ (for $\ell \geq 1$)

$$\tau_k[n+1|\ell] = \tau_k[n] + \varepsilon_\tau(\tau^{\mathcal{T}}[n|\ell] - \tau_k[n]), \tag{2.20}$$

this depends on the driving term $\tau^{\mathcal{T}}[n|\ell] = \frac{1}{\ell}\sum_{k=1}^{\ell} \tau_k[n]$, that is the mean values of the set $\tau^{\mathcal{T}}[n|\ell]$.

After one step-update there will be all the TO values $\tau_k[n+1] \sim N(\bar{\tau}[n], \sigma^2[n])$ for $k = 1, 2, \ldots, \ell$ with the same distribution as these are the set $\mathcal{T}[n]$, and the TO values of the remaining set $\mathcal{R}[n]$ have been updated according to (2.20), and these have a new conditional pdf: $\tau_k[n+1|\ell] \sim N(\bar{\tau}[n+1|\ell], \sigma^2[n+1|\ell])$ for $k = \ell + 1, \ldots, K$ with moments:

$$\begin{aligned} \bar{\tau}[n+1|\ell] &= (1-\varepsilon_\tau)\bar{\tau}[n] + \varepsilon_\tau E[\tau^{\mathcal{T}}[n|\ell]] = \bar{\tau}[n], \\ \sigma^2[n+1|\ell] &= \begin{cases} (1-\varepsilon_\tau)^2\sigma^2[n] + \frac{\varepsilon_\tau^2}{\ell}\sigma^2[n] & \text{for } \ell \geq 1 \\ \sigma^2[n] & \text{for } \ell = 0. \end{cases} \end{aligned} \tag{2.21}$$

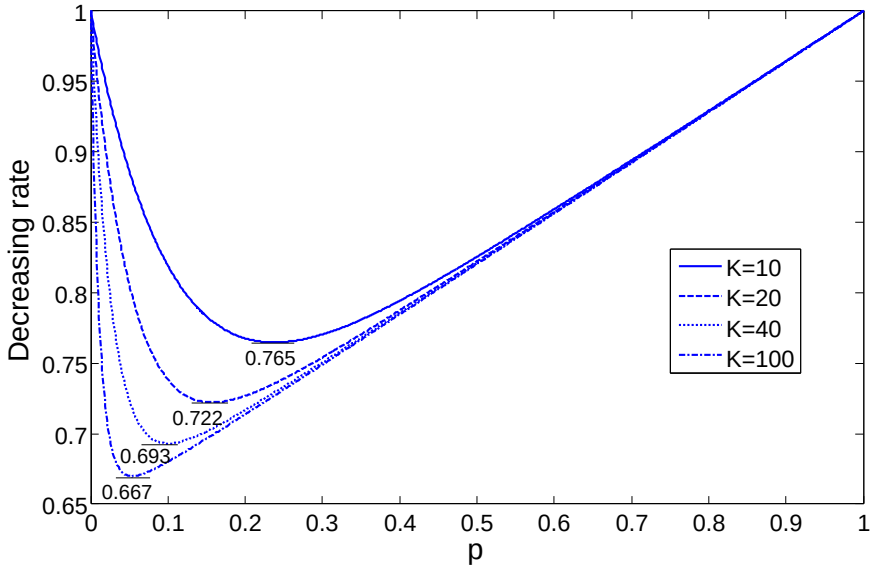


Figure 2.7: Decreasing rate vs. p for different K ($\varepsilon_\tau = 0.2$).

Since the mean value of every conditional pdf $\bar{\tau}[n+1|\ell]$ is independent on the update iteration n, the unconditional pdf can be assumed as Gaussian too, $\tau_k[n+1] \sim N(\bar{\tau}[n+1], \sigma^2[n+1])$ for $k = 1, 2, \ldots, K$. The variance of the unconditional pdf at iteration $n+1$ is

$$\begin{aligned}\sigma^2[n+1] =& \sigma^2[n]\text{Pr}[\ell = 0] + \sigma^2[n]\text{Pr}[\ell = K] \\ &+ \sum_{\ell=1}^{K-1} \left(\frac{\ell}{K}\sigma^2[n]\text{Pr}[K-\ell = |\mathcal{R}[n]|] \right. \\ &\left. + \frac{K-\ell}{K}\sigma^2[n+1|\ell]\text{Pr}[\ell = |\mathcal{T}[n]|] \right).\end{aligned} \tag{2.22}$$

When $\ell = 0$ all nodes are on receiving stage (there is no distribution of synchronisation signals) with $\text{Pr}[\ell = 0] = (1-p)^K$. When $\ell = K$ all nodes are transmitting (none is updating) with $\text{Pr}[\ell = K] = p^K$. The total variance decreases at each iteration to reach convergence state ($\sigma^2[n+1] < \sigma^2[n]$). Therefore, the decreasing rate is

$$\begin{aligned}&\frac{\sigma^2[n+1]}{\sigma^2[n]} = (1-p)^K + p^K \\ &+ \sum_{\ell=1}^{K-1} \left(\frac{\ell}{K} + \frac{K-\ell}{K}\left[(1-\varepsilon_\tau)^2 + \frac{\varepsilon_\tau^2}{\ell}\right] \right) \binom{K}{\ell} p^\ell (1-p)^{K-\ell}.\end{aligned} \tag{2.23}$$

For convergence, the time of convergence is minimised (or the convergence rate is maximised) when the decreasing rate (2.23) is minimised. Given K and ε_τ, the convergence rate can be maximised for a choice of the probability $\text{Pr}[k \in \mathcal{T}[n]] = p$ that minimises the total decreasing rate (2.23). Figure 2.7 shows the decreasing rate (2.23) versus

p for different number of nodes K. For any network configuration K the convergence time is minimised for $p < 0.5$, and for a large number of nodes (say $K = 40, 100$) the convergence rate is maximised when $p < 0.1$. Analysis shows that compared to the symmetric duplex scheduling ($p = 0.5$), the state of reception with few nodes transmitting the synchronisation state should be always privileged to guarantee a fast convergence to network synchronisation.

2.1.6 Numerical results

The numerical analysis is based on a mutually coupled fully connected network. In a dense network where $K \geq 10$, all nodes broadcast the same synchronisation waveform and transmission is simulated in form of frames, where each frame contains the training waveform (2.8) followed by OFDM symbols for payload. The synchronisation waveform is a CAZAC sequence with length $N = 64$ and root-index $u = 1$. Here, we consider $\alpha = 8$ for metrics (2.10). At step $n = 0$, TO for each node is uniformly distributed as $\tau[0] \sim \mathcal{U}(-\tau_{max}, -\tau_{max})$ and the CFO is normalised by f_s as Gaussian $f[0] \sim \mathcal{N}(0, f_{max}^2)$. The mean square dispersion errors (MSE), $MSE_{TO} = \frac{1}{K(K-1)} \sum_{k,i \neq k} (\tau_k - \tau_i)^2$ and $MSE_{CFO} = \frac{1}{K(K-1)} \sum_{k,i \neq k} (f_k - f_i)^2$ are the metric used here to measure the accuracy of TO and CFO synchronisation.

The coupling effect of TO-CFO is analysed in Figure 2.8 by 2D histogram (gray scale coding) for 1000 runs of Montecarlo simulation, $K = 10$ nodes and D-PLL parameters $\varepsilon_\tau = 0.4$ and $\varepsilon_f = 0.3$. Since duplexing strategy change the behaviour of the synchronisation convergence, here it is assumed that nodes act as full duplex ($\mathcal{T}[n] = \mathcal{R}[n]$). A detailed behaviour of the transition vs. synchronisation step n for one run is shown in upper-left histogram Figure 2.8 for both TO and CFO synchronisation. The TO-CFO coupled are analysed on the timing and frequency synchronisation by varying the initial dispersion $\tau_k[0]$ and $f_k[0]$. The behaviour of TO synchronisation (Figure 2.8.a) is not affected by variations on the CFO. On the other hand, the time of convergence of CFO synchronisation increases when the dispersion of TO increases as is shown in Figure 2.8.b due to the coupling effect of TO-CFO.

The RMSE of TO and CFO synchronisation vs. signal-to-noise ratio defined as σ_s^2/σ_w^2 with $\sigma_s^2 = N$ is in Figures 2.9 and 2.10. A master-slave (MS) topology with one node acting as reference agent for TO and CFO synchronisation is shown as reference. Distributed synchronisation for dense cooperative network ($K = 20$, fully connected) achieves a performance comparable with MS strategy except for degradation of 3 dB on ratio σ_s^2/σ_w^2 on threshold.

Duplex scheduling is evaluated and Figure 2.11 shows in a double scale both TO and CFO time of convergence for a fully-connected network (all-to-all connectivity) vs. the transmitting probability p for different number of nodes K and $\varepsilon_\tau = \varepsilon_f = 0.2$, the simulated network (dashed lines) and the analytic model (2.23) (solid lines) are analysed. In spite of the approximations, the decreasing rate (2.23) provides a good prediction of

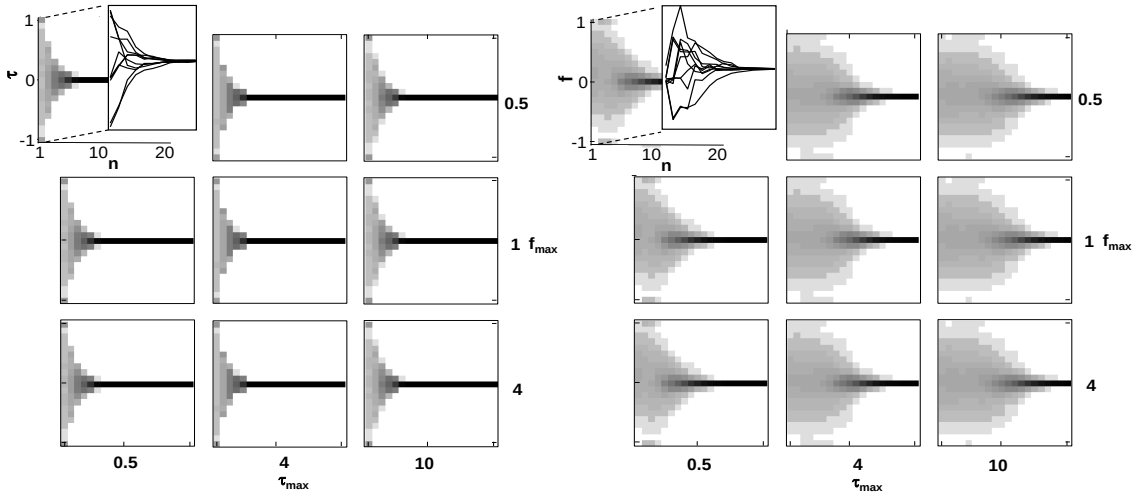


Figure 2.8: TO and CFO synchronisation varying initial dispersion $\tau_k[0] \sim U(-\tau_{max}, \tau_{max})$ and $f_k[0] \sim N(0, f_{max}^2)$ for $\tau_{max} = \{N/2, 4N, 10N\}$ and $f_{max} = \{0.5, 1, 4\}$, by 2D histogram (gray scale coding. $K = 10$, $N = 64$, step size: $\varepsilon_\tau = 0.4$, $\varepsilon_f = 0.3$). Detailed behaviour of transition toward convergence is shown in upper-left histogram ($\tau_{max} = 0.5$, $f_{max} = 0.5$).

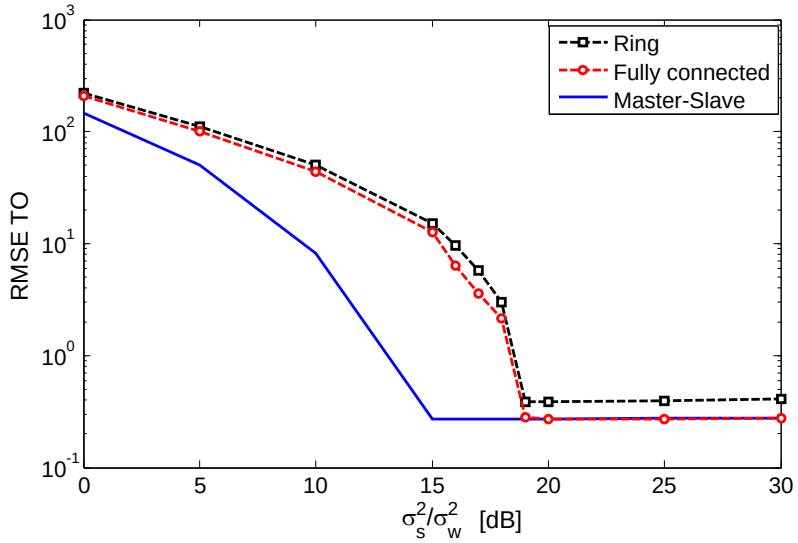


Figure 2.9: Root mean square error of TO vs. σ_s^2/σ_w^2 for a master-slave topology (lower bound), a fully connected (circle) for $K = 20$ and a ring topology (square) for $K = 40$ where the degree of each node is $d = 20$ nodes.

the time of convergence of TO and confirms that the minimum convergence time is for duplexing strategy with low transmitting probability p as in Figure 2.7. The CFO reaches

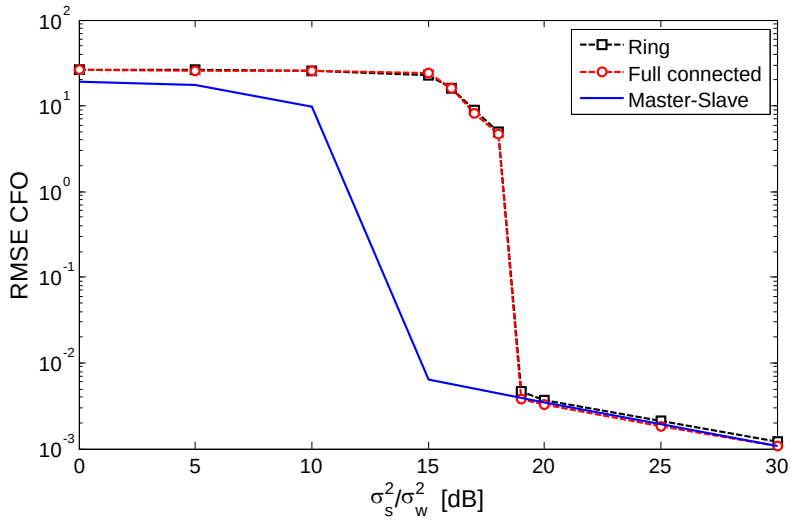


Figure 2.10: Root mean square error of CFO vs. σ_s^2/σ_w^2 for a master-slave topology (lower bound), a fully connected (circle) for $K = 20$ and a ring topology (square) for $K = 40$ where the degree of each node is $d = 20$ nodes.

a synchronisation status after TO convergence as CFO needs first to let both TO and integer CFO to converge, and then the fractional CFO can move last steps to global CFO convergence. The convergence rate of CFO is maximised for the same probability p that let TO reach its minimum time of convergence. Thus, if the TO time of convergence is minimised, the CFO time of convergence is minimised too, not only because CFO and TO synchronisations are sequential (fractional CFO follows the integer CFO that is paired with TO) but the synchronisation mechanism for CFO (Subsection 2.1.2) is the same as TO.

The fastest convergence is when network is full-duplex (lower bound), and this is used for the comparison in Figure 2.12. The dispersion of TO and CFO for full-duplex are compared with the half-duplex nodes for $K = 10$ nodes, using a random duplex scheduling with the optimum probability $p = 0.25$ to minimise the convergence time. Results in Figure 2.12 guarantee that optimised duplexing degrades the convergence time of TO and CFO synchronisation of 30–40% compared to the unrealistic full-duplex.

2.1.7 Conclusions

Joint time and frequency synchronisation scheme for uncoordinated networks with no reference was designed based on the CAZAC sequences as synchronisation frame that is the same for all nodes in the network, this makes the algorithm less complex by using just one correlator filter in receivers. The preamble decouple the effect of CFO on TO estimation and allow a joint timing and frequency synchronisation with accuracy

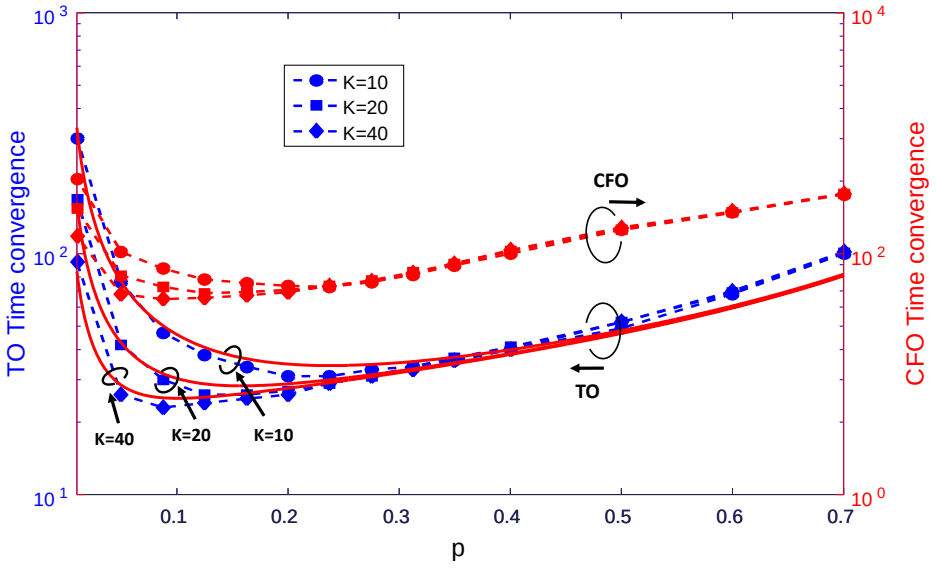


Figure 2.11: Convergence time of TO (left axis) and CFO (right axis) vs. p: simulations (dashed line) and analytic model (solid line) for different K and $\varepsilon_\tau = \varepsilon_f = 0.2$.

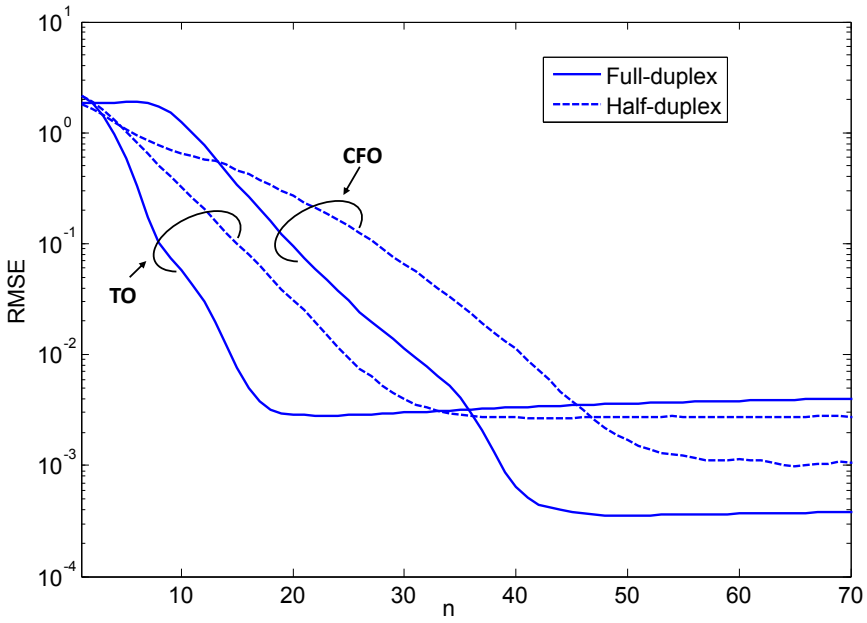


Figure 2.12: TO and CFO dispersion error of full-duplex network (lower bound) and optimised half-duplex ($p = 0.25$) vs. iterations for $K = 10$.

comparable to conventional master-slave synchronisation methods. For a random duplex scheduling strategy, it was proved that in a dense network, the minimum time of convergence can be reached with a lower transmission probability p. The optimum trans-

mission probability can be estimated by maximising the speed of the timing dispersion decreasing that depends on the number of nodes or graph connectivity and the convergence parameters ε_τ and ε_f. The fact that the network reaches a faster convergence with a lower transmission probability has the additional favourable advantage of reducing the power consumption during synchronisation.

2.2 Synchronisation algorithm evaluation

The experimental validation of the distributed synchronisation algorithm had been implemented in hardware on Ettus USRP N210 SDR hardware [9], with XCVR2450 daughter boards, programmed using GNU Radio, as part of HW demonstrator in WP5. The USRP hardware provides a flexible test system with signal bandwidths up to 20 MHz and centre frequencies between 2.4 and 2.5 GHz. As mentioned before, the fact that all nodes are assigned the same synchronisation signature reduces the number of required matched filters making it a suitable candidate for low cost hardware systems. The CFO and TO estimation and tracking capabilities of the algorithm has been investigated, also by validating whether the decoupled estimation of CFO and TO can be realised in practice [10].

For the hardware evaluation a centre frequency of 2.48 GHz was used, and the samples of the preamble were sent out at a rate of 1 MHz. N was set to equal 64 samples, giving a preamble length of 128 samples (128 µs duration). The experimental setup aims to investigate the estimation and tracking abilities of the algorithm, and not the consensus adaption. As such, the pool of transmitting nodes $\mathcal{T}[n]$ will remain constant, and one receiving node will be used to track and record the estimated CFO and TO from the superimposed waveforms. Up to three transmitting nodes were used in the experiments to investigate the super-positioning of preamble from the different transmitting nodes.

To implement the algorithm on hardware, it has been imposed slight modifications. The first was to interpolate the preamble data, as this eased the sampling of the signals at the receiver. A linear interpolator was used to increase the sample rate $1:4$, thus the actual sample rate of the USRPs was 4 Msample/s. The receiver was set to capture and process a window of samples ten times the length of the preamble, this size window was chosen as a trade-off between processing limitations and reducing the chances of capturing only part of a transmitted preamble at the receiver. For updating, the carrier frequency and timing of the receiver node ε_τ and ε_f were both set to 10^{-3}.

To estimate a relative TO the receiver node must know the minimum period of the preambles re-transmissions, this allows the receiver to estimate how much its own timing needs to be altered within this reference window. For speed of convergence a period of 64 samples, equivalent to N, was used. Depending on where the peaks of the correlations $r_{k,1}[\ell]$ and $r_{k,2}[\ell]$ occur relative to the centre of this window, the timing of the receiver can be altered.

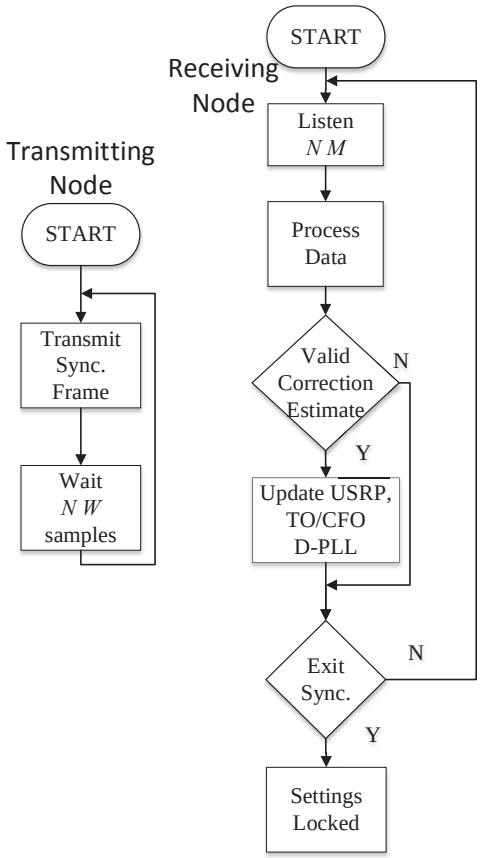


Figure 2.13: Flow-chart of the transmitting and receiving nodes.

The transmitters were set to transmit their preambles followed by a pause equivalent to 9 times the preamble length, this allowed the receiver to capture a preamble from each of the transmitters within each captured window of data. Figure 2.13 shows the flow-chart of the transmitting and receiving nodes, in the setup M was set to 10 and W to 9, the TO and CFO estimation algorithm is run in the 'Process Data' block. The output from the correlations $r_{k,1}[\ell]$ and $r_{k,2}[\ell]$ are used to verify the reception of a valid preamble, before any estimations are applied to the USRPs. For the tests involving multiple transmitters the preamble transmissions were synchronised, so that when altering the TO of the transmissions the delays were relative to one another. The USRP reference clocks were also run from a common source, this allowed an accurate CFO and TO to be set at each node, i.e. at each node is set the relative TO and CFO to be estimated.

During tests each USRP was connected to a laptop running GNU Radio. Figure 2.14 shows the test setup. The USRPs provide a flexible SDR platform, however most signal process is undertaken on the laptops, with the USRPs effectively acting as an up and down-converter for the IQ symbols sent from or to the laptops. Because of this approach to data processing, the USRPs are unable to benefit from many of the signal processing

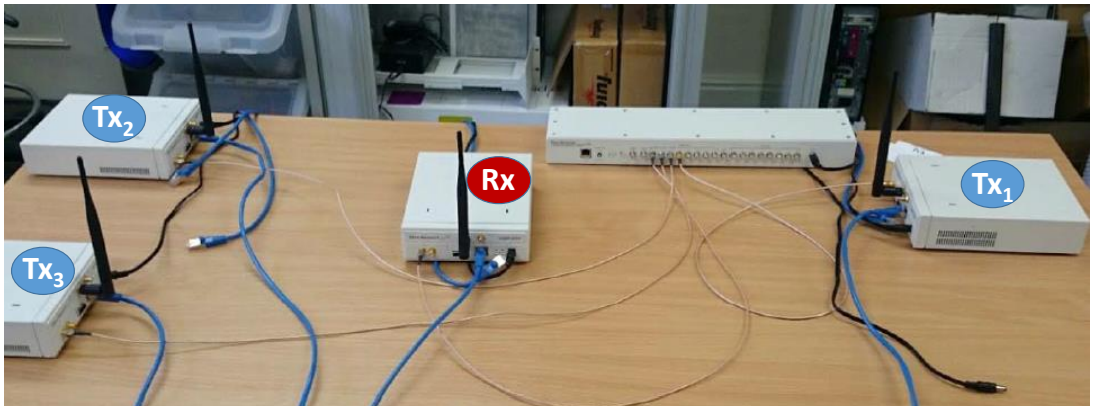


Figure 2.14: USRP hardware test set-up.

operations that can be efficiently implementable in hardware, but not in software, such as FIR filters.

In an optimised system, capable of efficiently implementing FIR filters in hardware, the matched filter operations could be off loaded onto the hardware. This would allow for cheaper, lower power processors to be used. As the number of matched filters in the distributed synchronisation algorithm is fixed at two, this allows the proposed system to be implemented on much lower complexity hardware while still allowing the size of the network to scale up according to the size necessary for modern dense networks.

Two experiments were undertaken to test different elements of the estimation algorithm. Each test was divided into stages where either the TO or CFO of the transmitting nodes was changed. The receiver would then process the received signals and adapt its time windowing and centre frequency. The receiver recorded the estimated CFO and TO, the next stage of the test was run only once the adaption of TO and CFO had converged to a stable consensus.

2.2.1 Test 1: Estimation and tracking ability, 2 Tx – 1 Rx

Test 1 investigated the estimation and tracking ability of the algorithm, and how it coped with a second transmitter joining part way through the tracking process. This test used two transmitting nodes and one receiver. Table 2.1 shows the ten stages of the test. For stages T1–T8 the second transmitter is switched off, and Tx_1 changes the CFO between -12.5 kHz, 0 kHz, 25 kHz and 50 kHz with respect to the carrier frequency of the receiver, and alters the TO of Tx_1 between zero and ten sample bins (10 µs). At stage T9 Tx_2 is turned on, initially with different settings to Tx_1, then in T10 Tx_2 alters it transmission to match Tx_1. On these stages, T9-T10, it is of interest to evaluate the stability of the algorithm when a node is added to the network.

Table 2.1: Test 1 sequences, involving Tx_1 and Tx_2, including approx start time of each stage.

Stage	Time Approx. (s)	Tx1 CFO (kHz)	Tx1 TO (Bins)	Tx2 CFO (kHz)	Tx2 TO (Bins)
T1	0	0	0	OFF	OFF
T2	9	0	10	OFF	OFF
T3	14	25	10	OFF	OFF
T4	22	-12.5	10	OFF	OFF
T5	32	-12.5	0	OFF	OFF
T6	39	50	0	OFF	OFF
T7	47	0	0	OFF	OFF
T8	63	25	10	OFF	OFF
T9	69	25	10	0	0
T10	79	25	10	25	10

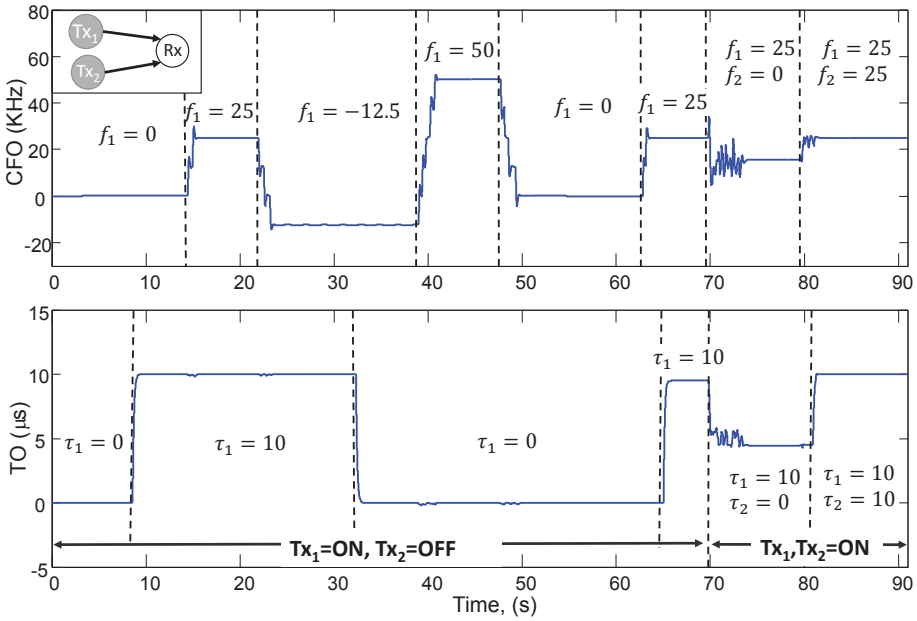


Figure 2.15: Test 1 results, showing the estimated CFO (top) and TO (bottom) at the receiving node for 2 nodes active on different stage and varying their TO and CFO parameters.

Table 2.2: Test 2 sequences, involving Tx_1, Tx_2 and Tx_3, including approx start time of each stage.

	Time	Tx1		Tx2		Tx3	
Stage	Approx. (s)	CFO (kHz)	TO (Bins)	CFO (kHz)	TO (Bins)	CFO (kHz)	TO (Bins)
T1	0	0	0	0	0	0	0
T2	22	25	10	0	0	0	0
T3	37	25	10	12.5	10	0	0
T4	52	25	10	12.5	10	37.5	10
T5	64	25	10	25	10	25	10

The results for Test 1 are in Figure 2.15. For TO tracking, when only Tx_1 is transmitting, the algorithm is seen to faithfully track the changes in transmitted signal. When Tx_2 is turned on, in stage T9, TO is estimated at around 5 µs, the algorithm estimates the relative TO error from the ensemble of both transmitted signals, which is half the TO of Tx_1 and Tx_2. Then in stage T10 when the two nodes transmissions are synchronised, the estimated TO increases to 10 µs. This test shows reliable tracking of the timing of the preamble for one and two transmitting nodes.

The algorithm is also shown to be capable of tracking the changes in the transmitter nodes CFO. At first stage, all transmitter nodes are operating at the same frequency, the CFO is estimated with a maximum error of 180 Hz, which for a 2.48 GHz centre frequency is an accuracy of 72.58 parts per billion (ppb). When Tx_1 and Tx_2 are operating at different frequencies the algorithm takes nearly 5 s to converge to a stable estimate of 15.57 kHz, which is slightly lower than the midpoint.

The two graphs of Figure2.15 validate experimentally the decoupled estimations of CFO and TO as described in Section 4.1.3, with the estimation of one remaining steady when the other is changed.

2.2.2 Test 2: Estimation and tracking ability, 3 Tx – 1 Rx

Test 2 evaluates the ability of the synchronisation algorithm to extract the relative TO and CFO error from the superimpose of the transmitted synchronisation waveform from three transmitting nodes. All nodes are initially synchronised, then one at a time change their properties before resynchronising in the final stage. Table 2.2 shows the five stages for this test.

The results of Test 2 are shown in Figure 2.16. At first and last stage of the test all transmitting nodes share the same configuration, so the estimation accuracy can only

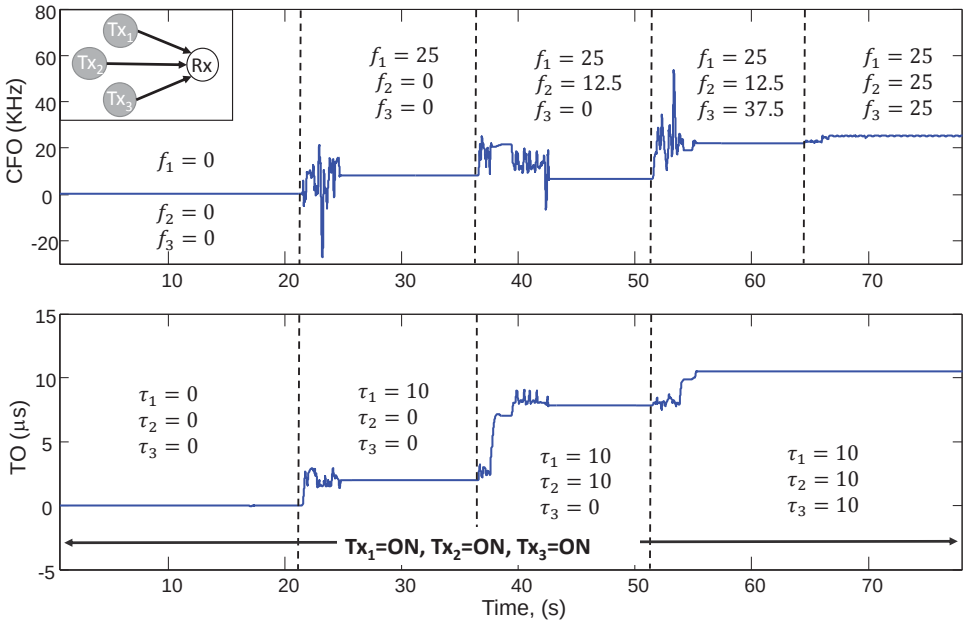


Figure 2.16: Test 2 results, showing the estimated CFO (top) and TO (bottom) at the receiving node for 3 nodes changing their CFO and TO parameters.

be compared here. However, it does give us insights into the behaviour of the algorithm when there are multiple transmitters with different configurations.

First looking at the TO estimations of the receiver, at stages T1 and T5 the algorithm correctly estimates a 10 sample bin difference. During the other stages the TO is again estimated somewhere between the TO of all the transmitter nodes. Moving from an estimated TO of 2 sample bin when two nodes TOs were set to zero and one node was set to ten. Then, in stage T3 the receiver estimated a TO of between 7 and 8, when there were two nodes with a TO of ten and one of zero.

The accuracy of the CFO estimation is measurable in stages T1 and T5. The CFO estimation error is between 140 and 250 Hz, which for the carrier frequency of 4.48 GHz is around 100 ppb. It is interesting to note that the estimated CFO between stages T2 and T3 actually decreased, even though the average carrier frequency of the transmitters increased, the reduction is only 1.6 kHz, and it was found that the signal received at the receiver node from Tx_2 was slightly weaker than from the other two nodes, this would account for the estimated CFO being effected more by changes in the other two nodes. It is also seen that the convergence time of the algorithm is significantly increased when increasing the number of nodes and varying configurations, often taking nearly 5 sec to reach a stable state.

The convergence time can be improved taking account some factors that can significantly reduce this time, one is the hardware implementation and processing delay, as an updated CFO may be applied to the USRPs after the next window of data has already

been captured, this can cause a small overestimation of the CFO and cause greater oscillations. The transient oscillations may also be reduced by using lower values of ε_τ and ε_f [11] in the iterative adaption, or by changing the frame structure for denser adaptation. In these tests the transmitters may be forced to have different and fixed CFO, and thus the receiver attempts to minimise an error function for which there may be several local minimums, in a fully adaptive system these local minimums would converge into an agreed global minimum.

2.2.3 Conclusions

The distributed synchronisation algorithm was implemented on software defined radios programmed in GNU radio. It was demonstrated the tracking and accuracy capability of the algorithm under a range of scenarios, including the super-positioning of synchronisation frames from multiple transmitters. When all transmitters were operating at the same carrier frequency and relative timings an estimation accuracy of carrier frequency was found to be within 100 part per billion, and the timing offsets were correctly estimated. These tests also validate the algorithms ability to decouple the timing and frequency estimates. The convergence time of the algorithm was found to significantly increase when transmitting nodes had different carrier frequency offsets. However, all nodes in a final system using the algorithm are likely to adapt their configurations, so would converge on a globally value of carrier frequency and timing.

2.3 Assessment of distributed algorithms for channel estimation and network localisation

In peer-to-peer cloud networks, self-learning of the network state is fundamental for the set-up of an efficient intra-cloud connectivity. In this section we investigate the performances of core consensus-based algorithms for self-learning of the network state presented in Deliverables D3.02 [4] and D2.31 [12] in experimental application scenarios. We analyse two specific application scenarios related to the distributed estimation of the channel parameters and of the geometric topology of the connectivity graph. In the former case, we are interested in the estimation of some environment-dependent parameters that rule the channel quality of all the links within the cloud, see the bivariate stochastic model presented in D3.02 [4], [13]. In the latter, each node of the cloud is required to infer the global properties of the whole network from local ranging measurement within the neighbourhood, i.e. the positions of all nodes regardless of link availability. An iterative Gauss-Newton location estimator is integrated into the consensus algorithms to handle the non-linearity of the measurement model. Performance limits are analysed and compared to those of conventional consensus algorithms, in terms of accuracy and convergence rate, using fundamental performance bounds as benchmarks.

The amount of signalling involved in consensus iterations is evaluated as being relevant for practical implementation within WP5 demonstrators.

2.3.1 Review of the bivariate channel model

The instantaneous power at time t, $\widetilde{\mathrm{RSS}}_{ij}(t)$ [W], for the link between nodes i and j with mutually distance D_{ij} can be modelled as [14]:

$$\widetilde{\mathrm{RSS}}_{ij}(t) = \widetilde{\mathrm{RSS}}_0 \left(\frac{D_0}{D_{ij}}\right)^{\gamma_\mathrm{P}} \tilde{n}_{ij}\tilde{\delta}_{ij}(t) \tag{2.24}$$

where $\widetilde{\mathrm{RSS}}_0$ is the received power at reference distance D_0, γ_P the path-loss index, $\tilde{n}_{ij}$ the log-normal shadowing term and $\delta_{ij}(t)$ the temporal fast fading. By converting the RSS in dB scale, $\mathrm{RSS}_{ij}(t) = 10\log_{10}\widetilde{\mathrm{RSS}}_{ij}(t)$, with proper mapping terms, we get:

$$\mathrm{RSS}_{ij}(t) = \mu_{\mathrm{P},0} - 10\gamma_\mathrm{P}\log_{10}\left(\frac{D_{ij}}{D_0}\right) + n_{ij} + \delta_{ij}(t), \tag{2.25}$$

where $\mu_{\mathrm{P},0}$ is the average received power in dBm at the reference distance D_0, $n_{ij} \sim \mathcal{N}(0, \sigma_\mathrm{P}^2)$ is the normal shadowing term with zero mean and standard deviation σ_P [dB], $\delta_{ij}(t) = 10\log_{10}\tilde{\delta}_{ij}(t) \sim \mathcal{LR}(0, K_{ij})$ is the temporal fading here assumed as log-Rician with zero mean and K-factor K_{ij} [dB]. The K factor measures the power ratio between the static and dynamic channel components, see D3.02 [4].

The model here proposed describes the changes of the RSS statistics, namely the mean $P_{ij} = \mathrm{E}_t[\mathrm{RSS}_{ij}(t)]$ and the variance $V_{ij} = \mathrm{Var}_t[\mathrm{RSS}_{ij}(t)] = \mathrm{Var}_t[\delta_{ij}(t)]$, over the D2D links (i, j) of the network, and particularly how they relate to the D2D distances D_{ij} (based on the network topology). In particular, the mean term P_{ij} accounts for the effects of fixed scattering/absorbing objects that influence the static multipath component [15], whereas the variance V_{ij} is due to moving scatterers/absorbers in the environment. According to the log-normal shadowing assumption for n_{ij} in (2.25), the spatial fluctuations of the RSS mean P_{ij} over link (i, j) are Gaussian:

$$P_{ij} \sim \mathcal{N}(\mu_\mathrm{P}\left(D_{ij}\right), \sigma_\mathrm{P}^2), \tag{2.26}$$

with spread σ_P depending on the specific environment and mean linearly decaying with the link log-distance according to the path-loss function:

$$\mu_\mathrm{P}\left(D_{ij}\right) = \mu_{\mathrm{P},0} - 10\gamma_\mathrm{P}\log_{10}\left(\frac{D_{ij}}{D_0}\right). \tag{2.27}$$

With regard to the RSS variance, no quantitative model has been proposed in the literature to represent this channel feature. Some insight can be gain from the experimental analysis in [15], which shows that the K factor K_{ij} is Gaussian distributed over the space,

and decreasing on average with the link log-distance; furthermore, K_{ij} is proved to be negatively correlated to P_{ij}. To analyse the RSS variance distribution we investigate the relationship $K_{ij} = K(V_{ij})$ between the RSS variance $V_{ij} = \int (r - P_{ij})^2 f_{\mathrm{RSS}_{ij}}(r) dr$ and the K factor K_{ij}, where $f_{\mathrm{RSS}_{ij}}(r)$ is the probability density function of the Log-Rice variable $\mathrm{RSS}_{ij}(t)$. By numerical integration, we get the function $K_{ij} = K(V_{ij})$, where the variance ranges from 0 dB in the LOS case, i.e. for $K_{ij} \longrightarrow +\infty$ dB, non-fading channel, to 30 dB in NLOS (for $K_{ij} \longrightarrow -\infty$ dB, Rayleigh fading). Typical K-factor observed in static network scenarios is above 2 dB, while for larger values it linearly scales with the log-variance $\Sigma_{ij} = \log(V_{ij})$. It follows that the RSS log-variance Σ_{ij} can be effectively modelled as Gaussian:

$$\Sigma_{ij} \sim \mathcal{N}(\mu_{\Sigma}\left(D_{ij}\right), \sigma_{\Sigma}^2), \tag{2.28}$$

with standard deviation σ_{Σ} [logdB] and mean $\mu_{\Sigma}\left(D_{ij}\right)$ linearly increasing with the link log-distance:

$$\mu_{\Sigma}\left(D_{ij}\right) = \mu_{\Sigma,0} - 10\gamma_{\Sigma} \log_{10}\left(\frac{D_{ij}}{D_0}\right), \tag{2.29}$$

where environment-dependent regressor parameters are $\mu_{\Sigma,0}$ and $\gamma_{\Sigma} \leq 0$, similarly to (2.27). Note that the parameters $\{\mu_{\Sigma,0}, \gamma_{\Sigma}, \sigma_{\Sigma}\}$ depend on the spatial density of the moving scatterers/absorbers in the space, e.g. $\sigma_{\Sigma} = 0$ for static environments without any moving objects, whereas it is $\sigma_{\Sigma} \neq 0$ and $\gamma_{\Sigma} \ll 0$ for high spatial density of moving objects).

The mean P_{ij} and the log-variance Σ_{ij} of the RSS of a D2D link (i, j) are thereby jointly Gaussian distributed,

$$\mathbf{y}_{ij} = \begin{bmatrix} P_{ij} \\ \Sigma_{ij} \end{bmatrix} \sim \mathcal{N}(\boldsymbol{\mu}\left(D_{ij}\right), \mathbf{Q}_{\mathrm{P}\Sigma}), \tag{2.30}$$

with mean value and covariance matrix:

$$\boldsymbol{\mu}\left(D_{ij}\right) = \begin{bmatrix} \mu_{\mathrm{P}}\left(D_{ij}\right) \\ \mu_{\Sigma}\left(D_{ij}\right) \end{bmatrix}; \quad \mathbf{Q}_{\mathrm{P}\Sigma} = \begin{bmatrix} \sigma_{\mathrm{P}}^2 & \rho\sigma_{\mathrm{P}}\sigma_{\Sigma} \\ \rho\sigma_{\mathrm{P}}\sigma_{\Sigma} & \sigma_{\Sigma}^2 \end{bmatrix}, \tag{2.31}$$

where $\rho < 0$ is the negative cross-correlation according to [15].

2.3.2 Distributed estimation of channel model parameters

According to the bivariate stochastic model presented in [4], the parameters that identify the propagation in the network area are the linear regressors $\boldsymbol{\theta} = [\mu_{\mathrm{P},0}, \gamma_{\mathrm{P}}, \mu_{\Sigma,0}, \gamma_{\Sigma}]^{\mathrm{T}}$ of the ensemble mean functions and the covariance matrix $\mathbf{Q}_{\mathrm{P}\Sigma}$, see Section 2.3.1 or [4] for details.

In order to test the distributed algorithm performances, a measurement campaign has been carried out at the third floor of the department DEIB of Politecnico di Milano with

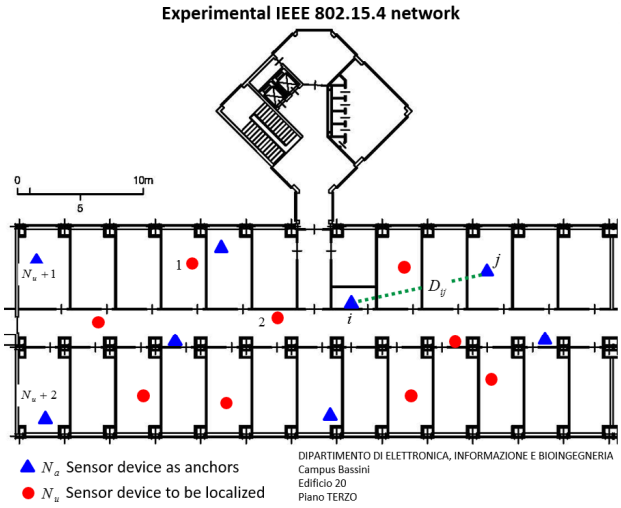


Figure 2.17: Sensor deployment for distributed network calibration and localisation (DEIB, Politecnico di Milano) over a floor of approximately $20\,\mathrm{m} \times 50\,\mathrm{m}$ rectangular area.

$N = 17$ devices deployed as shown in Figure 2.17), in accordance to the cloud architecture of the DIWINE project [16]. RSS measurements have been recorded during day-time in mixed LOS/NLOS conditions. Radio modules used for the experiments provide different programmable high-power modes with maximum transmit power of $18\,\mathrm{dBm}$, while the minimum received power is $-98\,\mathrm{dBm}$. A PC connected by serial interface to one of the devices collected all the measurements. For each link (i, j), RSS measurements are collected with sampling time $300\,\mathrm{ms}$ and over 3 pre-defined channels with centre frequencies $2.405\,\mathrm{GHz}$, $2.425\,\mathrm{GHz}$ and $2.45\,\mathrm{GHz}$, respectively (standard compliant channels 11,15 and 20). Considering all the frequencies, $M_0 = 15$ samples of RSS mean and variance (macro-parameters) $\{P_{ij}^{(m)}, \Sigma_{ij}^{(m)}\}_{m=1}^{M_0}$ are estimated for each link from independent RSS measurements (using RSS data-sets of 2000 samples).

For distributed estimation, we consider all the methods presented in [4][12], and listed here: 1) Simple Consensus on Local ML Estimates (D-SC) method [17], Weighted Consensus on ML Estimates (D-MLE) method [18] and Weighted Consensus with Accuracy Exchange (D-MLE-AE) method [19]. The estimate of $\mathbf{Q}_{\mathrm{P}\Sigma}$ is obtained once the consensus on $\boldsymbol{\theta}$ has been reached, by locally reconstructing the measurement errors and implementing a consensus algorithm adapted for the sample covariance matrix.

The performances of centralised and distributed algorithms are compared in terms of root mean square error (RMSE) of the estimate in Figure 2.18. The step size has been set to $\varepsilon = 0.99\varepsilon_{\max}$, with $\varepsilon_{\max} = 1/\Delta$ for D-SC and D-MLE methods and $\varepsilon_{\max} = 1$ for D-MLE-AE method. The RMSE is shown for all the methods vs. the number of iterations in Figure 2.18–(a) (for $\boldsymbol{\theta}$) and 2.18–(b) (for $\mathbf{Q}_{\mathrm{P}\Sigma}$). As expected, the D-SC method does not converge to the centralised estimate, as it does not exploit any information

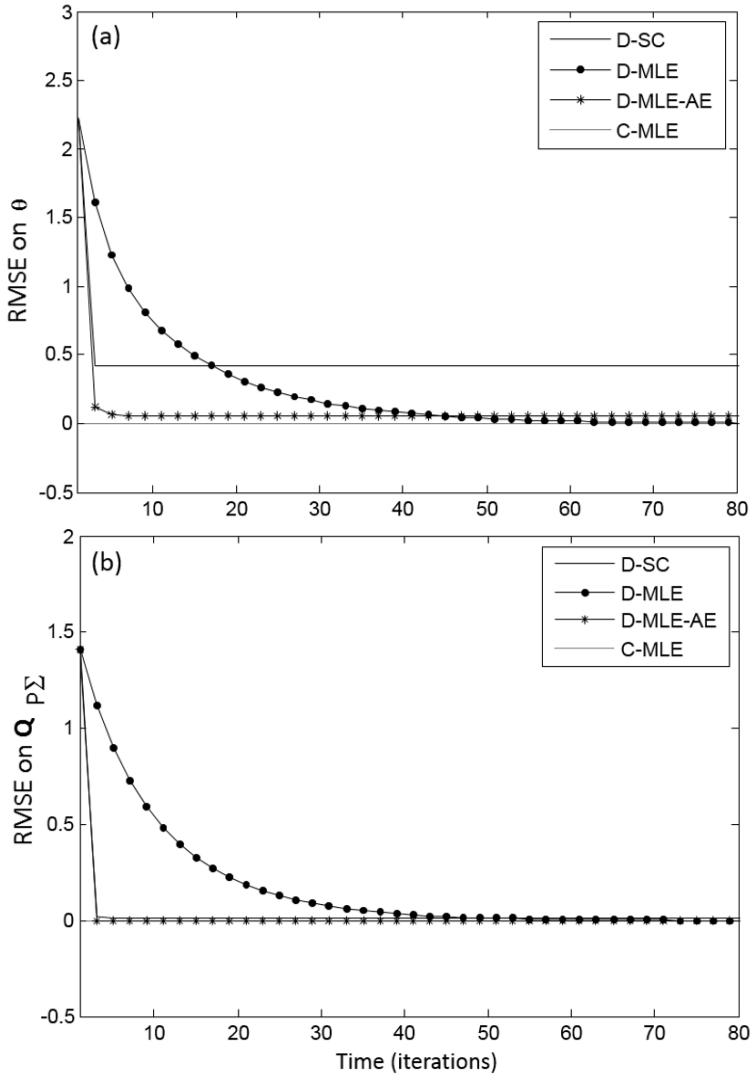


Figure 2.18: Error performance of distributed (D) consensus-based algorithms compared to the centralised approach (C-MLE) for parameter estimation $\boldsymbol{\theta}$ (a) and $\mathbf{Q}_{\mathrm{P}\Sigma}$(b).

on the accuracy of the estimates exchanged between nodes. On the other hand, the introduction of weighting in the data fusion allows the D-MLE method to converge to the centralised estimate, though with a pretty slow rate. The D-MLE-AE method does not converges exactly to the optimal performance, but it closely approaches it and converges faster than the other weighted consensus method. Though it requires a slightly higher communication overheard in the initialisation step to share the estimate covariances between nodes (($O(p^2/2)$ per link in subsequent iterations the overhead reduces to p parameters per link, as D-MLE, without any meaningful performance loss.

To conclude, the D-MLE-AE method is the one that provided the best trade-off of performances in terms of convergence rate, accuracy of the estimate and communication overhead. On the other hand, D-MLE guarantees optimal performance at convergence but with higher number of iterations.

2.3.3 Network localisation

An application that can highly benefit from the knowledge of the channel model recalled and estimated in previous section is network localisation, which aims to infer the network topology, i.e. the positions of all nodes, from local RSS observations. In this section we discuss the application proposing a distributed localisation approach based on the D-MLE-AE consensus method. The method D-MLE is not suited as it not fast enough (as proved in the previous subsection) and also it cannot handle under-determined local settings, see [4],[12].

We assume that the environmental-dependent parameters are now known at the nodes of the network, e.g. as estimated by a pre-calibration procedure carried out by a subset of anchor nodes of known positions. Recalling the wireless network structure in Figure 2.17 and denoting the location of node i by the Cartesian coordinates $\theta_i = [\theta_{xi}\ \theta_{yi}]^{\mathrm{T}} \in \mathbb{R}^2$, $i = 1, \ldots, N$, we assume that the nodes to be localised are indexed as $i = 1, \ldots, N_{\mathrm{u}}$, while the remaining $N_{\mathrm{a}} = N - N_{\mathrm{u}}$ are anchors with known positions. Distributed inference is here used for the estimation of the locations $\boldsymbol{\theta} = [\theta_1^{\mathrm{T}} \cdots \theta_{N_{\mathrm{u}}}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2N_{\mathrm{u}}}$ from D2D measurements $\mathbf{y}_{ij} = [P_{ij}\ \Sigma_{ij}]^{\mathrm{T}}$ of RSS mean and log-variance, see [4], taken over the active links (i, j) connecting nodes with $i \leq N_{\mathrm{u}}$ and $j \leq N_{\mathrm{u}}$.

Inference problem formulation

The inference problem for localisation is formulated expressing each measurement as a function of the unknown locations $\boldsymbol{\theta}$ as:

$$\mathbf{y}_{ij} = \mathbf{h}_{ij}(\boldsymbol{\theta}) + \mathbf{n}_{ij} \tag{2.32}$$

with function $\mathbf{h}_{ij}(\boldsymbol{\theta})$ relating the RSS observations $\mathbf{y}_{ij}$ to the node locations through the distances $D_{ij} = |\theta_i - \theta_j|$ according to the stochastic bivariate model proposed in [4], and $\mathbf{n}_{ij} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{\mathrm{P}\Sigma})$. Note that the function $\mathbf{h}_{ij}(\boldsymbol{\theta})$ is non-linear in the unknowns $\boldsymbol{\theta}$, making the estimation problem non linear. The set of measurements available at node i is the collection $\mathbf{y}_i = [\mathbf{y}_{ij_1}^{\mathrm{T}} \cdots \mathbf{y}_{ij_{d_i}}^{\mathrm{T}}]^{\mathrm{T}}$ of all the observations taken over the links to the neighbouring nodes $\mathcal{N}_i = \{j_1, \ldots, j_{d_i}\}$: $\mathbf{y}_i = \mathbf{h}_i(\boldsymbol{\theta}) + \mathbf{n}_i$, with $\mathbf{h}_i(\boldsymbol{\theta}) = [\mathbf{h}_{ij_1}^{\mathrm{T}}(\boldsymbol{\theta}) \cdots \mathbf{h}_{ij_{d_i}}^{\mathrm{T}}(\boldsymbol{\theta})]^{\mathrm{T}}$, $\mathbf{n}_i = [\mathbf{n}_{ij_1}^{\mathrm{T}} \cdots \mathbf{n}_{ij_{d_i}}^{\mathrm{T}}]^{\mathrm{T}} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_i)$ and $\mathbf{Q}_i = \mathbf{I}_{d_i} \otimes \mathbf{Q}_{\mathrm{P}\Sigma}$. Below we discuss the extension of the estimation methods presented in D3.02 [4] to the non-linear localisation problem.

Estimation methods

The global non-linear ML estimate of the locations $\boldsymbol{\theta}$ based on all the available P2P measurements (2.32) is:

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{N} \|\mathbf{y}_i - \mathbf{h}_i(\boldsymbol{\theta})\|^2_{\mathbf{Q}_i^{-1}}\,, \tag{2.33}$$

with $\|\mathbf{y}_i - \mathbf{h}_i(\boldsymbol{\theta})\|^2_{\mathbf{Q}_i^{-1}} = (\mathbf{y}_i - \mathbf{h}_i(\boldsymbol{\theta}))^{\mathrm{T}} \mathbf{Q}_i^{-1} (\mathbf{y}_i - \mathbf{h}_i(\boldsymbol{\theta}))$. The solution is not available in closed form but it can be evaluated numerically by the Gauss-Newton algorithm [20] with estimate refinements through iterations $k = 0, 1, \ldots$ as:

$$\hat{\boldsymbol{\theta}}(k+1) = \hat{\boldsymbol{\theta}}(k) + \left(\sum_{j=1}^{N} \mathbf{H}_j^{\mathrm{T}}(k) \mathbf{Q}_j^{-1} \mathbf{H}_j(k) \right)^{-1} \sum_{i=1}^{N} \mathbf{H}_j^{\mathrm{T}}(k) \mathbf{Q}_i^{-1} \left(\mathbf{y}_i - \mathbf{h}_i(\hat{\boldsymbol{\theta}}(k)) \right), \tag{2.34}$$

and Jacobian matrix $\mathbf{H}_i(k) = \left.\frac{\partial \mathbf{h}_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_i(k)} \in \mathbb{R}^{2d_i \times 2N_{\mathrm{u}}}$. The elements of $\mathbf{H}_i(k)$ are obtained by evaluating the derivatives of $\mathbf{h}_{ij}(\boldsymbol{\theta})$ with respect to the nth node location $\theta_n = [\theta_{xn}\ \theta_{yn}]^{\mathrm{T}}$, according to the stochastic bivariate model proposed in the D3.02 [4].

For distributed implementation of (2.33), we propose to employ the D-MLE-AE consensus-based method, adapting the processing to handle the non-linear nature of the estimation problem. We apply the consensus update rule to the linearised localisation model (2.32), at each node and for each consensus step, interleaving the iterations of the Gauss-Newton algorithm with those for consensus. Information exchange guarantees that every node can retrieve the positions of all nodes regardless of the connectivity. Specifically, at each iteration k, node i linearises the model $\mathbf{h}_i(\boldsymbol{\theta})$ around the previous step estimate $\hat{\boldsymbol{\theta}}_{i,\mathrm{D-MLE-AE}}(k)$, then it finds the ML solution to the linear problem and applies a consensus step based on information provided by neighbours. The iterative implementation can be summarised as follows:

$$\begin{aligned}
&\text{Linearization:} && \mathbf{H}_i(k) = \tfrac{\partial \mathbf{h}_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{i_{\mathrm{D-MLE-AE}}}(k)};\ \ \Delta\mathbf{y}_i(k) = \mathbf{y}_i - \mathbf{h}_i(\hat{\boldsymbol{\theta}}_{i_{\mathrm{D-MLE-AE}}}(k)) \\
&\text{Local estimate:} && \Delta\hat{\boldsymbol{\theta}}_i(k) = \left(\mathbf{H}_i^{\mathrm{T}}(k) \mathbf{Q}_i^{-1} \mathbf{H}_i(k) \right)^{-1} \mathbf{H}_i^{\mathrm{T}}(k) \mathbf{Q}_i^{-1} \Delta\mathbf{y}_i(k) \\
&\text{ML update:} && \hat{\boldsymbol{\theta}}_i(k) = \hat{\boldsymbol{\theta}}_{i_{\mathrm{D-MLE-AE}}}(k) + \Delta\hat{\boldsymbol{\theta}}_i(k) \\
&\text{Consensus:} && \hat{\boldsymbol{\theta}}_{i_{\mathrm{D-MLE-AE}}}(k+1) = \hat{\boldsymbol{\theta}}_{i_{\mathrm{D-MLE-AE}}}(k) + \varepsilon \textstyle\sum_{j \in \mathcal{N}_i} \mathbf{W}_j \left(\hat{\boldsymbol{\theta}}_j(k) - \hat{\boldsymbol{\theta}}_i(k) \right)
\end{aligned} \tag{2.35}$$

with weighting matrix proportional to the inverse of covariance

$$\mathbf{W}_j = \boldsymbol{\Gamma} \left(\mathbf{H}_j^{\mathrm{T}}(k) \mathbf{Q}_j^{-1} \mathbf{H}_j(k) \right),$$

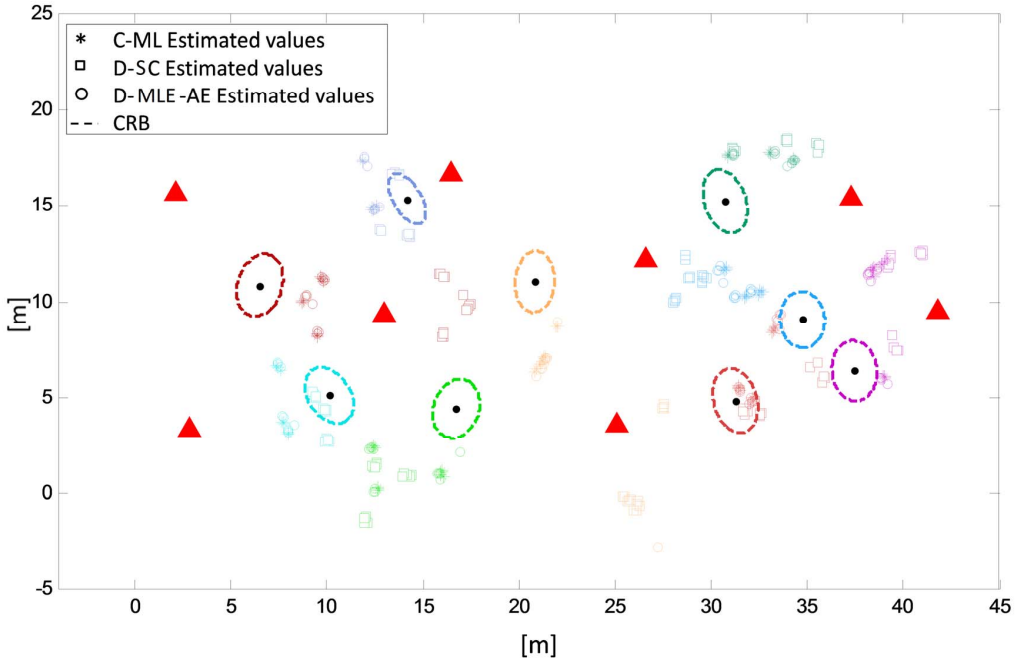


Figure 2.19: Scatter plot of the location estimates at each N_{u} nodes (different colours) for all the network-localisation algorithms: centralised (asterisk), D-SC (square) and D-MLE-AE estimates, compared to the CRB (black ellipse contour).

where $\mathbf{\Gamma}$ is a scaling factor [4] and initialisation to the local ML estimate at node i, i.e. $\hat{\boldsymbol{\theta}}_{i,\mathrm{D-MLE-AE}}(0) = \arg\min_{\boldsymbol{\theta}} \|\mathbf{y}_i - \mathbf{h}_i(\boldsymbol{\theta})\|^2_{\mathbf{Q}_i^{-1}}$. Note that in case of under-determined, the local estimate in (2.35) is replaced by the conventional pseudo-inverse solution for under-determined systems, but the D-MLE-AE is suitable also in this case. The exchange of $\mathbf{W}_j(k)$ can be limited to early steps and then reduced without any meaningful performance loss at latest iterations.

Experimental performance assessment

The performance of distributed network localisation is evaluated for the experimental network in Figure 2.17, with $N_{\mathrm{u}} = 9$ unknown nodes and $N_{\mathrm{a}} = 8$ anchors. The environment-dependent parameters have been chosen using the centralised estimates of the previous subsection in order to make a comparison between all the network-localisation algorithms. The step size has been set to $\varepsilon = 0.98/\Delta$ and $\varepsilon = 0.98$ respectively for AC and D-MLE-AE methods. Moreover, we consider the $M = 15$ independent measurements performed at each link for the estimation of the nodes' positions. In Figure 2.19, the scatter plot of the 15 location estimates for all the unknown nodes are evaluated for the network-localisation algorithms, i.e. the centralised approach and the distributed SC and D-MLE-AE methods, while the Cramer-Rao Bound (CRB) is shown as

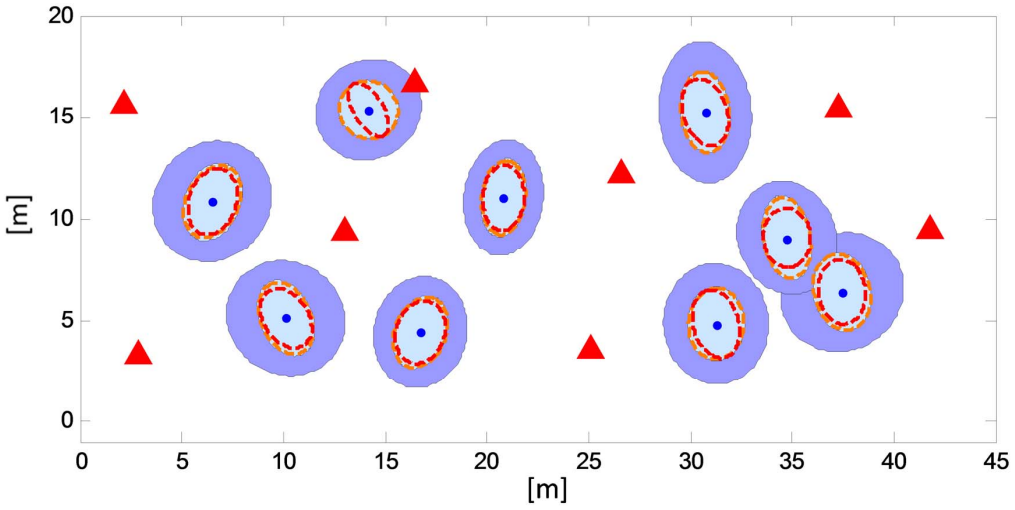


Figure 2.20: Performance in term of localisation error of network-localisation algorithms: D-SC (blue), D-MLE-AE (light blue), centralised estimate (orange contour) and CRB (red contour).

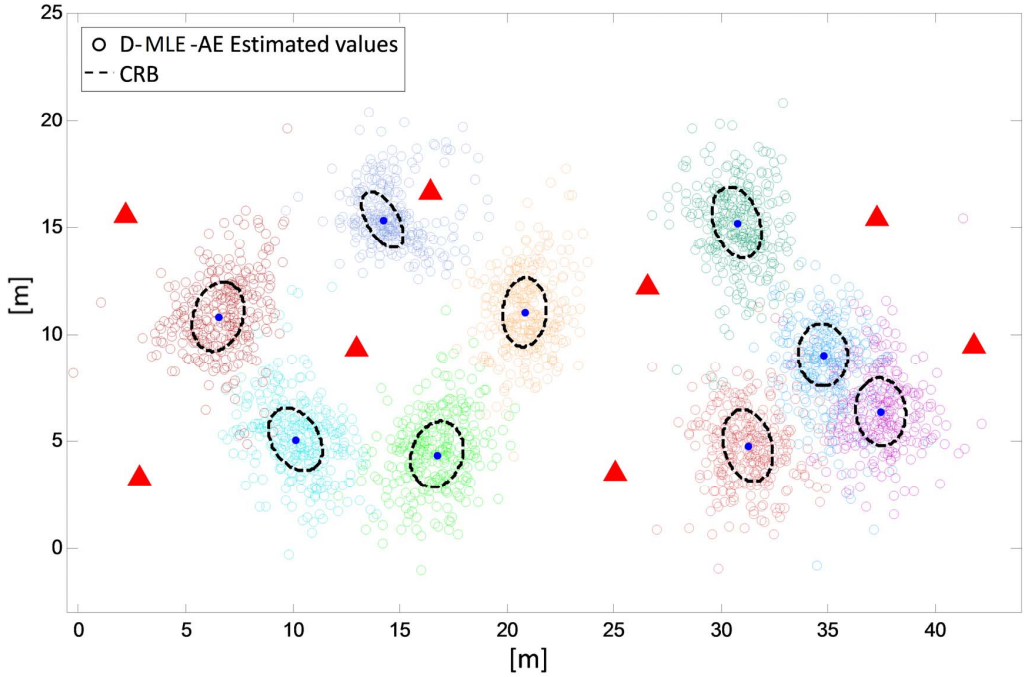


Figure 2.21: Scatter plot of the location estimates for D-MLE-AE compared to the CRB (black contour).

benchmark. It can be observed that the D-MLE-AE algorithm approaches the centralised ML estimate closer to the true positions of the nodes, while the D-SC method is quite far from it.

The performances of consensus algorithms are here analysed for a simulated scenario, designed as the experimental network in Figure 2.17 and with the knowledge of the environment-dependent parameters, previously estimated. In Figure 2.20, the performances are evaluated in terms of mean square error (MSE) of the estimate by averaging over 300 measurements. All the methods are compared by plotting – for each method and each unknown node – the $1\times$ standard deviation (1σ) error ellipse associated to the 2×2 MSE matrix $\mathrm{E}[(\hat{\theta}_i - \theta_i)(\hat{\theta}_i - \theta_i)^{\mathrm{T}}]$, which corresponds to 39level (under the Gaussian approximation). The D-MLE-AE method (light blue ellipse) is shown to outperform the D-SC method (blue ellipse) and to closely approach the centralised ML estimate (orange ellipse contour), while the CRB is shown as a reference (red ellipse contour). The difference between the CRB and the centralised ML is due to the non-linearity of the model (the bound is attained only asymptotically). The difference between the CRB and the centralised ML is due to the non-linearity of the model (the bound is attained only asymptotically). Moreover, in Figure 2.21, the scatter plot of the estimated values of each unknown node is illustrated for the D-MLE-AE method (coloured circle) and it is compared to the corresponding CRB (black ellypse contour).

2.3.4 Conclusions

In this section we evaluated the performance of different algorithms, introduced in D3.02 [4] and based on consensus approach, for distributed estimation in two application scenarios. In particular, we focused on two weighted consensus-based approaches (D-MLE and D-MLE-AE) for distributed identification of channel model parameters and distributed cloud network localisation with exchange of position accuracy information between nodes. To handle the non-linearity of the measurement model, an iterative Gauss-Newton algorithm has been embedded into the consensus procedure. The D-MLE method has been shown to reach the same performance of an equivalent centralised estimation. The D-MLE-AE method has been shown to closely attain the fundamental limit, and to provide meaningful performance gains in terms of convergence speed with respect to D-MLE, with moderate increase of information exchange between nodes. Furthermore, it turned out to be suited in critical scenarios with limited connectivity and underdetermined settings. The D-MLE-AE algorithm confirmed to be as the most suitable method for demonstration scenarios as implemented in D5.52 [2].

2.4 Distributed interference sensing and coordination for cloud scheduling

In this section we consider the problem of distributed spectrum sensing in multiple self-organising cloud networks sharing the same time-frequency resources [21],[22]. Each of the networks allocates autonomously radio resources so as to minimise mutual interference. Interference sensing is part of this cognitive framework where sensing devices,

or secondary users (SUs), exchange local estimates to cooperatively recognise and track the overall time-varying interference patterns caused by primary users (PUs). PUs are assumed to perform periodic transmission over pre-defined (but unknown to SUs) time-frequency hopped resources. Detection by the SUs is based on local processing and iterated exchanges of local decision with neighbours, so as to enable global fusion of sensed data as for an equivalent centralised approach. We propose a weighted-average consensus algorithm nested within a decision-directed procedure for distributed Bayesian detection of the PU spectrum occupancy. The distributed approach provides the estimate of the complete interference pattern to each SU regardless of the incomplete visibility at each node. Performance analysis is carried out both on simulated and real scenarios with mixed coexisting Wi-Fi and ZigBee devices, as a basis for the implementation into the CIMC demonstrator (WP5).

2.4.1 Introduction

Self-organised biological systems are characterised by individuals that make simple local decisions, while the exchange of those decisions collectively produces a global consensus on a complex behaviour. Recently, several studies have begun to apply bio-inspired techniques to wireless networks, focusing on synchronisation, content distribution, security, spectrum sharing and sensing [23]. The use of spectrum sensing is particularly crucial when multiple (and possibly heterogeneous) networks deployed in close proximity need to self-coordinate their access to a shared spectrum so as to avoid collisions.

A simplified but still relevant setting is considered in Figure 2.22 where two coexisting networks mutually sense their radio-frequency (RF) activity and desynchronise [24] the RF access to one another in order to avoid interference. The figure shows a single sensing stage, where, according to the cognitive jargon, the sensing network is referred to as secondary network while the monitored one is the primary network (roles interchange over time in the dynamic resource optimisation process). The users of the secondary network (SUs) act as interconnected sensing devices that cooperatively detect the time-frequency hopped resources of the primary users (PUs) and allocate their own resources so as to avoid cross-interference, e.g. accessing the spectrum holes of the primary network. Cooperation is crucial to detect the complete interference pattern of the primary network, as each SU has a partial visibility due to limited sensing range, mobility or shadowing/fading problems (as shown in the data sensed by SU_1 where RF activities of PU_3 and PU_4 are hidden).

In centralised approach to cooperative spectrum sensing [25]-[26], the SUs sense the spectrum individually, facing different channel conditions, and then transmit their sensed data to a fusion centre that makes the final decision about the spectrum occupancy based on all the observations. However, this can be very difficult under practical communication constraints [25], due to the extensive data to be propagated to the fusion centre over the dynamic steps of resource allocation. Focus of this section is on consensus algorithms

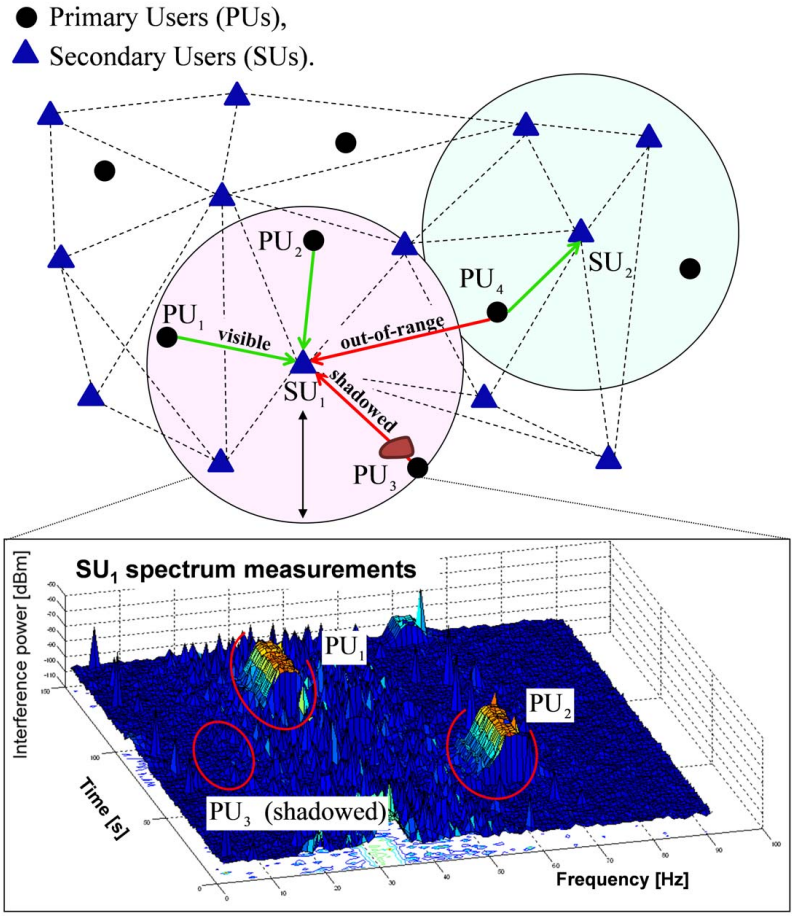


Figure 2.22: Cooperative spectrum sensing framework: PU_1 and PU_2 are sensed by SU_1 both in its radio coverage area (violet circle), whereas PU_3 (shadowed) and PU_4 (out of range) are hidden to SU_1. Below figure: example of RSS measurements captured by SU_1 over the time-frequency grid. Cooperation is pursued over the SU peer-to-peer links (dashed lines), to enable interference pattern detection, regardless of the limited visibility.

[17] that allow a distributed reconstruction of the time-varying pattern of the primary-network activity, without any central coordination, by exchange of local information among neighbouring SUs, i.e. within the coverage range. Energy detection is chosen as sensing technique, since it requires only received signal strength (RSS) measurements without any need of a-priori knowledge about the PU activity [27]. Differently from existing distributed algorithms that require local sharing of multiple spectrum measurements and are quite costly in term of communication overhead [28]-[29], we propose a novel weighted-average consensus algorithm where SUs exchange compact detection information instead of raw RSS data. Consensus is combined with an iterative decision-directed (DD) procedure for estimation of key PU interference parameters and for detection of the time-frequency activity pattern. In this section, weighting is designed so as

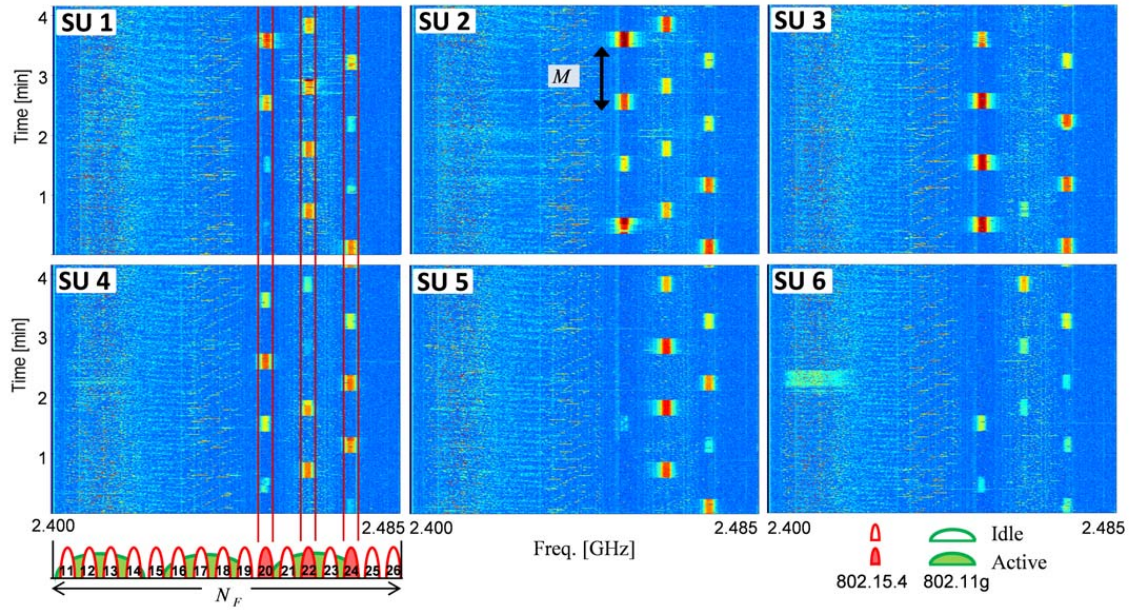


Figure 2.23: Spectrum sensing by 6 IEEE 802.15.4 devices: RSS measurements collected by each SU. Primary users perform periodic transmission tasks over the IEEE 802.15.4 channels 20, 22 and 24. Three IEEE 802.11g devices are also interfering in the same 2.4 GHz band.

to reach the same performance as for the centralised solution in few iterations and with a reduced amount of information exchange. Convergence is eased in dense networks thanks to the high degree of connectivity. The method is validated by experimental tests in a heterogeneous scenario with IEEE 802.15.4 and IEEE 802.11 devices.

The work is organised as follows. Signal modelling and background on Bayesian testing for spectrum sensing are in Sections 2.4.2 and 2.4.3. The DD method for estimation of the key interference parameters is proposed in Section IV and employed for cooperative detection in Section 2.4.5. Performance analysis is in Section 2.4.6, followed by the validation on real scenario in Section 2.4.7 and the concluding remarks in Section 2.4.8.

2.4.2 Spectrum sensing model (SSM)

We consider the coexistence of a primary and a secondary network, as depicted in Figure 2.22, see D5.52 [2] for applications to CIMC. The PUs transmit data over a subset of the available N_F frequencies, using a time-slotted transmission with slots repeated every M time samples, M denoting the frame duration (repetition period). A set of N_S SUs sense their transmission by reading the RSS over L periods, collecting an overall number of $N_F N_T$ RSS samples at each device, where $N_T = M \cdot L$ denotes the number of RSS samples per frequency.

In Figure 2.23, the occupation of the frequency-time resources by the primary network is exemplified, using experimental RSS data collected by $N_S = 6$ SUs for $N_F = 16$ channels and $L = 4$ time periods. In the example, three PUs are transmitting and hopping over three different frequencies. As illustrated in Figure 2.22, due to the limited sensing range of each SU device and dynamic fading/shadowing caused by mobility, each PU transmission is overheard by a subset of the SUs fragmented in time, leading to the incomplete observation of the transmission pattern as shown in Figure 2.23. The N_S SUs must engage in a cooperative decision process to detect the complete time-frequency mask and schedule their resources over the unused portion of the spectrum. Crucial is that every SU should converge to the same pattern of used resources.

Let $x_i(t, f)$ be the RSS, in logarithmic scale, of the signal sensed by SU i at time t on frequency f, with $t = 1, \ldots, N_T$, $f = 1, \ldots, N_F$ and $i = 1, \ldots, N_S$. The signal includes the interference of an active PU with probability P. This probability can be written as $P = P_a P_v$, where P_a is the probability that a PU is active in the considered time-frequency resource and P_v is the probability that it is visible to the SU, i.e. in the coverage area of the SU. According to the widely adopted lognormal power model, the RSS sample $x_i(t, f)$ can be approximated by a Gaussian random variable whose parameters depend on the absence or presence of the PU signal. These two disjoint hypotheses are denoted as $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. Note that, even if a PU is active, its signal can be observed by the SU in some periods and be hidden in others due to time-varying fading/shadowing conditions. The RSS sample $x = x_i(t, f)$ is:

$$x \sim \begin{cases} \mathcal{N}(\mu_1, \sigma_1^2), & \text{hypothesis } \mathcal{H}_1, \text{ with prob. } P \\ \mathcal{N}(\mu_0, \sigma_0^2), & \text{hypothesis } \mathcal{H}_0, \text{ with prob. } 1 - P \end{cases}. \tag{2.36}$$

Under the hypothesis $\mathcal{H}_0$, x models the power of the background noise at the SU receiver, Gaussian distributed with mean μ_0 and variance σ_0^2, with randomness due to measurement errors at the receiver equipment. On the other hand, under the hypothesis $\mathcal{H}_1$, x is the power of the PU signal measured at the SU, modelled as Gaussian with larger mean $\mu_1 > \mu_0$ (due to the interference) and variance $\sigma_1^2 > \sigma_0^2$ (due to shadowing).

The aim of this work is to detect the time-frequency resources (t, f) that are used by any user of the primary network and provide the related time-frequency transmission mask.

2.4.3 MAP testing for PU detection

Bayesian detection of the PU signal in the generic RSS sample x based on the maximum a posteriori (MAP) criterion is known to yield the likelihood ratio test (LRT) [30]:

$$\mathcal{L}(x) = \frac{p(x|\mathcal{H}_1)}{p(x|\mathcal{H}_0)} = \frac{G(x; \mu_1, \sigma_1^2)}{G(x; \mu_0, \sigma_0^2)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{1 - P}{P}. \tag{2.37}$$

For the specific spectrum sensing problem, using the model (2.36) with parameters $\boldsymbol{\theta} = [\mu_1, \mu_0, \sigma_1^2, \sigma_0^2, P]^{\mathrm{T}}$, the probability density function (pdf) of x is $p(x|\mathcal{H}_0) = G(x; \mu_0, \sigma_0^2)$

under the hypothesis $\mathcal{H}_0$, and $p(x|\mathcal{H}_1) = G(x;\mu_1,\sigma_1^2)$ under the hypothesis $\mathcal{H}_1$, where $G(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-(x-\mu)^2/2\sigma^2)$. The LRT can be implemented by the threshold detection:

$$x \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} S(\boldsymbol{\theta}), \tag{2.38}$$

with threshold $S = S(\boldsymbol{\theta})$ obtained as solution of:

$$\frac{G(S;\mu_1,\sigma_1^2)}{G(S;\mu_0,\sigma_0^2)} = \frac{1-P}{P}. \tag{2.39}$$

This leads to the quadratic equation $aS^2 + 2bS + c = 0$, with parameters $a = \sigma_1^2 - \sigma_0^2$, $b = \mu_1\sigma_0^2 - \mu_0\sigma_1^2$, $c = \mu_0^2\sigma_1^2 - \mu_1^2\sigma_0^2 - 2\sigma_1^2\sigma_0^2\ln(\frac{\sigma_1}{\sigma_0}\cdot\frac{1-P}{P})$, whose solution provides the threshold S.

Threshold based detection (2.38) can be easily applied once the model parameters $\boldsymbol{\theta}$ are known. In the following, we propose a cooperative approach for inferring $\boldsymbol{\theta}$ based on the sensed RSS data (Section 2.4.4) and estimating the interference pattern (Section 2.4.5).

2.4.4 Estimation of SSM parameters

In this subsection, an iterative DD method is proposed for estimation of the interference parameters $\boldsymbol{\theta}$. First, a single-node method is introduced as building block of the proposed methodology: the DD procedure is used separately by each SU i to obtain a local estimate $\hat{\boldsymbol{\theta}}_i$ from the local RSS dataset $\mathcal{X}_i = \{x_i(t,f) : t = 1,\dots,N_T, f = 1,\dots,N_F\}$, see examples of datasets in Figure 2.23. The method is then extended to a cooperative framework: the centralised solution provides a global estimate $\hat{\boldsymbol{\theta}}$ based on the complete dataset $\mathcal{X} \equiv \cup_{i=1}^{N_S}\{\mathcal{X}_i\}$, while the new method in Section 2.4.4-C allows to achieve the same result distributively by combining single-node DD processing with consensus iterations.

Non-cooperative (single-node) estimation

The parameters $\boldsymbol{\theta}$ can be estimated iteratively at the ith SU based on the local dataset $\mathcal{X}_i$ according to the DD method [31] here tailored to the model (2.36). Assume that $\hat{S}_i^{(k)} = S(\hat{\boldsymbol{\theta}}_i^{(k)})$ is the threshold obtained from the parameters' estimate $\widehat{\boldsymbol{\theta}}_i^{(k)} = [\hat{\mu}_{1,i}^{(k)}, \hat{\mu}_{0,i}^{(k)}, \hat{\sigma}_{1,i}^{2(k)}, \hat{\sigma}_{0,i}^{2(k)}, \hat{P}_i^{(k)}]^{\mathrm{T}}$ at kth iteration. Using the LRT (2.38) with $S = \hat{S}_i^{(k)}$, the RSS samples in the ith user dataset $\mathcal{X}_i$ can be partitioned into two subsets associated to the hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, $\mathcal{X}_{\mathcal{H}_1,i}^{(k)} = \{x \in \mathcal{X}_i : x \geq \hat{S}_i^{(k)}\}$ and $\mathcal{X}_{\mathcal{H}_0,i}^{(k)} = \mathcal{X}_i \setminus \mathcal{X}_{\mathcal{H}_1,i}^{(k)}$, respectively. The new parameters for the iteration $k+1$ are then obtained by computing

the sample means, variances and frequencies for the subsets $\mathcal{X}_{\mathcal{H}_1,i}^{(k)}$,$\mathcal{X}_{\mathcal{H}_0,i}^{(k)}$ as:

$$\begin{aligned}
\hat{\mu}_{1,i}^{(k+1)} &= \tfrac{1}{N_1^{(k)}} \textstyle\sum_{x\in\mathcal{X}_{\mathcal{H}_1,i}^{(k)}} x \\
\hat{\mu}_{0,i}^{(k+1)} &= \tfrac{1}{N_0^{(k)}} \textstyle\sum_{x\in\mathcal{X}_{\mathcal{H}_0,i}^{(k)}} x \\
\hat{\sigma}_{1,i}^{2(k+1)} &= \tfrac{1}{N_1^{(k)}} \textstyle\sum_{x\in\mathcal{X}_{\mathcal{H}_1,i}^{(k)}} \left(x-\hat{\mu}_{1,i}^{(k)}\right)^2 \\
\hat{\sigma}_{0,i}^{2(k+1)} &= \tfrac{1}{N_0^{(k)}} \textstyle\sum_{x\in\mathcal{X}_{\mathcal{H}_0,i}^{(k)}} \left(x-\hat{\mu}_{0,i}^{(k)}\right)^2 \\
\hat{P}_i^{(k+1)} &= \tfrac{N_1^{(k)}}{N_0^{(k)}+N_1^{(k)}}
\end{aligned} \tag{2.40}$$

where $N_1^{(k)} = |\mathcal{X}_{\mathcal{H}_1,i}^{(k)}|$ and $N_0^{(k)} = |\mathcal{X}_{\mathcal{H}_0,i}^{(k)}|$ are the cardinalities of the two subsets. In compact form, we rewrite (2.40) as:

$$\hat{\boldsymbol{\theta}}_i^{(k+1)} = \mathbf{g}(\mathcal{X}_i, \hat{S}_i^{(k)}, \hat{\boldsymbol{\theta}}_i^{(k)}), \tag{2.41}$$

with functions $\mathbf{g}(\,\cdot\,) = [g_1(\,\cdot\,), \cdots, g_5(\,\cdot\,)]^T$ defined according to the entries of (2.40). A new threshold $\hat{S}_i^{(k+1)} = S(\hat{\boldsymbol{\theta}}_i^{(k+1)})$ is calculated from (2.39) based on the new parameters and the process is repeated. The algorithm converges to an estimate $\hat{\boldsymbol{\theta}}_i^{(k)} = \hat{\boldsymbol{\theta}}_i$ when $\hat{P}_i^{(k)} = \hat{P}_i^{(k+1)} = \hat{P}_i^{(\infty)}$.

Centralised cooperative estimation

In a centralised cooperative approach, each SU is required to transmit the RSS observations to a fusion centre that can apply the iterative DD procedure (2.41) to the complete dataset $\mathcal{X}$ instead of the local one $\mathcal{X}_i$, so that:

$$\hat{\boldsymbol{\theta}}^{(k+1)} = \mathbf{g}(\mathcal{X}, \hat{S}^{(k)}, \hat{\boldsymbol{\theta}}^{(k)}). \tag{2.42}$$

The elements of the vectorial function $\mathbf{g}(\,\cdot\,)$ are defined as in Section 2.4.4 with $\mathcal{X}$ divided into the two subsets $\mathcal{X}_{\mathcal{H}_1}^{(k)} = \{x \in \mathcal{X} : x \geq \hat{S}^{(k)}\}$ and $\mathcal{X}_{\mathcal{H}_0}^{(k)} = \mathcal{X} \setminus \mathcal{X}_{\mathcal{H}_1}^{(k)}$ according to the global threshold $\hat{S}^{(k)} = S(\hat{\boldsymbol{\theta}}^{(k)})$ computed from the parameter estimate $\hat{\theta}^{(k)}$ of the previous iteration. The method is expected to provide a more accurate estimate of the SSM model as it combines all data from the N_S SUs.

Distributed cooperative estimation

In distributed spectrum sensing, data fusion is carried out in a fully decentralised way sharing information through the bidirectional peer-to-peer links of the SU network. This is modelled as an undirected graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with vertices $\mathcal{V} = \{1, \ldots, N_S\}$ representing the SUs and edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ representing the links. The set of neighbours for SU i is denoted as $\mathcal{N}_i = \{j|(j,i) \in \mathcal{E}\}$, the related node degree as $d_i = |\mathcal{N}_i|$ and the

maximum degree as $\Delta = \max_i d_i$. We denote as $\mathbf{A} = [a_{ij}]$ the $N_S \times N_S$ symmetric adjacency matrix that models the SU network connectivity, with $a_{ij} = 1$ if $(i,j) \in \mathcal{E}$ (i.e. if node j communicates with node i) and $a_{ij} = 0$ for any $(i,j) \notin \mathcal{E}$. The Laplacian matrix of the graph is $\mathbf{L} = \mathbf{D} - \mathbf{A}$, $\mathbf{D} = \text{diag}(d_1, \ldots, d_{N_S})$ being the degree matrix of $\mathcal{G}$.

Recalling the definition of $\mathbf{g}(\,\cdot\,)$ in (2.40)–(2.41), we observe that the centralised estimate $\mathbf{g}(\mathcal{X}, \hat{S}^{(k)}, \hat{\boldsymbol{\theta}}^{(k)})$ of the parameters in (2.42) can be seen as a weighted average of the N_S local estimates $\hat{\boldsymbol{\theta}}_i^{(k+1)} = \mathbf{g}(\mathcal{X}_i, \hat{S}^{(k)}, \hat{\theta}^{(k)})$ computed at the SU $i = 1, \ldots, N_S$ by averaging the RSS samples of the local dataset partitioned as $\mathcal{X}_i = \mathcal{X}_{\mathcal{H}_1,i}^{(k)} \cup \mathcal{X}_{\mathcal{H}_0,i}^{(k)}$ according to the global threshold $\hat{S}^{(k)}$. Namely, let θ be the generic element of the vector $\boldsymbol{\theta}$, the centralised estimate (2.42) can be expressed as:

$$\hat{\theta}^{(k+1)} = \frac{1}{\sum_{i=1}^{N_S} W_i^{(k)}} \sum\nolimits_{i=1}^{N_S} W_i^{(k)} \hat{\theta}_i^{(k+1)}, \tag{2.43}$$

with weights accounting for the different sample size at each SU, i.e. $W_i^{(k)} = \hat{P}_i^{(k)} = |\mathcal{X}_{\mathcal{H}_1,i}^{(k)}| / |\mathcal{X}_i|$ for θ representing any of the moments $\{\mu_1, \mu_0, \sigma_1^2, \sigma_0^2\}$ and $W_i^{(k)} = 1$ for $\theta = P$. This highlights that the global SSM estimate (2.42) can be computed in a distributed way through a weighted-average consensus approach [18], which is known to converge to the weighted average of the initial local estimates.

We thus propose to implement the DD procedure for threshold computation distributively, performing at each iteration the local computations (2.41) and then sharing information with neighbours through consensus. Remarkably, iterations for LRT are interleaved with those for consensus to favour the convergence. Namely, let $\hat{\boldsymbol{\theta}}_i^{(k+1)}(0) = \hat{\boldsymbol{\theta}}_i^{(k+1)}$ collect the parameter estimates obtained by node i from (2.41) at iteration $k+1$, a number of consensus iterations indexed as $q = 1, 2, \ldots, Q$ are performed before the next LRT iteration, according to the weighted average-consensus algorithm [18]:

$$\hat{\boldsymbol{\theta}}_i^{(k+1)}(q+1) = \hat{\boldsymbol{\theta}}_i^{(k+1)}(q) + \varepsilon \mathbf{W}_i^{(k)^{-1}} \sum_{j \in \mathcal{N}_i} (\hat{\boldsymbol{\theta}}_j^{(k+1)}(q) - \hat{\boldsymbol{\theta}}_i^{(k+1)}(q)), \tag{2.44}$$

with step size ε and 5×5 weighting matrix $\mathbf{W}_i^{(k)}$. As proved in [18], for $0 < \varepsilon < 2/\lambda_{\max}(\mathbf{W}^{(k)^{-1}} \widetilde{\mathbf{L}})$, with $\mathbf{W}^{(k)} = \text{blockdiag}(\mathbf{W}_1^{(k)}, \ldots, \mathbf{W}_{N_S}^{(k)})$ and $\widetilde{\mathbf{L}} = \mathbf{L} \otimes \mathbf{I}_5$, the algorithm converges to the weighted average of the local estimates:

$$\hat{\boldsymbol{\theta}}_i^{(k+1)}(\infty) = \left(\sum\nolimits_{j=1}^{N_S} \mathbf{W}_j^{(k)} \right)^{-1} \sum\nolimits_{i=1}^{N_S} \mathbf{W}_i^{(k)} \hat{\boldsymbol{\theta}}_i(0). \tag{2.45}$$

Taking into account (2.43), we set the weighting as $\mathbf{W}_i^{(k)} = \text{diag}(W_i^{(k)}, W_i^{(k)}, W_i^{(k)}, W_i^{(k)}, 1)$ so that (2.45) equals (2.42) and the distributed algorithm converges to the centralised estimate.

After Q consensus iterations, the five parameters $\hat{\boldsymbol{\theta}}_i^{(k+1)}(Q)$ are used to compute the new threshold $\hat{S}^{(k+1)}$, update the subsets $\mathcal{X}_{\mathcal{H}_1,i}^{(k+1)}$ and $\mathcal{X}_{\mathcal{H}_0,i}^{(k+1)}$, and repeat (2.41) and (2.44)

till convergence when $\hat{P}_i^{(k)} = \hat{P}_i^{(k+1)} = \hat{P}_i^{(\infty)}$. The parameter estimates at convergence will be denoted as $\hat{\theta} = \left[\hat{\mu}_1, \hat{\mu}_0, \hat{\sigma}_1^2, \hat{\sigma}_0^2, \hat{P}\right]^T$ and the threshold as $\hat{S}$.

2.4.5 Cooperative detection of interference

Aim of cooperative detection is the estimation of the $M \times N_F$ binary mask of the primary network, $\mathbf{B} = [b(m, f)]$, with elements defined as $b(m, f) = 1$ if any PU is transmitting on the time-frequency resource (m, f), and $b(m, f) = 0$ if no PU is allocated on that resource, $m = 1, \ldots, M$, $f = 1, \ldots, N_F$.

In both the centralised and distributed approaches, the iterative procedure in the previous subsection provides a threshold $\hat{S}$ and the related classification of the L-period dataset $\mathcal{X}_i$, for each user i, into the two classes $\mathcal{H}_0$ and $\mathcal{H}_1$. The result is a $LM \times N_F$ mask estimate $\mathbf{C}_i = [c_i(t, f)]$ with elements $c_i(t, f) = \mathrm{dec}_{\hat{S}}(x_i(t, f))$ and binary decision function defined as: $\mathrm{dec}_{\hat{S}}(x) = 1$ for $x \geq \hat{S}$ and $\mathrm{dec}_{\hat{S}}(x) = 0$ otherwise. However, detection performed on the dataset $\mathcal{X}_i$, even though based on the global threshold estimate, allows to sense only the PUs that are active in the area of SU i. For detection of the *overall spectrum mask*, we propose to proceed as follows.

In the *centralised approach* (Section IV-B), where the N_S masks $\mathbf{C}_i$ for $i = 1, \ldots, N_S$ are jointly available at the fusion center, cooperative decision on each time-frequency resource (m, f) is taken by evaluating the number of detected transmissions over the total number $N_S L$ (N_S users and L periods). The result is the soft-valued mask $\mathbf{Z} = [z(m, f)] \in \mathbb{R}^{M \times N_F}$ with elements $z(m, f) \in [0, 1]$ given by:

$$z(m, f) = \frac{1}{N_S L} \sum\nolimits_{i=1}^{N_S} \sum\nolimits_{\ell=1}^{L} c_i\left((\ell - 1)M + m, f\right). \tag{2.46}$$

Asymptotically, we should have $z(m, f) \to 0$ if the time-frequency resource (m, f) is free and $z(m, f) \to P_v$ if it is occupied by a PU, where P_v is the fraction of the observations belonging to $\mathcal{H}_1$. The final decision is thus obtained by comparing $z(m, f)$ with a threshold λ, as $\hat{b}(m, f) = \mathrm{dec}_{\lambda}(z(m, f))$, with λ selected based on the value of P_v. The estimated mask is $\hat{\mathbf{B}} = [\hat{b}(m, f)]$. If no a-priori information is available, we set $\lambda = 0.5$.

In the *distributed method* (Section IV-C), since each SU i can only access the local mask $\mathbf{C}_i$, it first computes the average of the local samples $c_i(t, f)$ over the periods; then it exchanges the average mask with neighbours and updates it iteratively by an average-consensus procedure similar to (2.44), with $\hat{\theta}$ replaced by the mask estimates $\mathbf{Z}_i$ and identity matrices as weights, till convergence to the global average (2.46) is reached.

2.4.6 Performance analysis

In this subsection, the performance of the proposed distributed detection method is compared to those of the centralised cooperative method and the non-cooperative (single-

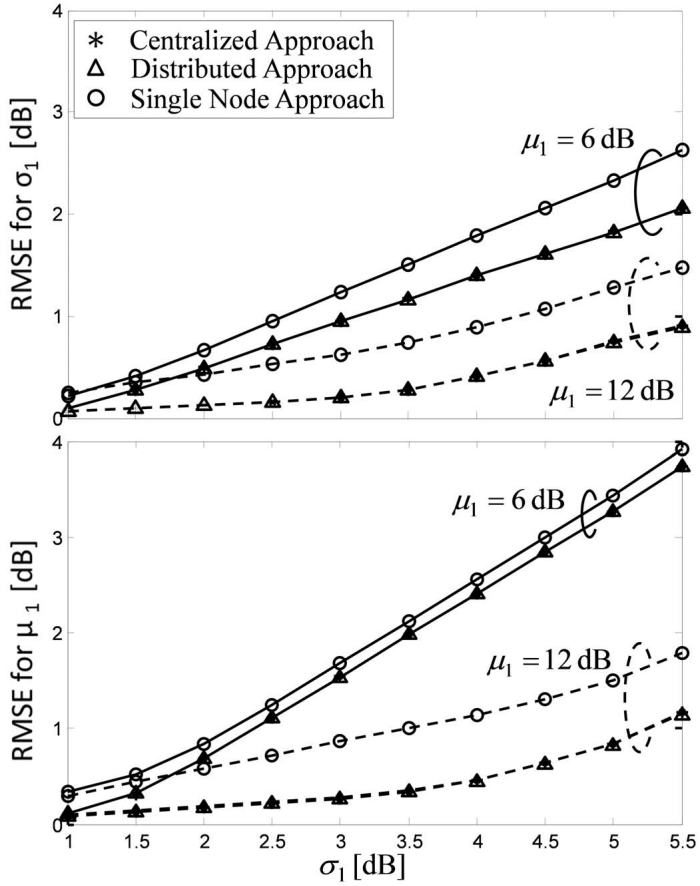


Figure 2.24: RMSE of the estimate for the interference parameters σ_1 (top) and μ_1 (bottom) vs. σ_1. RSS data is simulated using $\mu_1 = 6$ dB (solid line) and $\mu_1 = 12$ dB (dashed line), $\mu_0 = 0$ dB and $\sigma_0 = 1.5$ dB.

node) method considered as benchmarks. We analyse a scenario with a strongly connected network of $N_S = 10$ SUs that cooperatively sense the transmission of 3 PUs. The PUs transmit data over 3 of the $N_F = 15$ available carrier frequencies ($P_a = 0.2$). Each frame is composed of $M = 1$ sample repeated for $L = 20$ periods, for an overall number of $N_T = ML = 20$ samples. The SUs are assumed to have synchronised clocks during the sensing process and probability $P_v = 0.3$ of visibility of any transmitting PU. The N_T RSS samples collected by each SU are modelled as in (2.36) with parameters: $\mu_0 = 0$ dB and $\sigma_0 = 1.5$ dB for the background noise; $\mu_1 = \{6, 8, 12\}$ dB and σ_1 ranging in the interval $[1, 5.5]$ dB for the interference. For spectrum detection, the LRT algorithm is initialised using $P^{(0)} = 0.5$, $\mu_1^{(0)} = \mu_0^{(0)}$ equal to the sample mean of the available dataset ($\mathcal{X}$ for the centralised approach and $\mathcal{X}_i$ for the distributed one), $\sigma_1^{(0)}$ equal to the sample standard deviation of the same dataset, and $\sigma_0^{(0)}$ randomly distributed in the interval $[0.5, 2.5]$ dB. The interference binary mask is estimated as in Section 2.4.5 with threshold $\lambda = 0.5P_v$.

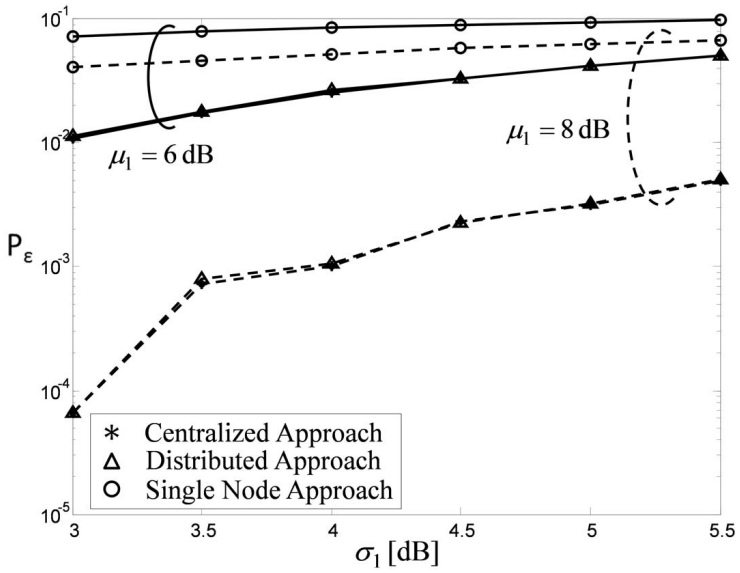


Figure 2.25: Probability of mis-classification P_ε vs. σ_1 for $\mu_1 = 6$ dB (solid line) and $\mu_1 = 8$ dB (dashed line).

Performances are evaluated in Figure 2.24 in terms of root mean square error (RMSE) of the SSM parameter estimates for the interference parameters σ_1 and μ_1. The RMSE for the noise parameters σ_0 and μ_0 is negligible. Convergence rate of the distributed algorithm depends on the connectivity graph. Performances in terms of the probability of mis-classification P_ε are in Figure 2.25. This probability can be expressed as a function of the probabilities of detection (P_d) and false alarm (P_{fa}) as $P_\varepsilon = P_a(1 - P_d) + (1 - P_a)P_{fa}$, which in the considered scenarios reduces to $P_\varepsilon \approx P_a(1 - P_d)$ since $P_{fa} \approx 0$. Performances are evaluated by averaging over 1000 measurements. In both figures the consensus-based distributed method is shown to outperform the non-cooperative one and to closely approach the centralised cooperative performance. For all methods, performance degrades for increasing σ_1 and/or decreasing μ_1, as the interference level by the PUs tends to be comparable with the background noise.

2.4.7 Experimental case study

In this section, we validate the cooperative spectrum detection on experimental data collected during an indoor measurement campaign at the third floor of the department DEIB of Politecnico di Milano. Experimental tests considered the coexistence of SU devices with both IEEE 802.15.4 (ZigBee) and IEEE 802.11g (Wi-Fi) compliant PU devices. The set-up consists of three ZigBee PU devices transmitting full data frames towards a coordinator device and three Wi-Fi PUs acting as infrastructure access points. The ZigBee PU transmitters are programmable devices configured to implement automatic power and gain adjustments based on the channel quality indication (CQI) and send acknow-

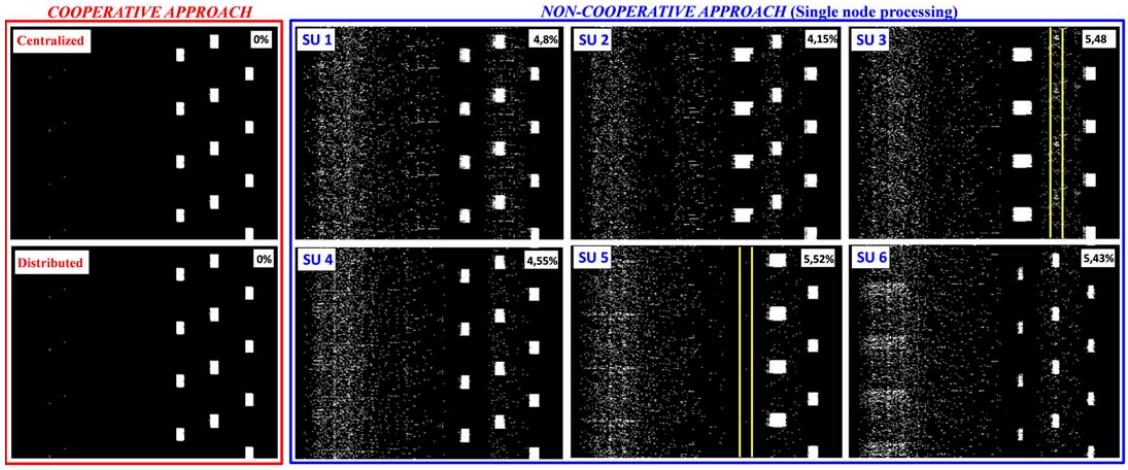


Figure 2.26: Time-frequency spectrum detection. Cooperative approaches (left box): centralised (top-left) and distributed (bottom-left) algorithms. Non-cooperative approach (right box): single node processing at node SU_i, $i = 1, \ldots, 6$. Yellow boxes highlight the mis-classification due to the effects of partial visibility for SU_3 and SU_5. The probability of mis-classification is shown in the top-right corner of each sub-figure.

ledged data frames with application-dependent duty cycle. They exchange data over predefined (but unknown to SUs) channels with centre frequencies $\{2.45, 2.46, 2.47\}$ GHz (standard compliant channels $\{20, 22, 24\}$). For each channel the occupied bandwidth is 3 MHz with nominal duty cycle of 30implementing cooperative spectrum sensing consist of PCs equipped with a portable spectrum analyser operating in the 2.4 GHz band. Power spectral measurements are taken with frequency steps of $\Delta f \simeq 333$ kHz (to cover the unlicensed $2.4 \div 2.495$ GHz band), resolution of 187.5 kHz and sampling time $\Delta t \simeq 536$ ms (dwell time of 1 ms). Measurements are then processed to extract the RSS information from which the relevant interference patterns can be tracked.

Time-frequency interference detection is implemented as in Section 2.4.5 with threshold $\lambda = 0.5$ (a-priori information is not available). Figure 2.26 shows the resulting $M \times N_F$ binary masks providing information about the time-frequency interference patterns caused by the PU primary networks. Binary masks are evaluated for all the considered algorithms: each subfigure highlights the most critical interference signals from the IEEE 802.15.4 devices. For all methods, we evaluate the probability of mis-classification P_ε defined as the percentage of the error with respect to the centralised approach (here considered as reference). Non cooperative detection is highly affected by errors due to partial visibility. The weighted-consensus method outperforms the non-cooperative one reaching the centralised detection. In addition, since the Wi-Fi interfering signals act as non-critical disturbances, they are considered as irrelevant by the detection process. Therefore, throughout cooperation between nodes, the critical PU spectrum occupancy is well reconstructed, even if some SUs are affected by limited sensing range and time-

varying fading/shadowing conditions.

2.4.8 Concluding remarks

In this section we proposed the use of weighted-average consensus for distributed detection of time-varying interference patterns. The consensus based algorithm is shown to reach the performance of centralised estimation strategy with reduced amount of information, since it is based on local computation and iterated sensing information exchange with neighbours. Interference pattern detection is validated based on an ad-hoc experimental measurement campaign to highlight a practical case study of unlicensed 2.4 GHz spectrum sharing for the implementation in CIMC demonstrator (WP5). The implementation in the demonstrator is detailed in D5.52 [2]

3 Relay processing and coding

3.1 Adaptive energy efficient scheme for QMF relaying

3.1.1 Introduction

In this work, the issues of energy consumption and adaptivity are considered for wireless cooperative relaying networks. We choose a specific relaying scheme called quantise-map-and-forward (QMF). The QMF relaying scheme, which is a version of compress-and-forward (CF), has been shown to be within a constant gap of the channel capacity [32] for certain unicast and multicast scenarios. From the point of view of the diversity-multiplexing tradeoff (DMT), QMF achieves the optimal DMT in full-duplex additive white gaussian noise (AWGN) wireless networks [33]. The DMT for half-duplex regimes is studied in [34] and [35]. In [35] several DMT regimes are identified as well as strategies to achieve the optimal tradeoffs.

The performance of QMF relaying is improved when the relays are equipped with multiple antennas [34]. However, the multiple antennas require multiple radio frequency (RF) chains at the transmitter and receiver sides of the relay resulting in increased energy consumption. The power consumption of the receive RF chains is dominated by the analog-to-digital converter (ADC) component, while at the transmit side it is dominated by the power amplifier (PA) and digital-to-analog converter (DAC). For QMF, multiple receive antennas can improve the performance of the relay which quantises the signal received from the source before forwarding it to the destination. However, as pointed out in [36], the ADC is the main bottleneck of the receiver, since it is costly and power-hungry. Therefore, it would be advantageous if a communication scheme could reduce the number of active ADCs and at the same time maintain its communication performance.

This work proposes an adaptive scheme wherein the number of active RF chains at the relay can be changed according to the channel conditions. Similar techniques are known, for example in the IEEE 802.11n standard [37]. While such approaches have been well investigated for point-to-point systems [38] [39], their extensions and application to more complex networks, such as cooperative networks are scarce [40].

In our case, the difficulty in enabling adaptivity for QMF relaying comes from joint decoding and successive interference cancellation (SIC) at the decoder. For this reason, we

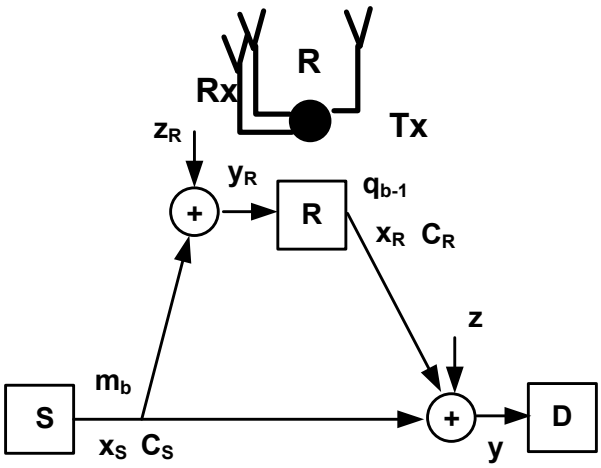


Figure 3.1: Relay channel.

choose to address the adaptivity problem by a combination of blind adaptation and cyclic redundancy check (CRC) codes. The main contributions of the work can be summarised as follows: 1) The adaptive cooperative communication scheme is defined within a control theoretic framework when the dynamics of frame error rate (FER) is unknown, where the control policy represents the number of active RF chains, 2) A heuristic method for blind adaptive policy construction is proposed, 3) The scheme is tested for the so-called asymmetric relay, i.e. when the number of receive and transmit RF chains are not the same, and 4) The scheme is simulated for finite-state as well as for block-fading Rayleigh channels.

3.1.2 Model of adaptive asymmetric cooperative relaying

We consider a simple cooperative network model consisting of three nodes, a source S, destination D and relay R [41] as shown in Figure 3.1. The destination D receives two versions of the source message, one directly from S, and the other from R. The role of the relay R is to help the destination D to decode the message coming from S by listening to the message transmitted by the source S and passing its version of the message to D. To improve its performance, the relay R can use more than one antenna as in [34]. However, this would lead to a larger energy consumption. In order to reduce the energy consumption of the relay circuitry, the number of receive and transmit RF chains can be different at R, implying an asymmetric architecture. Here, the number of receive RF chains is larger than the number of transmit chains. To further reduce the energy consumption of R, an adaptive scheme is proposed which takes into account the time varying nature of communication channels between nodes by choosing the appropriate number of active receive RF chains at the relay R.

To define the problem under consideration, it is important to look at the QMF scheme in

more detail [42]. When low-density parity-check (LDPC) codes are employed at S and R, a source message m_b is encoded by an LDPC code C_S at a time block b containing a codeword of N symbols, and transmitted so that D is able to receive all N symbols, while R receives fN symbols only, $0 < f < 1$. The relay R quantises and maps the received signal $y_R[b]$, where the mapping is another LDPC code C_R. The quantised version of the signal is denoted by q_b. During a given time block b, the relay R does not send the quantised message q_b to D, but the quantised message q_{b-1} which it received during the previous time block $b-1$. The idea can be understood by analysing the following equations, where SIC is used in addition to joint LDPC decoding of C_S and C_R

$$\tilde{y}[b-1] = h_{SD}[b-1]x_S(m_{b-1}) + \tilde{z}[b-1] \quad (3.1)$$

$$y[b] = h_{SD}[b]x_S(m_b) + h_{RD}[b]x_R(q_{b-1}) + z[b]. \quad (3.2)$$

The expression (3.1) relates to time block $b-1$, and it is a product of SIC as will be explained shortly. The expression (3.2) comes from the time block b, and it is a signal received at D when both S and R transmit at the same time. Here, $h_{SD}[b]$, $h_{SR}[b]$ and $h_{RD}[b]$ are the channel coefficients for three channels in Figure 3.1. These two equations are the inputs to the parallel LDPC decoder that decodes the codeword $x_S(m_{b-1})$, i.e. m_{b-1}, and the codeword $x_R(q_{b-1})$, i.e. an index q_{b-1} in the terminology of CF, which is side information for the decoding of the message m_{b-1}. Once $x_R(q_{b-1})$ is decoded, which produces an estimate $\hat{x}_R(q_{b-1})$, this estimate is subtracted from (3.2) giving $\tilde{y}[b] = h_{SD}[b]x_S(m_b) + \tilde{z}[b]$ which is a version of (3.1), but for a time block b. The second output of the LDPC decoder is an estimate $\hat{x}_S(m_{b-1})$, producing the estimate of m_{b-1}. Therefore, the message is decoded with a delay of one time block. A pictorial description of the decoding is shown in Figure 3.2 which illustrates the decoding dependence between different time blocks, i.e. codewords. From Figure 3.2, it is obvious that the decoding of a given codeword of block b will have an effect on the performance of subsequent blocks.

Remark 1. *The previous observation – illustrated by Figure 3.2 – means that an adaptive policy for cooperative networks has to be aware of the dependence in the decoding of different time blocks. In contrast, for the case when the decoding of the codewords is independent, the adaptation can be done independently for each time block.*

3.1.3 Adaptive policies for cooperative relaying and slow fading

The adaptive scheme proposed in this work can be seen as a control optimisation problem, more specifically, a tracking problem. The scheme adapts to channel conditions by maintaining the FER below a target FER, FER_T. This means that the achievable FER under given channel conditions has to be estimated. One possible way to estimate the FER is to introduce a CRC [43]. Then, the adaptive scheme operation can be illustrated as in Figure 3.3.

At the beginning of each fading block of length L_{FB} codewords, the initial number of active RF chains $u[0]$ is chosen. After the transmission, joint LDPC decoding and CRC

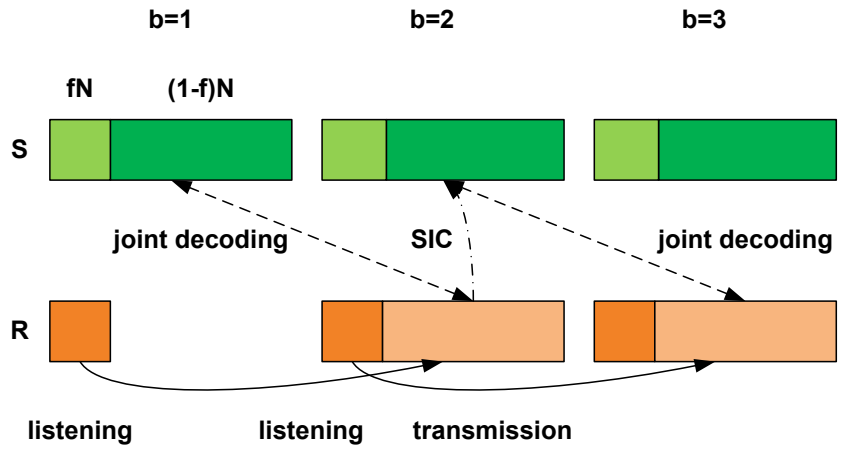


Figure 3.2: Joint decoding illustration for QMF relaying.

detection at D, the FER is estimated. If the FER is below FER_T, the number of active RF chains for the moment $b+1$ remains the same ($u[b+1] = u[b]$); if it is above, this number is changed to $u[b+1] \neq u[b]$. Following Remark 1, the choices of $u[b]$, $b = 0, \ldots, L_{FB}-1$, are not independent. Therefore, we define a vector or policy $u_k^b = (u[k]\, u[k+1] \ldots u[b])^T$ where $(.)^T$ denotes vector transpose. We also define channel state information (CSI) for the time block b as a vector $h[b] = (h_{SD}[b]\ h_{SR}[b]\ h_{RD}[b])^T$ and denote FER by $f[b]$.

The proposed adaptive cooperative relaying scheme is further defined, using a control theoretic framework. From this point of view, the policy u_k^b is a control policy, and $f[b]$ and $h[b]$ are state variables.

Definition 1. *Consider a discrete-time system described by $f[b]$ representing the FER, $h[b]$ the CSI, $y[b]$ the received signal, $u[b]$ a control for the time block b and the corresponding dynamics*

$$\begin{aligned} f[b+1] &= F_b(f_0^b, u_0^b, h_0^b) && (3.3)\\ h[b+1] &= H_b(h[b]) && (3.4)\\ y[b] &= h_{SD}[b]x_S(m[b]) + h_{RD}[b]x_R(q[b-1], u[b]) + z[b] && (3.5) \end{aligned}$$

such that the functions F_b and H_b are bounded and continuous. It is assumed that the state variable $f[b]$ can be estimated from the system output $y[b]$, $\tilde{f}[b] = C_{CRC}(y[b], h[b])$ by using appropriate CRC, C_{CRC}. Further, we define an asymptotic cost function in terms of a distance between the target FER_T, f_T, and current FER, $f[b]$, $d(f[b], f_T)$

$$J_{u_0^{N_c}} = \lim_{N_c \to \infty} \frac{1}{N_c} E\left\{ \sum_{b=0}^{N_c-1} d(f[b], f_T) + \varepsilon(u[b]) \right\} \quad (3.6)$$

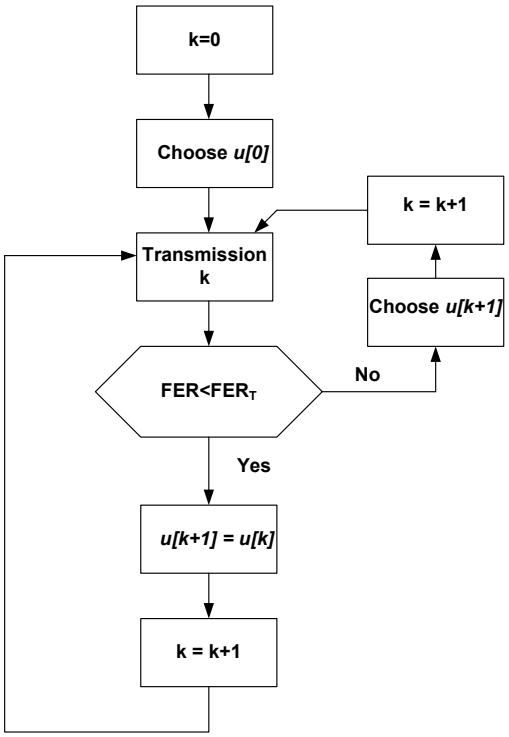


Figure 3.3: Block diagram of adaptive policy.

where N_c is the number of codewords or equivalently the time blocks during a transmission, and $\varepsilon(u[b])$ determines the cost of using $u[b]$ RF chains. The communication problem can be defined as an optimisation control problem

$$J_{u^*} = \min_{u_0^{N_c} \in \mathcal{U}} J_{u_0^{N_c}} \tag{3.7}$$

where $\mathcal{U}$ is the set of all admissible policies.

Remark 2. *(Definition 1): 1) For known dynamics F_b and H_b, the problem given in Definition 1 can be solved using dynamic programming methods, 2) For point-to-point networks, the communication systems do not need to use CRC codes to estimate FER; the known heuristic approach of channel abstraction can be employed to predict FER before the transmission and the appropriate communication scheme for the given CSI. This improves the throughput for the point-to-point case, 3) For cooperative communications, the abstraction heuristics are still under investigation, 4) From Figure 3.3 and* (3.6)*, one should note that $d(f[b], f_T)$ is defined such that it will imply FER $< FER_T$ condition with high probability.*

For cooperative communications, as briefly discussed in [44], the main obstacle is that one deals with multiple channels which makes the channel abstraction much more com-

plex. We leave for future research the investigation of channel abstraction for cooperative communications which may follow the approaches considered in [38].

Policy construction

As previously stated, this work follows another approach of estimating FER by means of a CRC code. However, although the estimation of $f[b]$ is possible after decoding each codeword, the dynamic programming will be difficult when the dynamics F_b is not known or is difficult to compute. Therefore, the so-called blind adaptive approach is proposed – when the function F_b is unknown – which needs to take into account the decoding dependence condition presented in Remark 1. For each fading block $(h[b]\ h[b+1] \dots\ h[b+L_{FB}-1])^T$ a vector policy $u_b^{b+L_{FB}-1}$ is constructed. Because of the slow fading assumption within one fading block of length L_{FB} codewords, it is true that $h[b] = h[b+1] = \dots\ = h[b+L_{FB}-1]$.

Hence, one possible way to construct the policy $u_b^{b+L_{FB}-1}$ is the following: the relay can utilise one RF chain $u[0] = 1$ at the start of the fading block, and then keeps increasing the number of RF chains for subsequent codewords within the fading block if the FER criterion is not met. However, because of the decoding dependence, the initial small number of antennas might adversely affect the decoding of the subsequent codewords leading to inadequate FER performance of the policy $u_b^{b+L_{FB}-1}$ as will be demonstrated later in Subsection 3.1.4. Therefore, depending on the network geometry, one should identify policies which will provide good FER performance and at the same time have satisfactory energy consumption proportional to $\sum_{i=b}^{b+L_{FB}-1} u[i]$. To make this procedure clear, three types of policies are introduced.

Definition 2. *(Policy types): 1) Type 1 policy keeps the number of antennas at the relay constant $u[b] = u[b+1] \dots = u[b+L_{FB}-1]$, e.g. $u_0^2 = (1,1,1)^T$, 2) Type 2 policy does not keep the number of antennas at the relay constant, e.g. $u_0^2 = (1,2,3)^T$, 3) Type 3 policy is an adaptive policy as illustrated in the flowchart in Figure 3.3.*

Remark 3. *(Policy ordering and lower bound on FER)*

Type 1 and Type 2 policies are fixed policies in the sense that the number of used antennas within one fading block will not adapt to channel conditions, i.e. for $u_0^1 = (1,4)^T$, $u[b]$ for $b = 0, 1$, are fixed regardless of channel conditions and current FER where $L_{FB} = 2$. However, $u_0^1 = (1,4)^T$ provides a lower bound in terms of FER for all policies which are of lower order comparing to $u_0^1 = (1,4)^T$ such as $\mathcal{Q} = \{(1,1),(1,2),(1,3)\}$, see Figure 3.4. Two policies $u_0^{b,1}$ and $u_0^{b,2}$ are partially ordered in $\leq_p$ sense, denoted by

$$(u^1[0], \dots, u^1[k], \dots, u^1[b]) \leq_p (u^2[0], \dots, u^2[k], \dots, u^2[b]),$$

if and only if $u^1[k] \leq u^2[k]$ for every k. Higher order policies have intuitively better FER performance, because then the relay provides more reliable side information to the destination.

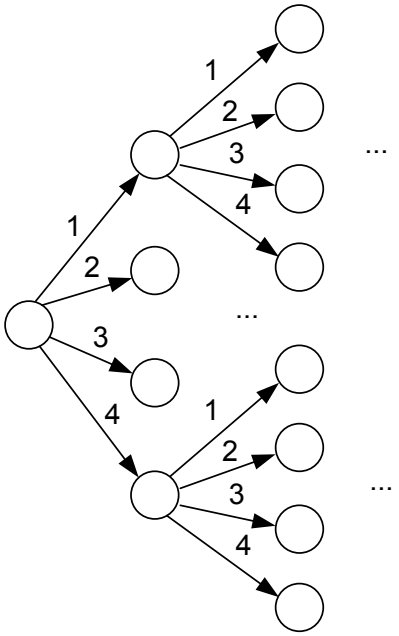


Figure 3.4: Policy graph.

A Type 3 policy is described by a set of policies which can be generated from a flowchart such as the one shown in Figure 3.3. The block 'Choose $u[k+1]$' in a specific realisation can be replaced, e.g. $u[k+1] = u[k]+1$. For instance, for $L_{FB} = 3$ start with $u[0] = 2$ and then keep increasing the number of antennas by 1 until FER_T is reached, resulting in a set of different possible policies $\mathcal{P} = \{(2,2,2),(2,2,3),(2,3,3),(2,3,4)\}$ (one of these policies is possible and will appear for corresponding channel conditions). We say that the Type 3 policy represented by $\mathcal{P}$ is derived from a Type 2 policy u_0^2 if u_0^2 is of a higher order compared to all policies belonging to $\mathcal{P}$, e.g. for $u_0^2 = (2,3,4)^T$.

The policy construction is then performed in the following manner: (i) Construct one or more Type 2 policies $u_b^{b+L_{FB}-1}$ which will serve as a lower bound in terms of FER for Type 3 adaptive policy, (ii) Check whether the Type 2 policies have acceptable FER performance and energy consumption as compared to Type 1 policies, (iii) Derive Type 3 adaptive policy from the chosen acceptable Type 2 ones.

One way to check the condition in step 2) of the policy construction is through simulations.

3.1.4 Case study for finite state and fading channels

To illustrate the benefits of the proposed scheme, several examples are presented where the scheme supports QMF relaying for three nodes and up to four receive RF chains at the relay receiver. The multiple antenna processing technique used at the receiver side

of the relay is maximum ratio combining (MRC). The source S and destination D are equipped with only one antenna.

It is assumed that C_S and C_R are Gallager error-correction LDPC codes (20000,10000) and (15000,5000), respectively. The notation (n, k) is used to describe the code, where n represents the codeword length and k is a sourceword length. The codes have similar performances in an AWGN channel as the codes used in [42]. At the destination D, joint message passing decoding is performed to decode both codes simultaneously. For this particular example, the joint iterative decoder applies 20 iterations. It is assumed that the average SNRs among the three nodes are such that $SNR_{SD} = SNR_{RD}$, while the SNR between the transmit antenna at S and one receive antenna at R, SNR_{SR} is equal to SNR_{SD}. In this situation multiple receive antennas at R help to overcome an unfavourable level of SNR_{SR}. However, this improvement comes at the cost of using multiple RF chains all the time which can be costly when the relay is battery powered.

As described in Figure 3.3, the adaptation of the transmission can be achieved with the goal of targeting a given FER. To track the change in FER, a sourceword at S is split into 50 frames, where each frame is encoded by a CRC code with 16 parity bits. After this, frames are collected and encoded into one C_S codeword. At the destination D, the messages coming from R and D are decoded jointly. This is followed by a CRC decoder used to estimate the FER. Throughout the transmission, the number of RF chains used at R is determined by the policy $u_b^{b+L_{FB}-1}$. In general the adaptation may use the FER of: 1) both messages, the main message sent by S and of the side information sent by R, or 2) the main message only. In this specific example, the latter approached is used.

Further, it is shown how to construct adaptive policies by using the approach explained in Subsection 3.1.3 to reduce the energy consumption of relay nodes.

Finite state channels

To gain insights into the proposed method, its performance for FSCs and binary phase shift keying (BPSK) modulation is first examined. The simulations under the FSC assumption should save simulation time – compared to fading channels – keeping in mind that the three node network can support six different channel conditions simultaneously. A two-state FSC model for all channels in the network is used as a first approximation [44], [45].

It is assumed that all six channels are slowly time-varying channels (six codewords per a fading block, $L_{FB} = 6$) where the six channel coefficients – due to four antennas at the relay – $h_{SR,1}$, $h_{SR,2}$, $h_{SR,3}$, $h_{SR,4}$, h_{SD} and h_{RD}- are chosen independently for each fading block. A 'good' state G of the two-state FSC model corresponds to a coefficient of unit power, while a 'bad' state B has a power 1/1.5. Thus, the channel coefficients take values from the set $\{1, 1/\sqrt{1.5}\}$. In addition to fading, the transmitted signals are subject to AWGN.

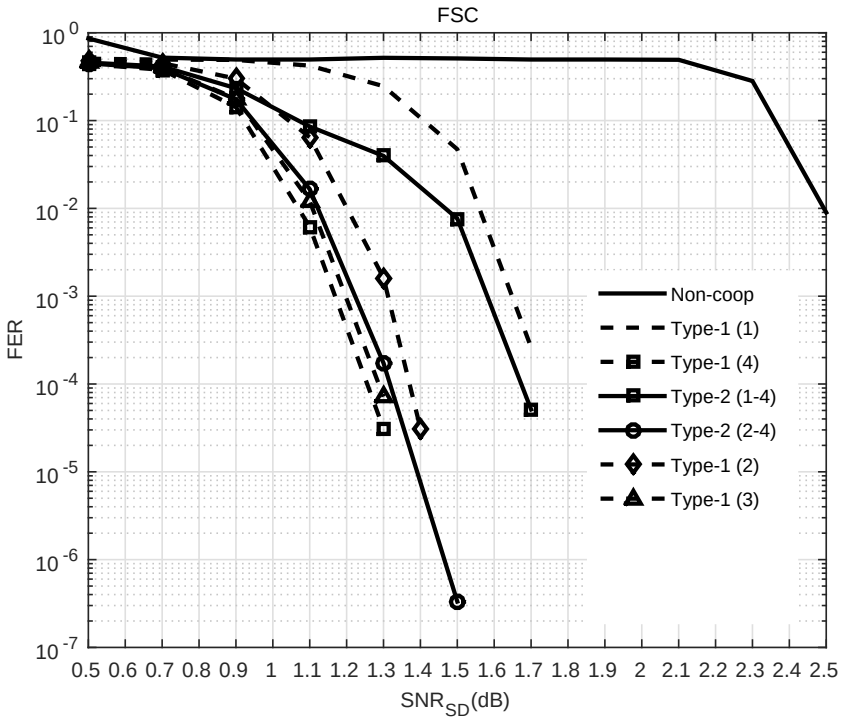


Figure 3.5: Performances of different Type 1 and Type 2 policies for QMF relaying with multiple antennas at the relay for BPSK modulation and FSC.

Figure 3.5 shows the FER performance for different Type 2 (1–4, 2–4) and Type 1 (1, 2, 3 and 4) relay receive RF chain policies. The Type 1 policies correspond to $(1,1,1,1,1,1)$, $(2,2,2,2,2,2)$, $(3,3,3,3,3,3)$ and $(4,4,4,4,4,4)$, while the Type 2 policies correspond to $(1,2,3,4,4,4)$ and $(2,3,4,4,4,4)$, respectively. The length of $u_b^{b+L_{FB}-1}$ is $L_{FB}=6$, and corresponds to the length of the fading block in terms of codewords.

The policy $(4,4,4,4,4,4)$ is the best Type 1 policy. It has a similar performance to $(3,3,3,3,3,3)$, is slightly better than $(2,2,2,2,2,2)$ and outperforms $(1,1,1,1,1,1)$ by 0.5 dB. Overall, it is by 1.5 dB better than non-cooperative transmission.

Out of the tested Type 2 policies, the best performance is shown by $(2,3,4,4,4,4)$. Its performance is close to the FERs of the Type 1 $(3,3,3,3,3,3)$ and $(4,4,4,4,4,4)$ as seen from Figure 3.5. On the other hand, for larger SNRs the performance of the second policy $(1,2,3,4,4,4)$ becomes close to the Type 1 policy $(1,1,1,1,1,1)$ showing a decreasing gain in spite of using the maximum number of RF chains for three time blocks. One difference between $(1,2,3,4,4,4)$ and $(2,3,4,4,4,4)$ is a starting point $u[0]$. Intuitively, this starting point affects the joint decoder at the destination D, since it determines the quality of side information coming from the relay R at the very start of the transmission. This is important since the first transmission affects the rest of transmissions due to the use of SIC and joint decoding. The price paid is that more energy is consumed by the

relay R with the policy $(2,3,4,4,4,4)$ compared to $(1,2,3,4,4,4)$.

As previously mentioned, the FER of $(2,3,4,4,4,4)$ serves as a lower bound for the derived algorithmic adaptive Type 3 policies described in Figure 3.3, where $u[0]=2$ and the block 'Choose $u[k+1]$' is represented by $u[k+1]=u[k]+1$. For example, $(2,3,4,4,4,4)$ gives a FER lower bound for $(2,2,2,3,4,4)$ or $(2,2,2,2,3,4)$ policies. Denote a Type 3 policy derived from $(2,3,4,4,4,4)$ by u^a. Table 3.1 shows u^a performance in terms of the average number of used antennas at R for different fixed target FERs, FER_T (10^{-2}, 10^{-3} and 10^{-4}) and corresponding SNRs. It can be seen that u^a is a compromise, since on average, it uses a larger number of RF chains for smaller SNRs to provide a good performance, but for larger SNRs it can reduce the number of RF chains to 2.

Table 3.1: Average number of antennas vs. SNR for Type 3 policy and fixed FER_T

FER_T 10^{-2}; SNR(dB)	1.15	1.2	1.25
av. num. of ant.	3.23	2.4	2
FER_T 10^{-3}; SNR(dB)	1.25	1.35	
av. num. of ant.	3.49	2	
FER_T 10^{-4}; SNR(dB)	1.37	1.4	
av. num. of ant.	2.99	2	

Block-fading channels

In this subsection, the simulations are extended for Rayleigh block-fading channels. Again, a slow block-fading scenario is considered where the six channel coefficients – including coefficients for four receive antennas at R – are independently chosen from a Rayleigh distribution at the beginning of a fading block and kept constant within the block. The length of the fading block is $L_{FB}=6$ codewords. The modulation at S and R is quadrature phase shift keying (QPSK).

Here, the same policies are tested as for FSCs. The results are shown in Figure 3.6. The Type 1 policy $(4,4,4,4,4,4)$ is the best out of all tested Type 1 policies; it outperforms $(2,2,2,2,2,2)$ by 0.8 dB, $(1,1,1,1,1,1)$ by 2 dB, and non-cooperative transmission by 4 dB for an FER of 2×10^{-2}. The $(3,3,3,3,3,3)$ policy shows slightly worse performance than $(4,4,4,4,4,4)$ so it was not shown in the figure.

The performance of the Type 2 policies for the Rayleigh block-fading channels are similar to their performances for FSCs. The policy $(2,3,4,4,4,4)$ is the best and again matches $(4,4,4,4,4,4)$ for most SNRs, however $(1,2,3,4,4,4)$ shows better performance for the Rayleigh fading, and it matches $(2,2,2,2,2,2)$ in contrast to the FSC case when it matches $(1,1,1,1,1,1)$. In addition, another policy $(1,2,2,2,2,2)$ (denoted by

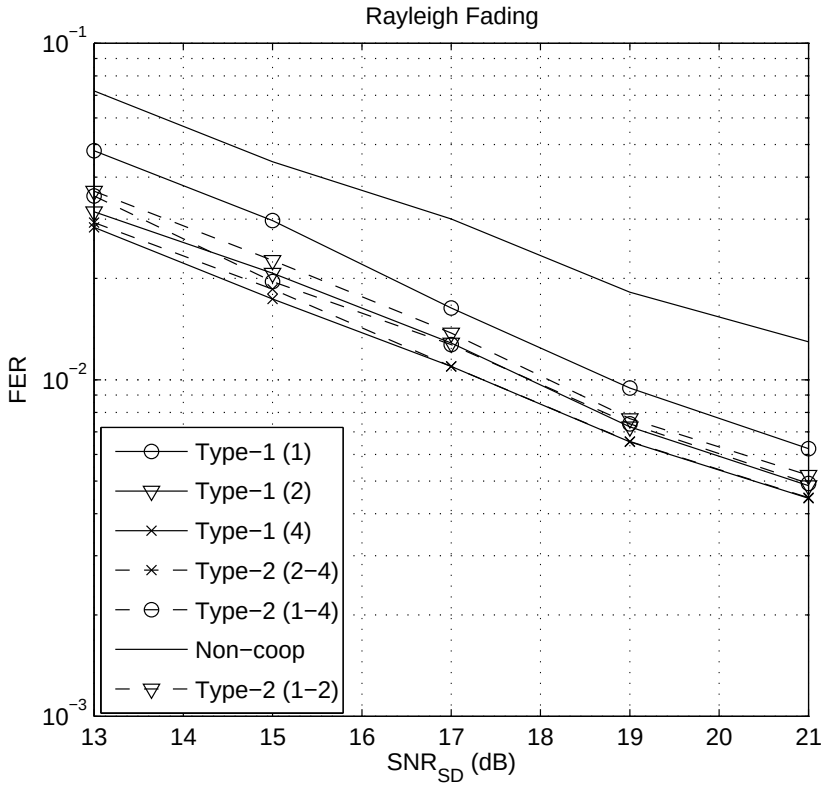


Figure 3.6: Performance of different Type 1 and Type 2 policies for QMF relaying with multiple antennas at the relay for QPSK modulation and block-fading Rayleigh channels.

Type 2 (1–2) in the figure) is tested, which has worse performance than $(2, 2, 2, 2, 2, 2)$ but by a small margin.

Therefore, one can choose from a variety of policies depending on the number of available RF chains and the performance requirements. For example, if the goal is to achieve the best performance and some reduction in energy consumption, then $(2, 3, 4, 4, 4, 4)$ is a possible choice. In the worst case scenario, a Type 3 policy derived from $(2, 3, 4, 4, 4, 4)$ will have 13% smaller energy consumption in the receiver chain than $(4, 4, 4, 4, 4, 4)$. On the other hand, if the energy consumption is more important then $(1, 2, 3, 4, 4, 4)$ or $(1, 2, 2, 2, 2, 2)$ could be more suitable. Similarly, in the worst case scenario, a Type 3 policy derived from $(1, 2, 2, 2, 2, 2)$ will require 9% less energy than $(2, 2, 2, 2, 2, 2)$. In more favourable scenarios for higher SNRs, the reduction can be up to 50%, this depends on channel conditions.

In the end, another question can be asked, which is: how will the performance change if the number of transmit antennas of the relay is increased. Figure 3.7, which compares the one and two relay transmit antennas cases, suggests that there could be a gain worth

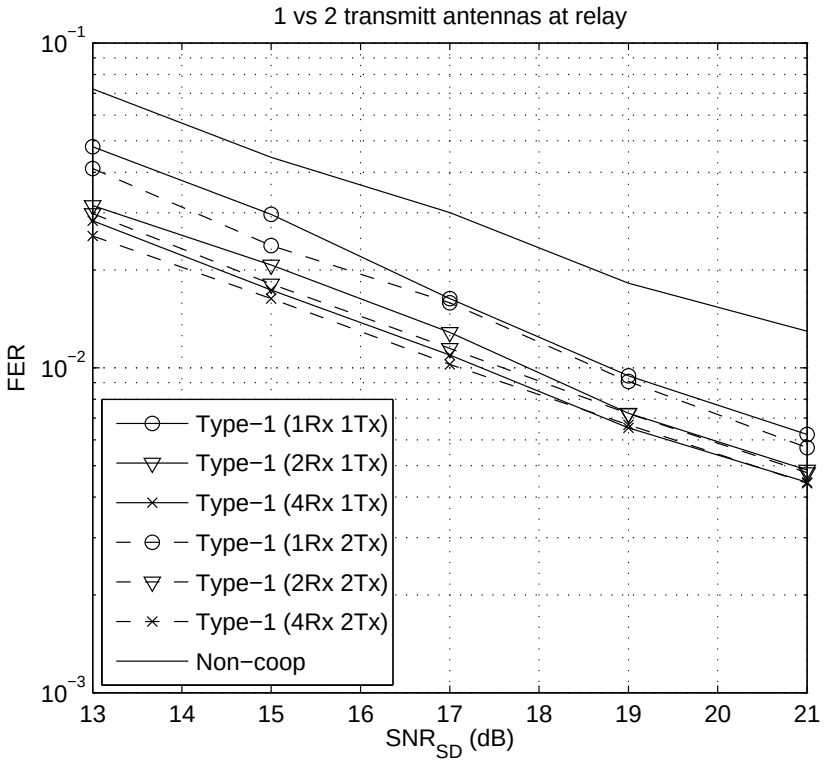


Figure 3.7: Performance of different Type 1 policies for QMF relaying with one vs. two transmit antennas at the relay for QPSK modulation and block-fading Rayleigh channels.

mentioning for one receive antenna at the relay, but the gain diminishes as the number of receive antennas increases. However, this additional RF chain will increase the energy consumption of the relay.

3.1.5 Conclusion

Problems of energy efficiency and adaptiveness are important issues that need to be addressed for cooperative wireless communication networks. In this work, an adaptive cooperative communication scheme is proposed to improve the energy efficiency of QMF by adapting the number of active RF chains at the relay to channel conditions; the goal is to keep the FER close to its target. However, cooperative communications often rely on complicated algorithms – such as QMF – so that physical layer channel abstraction is complex, making the adaptation difficult. Therefore, the proposed scheme is defined in a control theoretical framework, and a heuristic approach is proposed that combines a blind policy design and CRC codes. It is shown by simulation that considerable energy consumption reduction is possible for higher SNRs, while at the same time maintaining

good performance, using multiple antennas at relays, for lower SNRs. The scheme is tested for unequal RF chain numbers at the receive and transmit side of the relay. The results show that the asymmetric scheme can still deliver a very good performance gain.

Future research can address issues such as: 1) Whether a learning algorithm can be employed to capture FER dynamics and under which channel conditions, 2) Generalisation to multiple relays and other communication schemes such as CF and compute-and-forward.

3.2 Relay selection in cloud networks

3.2.1 Introduction

A technology regarded as the most promising for use in future cellular systems is physical layer relaying, in which an intermediate communication node termed relay node conveys a message from source to destination. A variety of relay techniques have been proposed and developed, ranging from simple amplify-and-forward (AF) techniques and analogue network coding methods to more complicated approaches of exploiting spatial diversity, including cooperative communication and so-called relay selection, the latter of which entails one or more relays being chosen out of a large group of nodes to retransmit a message from source to destination based on some desired performance criteria. It is of great interest to understand the implications of employing multicarrier signalling schemes such as orthogonal frequency-division multiplexing (OFDM) and related block-based single carrier schemes in advanced relay networks, particularly with reference to crucial properties that multi-carrier techniques exhibit in simpler point-to-point systems. The emphasis of the research project is to investigate and design new relay techniques that exploit the additional degrees of freedom provided by OFDM and related modulation schemes. Essentially, this project aims at how to enhance the reliability of wireless signals and meanwhile reduce the system complexity. Specifically, two aspects of the relay technology are investigated and analysed in details in the first year. They are related to an innovative selection scheme termed combined bulk/per-subcarrier selection (also known as combined bulk/per-tone selection, combined selection and joint selection etc.) [46–49].

Since the proposal of the cooperative network, the theoretical and implementation issues of relay networks have become important research topics in both industry and academia [50, 51]. As an effective way to exploit spatial and frequency diversity, relay selection has been touted as one of the most promising techniques in the next generation of cooperative networking [52]. For OFDM systems, bulk selection (a single relay is selected for transmission on all subcarriers) and per-subcarrier selection (selection is treated independently for each subcarrier, thus leading to transmission via multiple relays) have been proposed [53].

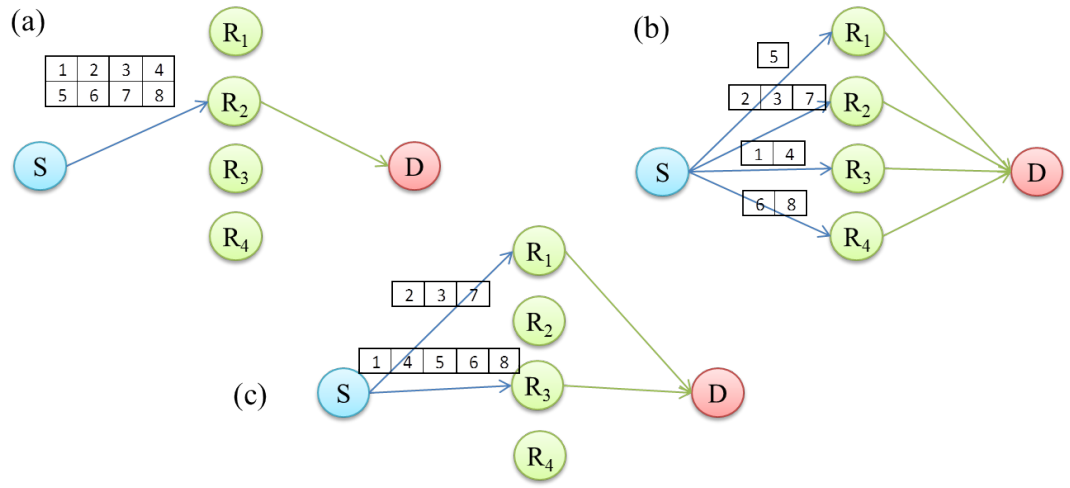


Figure 3.8: Illustration of (a) bulk, (b) per-subcarrier and (c) combined bulk/per-subcarrier relay selection strategies.

Obviously, bulk selection is simpler but has a poorer performance than per-subcarrier selection [54]. On the other hand, per-subcarrier selection has the optimal performance, but the system complexity is higher and thus a series of problems regarding synchronisation, implementation and channel state information (CSI) transmission are inevitable [48]. Combined selection can be regarded as a compromise scheme between bulk and per-subcarrier selections. To be specific, combined selection selects two relays first and then performs the per-subcarrier selection over these two relays [53]. Pictorial examples of the bulk, per-subcarrier and combined bulk/per-subcarrier selection strategies are illustrated in Figure 3.8.

In transmit antenna selection (TAS) scenarios, the OFDM system selecting only *two* out of the total available antennas can achieve the same outage and symbol-error probabilities as systems employing a per-subcarrier selection strategy in the high signal-to-noise ratio (SNR) regime [46]. In this report, we refer to this asymptotic behaviour as the *equivalence principle*.

It is natural to ask whether the combined selection strategy can be extended to relay selection. Clearly, the answer is "yes" to a certain degree. What is not obvious is whether the equivalence principle holds in relay selection systems, particularly for different forwarding schemes. Here, we prove that it does indeed hold for cooperative networks employing the three fundamental forwarding schemes: decode-and-forward (DF), fixed-gain (FG) AF, and variable-gain (VG) AF.

By the results presented in Section 3.2.2, the asymptotic outage performance of combined selection is identical to the performance of per-subcarrier selection at high SNR [55]. Hence, combined selection is regarded as the most appropriate selection scheme for cooperative networks containing a single user pair [55]. Meanwhile, we also pro-

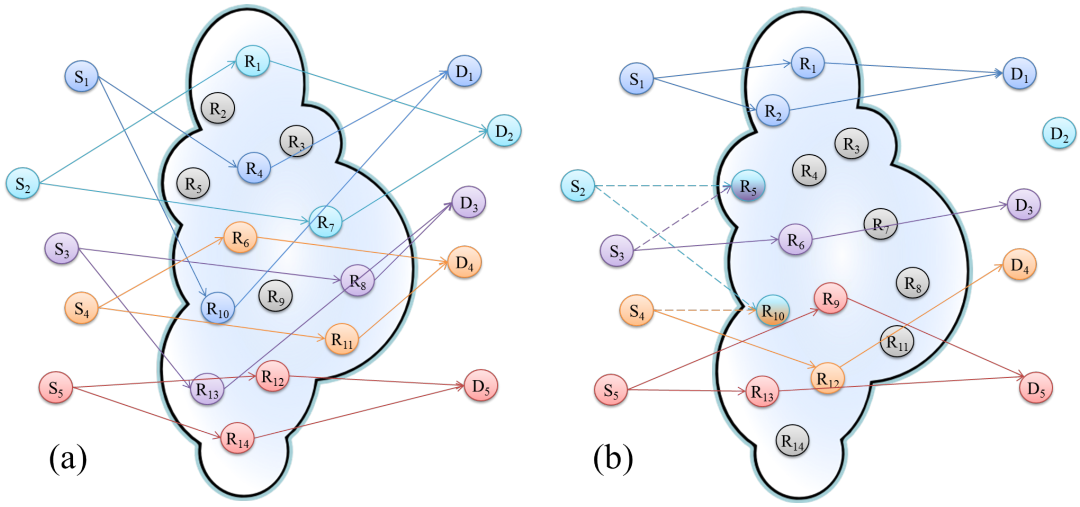


Figure 3.9: Illustration of combined bulk/per-subcarrier relay selection and transmission in the super dense cloud network equipped with multiple user pairs: (a) without contention, (b) with contention.

pose a methodology for the analysis the outage performance of combined selection over spatially correlated channels and carry out a series of preliminary investigations in Section 3.2.2.

However, in super dense cloud networks with multiple users, it would not necessarily be the case because another detrimental phenomenon called *selection contention* should be taken into consideration for such a multiuser scenario [56]. Therefore, in Section 3.2.3, we analyse the effects of outage and contention together by using a defined performance parameter termed the *bulk gain factor*. The main contribution of this work is to show that although the bulk selection scheme yields a worse outage performance, it would still be preferable in super dense cloud networks as long as a sufficiently large transmit power is provided, since it only demands one relay and has a smaller contention probability. Besides, the outage performance of AF relay in super dense cloud networks is analysed in particular in this section as well. All these results can be used as a design benchmark for physical layer applications/protocols used in super dense cloud networks. With these, it is possible to design a theory to further optimise super dense cloud networks by reducing contention probability. We finally conclude this report in Section 3.2.4.

3.2.2 Combined bulk/per-subcarrier relay selection in two-hop OFDM systems

In this section, we apply the concept of combined bulk/per-subcarrier selection to two-hop relay selection systems employing OFDM. The outage probability of the proposed strategy is analysed in the high SNR regime when DF, FG AF and VG AF are employed at the relays. We prove that the combined selection strategy achieves an outage probability equivalent to conventional per-subcarrier selection in the high SNR regime without

using the full set of available relays for selection. Moreover, we demonstrate through numerical simulations that this performance advantage holds when channels are spatially correlated. Besides, we also propose a methodology for the outage performance analysis of combined selection over spatially correlated channels and carry out a series of preliminary investigations.

Fundamentals

Relays and channels. For a typical OFDM system with a single source and destination, M relays and K subcarriers, if the combined selection strategy first chooses $L = 2$ out of M relays, the bulk selection, per-subcarrier selection and combined selection strategies can be illustrated in Figure 3.8. It is clear that for M relays and K subcarriers, $2MK$ channels exist in the space-frequency grid. For brevity, we tentatively assume that the $2MK$ channel coefficients are statistically independent, zero-mean, complex Gaussian (ZMCG) random variables with total variance μ_i, where $i \in \{1, 2\}$ corresponds to the first and second hop, respectively. The frequency-domain channel coefficients between the source and the mth relay for the kth subcarrier, $\forall\, m \in \mathcal{M} = \{1, 2, \ldots, M\}$ and $\forall\, k \in \mathcal{K} = \{1, 2, \ldots, K\}$, are denoted as $h(m, k)$, and similarly the channel coefficients between the mth relay and the destination for the kth subcarrier are denoted as $g(m, k)$. Meanwhile, without loss of generality, if we assume $h(m, k) \sim \mathcal{CN}(0, \mu_1)$ and $g(m, k) \sim \mathcal{CN}(0, \mu_2)$, it is well-known that $|h(m, k)|^2$ and $|g(m, k)|^2$ are distributed as $\Gamma(1, \mu_1)$ and $\Gamma(1, \mu_2)$[1].

We further assume that the CSI is perfectly estimated and shared among all communication nodes, and the relay network operates in a half-duplex protocol so that two orthogonal time slots are required for one complete transmission from source to destination. The noise statistics at the mth relay and at the destination are denoted by $n(m, k)$ and $\eta(k)$, respectively, and all $MK+K$ statistics are independent, identically distributed (i.i.d.) ZMCG random variables with variance $N_0/2$ per dimension.

Assuming equal bit and power allocation schemes are applied, the average transmit power per subcarrier at the source and at each utilised relay is denoted by P_t. Hence, the equivalent instantaneous end-to-end SNR[2] corresponding to the kth subcarrier and the mth relay using a DF protocol can be expressed as

$$\mathsf{SNR}(m, k) = \frac{P_t}{N_0} \min\left(|h(m, k)|^2, |g(m, k)|^2\right). \tag{3.8}$$

[1] $\mathcal{CN}(\,\cdot\,,\,\cdot\,)$ and $\Gamma(\,\cdot\,,\,\cdot\,)$ represent the Gaussian and the Gamma distributions, respectively.

[2] In fact, an outage in DF relaying networks depends on the minimum channel coefficient among the source-relay and the relay-destination links. Hence, we can employ the minimum channel coefficient as the equivalent channel quality indicator here.

It is also well known that for FG AF relaying, the instantaneous end-to-end SNR can be expressed as

$$\mathsf{SNR}(m,k) = \frac{G_F^2|h(m,k)|^2|g(m,k)|^2P_t}{(G_F^2|g(m,k)|^2+1)N_0}, \tag{3.9}$$

where $G_F = \sqrt{P_t/(P_t\mu_1+N_0)}$ is the per-subcarrier relay gain. Finally, for VG AF relaying with per-subcarrier relay gain factor $G_V = \sqrt{P_t/(P_t|h(m,k)|^2+N_0)}$, the instantaneous end-to-end SNR can be expressed as

$$\mathsf{SNR}(m,k) = \frac{|h(m,k)|^2|g(m,k)|^2P_t^2}{(|g(m,k)|^2P_t+|h(m,k)|^2P_t+N_0)N_0}. \tag{3.10}$$

Combined relay selection: An introduction. Now let us consider the selection process of the combined selection. It has been verified that the outage probability mainly depends on the worst channel over all subcarriers at high SNR and thus we should maximise the minimum channel coefficient over all subcarriers in order to minimise the outage probability [46]. Following this logic, the bulk selection criterion can be expressed as [56]

$$\hat{\mathcal{L}}_{bulk} = \arg\max_{\mathcal{L}\subseteq\mathcal{M}} \min_{k\in\mathcal{K}} \max_{l\in\mathcal{L}} \mathsf{SNR}(m,k) \tag{3.11}$$

where $\mathcal{L}$ identifies a pair of relays that can be used to carry out per-subcarrier selection. Clearly, $|\mathcal{L}| = 2$. Note that we could consider $|\mathcal{L}| > 2$, but this is not required to prove our main result, namely that the equivalence principle applies to relay selection systems.

For a given set of relays chosen according to (3.11), the optimal per-subcarrier selection criterion for subcarrier $k \in \mathcal{K}$ is given by [56]

$$\hat{l}_{pt}(k) = \arg\max_{l\in\hat{\mathcal{L}}_{bulk}} \mathsf{SNR}(m,k). \tag{3.12}$$

A note on the practicality of these selection processes is in order. In fact, it is straightforward to see that selection of this nature is practically achievable by adopting a well-designed timer-based scheme [48]. In this section, we consider the time interval within which two or more timers expire to be infinitesimal. Hence, we do not need to worry about the timer contention problem here. This is equivalent to assuming that the selection process is perfect, which is justified since we are concerned with theoretical performance limits.

Performance analysis

We adopt outage probability as a metric to measure the performance of OFDM systems employing combined relay selection. In what follows, we define the outage probability in the usual way:

$$F(s) = \mathbb{P}(\mathsf{SNR}(m,k) < s) \tag{3.13}$$

where s is an end-to-end SNR threshold. Our goal is to show that as $P_t/N_0 \to \infty$, the outage probability for cooperative networks employing combined selection (with $|\mathcal{L}| = 2$) tends to the outage probability for networks that use per-subcarrier selection over the relay set $\mathcal{L} = \mathcal{M}$. We consider each selection strategy in turn.

Combined selection. Denote the probability density function (PDF) corresponding to $F(s)$ as $f(s)$. Using *Lemma 1* in [46], we obtain the PDF of the outage probability of an OFDM relaying system employing combined bulk/per-subcarrier selection at high SNR:

$$f_{comb}(s) = MK[F(s)]^{M-1}[1 - F(s)]^{MK-M} f(s). \tag{3.14}$$

The corresponding CDF $F_{comb}(s)$ of $f_{comb}(s)$ can be used to quantify the asymptotic outage performance for the selection system.

DF relay

Now, let us focus on the DF relay case. Because $|h(m,k)|^2 \sim \Gamma(1,\mu_1)$ and $|g(m,k)|^2 \sim \Gamma(1,\mu_2)$, the PDF and CDF of $|h(m,k)|^2$ and $|g(m,k)|^2$, $\forall m \in \mathcal{M}$, $\forall k \in \mathcal{K}$ are

$$f(s) = e^{-s/\mu_i}/\mu_i \quad \Leftrightarrow \quad F(s) = 1 - e^{-s/\mu_i} \tag{3.15}$$

for $i = 1, 2$. Define the *normalised SNR* to be $\overline{\gamma} = P_t/N_0$. By applying the theory of order statistics [57], $F(s)$ in DF relaying networks with selection defined according to (3.8) is given by

$$F(s) = 1 - e^{-s(1/\mu_1 + 1/\mu_2)/\overline{\gamma}}. \tag{3.16}$$

Consequently, the PDF $f(s)$ corresponding to $F(s)$ in DF relaying networks is

$$f(s) = \frac{\mathrm{d}F(s)}{\mathrm{d}s} = \frac{1}{\overline{\gamma}} \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right) e^{-s\left(\frac{1}{\mu_1} + \frac{1}{\mu_2}\right)/\overline{\gamma}}. \tag{3.17}$$

According to (3.14), (3.16) and (3.17), the PDF of the worst channel at high SNR with the combined bulk/per-subcarrier selection is

$$f_{comb}(s) = \frac{MK}{\overline{\gamma}} \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right) \left[1 - e^{-s\left(\frac{1}{\mu_1} + \frac{1}{\mu_2}\right)/\overline{\gamma}} \right]^{M-1} \left[e^{-s\left(\frac{1}{\mu_1} + \frac{1}{\mu_2}\right)/\overline{\gamma}} \right]^{MK-M+1}. \tag{3.18}$$

Meanwhile, the system outage probability $F_{comb}(s) = \int_0^s f_{comb}(t)\mathrm{d}t$ can be expressed by an asymptotic approximation at high SNR ($\overline{\gamma} \to \infty$, or equivalently $s \to 0$). Employing a Taylor series expansion yields the asymptotic expression

$$F_{comb}(s) \simeq \left[\frac{K^{\frac{1}{M}} s(\mu_1 + \mu_2)}{\mu_1 \mu_2 \overline{\gamma}} \right]^M. \tag{3.19}$$

FG AF relay

As for the FG AF relay case, using (3.9), it can be shown that [58]

$$F(s) = 1 - 2e^{-\frac{s}{\mu_1\overline{\gamma}}}\sqrt{\frac{s(\mu_1+\frac{1}{\overline{\gamma}})}{\mu_1\mu_2\overline{\gamma}}}K_1\left(2\sqrt{\frac{s(\mu_1+\frac{1}{\overline{\gamma}})}{\mu_1\mu_2\overline{\gamma}}}\right) \tag{3.20}$$

where $K_1(\cdot)$ is the first order modified Bessel function of the second kind. This is the standard outage result for FG AF networks. Applying (3.14) and using the Puiseux series expansion of the Bessel function along with binomial and Taylor expansions of the algebraic and exponential terms in (3.20) yields the asymptotic outage expression

$$F_{comb}(s) \simeq \left[\frac{K^{\frac{1}{M}}s\ln(\overline{\gamma})}{\mu_2\overline{\gamma}}\right]^M. \tag{3.21}$$

VG AF relay

For VG AF relaying, the CDF of (3.10) has been proved to be [59]

$$F(s) = 1 - 2e^{-\frac{s}{\overline{\gamma}}\left(\frac{1}{\mu_1}+\frac{1}{\mu_2}\right)}\sqrt{\frac{s}{\mu_1\overline{\gamma}^2}\left(\frac{1}{\mu_2}+s\right)}K_1\left(2\sqrt{\frac{s}{\mu_1\overline{\gamma}^2}\left(\frac{1}{\mu_2}+s\right)}\right). \tag{3.22}$$

Using a similar approach to the FG AF case, we find that

$$F_{comb}(s) \simeq \left[\frac{K^{\frac{1}{M}}s(\mu_1+\mu_2)}{\mu_1\mu_2\overline{\gamma}}\right]^M \tag{3.23}$$

for large $\bar{\gamma}$ (equivalently, small s). Note that the expression given in (3.23) is exactly the same as that given in (3.19), which indicates the performance of VG AF relaying is identical to DF relaying network at high SNR. A related result was explored in [60]. However, the equivalence of DF and VG AF relay performance in the present context has not been reported in the literature until now.

Per-subcarrier selection. Considering the PDF of the worst subcarrier when applying per-subcarrier selection, we have that the PDF of the instantaneous end-to-end SNR is given by

$$f_{ps}(s) = MK\left[F(s)\right]^{M-1}\left\{1-\left[F(s)\right]^M\right\}^{K-1}f(s). \tag{3.24}$$

By repeating the derivations outlined in the previous subsection, the asymptotic expression (leading order term) for the CDF of the system employing per-subcarrier selection, denoted as $F_{ps}(s)$, can be easily calculated for all three kinds of relays. By carrying out these calculations, we find that the asymptotic expressions of $F_{ps}(s)$ with per-subcarrier selection for all three relaying protocols agree with the combined selection results given

in (3.19), (3.21), and (3.23). To see this result without going into the details of the specific calculations, one can observe (3.14) and (3.24) carefully, noting that both PDFs are dominated by the terms MK, $[F(s)]^{M-1}$ and $f(s)$ for the three relay protocols. Indeed, the differences in the expressions, notably $[1-F(s)]^{MK-M}$ and $\{1-[F(s)]^M\}^{K-1}$, are incorporated into higher order terms in the asymptotic expansion of the outage probability, and hence do not contribute at high SNR. Note, however, that these corrections will influence the rate of convergence of the combined selection performance to that of per-subcarrier selection. More strict and generalised mathematical proof is given in the next subsection.

Generalised theorem of equivalence principle. As we have proved the equivalence principle for three types of relays, one might wonder whether this equivalence principle can be further extended to other applications. To answer this question, a generalised equivalence principle can be stated as follows.

Proposition 1. *If the CDF of the i.i.d. end-to-end SNR, $F(s)$, can be expanded in the variable $\bar{\gamma}$ as*

$$F(s) = \sum_{i=i_0}^{\infty} c_i(s) \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i}{\theta}} [\ln(\bar{\gamma})]^r \tag{3.25}$$

where i_0 is an integer given by $i_0 = \arg\min_{n\in\mathbb{N}}\{c_n(s) \neq 0\}$; θ is a nonzero natural number; $\{c_i(s)\}$ represents a series of functions of s; $r \in \mathbb{N}$, then combined selection is able to achieve an outage probability equivalent to conventional per-subcarrier selection in the high SNR regime.

It can be proved as follows. From (3.25), we can determine $f(s)$ by

$$f(s) = \frac{\mathrm{d}F(s)}{\mathrm{d}s} = \sum_{i=i_0}^{\infty} \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i}{\theta}} [\ln(\bar{\gamma})]^r \frac{\mathrm{d}c_i(s)}{\mathrm{d}s}. \tag{3.26}$$

Therefore, according to (3.14), the asymptotic expression (leading order term) at $\bar{\gamma} \to \infty$ for the PDF $f_{comb}(s)$ can be determined by

$$\begin{aligned} f_{comb}(s) = MK &\left\{\sum_{i=i_0}^{\infty} c_i(s) \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i}{\theta}} [\ln(\bar{\gamma})]^r\right\}^{M-1} \left\{1 - \sum_{i=i_0}^{\infty} c_i(s) \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i}{\theta}} [\ln(\bar{\gamma})]^r\right\}^{MK-M} \\ &\times \left\{\sum_{i=i_0}^{\infty} \frac{\mathrm{d}c_i(s)}{\mathrm{d}s} \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i}{\theta}} [\ln(\bar{\gamma})]^r\right\} \sim MK[c_{i_0}(s)]^{M-1} \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i_0}{\theta}M} [\ln(\bar{\gamma})]^{rM} \frac{\mathrm{d}c_{i_0}(s)}{\mathrm{d}s}. \end{aligned} \tag{3.27}$$

Note, it is obvious that $c_{i_0}(0)$ must be zero in this case, since the asymptotic expression for $F(s)$ must be zero when $s = 0$. Otherwise, $F(0) \sim c_{i_0}(0)\,(1/\bar{\gamma})^{i_0/\theta}\,[\ln(\bar{\gamma})]^r \neq 0$,

$\forall \bar{\gamma} > 0$, which is against our definition of outage as given in (3.13). Hence, the asymptotic expression for $F_{comb}(s)$ is

$$F_{comb}(s) = \int_0^s f_{comb}(x)\mathrm{d}x \sim \left\{ K^{\frac{1}{M}} c_{i_0}(s) \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i_0}{\theta}} [\ln(\bar{\gamma})]^r \right\}^M . \tag{3.28}$$

According to (3.24), we can employ the same method to obtain the asymptotic expression for $F_{ps}(s)$ corresponding to $f_{ps}(s)$ at high SNR. It can be easily derived that $F_{ps}(s) \sim F_{comb}(s)$ and has exactly the same asymptotic expression as given in (3.28). As a consequence, we have proved that the asymptotic outage performances at high SNR produced by combined and per-subcarrier selections are the same, as long as their CDFs of the i.i.d. end-to-end SNR, $F(s)$, can be expanded as the form given in (3.25).

Numerical results

To verify our analysis in the last section, we employ Monte Carlo simulation methods to numerically study the outage performances of the worst subcarrier of OFDM systems employing per-subcarrier and combined bulk/per-subcarrier selection strategies respectively. Meanwhile, the asymptotic outage performance at high SNR is also taken into account in our simulations. In particular, we let $K = 8$, $s = 1$, $\mu_1 = \mu_2 = 2$, $N_0 = 1$ for all simulations and $M \in \{3, 5\}$ to observe the effect of M on the outage performance.

In addition, all details that have been given in previous sections are under the assumption of i.i.d. channels. However, this might not be the case in practice owing to the insufficient physical separations among relays [61]. Therefore, in order to examine the robustness and practicality of the combined selection strategy, we simulated its performance over spatially equally correlated channels as well[3]. To model the correlation phenomenon, we can construct the equally correlated Rayleigh fading channel by [62]

$$\begin{aligned} w(m,k) = &[\sqrt{1-\rho_i}x_i(m,k) + \sqrt{\rho_i}x_{i0}(k)] \\ &+ j[\sqrt{1-\rho_i}y_i(m,k) + \sqrt{\rho_i}y_{i0}(k)] \end{aligned} \tag{3.29}$$

where $\{w(m,k), \mu_i\} = \{h(m,k), \mu_1\}$ or $\{w(m,k), \mu_i\} = \{g(m,k), \mu_2\}$ corresponding to the first and the second hops; $i \in \{1, 2\}$; $j = \sqrt{-1}$; $x_i(m,k)$, $y_i(m,k) \sim \mathcal{N}(0, \mu_i/2)$ are i.i.d.; $x_{i0}(k)$, $y_{i0}(k) \sim \mathcal{N}(0, \mu_i/2)$ are i.i.d. and serve as a bridge to correlate all channels. Hence, $\forall m \neq n$ we have $\mathbb{E}\{w(m,k)w^*(n,k)\}/\sqrt{\mathbb{E}\{|w(m,k)|^2\}\mathbb{E}\{|w(n,k)|^2\}} = \rho$, which is the common cross-correlation coefficient over all channels. In simulations, we let $\rho \in \{0, 0.8\}$.

[3] We maintain independence in frequency as it mimics a block fading scenario, a common model in OFDM systems that utilise a resource block structure, e.g. LTE.

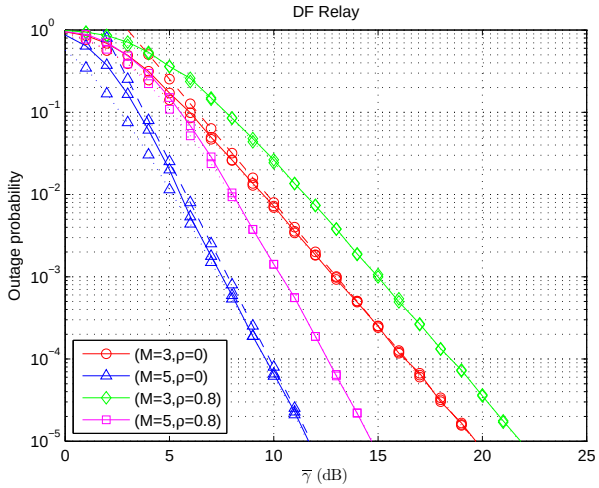


Figure 3.10: DF relay case: outage probability vs. SNR for per-subcarrier and combined bulk/per-subcarrier selection systems; (a) solid line: combined selection results, (b) dotted line: per-subcarrier selection results, (c) dashed line: asymptotic results.

The simulation results corresponding to the three relay protocols are presented in Figures 3.10, 3.11 and 3.12, respectively[4]. From these three figures, we can summarise some key points with respect to the combined selection strategy. First, for all three types of relays, it is clear that for the cases when $M = 3$ and $M = 5$, the outage probability of the system employing combined selection adheres to the equivalence principle over uncorrelated channels. Moreover, increasing the number of relays M will yield a better outage performance, again as expected. Note, however, that an increase in M does not mean that the number of *utilised* relays increases for the combined selection system since $|\mathcal{L}| = 2$. We also note from the figures that the equivalence principle appears to hold in correlated channels as well, thus suggesting that combined selection is a robust, practical solution for a broad range of applications. Finally, the asymptotic outage probability of the VG relaying scenario shown in Figure 3.12 is quite close to the performance of the DF relaying scenario as given in Figure 3.10, which agrees with our analysis (cf. (3.19) and (3.23)).

Outage performance analysis over spatially correlated channels

As the equivalence principle shown for spatially correlated channels, it is worth investigating further. As detailed in [62], we take a similar approach to transform a set of equally correlated random variables to a set of conditionally independent random variables, so that conventional analytical tools, e.g. order statistics can be applied to analyse them

[4]Here, more higher order terms for FG AF case are kept in order to illustrate the convergence between numerical and asymptotic results within a reasonable SNR range.

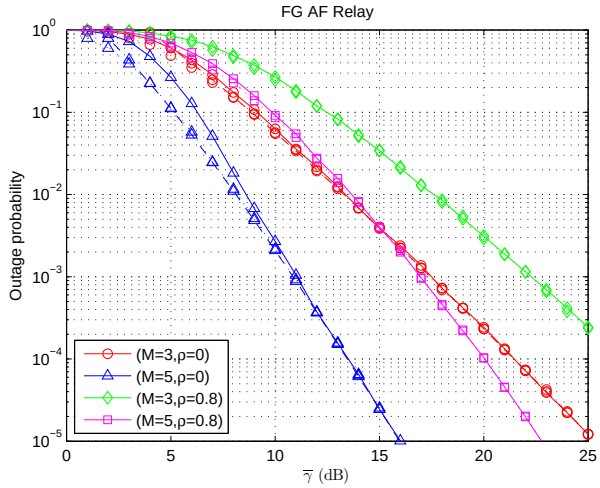


Figure 3.11: FG AF relay case: outage probability vs. SNR for per-subcarrier and combined bulk/per-subcarrier selection systems; (a) solid line: combined selection results, (b) dotted line: per-subcarrier selection results, (c) dashed line: asymptotic results.

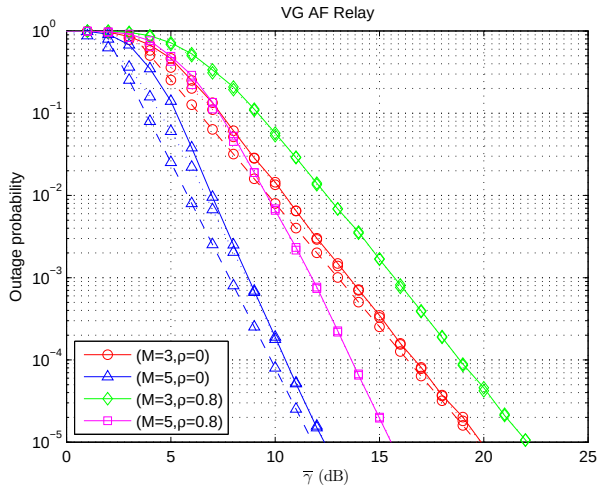


Figure 3.12: VG AF relay case: outage probability vs. SNR for per-subcarrier and combined bulk/per-subcarrier selection systems; (a) solid line: combined selection results, (b) dotted line: per-subcarrier selection results, (c) dashed line: asymptotic results.

effectively. To do so, we first assume that two references for each subcarrier $x_{i0}(k)$ and $y_{i0}(k)$ are fixed and have $x_{i0}(k) = X_{i0}(k)$ and $y_{i0}(k) = Y_{i0}(k)$. Therefore, the conditional distribution of $w(m,k)$ is $\mathcal{CN}(\sqrt{\rho_i}[X_{i0}(k) + jY_{i0}(k)], \mu_i(1-\rho_i))$. Consequently, given $x_{i0}(k) = X_{i0}(k)$ and $y_{i0}(k) = Y_{i0}(k)$, $|w(m,k)|^2 \sim \chi_2(\sqrt{\rho_i[X_{i0}^2(k) + Y_{i0}^2(k)]}, \mu_i(1-\rho_i))$. If we denote $T_i(k) = X_{i0}^2(k) + Y_{i0}^2(k)$, the conditional PDF and the conditional CDF

of $|w(m,k)|^2$ are given by [63]

$$f_w(s|T_i(k)) = \frac{1}{\mu_i(1-\rho_i)} e^{-\frac{s+\rho_i T_i(k)}{\mu_i(1-\rho_i)}} I_0\left(\frac{2\sqrt{\rho_i T_i(k) s}}{\mu_i(1-\rho_i)}\right) \tag{3.30}$$

and

$$F_w(s|T_i(k)) = 1 - Q\left(\sqrt{\frac{2\rho_i T_i(k)}{\mu_i(1-\rho_i)}}, \sqrt{\frac{2s}{\mu_i(1-\rho_i)}}\right), \tag{3.31}$$

where $I_0(\cdot)$ is the zero order modified Bessel function of the first kind; $Q(\cdot,\cdot)$ is the first order Marcum Q function.

Meanwhile, the PDF and CDF of $T_i(k)$ can be obtained as:

$$f_{T_i}(T_i(k)) = \frac{1}{\mu_i} e^{-\frac{T_i}{\mu_i}} \Leftrightarrow F_{T_i(k)}(T_i(k)) = 1 - e^{-\frac{T_i(k)}{\mu_i}}. \tag{3.32}$$

Denote $\mathbf{T}_1 = \{T_1(1), T_1(2), \ldots, T_1(K)\}$ and $\mathbf{T}_2 = \{T_2(1), T_2(2), \ldots, T_2(K)\}$. Because $\forall i \in \{1,2\}$ and $k \in \mathcal{K}$, $T_i(k)$ are mutually independent, we thereby can derive the joint PDF corresponding to $\mathbf{T}_1$ and $\mathbf{T}_2$ by

$$f_{\mathbf{T}}(\mathbf{T}_1, \mathbf{T}_2) = \left(\frac{1}{\mu_1\mu_2}\right)^K \prod_{k=1}^{K} \left(e^{-\frac{T_1(k)}{\mu_1}} e^{-\frac{T_2(k)}{\mu_2}}\right). \tag{3.33}$$

Accordingly, we can denote the conditional *a priori* outage probability as $F(s|T_1(k), T_2(k))$. We can also denote the conditional *a posteriori* outage probabilities for combined selection and per-subcarrier selection as $F_{comb}(s|\mathbf{T}_1, \mathbf{T}_2)$ and $F_{ps}(s|\mathbf{T}_1, \mathbf{T}_2)$ respectively.

We present the all-important contribution here. A generalised equivalence principle can be stated as follows.

Proposition 2. *If the conditional CDF of the end-to-end SNR, $F(s|T_1(k), T_2(k))$, can be expanded in the variable $\bar{\gamma} = P_t/N_0$ as*

$$\begin{aligned} F(s|T_1(k), T_2(k)) &= \sum_{i=i_0}^{\infty} c_i(s|T_1(k), T_2(k)) \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i}{\theta}} [\ln(\bar{\gamma})]^r \\ &\sim c_{i_0}(s|T_1(k), T_2(k)) \left(\frac{1}{\bar{\gamma}}\right)^{\frac{i_0}{\theta}} [\ln(\bar{\gamma})]^r, \end{aligned} \tag{3.34}$$

where i_0 is an integer given by $i_0 = \arg\min_{n\in\mathbb{N}} \{c_n(s|T_1(k), T_2(k)) \neq 0\}$; θ is a nonzero natural number; $\{c_i(s|T_1(k), T_2(k))\}$ represents a series of functions of s, given $T_1(k), T_2(k)$; $r \in \mathbb{N}$, then combined selection is able to achieve an outage probability equivalent to conventional per-subcarrier selection as $\bar{\gamma} \to \infty$.

This proposition can be similarly proved as the one proposed for uncorrelated channels. Then, by integrating the correlated terms, we can have the form of the asymptotic expression of outage probability given by combined and per-subcarrier selection below:

$$\{F_{comb}(s), F_{ps}(s)\} \sim K\left(\frac{1}{\mu_1\mu_2}\right)\left\{\left(\frac{1}{\bar{\gamma}}\right)^{\frac{i_0}{\theta}}[\ln(\bar{\gamma})]^r\right\}^M \int_0^\infty\int_0^\infty \{c_{i_0}(s|T_1,T_2)\}^M\left(e^{-\frac{T_1}{\mu_1}}e^{-\frac{T_2}{\mu_2}}\right)\mathrm{d}T_1\mathrm{d}T_2. \tag{3.35}$$

We apply (3.35) to DF, FG AF and VG AF relay over spatially correlated channels and obtain the following asymptotic expressions for outage probability:

$$\{F_{comb}(s), F_{ps}(s)\} \sim K(1-\rho_1)(1-\rho_2)\left(\frac{s}{2\bar{\gamma}}\right)^M \sum_{m=0}^{M}\binom{M}{m}\frac{\left(\frac{1}{\psi_1}\right)^{M-m}\left(\frac{1}{\psi_2}\right)^m}{[1+\rho_1(M-m-1)][1+\rho_2(m-1)]}, \tag{3.36}$$

$$\{F_{comb}(s), F_{ps}(s)\} \sim \frac{K(1-\rho_1)(1-\rho_2)}{[1+\rho_1(M-1)][1+\rho_2(M-1)]}\left[\frac{\mu_1 s}{4\psi_1\psi_2\bar{\gamma}}\ln(\bar{\gamma})\right]^M, \tag{3.37}$$

and

$$\{F_{comb}(s), F_{ps}(s)\} \sim K(1-\rho_1)(1-\rho_2)\left(\frac{s}{2\bar{\gamma}}\right)^M \sum_{m=0}^{M}\binom{M}{m}\frac{\left(\frac{1}{\psi_1}\right)^{M-m}\left(\frac{1}{\psi_2}\right)^m}{[1+\rho_1(M-m-1)][1+\rho_2(m-1)]}. \tag{3.38}$$

Note, the unconditional outage probability of VG AF relay systems (c.f. (3.38)) is exactly the same as that given in DF relay systems (c.f. (3.36)), which aligns with the numerical results presented in [60]. Numerical results are also shown in Figure 3.13, 3.14 and 3.15 to very out analysis and the effectiveness of this methodology.

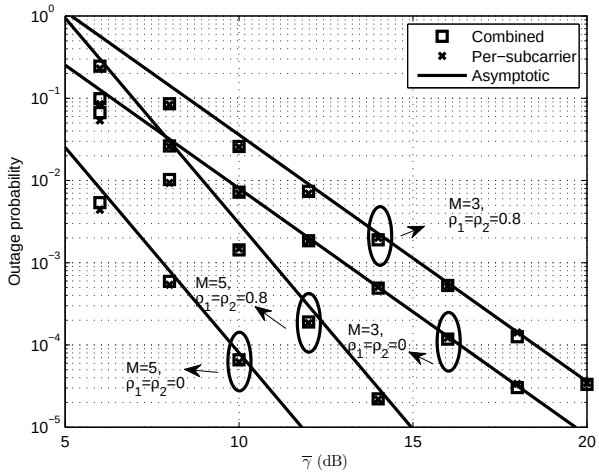


Figure 3.13: Two-hop DF relay selection case: outage probability vs. SNR for per-subcarrier and combined bulk/per-subcarrier selection systems.

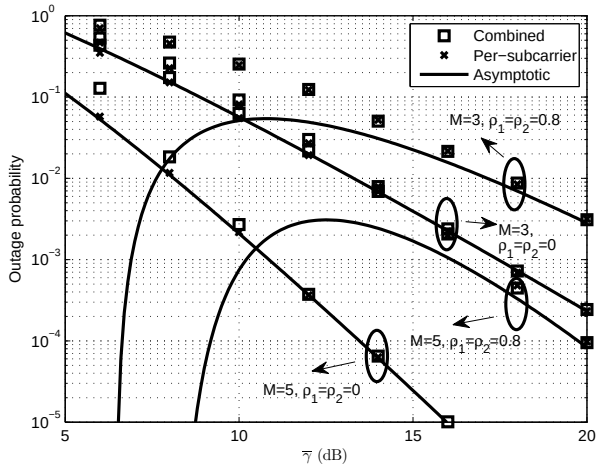


Figure 3.14: Two-hop FG AF relay selection case: outage probability vs. SNR for per-subcarrier and combined bulk/per-subcarrier selection systems.

Summary

In this section, we considered the strategy of combined bulk/per-subcarrier relay selection for DF, FG AF, and VG AF cooperative relay networks. Through analysing the outage probability at high SNR, we have shown that combined relay selection satisfies the *equivalence principle*, i.e. it achieves the same performance at high SNR as per-subcarrier selection, but where only two out of M relays are active rather than the full set. Furthermore, a generalised situation without specifying the relaying protocol is also analysed in this section, and we proved that the equivalence principle is generally valid as long as

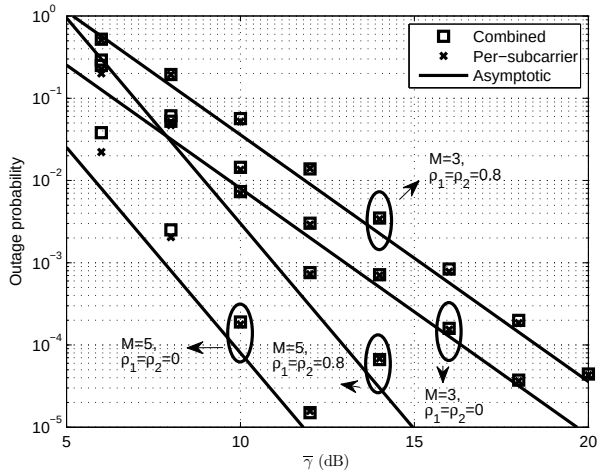


Figure 3.15: Two-hop VG AF relay selection case: outage probability vs. SNR for per-subcarrier and combined bulk/per-subcarrier selection systems.

the CDF of the i.i.d. end-to-end SNR, $F(s)$, can be expanded as the form given in (3.34). Hence, it is possible to achieve very good performance at a reduced system complexity. Numerical results validated our analysis and furthermore demonstrated that the equivalence principle appears to hold in spatially correlated channels as well. Moreover, we also proposed a systematic methodology to analyse the outage performance of combined selection over spatially correlated channels and prove its effectiveness.

3.2.3 Comparison of multicarrier relay selection schemes in super dense cloud networks

In this section, we critically compare the performances of bulk, per-subcarrier and combined selection schemes in super dense cloud networks. Specifically, their outage and contention probabilities are critically appraised. To analyse the effects of outage and contention together, a performance parameter termed the *bulk gain factor* is defined and employed. Finally, we find that although the bulk selection scheme yields a worse outage performance, it would still be preferable in super dense cloud networks as long as a sufficiently large transmit power is employed, since it only demands one relay and has a lower contention probability. Note, this conclusion is constructed based on the fact that when the bulk selection scheme is applied, the contention probability is still much higher than the outage probability in super dense cloud networks. Our results serve as a design benchmark for physical layer applications/protocols used in super dense cloud networks. Also, with these results, we provide a theoretical basis to further optimise super dense cloud networks by reducing contention probability. Besides, the outage performance of AF relay in super dense cloud networks is analysed in particular in this section as well.

Fundamentals

Two-hop OFDM relay system. In this section, we consider a multiuser OFDM-based two-hop relay system in which the set of U *user pairs* (source-destination pairs) is denoted by $\mathcal{U}$; the set of M relays is denoted by $\mathcal{M}$; the set of N subcarriers is denoted by $\mathcal{N}$. Meanwhile, the channel coefficient between the uth source and the mth relay for the nth subcarrier is denoted by $h(u, m, n)$. Similarly, the channel coefficient between the mth relay and the uth destination for the nth subcarrier is denoted by $g(u, m, n)$, $\forall u \in \mathcal{U}$, $m \in \mathcal{M}$, $n \in \mathcal{N}$. Furthermore, we assume that $\forall u \in \mathcal{U}$, $m \in \mathcal{M}$, $n \in \mathcal{N}$, both $h(u, m, n)$ and $g(u, m, n)$ are i.i.d. as $\mathcal{CN}(0, \mu_1)$ and $\mathcal{CN}(0, \mu_2)$, respectively. The additive noise at the mth relay and the uth destination on the nth subcarrier are denoted by $v(m, n)$ and $\eta(u, n)$, which are i.i.d. as $\mathcal{CN}(0, N_0)$. Finally the end-to-end instantaneous SNR on the nth subcarrier, for the uth user pair and forwarded by the mth relay is denoted by $\mathsf{SNR}(u, m, n)$.

Also, we assume all communication nodes in this multiuser two-hop OFDM relay system are perfectly synchronised in both time and frequency domains. The transmit power per subcarrier at all nodes is identical and denoted by P_t. The entire network operates in a half duplex manner and thus two orthogonal time slots are required for each complete transmission.

Relay selection in super dense cloud networks. *Bulk selection* For the uth source, bulk selection scheme only selects *one* out of M relays in the relay network according to the selection criterion [48]

$$\mathcal{L}_u^{bulk} = \arg \max_{m \in \mathcal{M}} \min_{n \in \mathcal{N}} \mathsf{SNR}(u, m, n) \tag{3.39}$$

where $\mathcal{L}_u^{bulk}$ is the set (of cardinality one) denoting the selected relay for the uth source.

Per-subcarrier selection For the uth source, per-subcarrier selection scheme selects L relays from M relays in a per-subcarrier manner and $1 \leq L \leq N$. Therefore, for the uth source, the relay corresponding to the nth subcarrier is individually selected by [48]

$$l_u^{ps}(n) = \arg \max_{m \in \mathcal{M}} \mathsf{SNR}(u, m, n) \tag{3.40}$$

where $l_u^{ps}(n)$ is the index of the selected relay for the uth source corresponding to the nth subcarrier. Then, this selection process will be repeatedly applied for all subcarriers and finally L relays are selected. The set of all L selected relays is denoted by $\mathcal{L}_u^{ps}$.

Combined selection For the uth user, the combined selection scheme first selects *two* out of M relays according to the criterion [56]

$$\mathcal{L}_u^{comb} = \arg \max_{\mathcal{L}_2 \subseteq \mathcal{M}} \min_{n \in \mathcal{N}} \max_{l \in \mathcal{L}_2} \mathsf{SNR}(u, l, n) \tag{3.41}$$

where $\mathcal{L}_u^{comb}$ is the set consisting of two optimally selected relays and $|\mathcal{L}_u^{comb}| = 2$; $\mathcal{L}_2$ identifies a pair of relays that can be employed to carry out per-subcarrier selection and $|\mathcal{L}_2| = 2$. Obviously, we have $M(M-1)/2$ available options of $\mathcal{L}_2$.

Then, per-subcarrier selection is performed over the two selected relays in $\mathcal{L}_u^{comb}$ for each subcarrier by

$$l_u^{comb}(n) = \arg \max_{\in \mathcal{L}_u^{comb}} \mathsf{SNR}(u, l, n) \tag{3.42}$$

where $l_u^{comb}(n)$ is the index of the selected relay corresponding to the nth subcarrier.

In order to facilitate the following analysis, we use $\mathcal{L}_u$ to denote the set of selected relays for the uth user pair when discussing a general case of relay selection. Also, we use $L = |\mathcal{L}_u|$ to denote the magnitude of the subset $\mathcal{L}_u$.

Outage. We define the *a priori* outage probability, without conditioning on the selection process, for the uth user's nth subcarrier forwarded via the mth relay as

$$F(s) = \mathbb{P}(\mathsf{SNR}(u, m, n) < s), \tag{3.43}$$

where s is an end-to-end SNR threshold. The probability density function (PDF) corresponding to $F(s)$ is denoted as $f(s)$. We consider outage at a network level and define the total network outage event of a multiuser OFDM relay system after selection in the following manner.

Definition 3. *A total network outage event will occur when*

$$\min_{u \in \mathcal{U}} \min_{n \in \mathcal{N}} \max_{l \in \mathcal{L}_u} \mathsf{SNR}(u, l, n) < s. \tag{3.44}$$

Furthermore, we also define the system outage probability by

$$P_o(s) = \mathbb{P}\left(\min_{u \in \mathcal{U}} \min_{n \in \mathcal{N}} \max_{l \in \mathcal{L}_u} \mathsf{SNR}(u, l, n) < s \right). \tag{3.45}$$

Contention. In the multiuser scenario, multiple source-destination pairs wish to communicate simultaneously via the relay cluster in a frequency-division multiple access (FDMA) manner. In this extension, it is assumed that a relay is only able to aid the transmission of a single user pair. An immediate consequence of this is that if multiple users contend for the same relay, the selection mechanism will fail, as illustrated in Figure 3.9. We can define the selection contention event in the following way [56]

Definition 4. Selection contention *is said to have occurred if*

$$\bigcap_{u=1}^{U} \mathcal{L}_u \neq \varnothing. \tag{3.46}$$

In other words, in order to prevent selection contention, all U bulk selection subsets $\mathcal{L}_u$ produced by (3.39), (3.40) or (3.41), $\forall u \in \mathcal{U}$ should be disjoint. The contention probability is thereby defined as

$$P_c = \mathbb{P}\left(\bigcap_{u=1}^{U} \mathcal{L}_u \neq \varnothing\right). \tag{3.47}$$

System failure. As introduced previously, there exist two mechanisms that could collapse a multiuser OFDM relay system: outage events or contention events. To consider both effects together, we define the event of system failure as the union of the contention and outage events as defined in Definitions 3 and 4. Therefore, the system failure probability is given by

$$P_f = P_o(s) + P_c - P_o(s)P_c. \tag{3.48}$$

Super dense cloud network. Following the definition of a super dense cloud network given in [56], we propose a more robust definition of a super dense cloud network here.

Definition 5. *A network is said to be super dense when*

$$M \gg 2U^2 \tag{3.49}$$

and

$$M \gg N^2/2. \tag{3.50}$$

We adopt these two relations throughout this section. The motivation for (3.50) will become clear below.

Performance analysis

Outage performance. In this subsection, we temporarily omit the effect of selection contention and only analyse the outage performance of these three selection schemes.

Bulk selection By the theory of order statistics [57], it is straightforward to derive the outage probability of an OFDM relay system equipped with bulk selection to be

$$P_o^{bulk}(s) = 1 - \left\{1 - \left[1 - (1 - F(s))^N\right]^M\right\}^U. \tag{3.51}$$

Per-subcarrier selection Similarly, we can obtain the system outage probability produced by per-subcarrier selection by

$$P_o^{ps}(s) = 1 - \left[1 - (F(s))^M\right]^{NU}. \tag{3.52}$$

Combined selection According to *Lemma 1* in [46], for the uth user, the PDF of the outage probability at high SNR is

$$f_u(s) = MN(F(s))^{M-1}(1-F(s))^{MN-M}f(s). \tag{3.53}$$

The corresponding cumulative distribution function (CDF) is thus

$$F_u(s) = \int_0^s f_u(t)\mathrm{d}t = MN\mathrm{B}_{F(s)}\left(M, 1+M(N-1)\right) \tag{3.54}$$

where $\mathrm{B}_x\left(a,b\right)$ is the incomplete beta function and can be expressed as

$$\mathrm{B}_{F(s)}\left(M, 1+M(N-1)\right) = \int_0^{F(s)} t^{M-1}(1-t)^{M(N-1)}\mathrm{d}t. \tag{3.55}$$

Therefore, using order statistics, we can obtain the CDF of the system outage event at high SNR by

$$P_o^{comb}(s) = 1-\left[1-F_u(s)\right]^U = 1-\left[1-MN\mathrm{B}_{F(s)}\left(M, 1+M(N-1)\right)\right]^U. \tag{3.56}$$

Asymptotic outage performance In order to study the quantitative relation among P_f, $P_o(s)$ and P_c, we need to specify a relaying protocol so that a certain $F(s)$ can be given. In this section, we choose the DF relaying protocol as an example to examine the quantitative relation among P_f, $P_o(s)$ and P_c. The reason for choosing DF relaying is because of its simplicity and good outage performance compared to other types of relaying. The outage probability $F(s)$ in the DF relay network is given by

$$F(s) = 1-e^{-s(1/\mu_1+1/\mu_2)/\overline{\gamma}}, \tag{3.57}$$

where $\overline{\gamma} = P_t/N_0$ is the normalised SNR.

Therefore, by the Taylor series expansion, we can expand the system outage probability produced by bulk selection given in (3.51) at high SNR ($\overline{\gamma}\to\infty$) as

$$P_o^{bulk}(s) < \widetilde{P}_o^{bulk}(s) \sim \left[\frac{U^{\frac{1}{M}}N(\mu_1+\mu_2)s}{\mu_1\mu_2\overline{\gamma}}\right]^M \tag{3.58}$$

where $\widetilde{P}_o(s)$ represents the asymptotic expression pertaining to $P_o(s)$ for a large $\overline{\gamma}$.

Similarly, employing the Taylor series expansion yields the asymptotic expression of (3.52)

$$P_o^{ps}(s) < \widetilde{P}_o^{ps}(s) \sim \left[\frac{U^{\frac{1}{M}}N^{\frac{1}{M}}(\mu_1+\mu_2)s}{\mu_1\mu_2\overline{\gamma}}\right]^M. \tag{3.59}$$

Also, it is straightforward to obtain the asymptotic expression of (3.56) by the same method

$$P_o^{comb}(s) < \tilde{P}_o^{comb}(s) \sim \left[\frac{U^{\frac{1}{M}} N^{\frac{1}{M}} (\mu_1 + \mu_2) s}{\mu_1 \mu_2 \overline{\gamma}}\right]^M . \tag{3.60}$$

The same asymptotic expressions obtained by using per-subcarrier selection and combined selection (cf. (3.59) and (3.60)) prove that the outage performances of per-subcarrier selection and combined selection are identical at high SNR, even for such a multiuser scenario in super dense cloud networks. Therefore, assuming the same network configuration, we have

$$P_o^{pt}(s) < P_o^{comb}(s) < P_o^{bulk}(s). \tag{3.61}$$

Contention performance. For each user, L relays will be selected and thus UL relays will be selected in total if selection contention does not occur. Therefore, the relation $M \geq UL$ should always be satisfied. Otherwise, it is impossible to circumvent selection contention. According to our definition of a super dense cloud network (cf. (3.49) and (3.50)), $M \geq UL$ can always be met in super dense cloud networks.

For bulk selection, $L = 1$ and the contention probability is thereby

$$P_c^{bulk} = 1 - \prod_{u=1}^{U} \frac{M - u + 1}{M} = \frac{U(U-1)}{2M} + O\left(\frac{1}{M^2}\right). \tag{3.62}$$

For combined selection, $L = 2$ and the contention probability is thereby [56]

$$P_c^{comb} = 1 - \prod_{u=1}^{U} \frac{(M - 2u + 1)(M - 2u + 2)}{M(M-1)} = \frac{2U(U-1)}{M} + O\left(\frac{1}{M^2}\right). \tag{3.63}$$

However, the number of selected relays L by per-subcarrier selection is not so obvious, since L is a variable varying between 1 and N in this scenario [64]. To facilitate our analysis pertaining to contention probability when per-subcarrier selection is applied, we have the following lemma.

Lemma 1. *According to Definition 5, we have $L \approx N$ for the per-subcarrier selection scheme.*

This lemma can be proved as follows. The probability that the number of selected relays is equal to the number of subcarriers is

$$\mathbb{P}(L = N) = \frac{\binom{M}{N} N!}{M^N} = 1 - \frac{N^2 - N}{2M} + O\left(\frac{1}{M^2}\right). \tag{3.64}$$

Therefore, assuming $M \gg N^2/2$ (cf. (3.50)), we have

$$\mathbb{P}(L = N) \approx 1. \tag{3.65}$$

As a consequence, the approximation of L is given by

$$L \approx N. \tag{3.66}$$

From *Lemma 1*, the contention probability of per-subcarrier selection can be approximated to

$$P_c^{ps} \approx 1 - \prod_{u=1}^{U} \prod_{l=1}^{N} \frac{M - Nu + l}{M + 1 - l} = \frac{N^2 U(U-1)}{2M} + O\left(\frac{1}{M^2}\right). \tag{3.67}$$

Compared to the asymptotic expressions given in (3.62), 3.63 and (3.67), it is obvious that $\forall N > 2$, we have

$$P_c^{bulk} < P_c^{comb} < P_c^{pt} \tag{3.68}$$

Simulations

Simulation configurations. To verify the analysis in the previous section for both outage and contention, we mainly employ analytical plots to study the quantitative relations among $P_o(s)$, $\overline{\gamma}$ and M[5]. One should note that because the correctness and accuracy of the models that were used in the analysis have been numerically verified in [55, 56], we can directly employ them to examine the quantitative relations without performing numerical simulations.

Three simulations are carried out to examine:

1. the quantitative relation among outage probability $P_o(s)$, normalised SNR $\overline{\gamma}$ and the number of relays M for bulk and per-subcarrier selections;

2. the relation among contention probability P_c, the number of relays M and the number of users U for bulk and combined selections;

3. the priority of bulk selection in terms of the number of relays M and the normalised SNR $\overline{\gamma}$.

As for the configurations of constant parameters, we set $N = 8$, $\mu_1 = \mu_2 = 1$, $N_0 = 1$ and $s = 1$.

[5]Most figures in this section are generated by theoretical expressions, rather than numerical results.

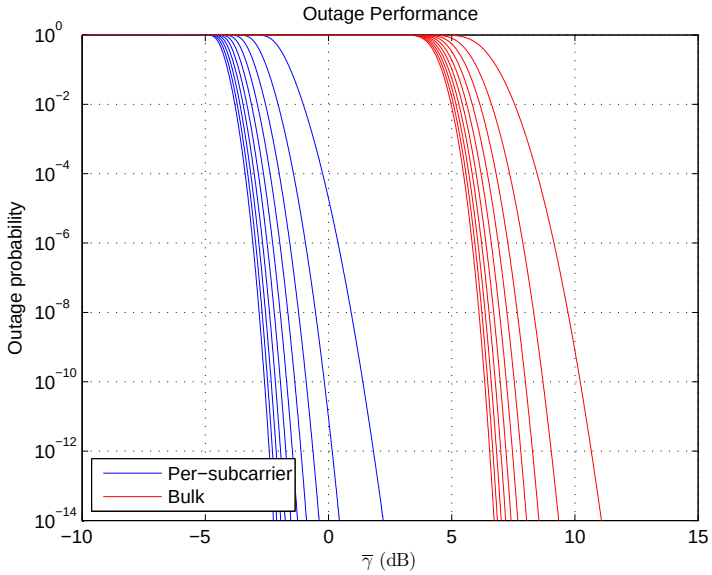


Figure 3.16: Analytical outage performances of per-subcarrier and bulk selections. The different curves correspond to $M \in \{100, 200, \cdots, 1000\}$ running from top to bottom.

Simulation of outage. As we analyse above in (3.61), the optimal outage performance is given by per-subcarrier selection. Therefore, to observe the outage performance of bulk selection, it is only necessary to compare it with the outage performance given by per-subcarrier selection.

For $M \in [100, 1000] \cap \mathbb{Z}$ and $U = 5$, the expressions of bulk and per-subcarrier selection schemes are presented in Figure 3.16. From this figure, it is clear that as long as the normalised SNR $\overline{\gamma}$ can be larger than approximate 6 dB[6], the outage probability $P_o^{bulk}(s)$ produced by bulk selection is sufficiently small. These analytical plotted results indicate that although per-subcarrier selection has a much better outage performance than bulk selection, the outage performance produced by bulk selection is still satisfactory in the super dense cloud network where M is sufficiently large, provided $\overline{\gamma} \gtrapprox 6$ dB. Therefore, these results provide an approach to further reduce the OFDM relay system's complexity by decreasing the number of demanded relays from two to one.

Simulation of contention. Now, we should examine the order of magnitude of contention probability in the super dense cloud network. As we analysed above in (3.68), the contention performance of per-subcarrier selection is much worse than the contention performance of combined selection so that we can simply compare the contention performances between bulk and combined selection. Let $U \in \{5, 10\}$ and vary

[6]This value depends on U, N, M, s, μ_1 and μ_2 as well as the required satisfactory outage probability.

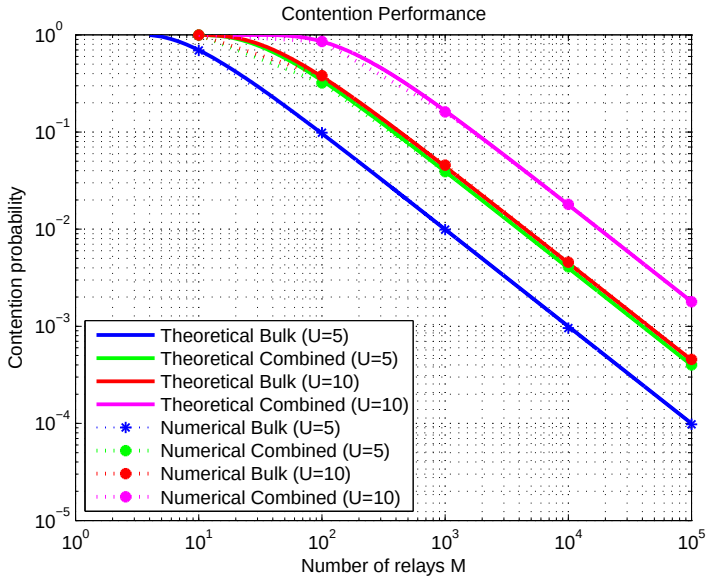


Figure 3.17: Analytical and numerical contention performance, given $U = 5, 10$.

$M \in \{10^1, 10^2, \cdots, 10^5\}$ for numerical simulations. The simulation results are illustrated in Figure 3.17. From this figure, two key points can be summarised. First, quantitatively, $P_c \gg P_o(s)$ is valid for bulk selection in super dense cloud networks, provided $\overline{\gamma} \gtrapprox 6\,\text{dB}$. Second, the advantage of bulk selection over combined selection on contention performance is obvious. Therefore, considering the outage performance provided by bulk selection is still good enough, it is the most appropriate selection scheme in super dense cloud networks as long as a sufficient transmit power is provided.

Simulation of system failure. To provide a more reliable result to verify the advantage of bulk selection over per-subcarrier and combined selections in super dense cloud networks, we compare the system failure performances provided by these three selection schemes in this subsection. To critically compare the system failure performance, we need to define a variable α termed *bulk gain factor* by

$$\alpha := 10\log_{10}\left(\frac{P_o^{ps}(s) + P_c^{comb} - P_o^{ps}(s)P_c^{comb}}{P_f^{bulk}}\right). \tag{3.69}$$

From the definition of bulk gain factor, we can see that the system failure probability produced by bulk selection is compared to the combination of the outage probability produced by per-subcarrier selection and the contention probability produced by combined selection. This is an original way to compare three selection schemes jointly, by which the priority of bulk selection over the other two selection schemes in super dense cloud networks can be critically analysed. Specifically, according to (3.61) and (3.68),

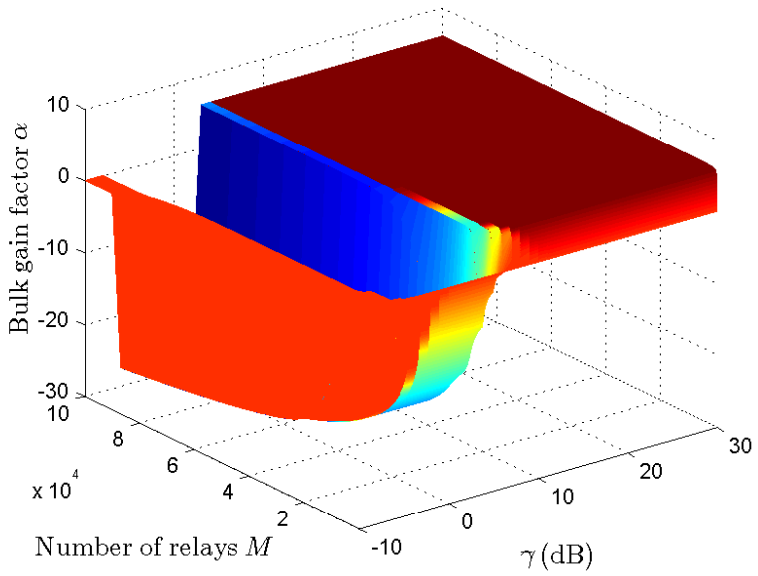


Figure 3.18: Analytical relation among α, $\overline{\gamma}$ and M, given $U = 10$. 1) Brown region: bulk selection is preferable; 2) blue region: bulk selection is not preferable; 3) red region: no much difference among all three selection schemes

$10\log_{10}(P_f^{ps}/P_f^{bulk}) > \alpha$ and $10\log_{10}(P_f^{comb}/P_f^{bulk}) > \alpha$ are always valid in the super dense cloud network. Hence, when $\alpha > 0$, it is clear that bulk selection is preferable to both combined and per-subcarrier selections, since $10\log_{10}(P_f^{ps}/P_f^{bulk}) > \alpha > 0$ and $10\log_{10}(P_f^{comb}/P_f^{bulk}) > \alpha > 0$. The relation among α, $\overline{\gamma}$ and M is simulated and shown in Figure 3.18, given $U = 10$. To be clear, the projection of α in the $\overline{\gamma} - M$ plane is illustrated in Figure 3.19 as well. From these figures, we can see that above $\overline{\gamma} \approx 6$ dB, bulk selection scheme will have an evident advantage over other two selection schemes in terms of system failure performance. However, with the increase of $\overline{\gamma}$, this advantage will not be enhanced accordingly, while it will maintain at a certain level for a given M. Meanwhile, decreasing M will lead α being increased by an insignificant amount.

Outage performance analysis of AF relays in super dense cloud networks

As the case of DF relay has been analysed above, we now analyse the case of FG AF relay in super dense cloud networks in terms of outage performance. As long as its outage performance is obtained, the failure performance can be obtained by exactly the same way, and we will not detail it in this section.

Again, an outage on a particular tone is considered to have occurred if the corresponding end-to-end SNR is lower than s. Thus, for any given subcarrier, the *a priori* outage

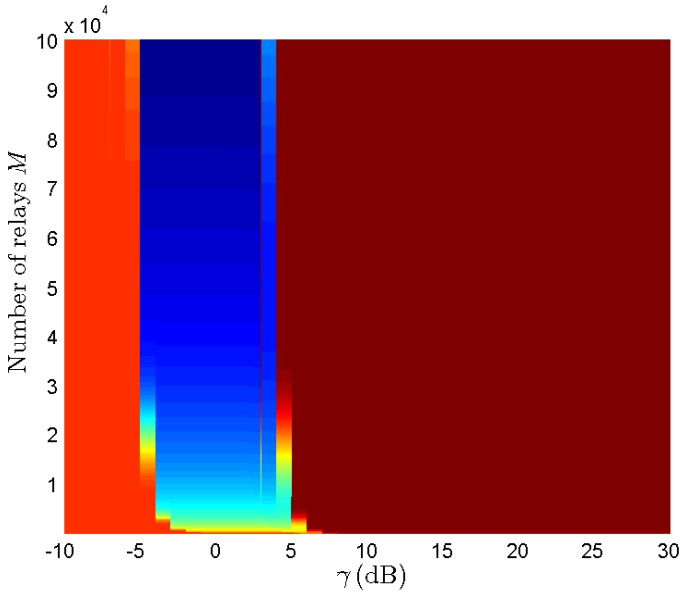


Figure 3.19: Projection of α in x-y plane (analytical relation), given $U = 10$. 1) Brown region: bulk selection is preferable; 2) blue region: bulk selection is not preferable; 3) red region: no much difference among all three selection schemes

probability, i.e. the outage probability not conditioned on any form of relay selection having taken place, can be alternatively expressed as

$$F(\xi) = \mathbb{P}\left(\overline{\mathsf{SNR}}(u,m,n) < \xi\right) \tag{3.70}$$

where

$$\xi = s/\bar{\gamma} \tag{3.71}$$

and $\overline{\mathsf{SNR}}(u,m,n) = \mathsf{SNR}(u,m,n)/\bar{\gamma}$ is the end-to-end SNR normalised by the average end-to-end SNR $\bar{\gamma}$. Thus, the function $F(\,\cdot\,)$ is now defined as the cumulative distribution function for the normalised end-to-end SNR.

To study the performance of bulk/per-subcarrier relay selection in super dense cloud networks, we consider the worst-case outage probability of a user. It is known that in bulk/per-subcarrier selection schemes, the worst possible SNR is the Mth worst out of the MN space-frequency channels available, see Lemma1 of [46]. Let $\mathcal{E}$ denote the event that the subcarrier in question corresponds to the Mth worst. We require the density function of the Mth worst normalised end-to-end SNR conditioned on $\mathcal{E}$. Using the theory of order statistics [46, 57], this density can be written as

$$f_{\min}(x|\mathcal{E}) = M\binom{MN}{M}F(x)^{M-1}(1-F(x))^{M(N-1)}f(x) \tag{3.72}$$

where $f(x)$ is the unconditional probability density function for the normalised end-to-end SNR on any given subcarrier. Clearly, the functions f and F depend on the particular

relaying scheme in question, e.g. fixed-gain AF or variable-gain AF. For now, we refrain from defining these functions for the sake of generality. In any case, we will see that our results are universal, i.e. they apply to all relay schemes.

Focusing on the nth single subcarrier, where n can be arbitrarily chosen, we note that the *a posteriori*[7] outage probability $P_o(n)$ satisfies the inequality

$$P_o(n) \leq \overline{P}_o(\xi) = \int_0^{\xi} f_{\min}(x|\mathcal{E})\,\mathrm{d}x. \tag{3.73}$$

Substituting (3.72) into this equation and making the change of variable $y = F(x)$ leads to

$$\overline{P}_o(\xi) = M\binom{MN}{M}\sum_{k=0}^{M(N-1)}\binom{M(N-1)}{k}\frac{(-1)^{-k}}{M+k}F(\xi)^{M+k} \tag{3.74}$$

which can be written succinctly as

$$\overline{P}_o(\xi) = \binom{MN}{M}F(\xi)^M {}_2F_1(M, -M(N-1); M+1; F(\xi)) \tag{3.75}$$

where ${}_2F_1(a, b; c; z)$ is the Gauss hypergeometric function.

For a super dense cloud network, M is large. The inherent diversity in the network effectively pushes this bound very low for certain parameterisations of N and $F(\xi)$, thus implying the actual outage probability of a system with no contention is negligible. To visualise this situation, consider the case where $M = 100$. A graph depicting the bound for this case is illustrated in Figure 3.20. The shaded regions in this graph indicate the magnitude of the outage bound $\overline{P}_o(\xi)$, with darker regions corresponding to lower probabilities. This figure suggests that, for a given N, there is a critical probability, call it $F_N(\xi)$, above which the bound is no longer tight, which further implies the outage probability of the system could be poor. The critical value of the average SNR corresponding to $F_N(\xi)$ is given by

$$\bar{\gamma} = \frac{s}{F^{-1}(F_N(\xi))}. \tag{3.76}$$

Progress can be made towards characterising the critical probability $F_N(\xi)$ by considering the leading order of (3.75). Letting M grow large, we can expand the gamma functions and retain the leading order of each term in the hypergeometric series to obtain

$$\overline{P}_o(\xi) = e^{-E(M,N,\xi)}(1 + O(1/M)) \tag{3.77}$$

[7]The term *a posteriori* refers to the fact that we now focus on the outage probability on a selected relay/subcarrier pair, i.e. after combined bulk/per-subcarrier relay selection has taken place.

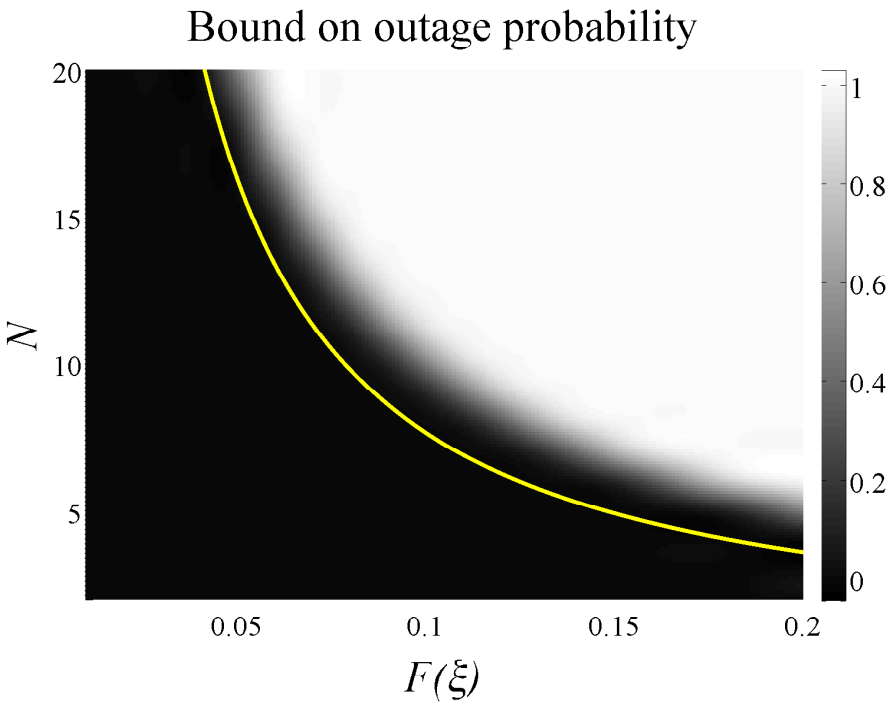


Figure 3.20: Outage probability bound $\overline{P}_o(\xi)$ given in (3.75) as a function of the number of subcarriers N and the independent subcarrier outage probability $F(\xi)$ with $M = 100$. The dark areas indicate regions where the outage is small/negligible, while the light regions show where the bound approaches one. The yellow line denotes the asymptotic (in M) phase transition described by (3.80) where the bound switches from decaying exponentially with M to increasing rapidly to one.

where

$$E(M,N,\xi) = \frac{1}{2}\log M - M\Bigg(\log N - (N-1)F(\xi) + \log F(\xi) - (N-1)\log\left(1-\frac{1}{N}\right)\Bigg) + 1 + \frac{1}{2}\log\left(1-\frac{1}{N}\right) + \frac{1}{2}\log 2\pi. \quad (3.78)$$

For the outage bound to decrease with M, we must have $E(M,N,\xi) > 0$. In the limit of large M, this condition becomes

$$\log N - (N-1)F(\xi) + \log F(\xi) - (N-1)\log\left(1-\frac{1}{N}\right) < 0. \quad (3.79)$$

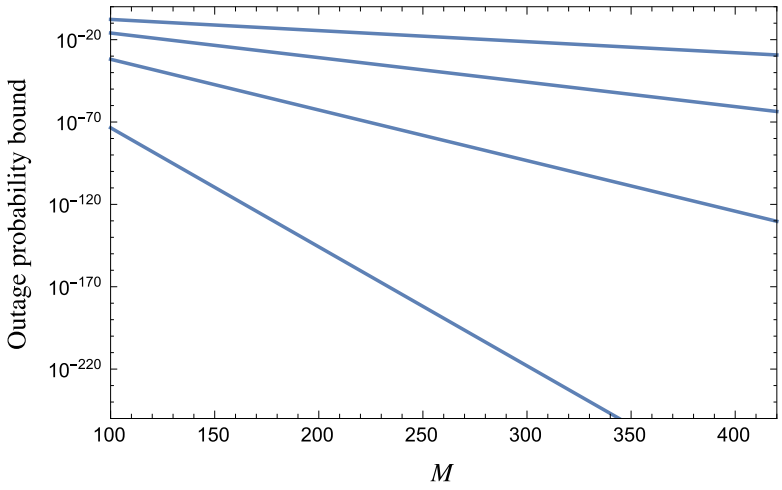


Figure 3.21: Relationship between $\overline{P}_o(\xi)$ and M for $N = 8$. The different curves correspond to $F(\xi) \in \{0.01, 0.03, 0.05, 0.07\}$ running from bottom to top.

Focusing on the critical point where this inequality is just satisfied, i.e. it is an equality, will yield a relation between $F(\xi)$ and N. In this case, solving for $F(\xi)$ gives the critical probability

$$F_N(\xi) = -\frac{1}{N-1} W\left(-\left(1-\frac{1}{N}\right)^N\right) \tag{3.80}$$

where $W(z)$ is the solution to $We^W = z$ (Lambert's function). Eq. (3.80) is plotted in Figure 3.20, where it can be seen that it provides an excellent approximation to the phase transition mentioned above. Trivially, this expression combined with (3.76) gives a lower bound on the average end-to-end SNR, above which an increase in M will result in an arbitrarily small outage probability for the system.

To illustrate the behaviour of $\overline{P}_o(\xi)$, as defined by (3.75), when M varies, we have plotted the outage bound against M for various values of $F(\xi)$ in Figure 3.21. From this figure, the increase in the slope of the bound as $F(\xi)$ increases is apparent. Additionally, the very low values of $\overline{P}_o(\xi)$ that can be attained are also clear.

As a final point of discussion, we will provide a feeling for what values the critical average end-to-end SNR might take for a practical system. Consider the case where $N = 8$ and FG AF relaying is used for each selected relay-subcarrier pair. Assume $s = 1$ and $P_r = P_t$ for this example, so that $\xi = 1/\bar{\gamma}$. We can invoke (3.80) to obtain:

$$F_8(\xi) = 0.0964. \tag{3.81}$$

Furthermore, the closed-form expression of $F(\xi)$ for FG AF relaying given the parameters above is [58]

$$F(\xi) = 1 - 2e^{-\frac{s}{\bar{\gamma}}} \sqrt{\frac{C\,s}{\bar{\gamma}^2}} K_1\left(2\sqrt{\frac{C\,s}{\bar{\gamma}^2}}\right) \tag{3.82}$$

where $C = \bar{\gamma}/G^2$. Using (3.80), we find that the critical average SNR value is

$$\bar{\gamma} \approx 17 \text{ dB}. \tag{3.83}$$

Consequently, we can be sure that for systems where $\bar{\gamma} > 17$ dB, the outage probability can be made arbitrarily small by choosing M to be large enough.

Summary

In this section, we critically compared the performances of bulk, per-subcarrier and combined selection schemes in super dense cloud networks. The main contribution of the work in this section is that we found that although the bulk selection scheme yields a worse outage performance, it would still be preferable in super dense cloud networks, since it only demands one relay and has a smaller contention probability which dominates the system failure probability as long as a sufficiently large transmit power can be provided. This result provides an approach to further optimise the system failure performances and simultaneously reduce the system complexities of the applications/protocols used in super dense cloud networks. Besides, the outage performance of AF relay in super dense cloud networks was analysed in particular in this section as well, and we obtained a tight upper bound on the outage probability produced by combined selection for AF relay in super dense cloud networks.

3.2.4 Conclusions

In this report, we presented the work which has been done during the DIWINE funding period. The completed work presented in Sections 3.2.2 and 3.2.3 can be summarised as follows. We extended the concept of combined selection scheme from transmit antenna selection to relay selection. In particular, we mathematically proved the equivalence principle of combined selection for DF, FG AF and VG AF relay networks when channels are i.i.d. and we illustrated that the equivalence principle can be held for correlated channels by numerical simulations. We proposed a systematic methodology to analyse the outage performance of combined relay selection over spatially correlated channels. We applied combined relay selection in super dense cloud networks and defined the selection contention problem. We deduced the closed-form expression of contention probability of combined selection. We proposed a new and specific definition of the super dense cloud network. Furthermore, we defined the system failure probability to analyse the reliability and performance of a communication system when selection contention exists. In this context, we found out that outage is not a problem for the communication system in super dense cloud networks, compared to contention problem. We also found out that bulk selection may be the most appropriate selection scheme being used in super dense cloud networks, since the it only demands a single relay. Furthermore, we derived upper and lower bounds on outage probability produced by combined selection

and proved its tightness in super dense cloud networks. We defined the bulk gain factor and used it as a powerful tool to compare the failure performance of bulk, per-subcarrier and combined selection schemes. We outlined the essential method to solve the selection contention problem and we analysed the outage performance of combined selection of FG AF relay in super dense cloud networks.

A detailed presentation of the results of these activities can be found in [56, 65]. In particular, in [56] we extend this strategy to the case in which a set of user pairs communicate via a cluster of intermediary AF relaying nodes. We obtain an upper bound on the outage probability of the network, and a closed-form expression for the multiuser contention probability (the probability that distinct user pairs select the same relay). Consequently, we are able to construct an upper bound on the failure probability of the network, i.e. the probability that either an outage or contention occurs. We show that there exists a critical average end-to-end SNR, above which the outage probability decays exponentially with the number of available relays. In this case, the network failure probability is dominated by the contention probability. In [66], we critically compare the performances of bulk, per-subcarrier and combined selection schemes in super dense cloud networks. Specifically, their outage and contention probabilities are critically appraised. Finally, we find that although the bulk selection scheme yields a worse outage performance, it would still be preferable in super dense cloud networks as long as a sufficiently large transmit power is employed, since it only demands one relay and has a lower contention probability. In [67], we mathematically prove that the combined selection strategy is able to achieve an optimal outage probability equivalent to conventional per-subcarrier selection in the high SNR regime without using the full set of available relays for selection. This proof is valid for DF, FG AF and VG AF relays over i.i.d. channels. Moreover, in [65], we further prove the validity of this equivalence principle over spatially correlated channels.

3.3 Degradedness of fast fading Gaussian multiple-antenna wiretap channels with statistical CSIT

3.3.1 Introduction

The key-based enciphering is traditionally used to ensure the security of data transmission. However the key distributions and managements may be challenging tasks [68] for secure wireless communication systems due to the additional control signalling and feedback channel management. On the other hand, the physical-layer security introduced in [69][70] is appealing due to its keyless nature. One of the fundamental problems for physical-layer security is characterising the secrecy capacity for wiretap channels. The secrecy capacity is the maximum achievable secrecy rate between a transmitter and a legitimate receiver, with a secrecy constraint imposed to avoid information leakage to

an eavesdropper[8][69][70]. Further enhancements are attainable by employing multiple antennas at each node, e.g. in [71–73]. In the wireless environments where each node has a single antenna, the time-varying characteristics of fading channels can also be exploited to enhance the secrecy capacity[74]. However, to show the secrecy capacity results of the above works [71–74], at least perfect knowledge of the Bob's channel state information at the transmitter (CSIT) is required. Due to the limited feedback bandwidth and the delay caused by the channel estimation, Alice may not be able to track the channel coefficients if they vary rapidly. Thus, for fast fading channels, it is more practical to consider the case with only partial CSIT of Bob's channel [75–80]. In this case, when Alice has multiple antennas, only the special case that when both Bob's and Eve's channels are Rayleigh distributed (but with different statistics) and the entries of each channel vector are i.i.d., respectively, the ergodic secrecy capacity is known [78]. This result is further extended to the cases in which Bob and Eve have multiple antennas with total and per-antenna power constraints [81]. For an extreme case that the channel gains are unknown for all parties with fast Rayleigh faded Bob's and Eve's channels, capacity can be achieved by discrete channel input [82]. But for more general settings, e.g. Bob's and Eve's channels do not belong to the same type of distribution, only some lower and upper bounds of the secrecy capacity are known [76][79][83][84]. More specifically, although the general secrecy capacity formula was shown in [70], the optimal selection of the auxiliary random variable and channel prefixing in this formula are *still unknown* for the partial CSIT cases in general.

The existing characterisations of secrecy capacities of Gaussian wiretap channels highly depend on the knowledge of CSIT, even though the secrecy capacities can be solved for some rare cases that there is only partial or no CSIT [78][81][82]. However, when there is only statistical CSIT under fast fading, the ergodic secrecy capacity is unknown in general. In this work we investigate this problem from the channel orders point of view. More specifically, even if the instantaneous CSIT is not available, we may still be able to derive the positive secrecy capacity according to the *channel orders*, which are measured *stochastically*. In this work we characterise the ergodic secrecy capacity of the fast fading wiretap channel with multiple antennas at Alice and multiple antennas at both Bob and Eve. We assume that there is only statistical CSIT of both channels at Alice. The main contribution of this work is as follows. We characterise the relation between the existence of the equivalent degraded wiretap channel and the usual multivariate stochastic order. Several commonly considered types of wiretap channels with practical channel distributions are investigated and the conditions to attain the degradedness are derived. These characterisations of ergodic secrecy capacities of fast fading multiple antennas wiretap channels with statistical CSIT are missing in the literature. But they are important since first, in practice there is no reason for Alice to have perfect CSI of Eve's channel. Second, the proposed scheme provides a simple way to determine the ergodic secrecy capacity even if only statistical CSIT is available, which highly simplifies

[8] In the following to simplify the expression we denote the transmitter, the legitimate receiver and eavesdropper by Alice, Bob and Eve, respectively.

the practical system design, e.g. the optimisation over prefixing can be avoided.

Notation: In this work, upper case normal letter denotes random variables, upper case bold letters are either random vectors or matrices, which will be defined when they are first mentioned; lower case bold letters denote vectors. The mutual information between two random variables is denoted by $I(.;.)$. The complementary cumulative density function (CCDF) is denoted by $\bar{F}_X(x) = 1 - F_X(x)$, where $F_X(x)$ is the CDF of X. And we denote the probability mass function (PMF) and probability density function (PDF) by P and f, respectively. $E[.]$ denotes the expectation; $H(.)$ and $h(.)$ denote the entropy and differential entropy, respectively. $X \sim F$ denotes that the random variable X follows the distribution F. vec($\boldsymbol{X}$) concatenates the column vectors of $\boldsymbol{X}$ as a super vector. $\boldsymbol{I}_n$ is the identity matrix with dimension n. The logarithm used in the work is with base 2.

The rest of the work is organised as follows. In Subsection 3.3.2 we introduce the preliminaries. In Subsection 3.3.3 we discuss our main results, i.e. the conditions to have degraded wiretap channel when channel enhancement is not used. Finally, Subsection 3.3.4 concludes this work.

3.3.2 Preliminaries

In this section we first introduce the underlying background knowledge for this work. Then we introduce the known results.

Basic definitions and properties

We first introduce the following definitions, which are important to derive the main results in this work.

Definition 6. *[85, (1.A.2)] For random variables X and Y, $X \geq_{st} Y$, if $\bar{F}_X(x) \geq \bar{F}_Y(x)$, for all x.*

Lemma 2. *[86, Lemma 9.2.1] Let F and G be continuous distribution functions. If X has distribution F then the random variable $G^{-1}(F(X))$ has distribution G.*

Proposition 3. *[86, Proposition 9.2.2] If $X \geq_{st} Y$, then there exist random variables X' and Y' having the same distributions as X and Y, respectively, such that $P(X' \geq Y') = 1$.*[9]

[9] The proof of this result is restated as follows. Let $U \sim Unif(0,1)$ and F be a distribution function. From Lemma 2 we know that $F^{-1}(U) \sim F$, where F^{-1} is the inverse mapping. Furthermore, if $X \geq_{st} Y$, $X \sim F$ and $Y \sim G$, we can form X' and Y' having the same distributions as X and Y, respectively, by $X' = F^{-1}(U)$ and $Y' = G^{-1}(U)$ as shown in Figure 3.22. In addition to the assumption $X \geq_{st} Y$, i.e. $F(x) \leq G(x)$, $\forall x$, it is clear that $P(X' \geq Y') = 1$.

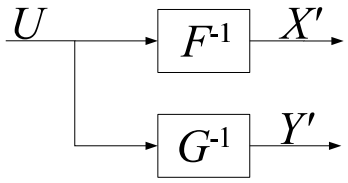


Figure 3.22: The explicit generation of $X' \sim F$, $Y' \sim G$, and $P(X' > Y') = 1$, where $U \sim \text{Unif}(0, 1)$.

The way to find X' and Y' in Proposition 3 is coined as *coupling* [86]. The importance of the technique of coupling is that, in the original model, even we know $X \geq_{st} Y$, the order of x and y may be different in each realisation, which makes the justification of degradedness not easy. However, by the same marginal property of wiretap channels, as long as we can find the random variables X' and Y' which follow the same distributions of X and Y, respectively, and for each realisation $x' > y'$, we attain the degradedness. Proposition 3 can be directly extended to vector case [85, Theorem 6.B.1] coined as the *usual multivariate stochastic order*, where in the corresponding expressions $P(\boldsymbol{X}' \geq \boldsymbol{Y}') = 1$, the inequalities are element-wise.

In addition, since the capacities of wiretap channels only depend on the marginal distributions, in the following discussions and derivations, the *degraded* means the stochastically degraded [68]. Now we introduce the usual stochastic ordering followed by the technique of coupling.

Single antenna case

In this section we review the relation between the existence of the equivalent degraded wiretap channel and the usual stochastic order between Bob's and Eve's channels. This relation is helpful to check the existence of the ergodic secrecy capacity. The considered fast fading wiretap channel with single antenna at all nodes can be specialised from [70] as

$$Y_r = \sqrt{H_r} X + Z_r, \tag{3.84}$$

$$Y_e = \sqrt{H_e} X + Z_e, \tag{3.85}$$

where H_r and H_e are real-valued non-negative independent random variables denoting the square of Bob's and Eve's fading channels with complementary cumulative density functions (CCDF) $\bar{F}_{H_r}$ and $\bar{F}_{H_e}$, respectively. X is the channel input. Without loss of generality, we consider the channel input power constraint as $E[X^2] \leq P_T$. Due to the assumption of full CSI at the receivers, cases with negative Y_r and Y_e can be de-rotated at the receiver. Since this phase rotation is independent to other variables, we can form the equivalent channel model as (3.84) and (3.85) without changing the capacity. The

noises Z_r and Z_e at Bob and Eve, respectively, are independent additive white Gaussian noises (AWGN) with zero mean and unit variance. We assume that Alice only knows the distributions but not the instantaneous realisations of H_r and H_e.

From the proof of Lemma 1 in [87], we can find that with proper selection of the joint distribution, we attain the result that even without perfect CSIT, the comparison of the CCDFs of the fading channels may reflect on the order of the channel magnitudes in probability. More specifically, we can summarise that if $H_r \geq_{st} H_e$, there exists an equivalent (in the sense of having the same ergodic secrecy capacity) wiretap channel formed by H'_r and H'_e, such that $P(H'_r \geq H'_e) = 1$.

Remark 4. *The above discussion can be easily extended from real to complex cases. Assume the noises at Bob and Eve now are circularly symmetric complex Gaussian noises with zero means and unit variances. Since both Bob and Eve know their CSI, the phase rotation due to the complex channel can be compensated at Bob and Eve without changing the secrecy capacity. After this compensation we can form an equivalent complex wiretap channel as*

$$\begin{aligned} Y'_{r,I} + iY'_{r,Q} &= \sqrt{H_r}X_I + Z'_{r,I} + i(\sqrt{H_r}X_Q + Z'_{r,Q}) \\ Y'_{e,I} + iY'_{e,Q} &= \sqrt{H_e}X_I + Z'_{e,I} + i(\sqrt{H_e}X_Q + Z'_{e,Q}), \end{aligned} \tag{3.86}$$

where $X = X_I + iX_Q$, $Z'_r = Z'_{r,I} + iZ'_{r,Q}$ and $Z'_e = Z'_{e,I} + iZ'_{e,Q}$ are the rotated version of Z_r and Z_e. Since Z_r and Z_e are circularly symmetric complex Gaussian noises, Z'_r and Z'_e have the same distributions of Z_r and Z_e, respectively. Then it can be seen that the in-phase and quadrature channels form a pair of identical parallel real wiretap channels as shown in (3.84) *and* (3.85).

3.3.3 Multiple antennas without channel enhancement

Based on the observation from single antenna cases, we aim to extend the description of this relation to cases in which Bob and Eve have multiple antennas, i.e. the case with multiple-input multiple-output multiple-antenna eavesdropper (MIMOME). Without loss of generality, we assume all nodes are equipped with the same number of antennas n_T. The received signals at Bob and Eve can then be respectively expressed as

$$\boldsymbol{Y_r} = \boldsymbol{H_r}\boldsymbol{X} + \boldsymbol{Z_r}, \tag{3.87}$$

$$\boldsymbol{Y_e} = \boldsymbol{H_e}\boldsymbol{X} + \boldsymbol{Z_e}, \tag{3.88}$$

where $\boldsymbol{Z_r} \sim \mathsf{CN}(0, \boldsymbol{I}_{n_T})$ and $\boldsymbol{Z_e} \sim \mathsf{CN}(0, \boldsymbol{I}_{n_T})$, $\boldsymbol{H_r}$ and $\boldsymbol{H_e} \in \mathbb{C}^{n_T \times n_T}$ with entries varying for each code symbol. For the MIMOME case, we apply the vector version of Proposition 3 to the eigenvalues of $\boldsymbol{H_r}^{-1}\boldsymbol{H_r}^{-H}$ and $\boldsymbol{H_e}^{-1}\boldsymbol{H_e}^{-H}$, which are real.

From random matrix theory we know that the probability of the random matrix to be full rank approaches 1 when the dimension of the matrix increases[10]. This implies that the probability of $\mathbf{H}^{-1}\mathbf{H}^{-H}$ having finite eigenvalues approaches 1. On the other hand, we can also construct an alternative channel with full rank which does not change the capacity [72]. Thus, we may assume the channel matrices are invertible when n_T is large enough but not infinity. In addition with the assumption of full channel state information at receiver (CSIR), we can normalise (3.87) and (3.88) equivalently as

$$\boldsymbol{Y}_{\boldsymbol{r}}' = \boldsymbol{X} + \boldsymbol{Z}_{\boldsymbol{r}}', \tag{3.89}$$

$$\boldsymbol{Y}_{\boldsymbol{e}}' = \boldsymbol{X} + \boldsymbol{Z}_{\boldsymbol{e}}', \tag{3.90}$$

where $\boldsymbol{Z}_{\boldsymbol{r}}' \sim \mathsf{CN}(0,\, \boldsymbol{A})$ and $\boldsymbol{Z}_{\boldsymbol{e}}' \sim \mathsf{CN}(0,\, \boldsymbol{B})$, $\boldsymbol{A} \triangleq \boldsymbol{H}_{\boldsymbol{r}}^{-1}\boldsymbol{H}_{\boldsymbol{r}}^{-H}$, $\boldsymbol{B} \triangleq \boldsymbol{H}_{\boldsymbol{e}}^{-1}\boldsymbol{H}_{\boldsymbol{e}}^{-H}$. For full CSIT and full CSIR cases, to make the Markov chain $\boldsymbol{X} \to \boldsymbol{Y}_{\boldsymbol{r}}' \to \boldsymbol{Y}_{\boldsymbol{e}}'$ valid, i.e. it is a (stochastically) degraded wiretap channel, the constraint $\boldsymbol{B} - \boldsymbol{A} \succcurlyeq \mathbf{0}$ is sufficient[11]. In the considered scenario we have full CSIR but only statistical CSIT, so we aim to construct an equivalent degraded channel such that $P(\boldsymbol{B}' - \boldsymbol{A}' \succcurlyeq \mathbf{0}) = 1$ according to the coupling. In the following we aim to find the relation of the degradedness and the stochastic order among the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{B}$. Note that in [85, Theorem 6.B.1] the usual stochastic order in vector (but not matrix) version is considered, where in the expression of $\mathrm{vec}(\boldsymbol{B}) \leq_{st} \mathrm{vec}(\boldsymbol{A})$, the inequality is element-wise, i.e. $b_i' \leq a_i',\ \forall\, i$, for $P(\mathrm{vec}(\boldsymbol{B}') \leq \mathrm{vec}(\boldsymbol{A}')) = 1$. However, we can not directly apply the multivariate usual stochastic order to our scenario since it does not guarantee the positive definiteness of $\boldsymbol{B}' - \boldsymbol{A}'$. Instead, it is sufficient to check the stochastic order of the eigenvalues of $\boldsymbol{A} - \boldsymbol{B}$, i.e. $\Lambda_{\boldsymbol{B}} \geq_{st} \Lambda_{\boldsymbol{A}}$, to attain the existence of $\boldsymbol{A}'$ and $\boldsymbol{B}'$ such that $P(\boldsymbol{B}' - \boldsymbol{A}' \succcurlyeq \mathbf{0}) = 1$ after using coupling.

We first transform $\boldsymbol{B}' - \boldsymbol{A}' \succcurlyeq \mathbf{0}$ into the form such that we can simply connect it to the eigenvalues by the following Lemmas.

Lemma 3. *([89, 10.50(b)]) Let $\boldsymbol{Y} \succ \mathbf{0}$ and Hermitian, and $\boldsymbol{X} \succcurlyeq \mathbf{0}$ and Hermitian. $\boldsymbol{Y} - \boldsymbol{X} \succcurlyeq \mathbf{0}$ if and only if the eigenvalues of $\boldsymbol{X}\boldsymbol{Y}^{-1}$ all satisfy $\lambda_i \leq 1$.*

We then use the following Lemma to connect the eigenvalues of $\boldsymbol{X}\boldsymbol{Y}^{-1}$ to those of $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Lemma 4. *([90, Theorem 9H.1]) If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are $n \times n$ positive semidefinite Hermitian matrices, then*

$$\lambda_{max}(\boldsymbol{X}\boldsymbol{Y}) \leq \lambda_{max}(\boldsymbol{X})\lambda_{max}(\boldsymbol{Y}). \tag{3.91}$$

[10] From [88, Theorem 1.3] we know that $P(n\sigma_n(\boldsymbol{H})^2 \leq t) = 1 - e^{-t^2/2-t} + O(n^{-c})$, where $\boldsymbol{H}$ is $n \times n$ with i.i.d. entries with zero mean and unit variance, σ_n denotes the least singular value of $\boldsymbol{H}$, $c > 0$. Therefore, when t is small, the RHS approaches zero, which gives us the desired property.

[11] The reason that it is not necessary is, we may be able to use the channel enhancement scheme to obtain a degraded channel with $\boldsymbol{B} \not\succeq \boldsymbol{A}$.

Then from Lemma 3 and Lemma 4 we can derive the following theorem, which stochastically compares the minimum and maximum eigenvalues of the covariance matrices of Bob's and Eve's channels.

Theorem 1. *A sufficient condition to have a degraded MIMOME wiretap channel is*

$$\lambda_{min}(\boldsymbol{H_r}\boldsymbol{H_r}^H) \geq_{st} \lambda_{max}(\boldsymbol{H_e}\boldsymbol{H_e}^H). \tag{3.92}$$

Remark 5. *To have a degraded wiretap channel,* (3.92) *is a stringent condition to satisfy. The reasons are: 1)* $\boldsymbol{A} \preccurlyeq \boldsymbol{B}$ *may not be necessary for the existence of the degraded wiretap channel. More specifically, for the full CSIT case, for arbitrary covariance matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$*, [72] proves that such channel can be transformed into a degraded one by the channel enhancement technique; 2) the usual stochastic ordering is sufficient but not necessary, which can be seen from the SISOSE case [87].*

Example 1: Consider a $2 \times 2 \times 2$ Gaussian wiretap channel. A Rician fading channel with factor K can be described by

$$\boldsymbol{H} = \frac{\sqrt{K}}{\sqrt{1+K}}\bar{\boldsymbol{H}} + \frac{1}{\sqrt{1+K}}\boldsymbol{H}_w, \tag{3.93}$$

with mean $\sqrt{K}/\sqrt{1+K}\bar{\boldsymbol{H}}$ and covariance matrix $\boldsymbol{I}_N/(1+K)$, where $\boldsymbol{H}_w$ is a 2×2 random matrix with i.i.d. $\mathsf{CN}(0,1)$ entries. From [91, Theorem 1, Theorem 2] we know that the CDFs of the minimum and maximum eigenvalues of a noncentral Wishart matrix $\boldsymbol{H}\boldsymbol{H}^H$ are respectively as

$$F_{min}(x) = 1 - \frac{\det(\boldsymbol{\Psi}(x))}{\det(\boldsymbol{\Psi}(0))},\ F_{max}(x) = \frac{\det(\boldsymbol{\Xi}(x))}{\det(\boldsymbol{\Psi}(0))}, \tag{3.94}$$

where

$$\{\boldsymbol{\Psi}(x)\}_{i,j} = \begin{cases} 2^{(2i-s-t)/2}Q_{s+t-2i+1,t-s}(\sqrt{2\lambda_j}, \sqrt{2x}), \\ \qquad\qquad\qquad\qquad j = 1, \ldots, L \\ \Gamma(t+s-i-j+1, x),\ j = L+1, \ldots, s, \end{cases}$$

$$\{\boldsymbol{\Xi}(x)\}_{i,j} = \begin{cases} 2^{(2i-s-t)/2}\left[Q_{s+t-2i+1,t-s}(\sqrt{2\lambda_j}, 0) - \right. \\ \qquad\qquad \left. Q_{s+t-2i+1,t-s}(\sqrt{2\lambda_j}, \sqrt{2x})\right], \\ \qquad\qquad\qquad\qquad j = 1, \ldots, L \\ \gamma(t+s-i-j+1, x),\ j = L+1, \ldots, s, \end{cases}$$

where L is the number of nonzero eigenvalues $\{\lambda_j\}$ of $K\bar{\boldsymbol{H}}\bar{\boldsymbol{H}}^H$, $s = \min(n_T, n_R)$, $t = \max(n_T, n_R)$, n_T and n_R are the numbers of transmit and receive antennas, respectively; $Q_{p,q}(a,b)$ is the Nuttall Q-function, Γ and γ are the upper and lower incomplete gamma functions, respectively. We set $(\lambda_1, \lambda_2) = (2.1K, 1.9K)$, where K is the Rician K-factor.

From numerical results we can find that if we set the K-factor of Bob's channel as 15, then that value of Eve's channel shall not be larger than 6.9 to satisfy (3.92). We then show the ergodic secrecy capacity versus different K-factors in Figure 3.23 for cases with Bob's K-factor larger or equal to 15 such that (3.92) is satisfied under different transmit power P_T. Since there is no analytical solution of the input covariance matrix for such channel, we exhaustively search the 2×2 optimal input covariance matrix. The number of random channel realisations is 10^4. From this figure we can find that the ergodic secrecy capacity increases with increasing Bob's K-factor, which means under fixed Eve's channel, the reduction of the uncertainty of Bob's CSI at Alice indeed improves the ergodic secrecy capacity. In addition, we can also observe that due to imperfect CSIT, the efficiency of increasing transmit power to increase the ergodic secrecy capacity is low. Moreover, we can see that the ergodic secrecy capacity of $P_T = 500$ is quite close to that with infinite P_T. Note that when Bob's K-factor approaches infinity, Alice knows Bob's CSI perfectly.

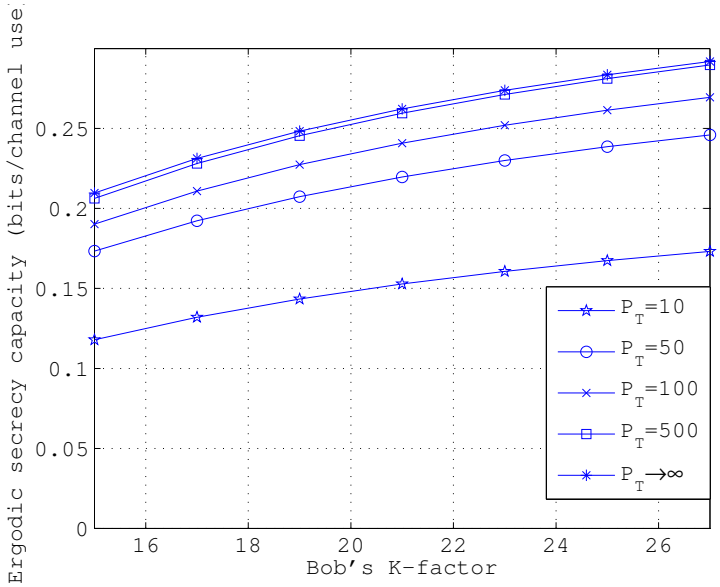


Figure 3.23: The ergodic secrecy capacities versus K-factors of Bob's channel when that of Eve's channel is set as 6.9 under different transmit power (linear).

Remark 6. *Note that when the fast fading wiretap channel with only statistical CSIT is verified as a degraded one, i.e. there exists $\boldsymbol{A}'$ and $\boldsymbol{B}'$ such that $\boldsymbol{A}' \succcurlyeq \boldsymbol{B}'$ for each channel realisation, where $\boldsymbol{A}'$ and $\boldsymbol{B}'$ are the covariance matrices of the equivalent noises at Bob and Eve. Then by [92, Proposition 1] we know that it is a convex problem. For full CSIT cases we can use convex optimisation tools to solve it numerically or some partial analytical results can be seen in [92, Theorem 2][93] [94], etc.*

In the following we show another sufficient condition to have a degraded channel.

Theorem 2. *Let $\boldsymbol{U_r}\boldsymbol{D_r}\boldsymbol{V_r}^H$ and $\boldsymbol{U_e}\boldsymbol{D_e}\boldsymbol{V_e}^H$ be the singular value decompositions (SVD) of $\boldsymbol{H_r}$ and $\boldsymbol{H_e}$, respectively. Assume $\boldsymbol{V_r}$ is independent to $\boldsymbol{D_r}$ and $\boldsymbol{U_r}$, and $\boldsymbol{V_e}$ is independent to $\boldsymbol{D_e}$ and $\boldsymbol{U_e}$. Also assume that $\boldsymbol{V_r}^H$ and $\boldsymbol{V_e}^H$ have the same distribution. If $\boldsymbol{D_r} \geq_{st} \boldsymbol{D_e}$, then there exists an equivalent degraded wiretap channel.*

Remark 7. *The requirement of $\boldsymbol{V_r}$ being independent to $\boldsymbol{D_r}$ and $\boldsymbol{U_r}$ in the proof of Theorem 2 is valid for the case that if the channel matrix is with i.i.d. Gaussian entries, then we can use the LQ decomposition (LQD) to get the right singular vector $\boldsymbol{V_r}$ and $\boldsymbol{V_e}$ which are the $\boldsymbol{Q}$ of the LQD and are independent of the $\boldsymbol{L}$ [95, Theorem 2.3.18]. And the random matrix $\boldsymbol{Q}$ follows the isotropic distribution (i.d.) with pdf [96]*

$$f(\boldsymbol{Q}) = \frac{\prod_{k=1}^{n_T}\Gamma(k)}{\pi^{\frac{n_T(n_T+1)}{2}}}\delta(\boldsymbol{Q}^H\boldsymbol{Q} - \boldsymbol{I}_{n_T}), \tag{3.95}$$

where Γ is the gamma function and δ is the delta function. Thus, if $\boldsymbol{H_r}$ and $\boldsymbol{H_e}$ are random matrices with i.i.d. Gaussian entries, $\boldsymbol{Q_r}$ and $\boldsymbol{Q_e}$ after LQD are both i.d. having the same distribution. Then the requirements of Theorem 2 are automatically satisfied.

In the following we consider another condition on the structure of the random matrix. We prove that if the channels can be decomposed into i.d. unitary matrices, the MIMOME wiretap channel is equivalent to a degraded one.

Theorem 3. *Let $\boldsymbol{H_r} = \boldsymbol{\Sigma_r}^{1/2}\boldsymbol{H_1}$, $\boldsymbol{H_e} = \boldsymbol{\Sigma_e}^{1/2}\boldsymbol{H_2}$. If $\boldsymbol{H_1}$ and $\boldsymbol{H_2}$ are i.d. and $\boldsymbol{\Sigma_r} \succeq \boldsymbol{\Sigma_e} \succ \boldsymbol{0}$, then it is equivalent to a degraded wiretap channel.*

Remark 8. *For all Theorems 1, 2, and 3, we transform the original channels to an equivalent one, such that for all code symbols (channel realisations) the channels are degraded.*

Remark 9. *The constraint $\boldsymbol{\Sigma_r} \succeq \boldsymbol{\Sigma_e}$ in Theorem 3 may be relaxed to $\boldsymbol{\Sigma_r} \not\succeq \boldsymbol{\Sigma_e}$ by deterministic channel enhancement.*

The channel enhancement argument [72], which is originally designed for channels with full CSIT, may be applied to the considered model where the transmitter only has statistical CSIT and the channels are fast faded. In the following we illustrate one example.

Example 2: For the wiretap channel

$$\boldsymbol{Y_r} = \boldsymbol{H}\boldsymbol{\Sigma_r}^{1/2}\boldsymbol{X} + \boldsymbol{Z_r}, \tag{3.96}$$

$$\boldsymbol{Y_e} = \boldsymbol{H}\boldsymbol{\Sigma_e}^{1/2}\boldsymbol{X} + \boldsymbol{Z_e}, \tag{3.97}$$

assume that the fading channel $\boldsymbol{H}$ has realisations $\{\boldsymbol{H}_0, \{\boldsymbol{A}\boldsymbol{H_0}\} : \boldsymbol{A} \in U(n_r^2)\}$, where $U(n)$ is the unitary group with degree n. The distribution of $\boldsymbol{H}$ can be more general than that in Example 2. Then it can be easily seen that we can apply the channel enhancement to the pair of channel realisations $(\boldsymbol{H}_0\boldsymbol{\Sigma_r}^{1/2}, \boldsymbol{H}_0\boldsymbol{\Sigma_e}^{1/2})$ to achieve the secrecy capacity. A simple way to see it is that the receivers know $\boldsymbol{A}$ and multiplying $\boldsymbol{Y_r}$ and $\boldsymbol{Y_e}$ by unitary $\boldsymbol{A}$ does not change the capacity.

3.3.4 Conclusion

In this work we characterise the relation between the usual multivariate stochastic order and the degradedness among the legitimate receiver's and eavesdropper's channels. We consider the transmitter is with multiple antennas and the legitimate receiver and the eavesdropper are both with multiple antennas under fast fading with statistical CSIT. More specifically, based on the technique of coupling we derive criteria to check the degradedness of several commonly considered wiretap channels. One example of a Rician MIMOME $2 \times 2 \times 2$ channel is illustrated where the secrecy capacity can be derived under different K-factors.

4 Cloud self organisation

4.1 Cloud resource scheduling by distributed interference coordination

4.1.1 Introduction

In the last years, new generation wireless communication systems, such as 3GPP Long Term Evolution (LTE) [97], have been developed to face the increasing demand for high data rate and to support heterogeneous traffic with different quality of service (QoS) requirements. With the rapid growth of smart mobile devices, improving the spectrum usage has become crucial to overcome the limited bandwidth resources of cellular networks. In this context, D2D communications enabled by self-organised cloud network can be considered as a promising technology for enhancing the spectral efficiency and thus increasing the system capacity.

One fundamental issue is the spectrum sharing strategy. Proposed schemes in the literature can be divided into two categories: (i) underlay, where D2D users share the time/frequency resources occupied by cellular users [98]; (ii) overlay, based on the use of orthogonal time/frequency resources between cellular and D2D users. In D2D communication underlaying cellular network, the spectrum efficiency is improved by sharing the resources of cellular user equipments (UE) with D2D users; the interference has to be properly controlled by the BS so that the D2D transmission does not interfere with the UEs decreasing the capacity of the cellular network [99]. On the other hand, in overlaying cellular network, the D2D users can communicate bi-directionally with each other while assisting the two-way communication between the BS and the UE [100]. This is the approach considered in this section as a potential application of the DIWINE concept for self-organising cloud network. In this context, spectrum sharing resembles on the same concept of cognitive radio networks where secondary users may access the spectrum if primary users are not active or they do not cause unacceptable interference [21], except that none of the users has privileges.

Most of the works in this areas [98]–[100] do not cover multi-class service requirements and assume that one D2D user can only share resources with one UE limiting the throughput demands of high-rate services. However minimal QoS for D2D services also needs to be guaranteed. In [101], D2D users reuse the uplink resources of multiple UEs to improve the spectrum utilisation and to achieve the requested data rate based on power control of D2D transmitters. The authors in [102] propose a resource allocation method based on the maximisation of the transmission capacity by scheduling D2D users

to different frequency bands. The resource allocation of UEs and D2D users is usually coordinated by the BS, where information exchange and signalling between BS-D2D is needed for channel quality information (CQI) and/or positioning of the devices. D2D resource scheduling has been widely investigated assuming a centralised coordinator, while a distributed resource management system which guarantees multi-class service requirements is the open issue addressed here.

In this section, we focus on resource allocation for D2D communications based on LTE-Advanced networks, which enable flexible allocation of the physical resource blocks (PRBs) in the time-frequency (TF) domain [103]. In LTE systems, the radio resources are usually allocated by a base station (BS), such as an evolved Node B, and the basic requirement is to provide the TF resources, i.e. the PRBs, without spectrum collisions. However, for high density of D2D users, as in emerging 5G cellular networks [104], spectrum allocation with limited control by the BS is crucial to manage resources according to the traffic load intensity of D2D links. Aim of this work is to provide a new distributed resource management scheme, where the D2D links iteratively refine their allocation over a shared spectrum, making it increase/decrease based on the sensed intra-cluster interference and the QoS requirements of different class of services, till a steady state of the network is reached. In the proposed algorithm, each D2D scheduler first senses the link quality to drop strongly interfered RBs and mitigate the mutual interference with other D2D users. Then, it extends or reduces the set of allocated RBs to meet the QoS requirements. The selection of the resources to be added or released in the allocation is performed by minimising the boundary extension of the TF spectrum region. This criterion is selected to avoid too much indented or fragmented region allocations with large boundary areas that are likely to be affected by cross-interference due to TF jitter [5]. The proposed algorithm is able to maximise the D2D cluster throughput without any centralised coordination and for heterogeneous traffic with different kinds of service.

The work is organised as follow: Subsection 4.1.2 introduces the network scenario, as well as the traffic and interference models. The resource allocation algorithm is described in Subsection 4.1.3. Numerical results are in Subsection 4.1.4 for two network scenarios with homogeneous and heterogeneous traffics and for different QoS requirements. Concluding remarks are in Subsection 4.1.5.

4.1.2 Problem setting

Network scenario

We consider the single-cell scenario in Figure 4.1 consisting of one BS, a number of UEs and several D2D users that are grouped into clusters (or sub-networks). The system employs orthogonal multi-carrier transmission, such as orthogonal frequency division multiplexing (OFDM), to support multiple access for both cellular and D2D communications. We assume that the BS can allocate the UEs and the D2D clusters onto orthogonal

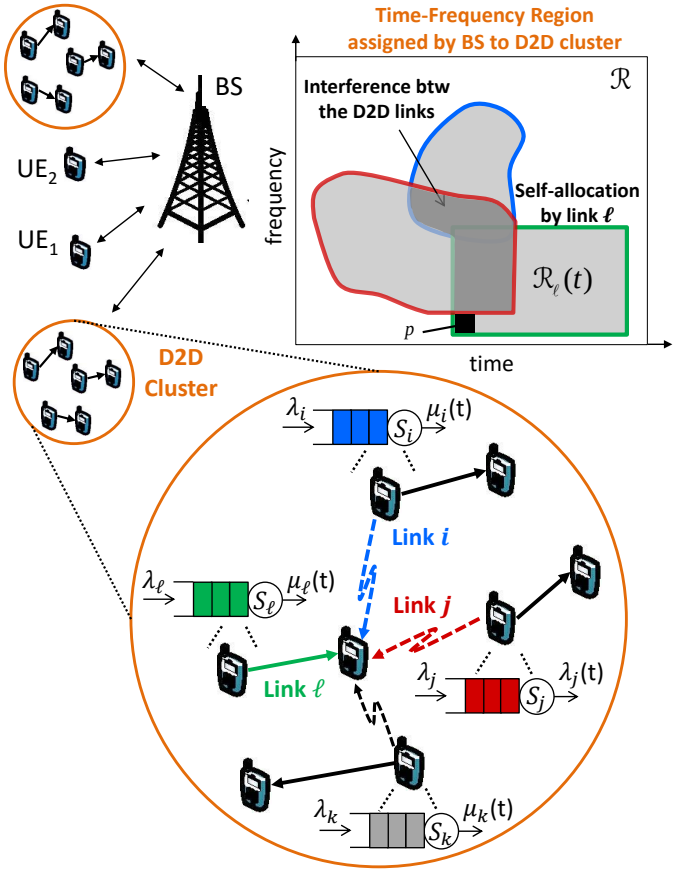


Figure 4.1: Network and interference scenario for D2D communications overlaying a cellular network. The traffic model consists of a set of L independent queues (one per D2D link), with traffic arrival rates $\{\lambda_\ell\}_{\ell=1}^{L}$ and time-varying service rates $\{\mu_\ell(t)\}_{\ell=1}^{L}$.

spectrum resources with no mutual interference [99]. On the other hand, each cluster is assigned a TF region that is shared among all the D2D links in the cluster resulting in possible intra-cluster interference. Inter-cell interference, i.e. originated from different cells, is assumed to be negligible as mitigated by means of frequency reuse, power control or resource allocation strategies for cellular systems [105].

This work is focused on resource management within a single cluster of L D2D links sharing a TF region $\mathcal{R}$. Following the LTE air protocol, the region is assumed to be composed by a number of PRBs, each occupying 1 slot of 0.5 ms in the time domain and a bandwidth of 180 kHz in the frequency domain. The minimum allocation to a single user is a sub-frame of 1 ms composed by 2 PRBs. We thus consider the set of two consecutive PRBs as the unitary resource block (RB) for allocation. The D2D links have to self coordinate their allocation over $\mathcal{R}$ so as to minimise the cross interference. The interference depends on the distance between the D2D users, thereby spatial reuse of some TF resources may be tolerated if needed to satisfy QoS requirements.

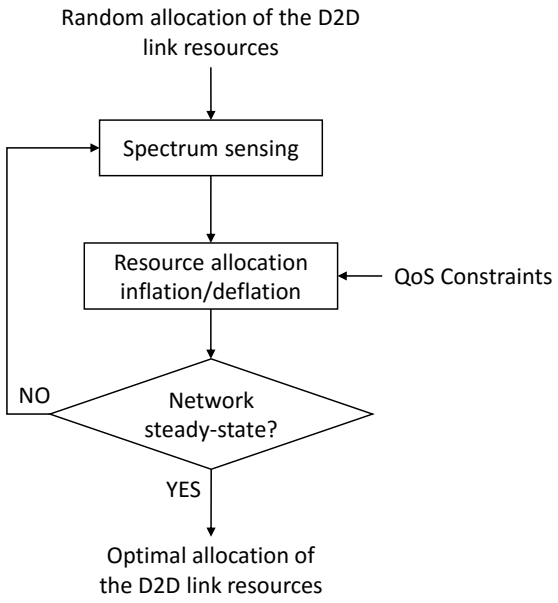


Figure 4.2: Flow diagram of the proposed scheduling algorithm.

Traffic and interference modelling

Consider a packet traffic-heterogeneous network where each of the D2D users needs to manage a specific traffic with a related QoS requirement. Traffic can range from VoIP, characterised by low-speed, e.g. 16 kbps, and strict latency (< 50 ms), to HTTP service with higher bit-rate, or FTP which is a best-effort service with high bit-rate and low delay requirement [106].

A distributed scheduling approach is considered where L schedulers, one per each D2D link in the cluster (see Figure 4.1), act independently trading between the incoming packet arrivals and the service capability to comply with the QoS constraints. Each local scheduler has to self-adapt the allocation without explicit signalling with the other D2D users. The traffic model consists of a set of L independent queues, one per each D2D link (as detailed in Figure 4.1), with traffic arrivals modelled as Poisson distributed with rate λ_ℓ for link ℓ, $\ell = 1, \ldots, L$, and an infinite buffer (queue) with first-in-first-out (FIFO) serving policy. The service data rate at scheduler ℓ is characterised by a time-varying offered throughput $\mu_\ell(t)$, which dynamically changes to guarantee the requested throughput λ_ℓ that satisfies the QoS. Traffic and service rates are measured in term of bps to simplify the reasoning.

Each scheduler iteratively allocates the TF resources to reach a stability condition that complies with the QoS of its own traffic class, according to the throughput requirement and the interference generated from the other D2D links. The scheduling algorithm – same for all links – is designed so as to react to the mutual interference by reducing/augmenting the resources, as shown in Figure 4.2. The metric used for the ad-

aptation to the interference level at each time-step $t = 1, 2, \ldots$ is the signal to interference plus noise ratio (SINR) observed over the different RBs. Assuming the transmission power to be equal for all the D2D users, the power $P_{i,j}$ of the signal transmitted by D2D user i and received by D2D user j is modelled as:

$$P_{i,j} = P_{ref} \left(\frac{d_{ref}}{d_{i,j}} \right)^{\gamma} Y_{i,j}, \tag{4.1}$$

where P_{ref} is the received power at a reference distance d_{ref}, $d_{i,j}$ the distance between the two users, γ the path loss exponent and $Y_{i,j}$ a random scaling term accounting for Rayleigh fading, assumed to be constant within one time frame and varying over different time frames. Let p be any generic RB in $\mathcal{R}$, $\mathcal{R}_\ell(t)$ the region allocated by the ℓth scheduler at iteration t (see Figure 4.1), and $\mathcal{I}_\ell(t)$ the set of D2D transmitters interfering with the D2D link ℓ at iteration t. The SINR measured over the D2D link ℓ on resource p at time t can be defined as:

$$\text{SINR}_\ell^p(t) = \frac{P_{\text{tx}_\ell,\text{rx}_\ell}}{P_n + \sum\limits_{i \in \mathcal{I}_\ell(t)} P_{i,\text{rx}_\ell}}, \tag{4.2}$$

where $P_{\text{tx}_\ell,\text{rx}_\ell}$ is the received signal power for the link ℓ between the transmitting node tx_ℓ and the receiving node rx_ℓ, P_n is the power of the additive white Gaussian noise (AWGN), while the second term in the denominator represents the overall interference power. The throughput offered by the pth RB, $\mu_\ell^p(t)$, is computed as a function of the instantaneous SINR [107]:

$$\mu_\ell^p(t) = B\log_2(1 + \text{SINR}_\ell^p(t)), \tag{4.3}$$

where B is the RB bandwidth. The total throughput offered to link ℓ is then computed as:

$$\mu_{\mathcal{R}_\ell}(t) = \sum_{p \in \mathcal{R}_\ell} \mu_\ell^p(t), \tag{4.4}$$

accounting for all the resources in the region $\mathcal{R}_\ell(t)$.

4.1.3 Self-adaptive resource allocation algorithm

A distributed scheduling algorithm is proposed in this subsection where all the link schedulers act independently mutually sensing one another. Each link scheduler iteratively refines the allocation based on the locally sensed intra-cloud interference, reallocating the collided TF resources towards the non-interfered spectrum regions, till the QoS is fulfilled and an equilibrium with the other links is reached. At iteration t, the ℓth scheduler updates its own allocation $\mathcal{R}_\ell(t)$ as described in the following, making it grow/reduce in order to (i) minimise the cross interference with other D2D links (Subsection 4.1.3) and (ii) select the resources to be included/released so as to guarantee the QoS and minimise the loss of efficiency due to TF jitter (Subsection 4.1.3).

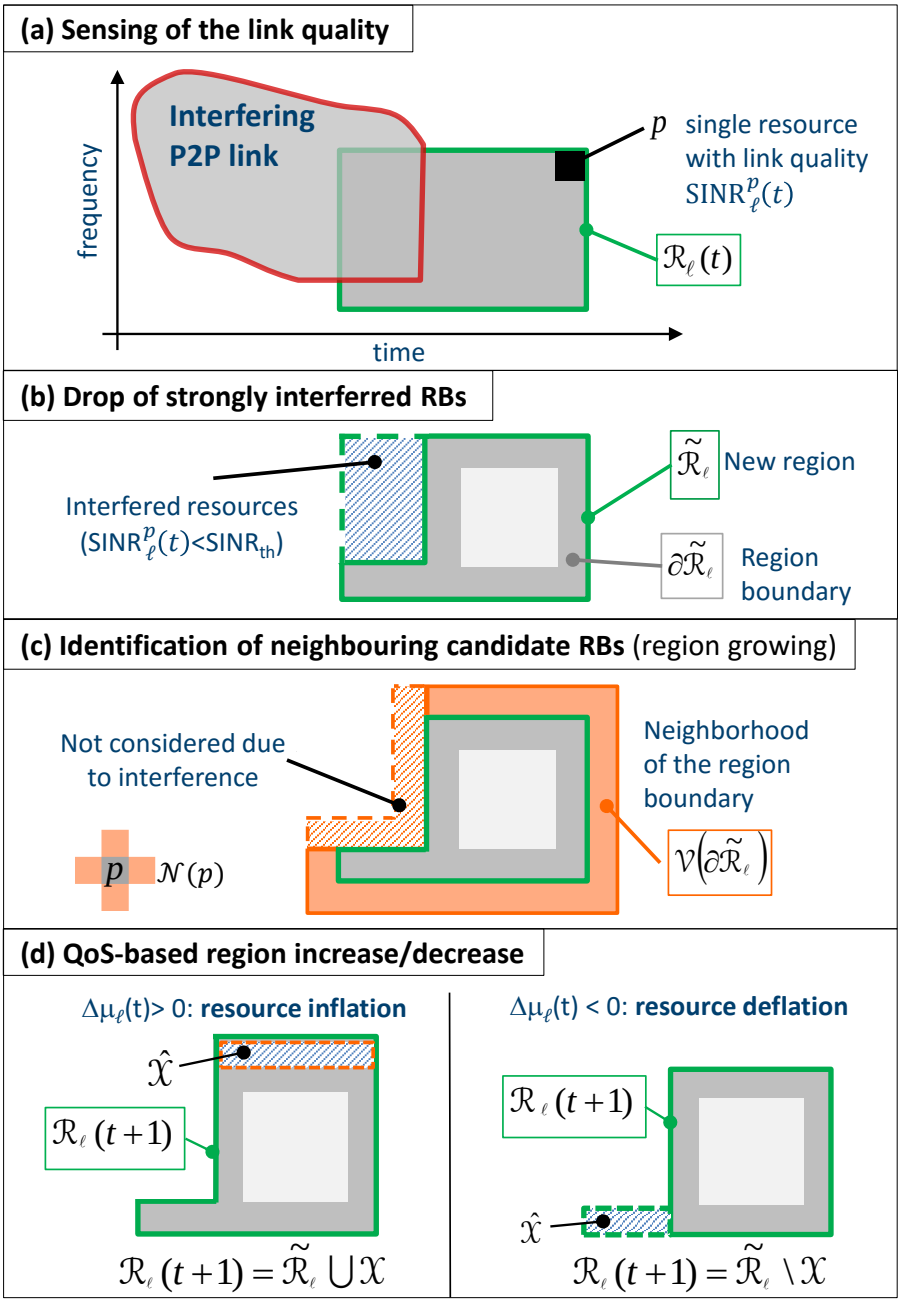


Figure 4.3: Resource allocation procedure at iteration t: a–b) Sensing of the link quality and drop of strongly interfered RBs; c–d) Identification of candidate resources for possible region growing and region increase/decrease based on throughput requirement.

Spectrum sensing and cross-interference minimisation

Consider the scheduling process for the D2D link ℓ at iteration t. As illustrated in Figure 4.3-(a), since the scheduling is not centrally coordinated some TF resources $p \in \mathcal{R}_\ell(t)$ might have been simultaneously allocated by different links which, if spatially co-located, give rise to a strong interference in the overlapped spectrum region. To mitigate this mutual interference, the scheduler ℓ senses the link quality in the area of the previous allocation $\mathcal{R}_\ell(t)$, searching for the strongly interfered RBs. Any RB $p \in \mathcal{R}_\ell(t)$ with SINR value lower than a certain threshold SINR_{th}, i.e. $\text{SINR}_\ell^p(t) \leq \text{SINR}_{\text{th}}$, is dropped by the scheduler, see Figure 4.3-(b). The updated set for link ℓ is then $\tilde{\mathcal{R}}_\ell = \{p \in \mathcal{R}_\ell(t) : \text{SINR}_\ell^p(t) \geq \text{SINR}_{\text{th}}\}$ with boundary denoted as $\partial\tilde{\mathcal{R}}_\ell$.

QoS-based self-adjustment of the TF resources

Once the allocation has been adjusted to avoid mutual interference, the set of RBs allocated to link ℓ, $\tilde{\mathcal{R}}_\ell$, has to be extended or reduced to meet the QoS requirement. The throughput variation that is needed to satisfy the request λ_ℓ is determined as:

$$\Delta\mu_\ell(t) = \lambda_\ell - \mu_{\tilde{\mathcal{R}}_\ell}(t), \tag{4.5}$$

where $\mu_{\tilde{\mathcal{R}}_\ell}(t)$ is the total throughput offered to link ℓ by the current allocation $\tilde{\mathcal{R}}_\ell$ computed as in (4.4). For $\Delta\mu_\ell(t) > 0$ (region inflation), the local scheduler has to incorporate new resources $\mathcal{X}$ in the link allocation, see Figure 4.3-(c). These are selected in the neighbourhood of the current region boundary, excluding the interfered regions, i.e. in the spectrum area $\mathcal{V}(\partial\tilde{\mathcal{R}}_\ell) = \{v : v \in \{\mathcal{N}(p) \setminus \tilde{\mathcal{R}}_\ell\} \wedge p \in \partial\tilde{\mathcal{R}}_\ell \wedge \text{SINR}_\ell^v(t) \geq \text{SINR}_{\text{th}}\}$, $\mathcal{N}(p)$ denoting the set of the 4-nearest neighbours of the RB p. On the other hand, for $\Delta\mu_\ell(t) < 0$ (region deflation), the region has to be reduced by releasing some RBs $\mathcal{X}$ from the region boundary $\partial\tilde{\mathcal{R}}_\ell$. The updated region is, as in Figure 4.3-(d),

$$\mathcal{R}_\ell(t+1) = \begin{cases} \tilde{\mathcal{R}}_\ell(t) \cup \mathcal{X} & \text{if } \Delta\mu_\ell(t) > 0 \\ \tilde{\mathcal{R}}_\ell(t) \setminus \mathcal{X} & \text{if } \Delta\mu_\ell(t) < 0 \end{cases}. \tag{4.6}$$

The selection of the resource set $\mathcal{X}$ to be incorporated ($\mathcal{X} \subseteq \mathcal{V}(\partial\tilde{\mathcal{R}}_\ell)$) or released ($\mathcal{X} \subseteq \partial\tilde{\mathcal{R}}_\ell$) is performed so as to minimise the boundary extension. This criterion is selected to avoid too much indented or fragmented region allocations with large boundary areas that are likely to be affected by cross-interference due to TF jitter; guard bands/intervals are typically prescribed to avoid such interference but this is known to reduce the spectrum efficiency. The length of the boundary $\partial\mathcal{R}_\ell(t+1)$ can be conveniently computed as the cardinality, i.e. the number of RBs, of the set $\mathcal{F}(\partial\mathcal{R}_\ell(t+1)) = \{f : f \in \{\mathcal{N}(p) \setminus \mathcal{R}_\ell(t+1)\} \wedge p \in \partial\mathcal{R}_\ell(t+1)\}$, which is the set of external neighbours of the boundary of the new region $\mathcal{R}_\ell(t+1)$. Thereby, optimal selection of the resource set is

Table 4.1: Simulation system parameters.

Parameter	**Value**
RB structure	12 sub-carriers and 1 sub-frame
RB frequency bandwidth	180 kHz
Time frame	10 ms
Total number of available RBs	100
Number of D2D links	10, 15, 20, 30, 50
Reference distance	10 m
Max. D2D distance	300 m
Path-loss exponent	2
P_{ref}/P_n	20 dB
SINR threshold, SINR_{th}	10 dB

Table 4.2: Simulation traffic parameters.

Average data rate (QoS)	**Value**
HTTP traffic	512 kbps
FTP traffic	360 kbps
VoIP traffic	64 kbps

obtained by minimising the ratio between the number of RBs in the boundary $\partial\mathcal{R}_\ell(t+1)$ and the number of RBs in the region $\mathcal{R}_\ell(t+1)$ evaluated as:

$$\hat{\mathfrak{X}} = \arg\min_{\mathfrak{X}} \frac{|\mathcal{F}(\partial\mathcal{R}_\ell(t+1))|}{|\mathcal{R}_\ell(t+1)|}, \tag{4.7}$$

subject to the constraint of throughput fulfillment:

$$\sum_{p\in\mathcal{R}_\ell(t+1)} \mu_\ell^p(t) \geq \mu_{\tilde{\mathcal{R}}_\ell}(t) + \epsilon\Delta\mu_\ell(t), \tag{4.8}$$

where $0 < \epsilon < 1$ is a step size parameter. Iterations of the proposed algorithm are repeated till $|\Delta\mu_\ell(t) - \Delta\mu_\ell(t-1)| < \delta$. A summary of the scheduling steps is shown in Figure 4.3.

4.1.4 Numerical results

The performance analysis is carried out in a LTE framework, by simulating a dense D2D cluster with parameters as in Table 4.1. The performance is evaluated through Monte Carlo simulations using Matlab software. The proposed scheduling strategy is analysed for two different network scenarios: (i) homogeneous traffic, where all D2D users manage FTP service, and (ii) heterogeneous traffic, where 50

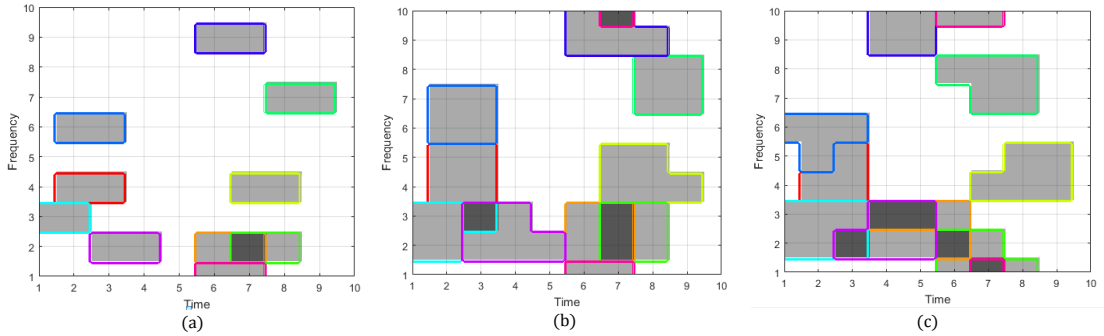


Figure 4.4: Self-adaptive resource allocation of $L = 10$ D2D schedulers (indicated by colours): a) random network state initialisation; b) D2D self-adjusting resource allocation at iteration $t = 4$; c) convergence after compliance of QoS requirements ($t = 10$).

In Figure 4.4, an example of successive refinements of the resource allocation by the distributed scheduling algorithm is illustrated for $L = 10$ D2D links. The D2D mutual interference is simulated according to a disc model with maximum distance $d_{max} = 300$ m. As initial state, the D2D users randomly allocate a fixed amount of resources, i.e. 2 RBs, in frequency and time (Figure 4.4-(a)). Then, the local schedulers dynamically adapt their allocations as illustrated in Subsection 4.1.3 (Figure 4.4-(b)) till they reach a steady state that satisfies the QoS requirements (Figure 4.4-(c)). The algorithm implements spectrum reuse, if requested, based on the sensed-interference level between the D2D links.

The metric used for evaluation is the total normalised requested throughput (QoS need) for the D2D cluster defined as $\xi(t) = \sum_\ell \mu_{\mathcal{R}_\ell}(t) / \sum_\ell \lambda_\ell$, as the ratio between the total offered throughput (for all the D2D links) and the total requested throughput. Figure 4.5 compares the total normalised throughput $\xi(t)$ of the distributed and centralised scheduling approaches over the iterations t, for $L = \{10, 50\}$ D2D links and for both single (FTP traffic) and multi-class (Table 4.2) services. In the *centralised scheduling* the allocation of the D2D cluster resources over the available spectrum region is performed centrally by the BS, which assigns to each D2D link the amount of resource that satisfies the QoS requirements, reusing the spectrum if needed.

From the results in Figure 4.5, it can be noticed that the distributed algorithm converges to the centralised performance in few iterations. As expected, the normalised throughput $\xi(t)$ at convergence is lower for increasing number of D2D links and the convergence time increases with the number of links. Notice that the distributed algorithm converges faster in homogeneous traffic conditions than in heterogeneous ones due to the management of different service classes (QoS requirements).

Figure 4.6 shows the total normalised throughput versus the loading factor η, which evaluates the overall fraction of occupied TF region. Clearly, the loading factor increases for increasing number of users. The performance is evaluated for both homogeneous

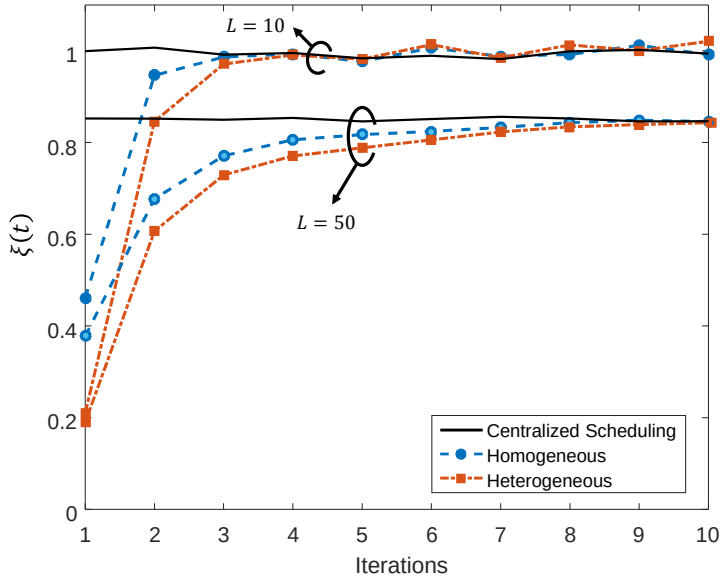


Figure 4.5: Normalised system throughput $\xi(t)$ vs. iterations (t) for $L = 10, 50$ D2D links. Performance of the distributed algorithm is evaluated for both homogeneous and heterogeneous traffic networks and it is compared to the centralised approach.

and heterogeneous traffic and the centralised approach is shown as reference. It can be observed that for small number of links ($L = 10, 15, 20$), η is less or close to 1 as the D2D users are able to self-allocate their resources and reach the QoS targets without any spectrum overlap, as the total required throughput is below the one offered by the spectrum region $\mathcal{R}$. On the other hand, for larger D2D clusters ($L = 30$), the loading factor at convergence is $\eta > 1$, as the D2D users have to spatially reuse the resource blocks to reach their QoS targets. It can be noticed that the total offered throughput decreases with respect to the requested one as the number of D2D links increases.

Figure 4.7 shows the normalised system throughput versus the loading factor η, for different values of the maximum distance of the interferers to the D2D receiver $d_{max} = 100, 300$ m. Decreasing d_{max} is equivalent to enlarge the spatial density of the D2D links in the cluster. In fact, if d_{max} is reduced from 300 m to 100 m, assuming a fixed number of links, i.e. for a fixed η value, this corresponds to an increase 9 times the spatial density for $d_{max} = 300$ m. Performance is evaluated for homogeneous and heterogeneous traffic networks and the centralised approach is shown as reference. It can be notice that the normalised system throughput for $d_{max} = 100$ m has a decay rate higher than the one for $d_{max} = 300$ m, due to the increased spatial density of the D2D links in the cluster. The proposed method reaches a performance close to the centralised approach.

The D2D cluster service satisfaction is investigated in Figure 4.8 vs. the number of D2D links L for both homogeneous and heterogeneous traffic-network. The QoS satisfaction is measured as the ratio between the number of D2D pairs meeting the QoS requirements

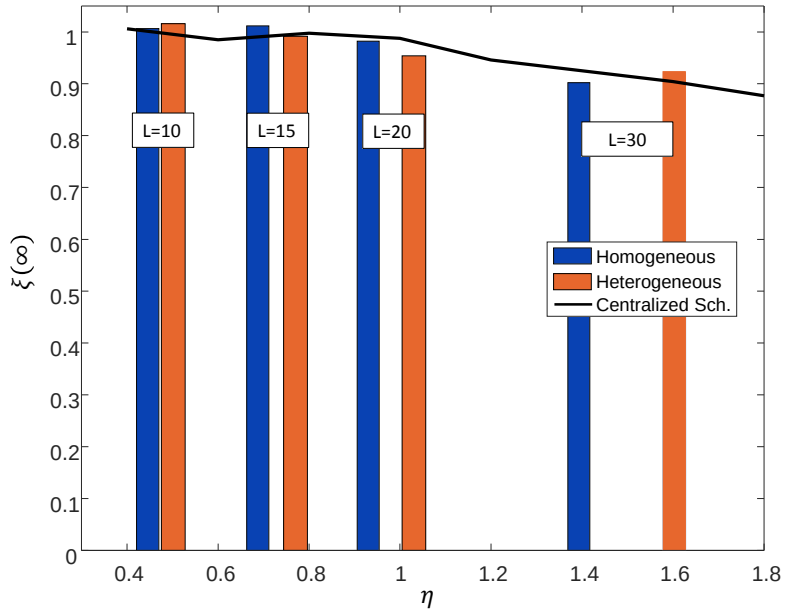


Figure 4.6: Normalised system throughput at convergence $\xi(t \to \infty)$ vs. loading factor $\eta = \sum_\ell |\mathcal{R}_\ell(\infty)|/|\mathcal{R}|$ for varying number of D2D links, for both homogeneous and heterogeneous traffic network, compared to the centralised approach.

and the total number of D2D pairs in the network. The numerical results show that the proposed distributed algorithm allows the D2D cluster to reach a high level of QoS satisfaction even when the number of D2D links increases ($L = 30$), as the allocation of TF resources is based on QoS constraints.

4.1.5 Conclusion

In this section we proposed a new distributed scheduling algorithm where multiple D2D links, sharing the same spectrum, self-adapt their resource allocation by dynamically inflating/deflating their time-frequency region based on the sensed interference level and the QoS requirements, without any central coordination. The method was shown to closely reach the performance of the centralised approach, both for single and multiple service classes, i.e. for homogeneous and heterogeneous traffic scenarios, managing the mutual interference and satisfying the QoS requirements of all D2D links in a fully distributed way. Moreover, numerical analysis shows that the distributed resource scheduling method can achieve the same performances of centralised optimisations when evaluated in term of system capacity, even in case of spectrum reuse.

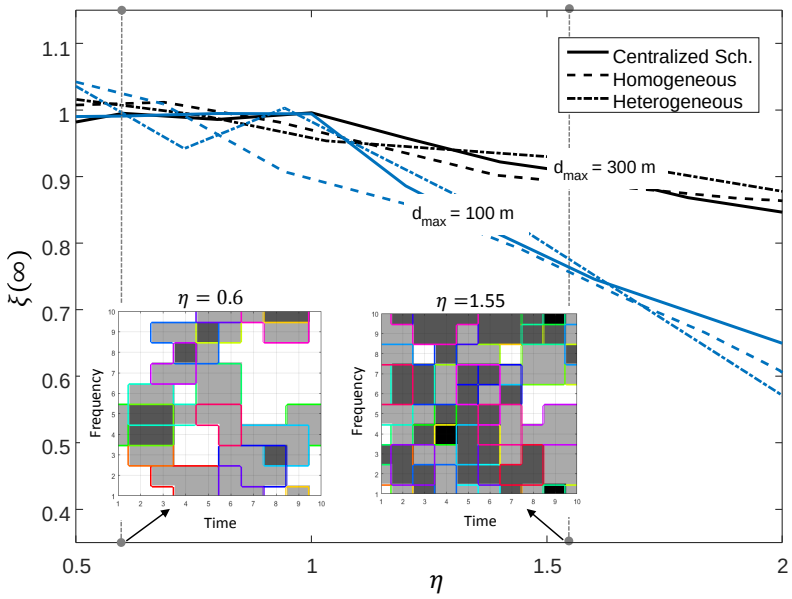


Figure 4.7: Normalised system throughput at convergence $\xi(t \to \infty)$ vs. loading factor η for varying maximum distance of the interferers to the D2D receiver ($d_{max} = 100, 300$ m), for homogeneous and heterogeneous traffic networks compared to the centralised approach. One sample of the distributed scheduling algorithm is shown for $\eta = 0.6, 1.55$.

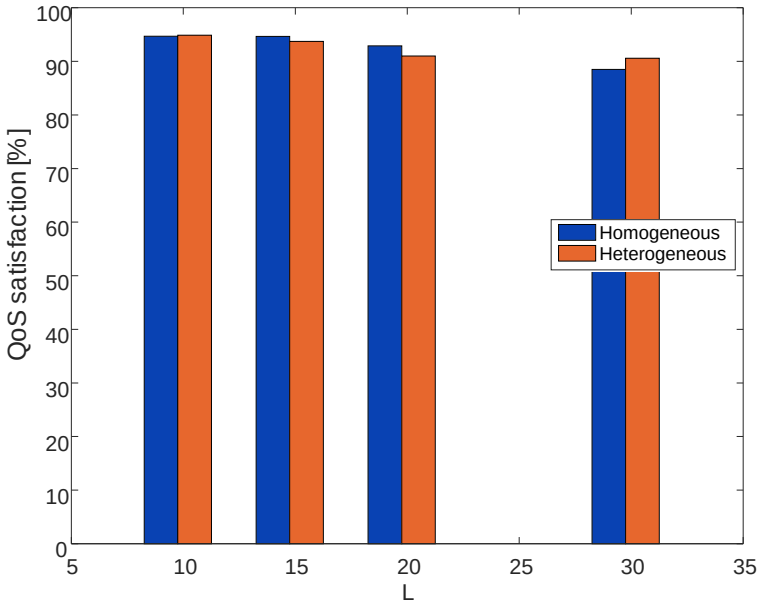


Figure 4.8: QoS satisfaction for the D2D cluster vs. the number L of D2D links for homogeneous and heterogeneous traffic scenarios.

4.2 Distributed learning process for HNC selection – Advanced scenarios

A Distributed Learning Algorithm (DLA) – a protocol for selection of Wireless Physical Layer Network Code (WPLNC) in DIWINE cloud scenarios – was described and its properties discussed in deliverable D3.02 [4]. The algorithm was successfully transferred for demonstration in real world scenario to WP5, see D5.41 [108]. This deliverable extents discussion about utility functions related to actual channel conditions and summarises changes in necessary information exchange among nearest neighbours. Advanced scenarios, that goes beyond the capability of HW demonstration, assuming unfriendly and/or malicious relays are also presented, analysed and discussed.

4.2.1 Utility function related to instantaneous channel conditions

The DLA originally uses relay output cardinality as a target for optimisation while guaranteeing invertibility of WPLNCs at all final destinations. This utility function relates to energy efficiency (UT1), since minimal cardinality WPLNCs maximally conserve energy, it also multiplies network throughput (UT1) due to WPLNC gains. The algorithm with cardinality utility works in distributed network model (UT4). On the other hand the overall performance can be significantly affected by a wireless channel. Particular channel conditions (attenuation and/or phase rotation) may have destructive impact on particular WPLNCs due to ambiguous signal overlaps, for details see D4.02 [109] or [110]. It is possible to cope with instantaneous channel conditions by modification of DLA utility function.

First of all it is important to note that randomly selected values of the utility function – which is the case for utility function directly related to random wireless channel – will with high probability violate potential property of underlying game [111]. It is the potential property that guarantees nice behaviour of DLA – existence of Nash equilibria, convergence of the process etc. Thus, any criterion that is related to random channel can be used only as an additional measure for WPLNC selection. In D3.02 a hierarchical minimal distance was discussed as a candidate additional criterion for selection of WPLNCs. It was shown that minimal hierarchical distance variations are zero or almost negligible for many WPLNCs and channel conditions, see Figures 4.9 and 4.11, thus it can be hardly used as a WPLNC selection metric.

To correctly cope with wireless channels the metric must reasonably describe performance behaviour. A possible choice is a complete hierarchical spectrum, however, due to its complexity this is barely tractable solution. In [112] a Symbol Error Rate (SER) is used to distinguish among the minimal WPLNCs. A WPLNC with the minimal SER in actual channel conditions is used. Whenever a minimal WPLNC is going to be selected by the DLA a WPLNC minimising actual SER is chosen. A newly added SER-based criterion is

$$\mathrm{MIN}_{\mathrm{new}} = \min\{\widehat{\mathrm{SER}}\}, \tag{4.9}$$

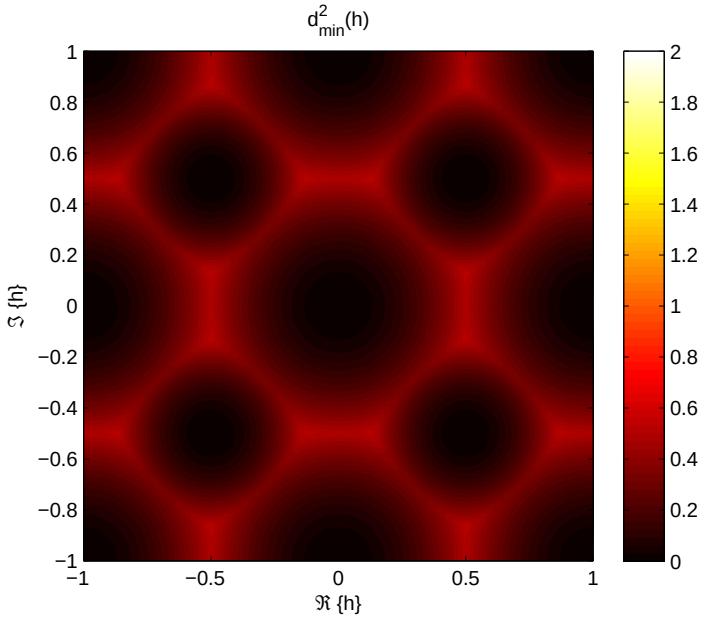


Figure 4.9: Minimal hierarchical distance – a minimal WPLNC.

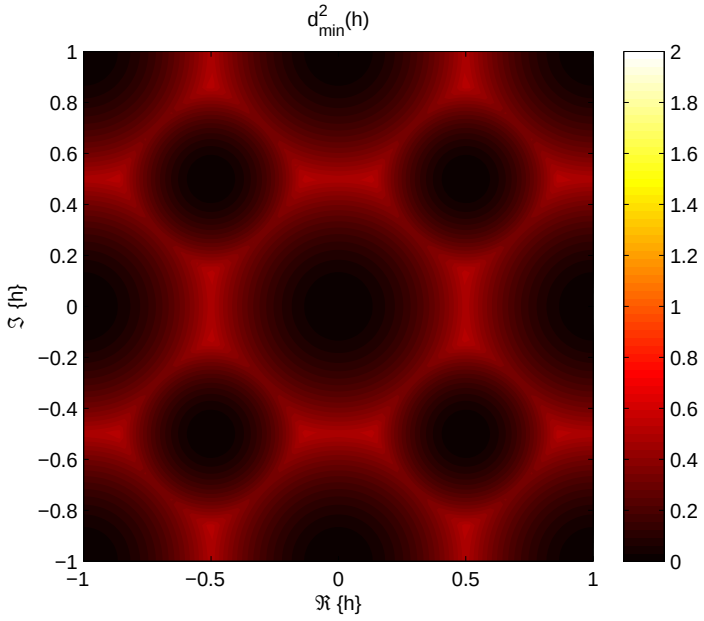


Figure 4.10: Minimal hierarchical distance – an extended WPLNC.

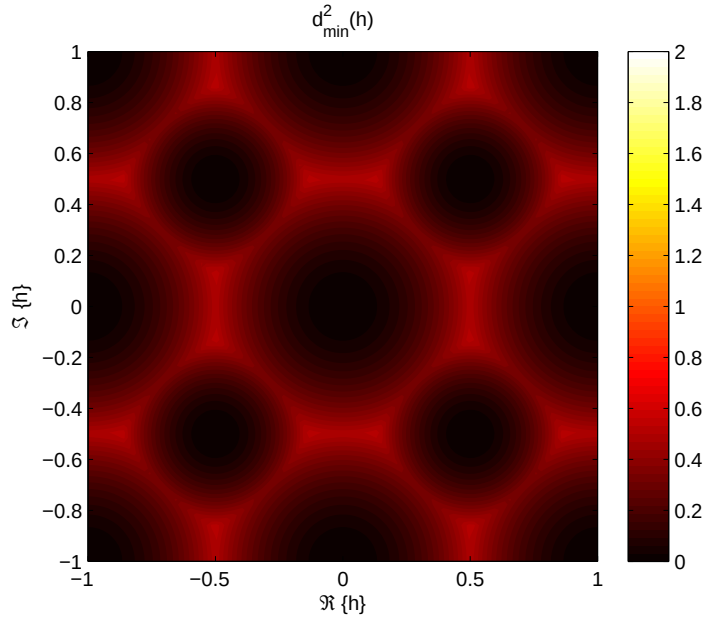


Figure 4.11: Minimal hierarchical distance – a full WPLNC.

where MIN_{new} is newly selected minimal WPLNCs, $\widehat{\text{SER}}$ is actual SER estimate and minimisation is done over all available invertible minimal WPLNCs.

The instantaneous SER computation is based on known channel estimates provided by the pilot signals. Again, since SER performance differences are negligible for many WPLNCs, especially in the case of full and extended WPLNCs, the SER is used as the additional criterion only when the minimal WPLNCs are achieved by DLA. From a measurement campaign carried within WP5 we have preselected WPLNCs with typically good SER performance across various channel realisations. These WPLNCs made a list of available actions for DLA.

Figure 4.12 shows an uncoded SER in 2 source 1 relay network which is a natural WPLNC building block of more complex networks. SER is measured at the relay, since any error in WPLNC mapping can be hardly corrected by the rest of the network especially when no or weak error correction coding is used. A system with newly proposed SER-based additional criterion is compared to the former DLA solution. Both numerical as well as USRP based results are shown. USRP results were obtained within WP5 testing.

4.3 Non-cooperating relays

In this section we focus on relays that do not fully follow the idea of the cloud concept. Particularly, we are interested in two cases – intentionally malicious relays, that wish to cause harm to the communication and intentionally over-selfish relays, that do not

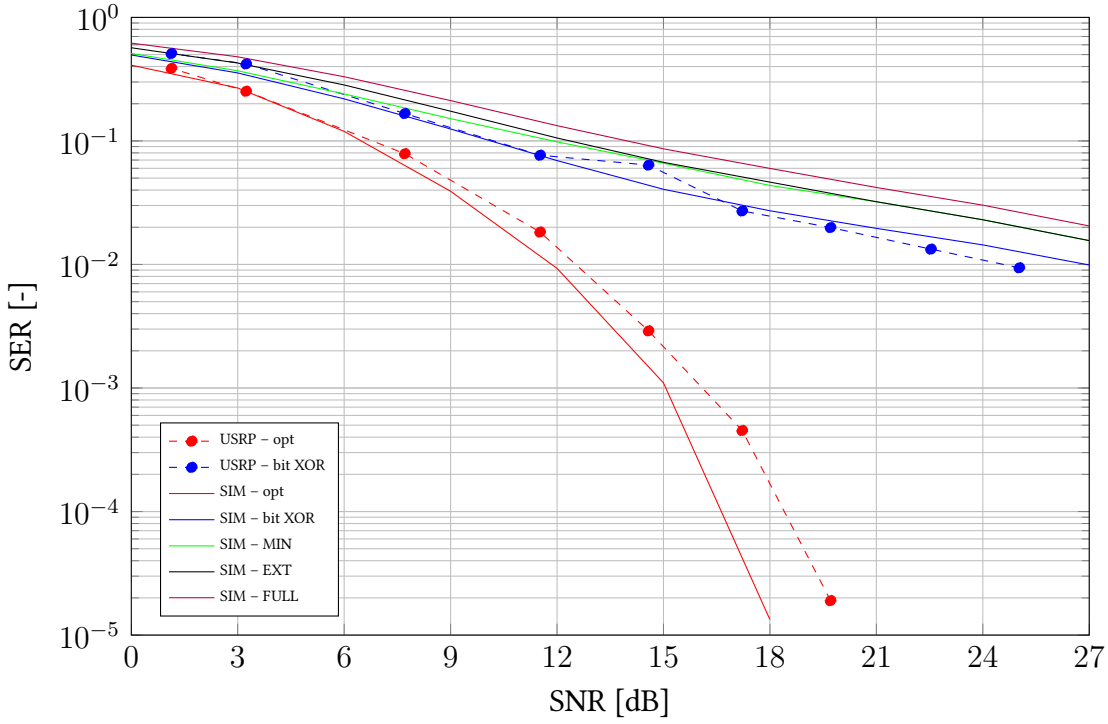


Figure 4.12: Uncoded SER performance of various WPLNC mappings. Both simulation and HW experiments results are shown.

follow altruistic cooperation and wish to seize maximum in emergency states such as emptying battery.

It is important to highlight that this section goes beyond the capabilities of demonstration platform, so the content presented here is mainly theoretical, not verified in real-world scenarios. Different network models as well as different utility functions and information knowledge are assumed compared to the DLA. Also a protocolised form suitable for real implementation is missing yet.

4.3.1 Intentionally malicious relay

The most of current research work optimistically assumes that all the relays altruistically cooperate on their task of source to destination communication. We have introduced a relay node with aims to make a disorder among the other relays. However, the type – either friendly or malicious – of relay behaviour is only its private information unknown to the other relays nor the destinations. By selecting its own WPLNC mapping the malicious node attacks the friendly behaving relays as well as the destinations. All the relays – both friendly and malicious – are rewarded/penalised according to how well their perform their appropriate tasks.

We describe the scenario as a static (single shot) incomplete information game among the relays and study it from the game theoretic perspectives. We mostly focus on conditions for existence of particular equilibrium points given the probability of malicious node existence and the valuation of the utility functions.

Since the WPLNC is quite sensitive to the selection of mapping functions the maliciously behaving node may cause significant damage to the network simply by selecting improper WPLNC mapping (obviously also by transmitting random data, etc.). Coexistence with the evil node is especially important for security critical applications.

The presence of maliciously behaving node among the fair players is widely analysed in the area of sensor networks. Game theoretical approach to coexistence in point to point scenario is presented in [113]. Relay networks (with single or multiple relay) with malicious nodes are in [114, 115] however both are single source cases. Our work presented here and originally in [116] extends the issues of malicious relays to multi-node cloud networks.

System model

A network of interest consists of two independent sources $\mathcal{S}_1, \mathcal{S}_2$, two relays $\mathcal{R}_R, \mathcal{R}_C$, where subscript R stands for row and C for column which is a very classical way to distinguish the players, and a destinations $\mathcal{D}$, see Figure 4.13. This is a multi-relay and wireless extension of the well-known butterfly network [117]. It is also the simplest network that can illustrate the issues studied here. The networks with more nodes are

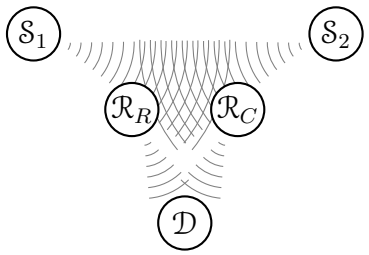


Figure 4.13: Two relay network – network of interest for malicious scenarios.

obviously possible but their analysis is impractical however the results shown can be straightforwardly extended.

The most important aspect is that we assume that the relay $\mathcal{R}_C$ is always “friendly" – its aim is always to support source to destination communication and it is rewarded for this task. However the relay $\mathcal{R}_R$ is not always friendly, with some probability it can behave maliciously. When this relay is malicious its aim is to cause harm to both $\mathcal{R}_C$ and the destinations by using such a processing that prevents $\mathcal{D}$ from recovery of the source data. Also the malicious relay is rewarded for the harm caused. Rewards and utility functions will be presented later. The relay $\mathcal{R}_C$ does not know the type of the $\mathcal{R}_R$ while $\mathcal{R}_R$ itself knows it.

Note that the transmissions of the source nodes are not orthogonally separated, i.e. relay $\mathcal{R}_i,\ i \in \{R, C\}$ receives superposition of both sources. For $\mathcal{R}$ to $\mathcal{D}$ communication it is assumed for the sake of simplicity that the individual transmissions can be obtained separately at the $\mathcal{D}$. The half-duplex assumption forces $\mathcal{S}$ to $\mathcal{R}$ and $\mathcal{R}$ to $\mathcal{D}$ transmissions to occur in two consecutive time slots.

Since WPLNC is assumed, each relay observes a superposition of incoming signals from both sources. Those observations may differ due to different parameters of the wireless links. Instead of distinguishing both individual sources the relay works with the overlapped signal as a whole. Relay processing – a decode and forward type – is generally given by a mapping

$$f_i : c_1 \times c_2 \to c_i, \tag{4.10}$$

where $i \in \{R, C\}$ denotes the relay, c_1, c_2 and c_i are symbols transmitted by $\mathcal{S}_1, \mathcal{S}_2$ and $\mathcal{R}_i$ respectively. Notice that for simplicity we assume mapping over individual symbols not over whole codewords, which is also possible. The mapping is allowed to be many-to-one. In fact to utilise the gain of WPLNC it must be the many-to-one mapping.

When the WPLNC mapping f_i given by Eq.(4.10) is a *linear* one then the WPLNC operation of the relay can be described by matrix multiplication

$$\bar{c}_i = \mathbb{X}_i \begin{bmatrix} \bar{c}_1 \\ \bar{c}_2 \end{bmatrix}, \tag{4.11}$$

where $i \in \{R, C\}$ denotes the relay, $\bar{c}_1, \bar{c}_2$ and $\bar{c}_i$ are binary representations of symbols transmitted by $\mathcal{S}_1, \mathcal{S}_2$ and $\mathcal{R}_i$, respectively and $\mathbb{X}_i$ is binary representation of WPLNC mapping, particular examples will follow later.

As an example we show the matrix description of the following WPLNC mappings. Notice that we assume the source symbol cardinality $|c| = 4$, i.e. 2 bits representation.

$$\mathbb{X}_i^1 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad \mathbb{X}_i^2 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix},$$
$$\mathbb{X}_i^3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbb{X}_i^4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.12}$$

where $\mathbb{X}_i^1$ and $\mathbb{X}_i^2$ are the minimal WPLNC mappings (the first one is the well known XOR mapping), $\mathbb{X}_i^3$ is the extended and $\mathbb{X}_i^4$ is the full mapping. Notice that cardinality category (MIN, EXT, FULL) directly corresponds with the number of (linear independent) rows of $\mathbb{X}_i$ since it defines the number of bits of the output symbols.

According to the assumptions every destination node observes c_R and c_C symbols separately and wants to recover original source symbols c_1, c_2 from them. This is possible if and only if it is allowed by the properties of both relay mappings f_R and f_C, i.e. if and only if an inverse mapping exits at the destination $\mathcal{D}_j$:

$$f_j^{-1} : c_R \times c_C \to c_1 \times c_2. \tag{4.13}$$

If this inverse mapping exists for f_R, f_C we call the pair of them an *invertible pair*, otherwise it is a *non-invertible* pair.

In terms of presented matrix description (for linear WPLNC mappings only) the pair of mapping is invertible at $\mathcal{D}_j$ if and only if there exist a matrix $\mathbb{X}_j^{-1}$ such that

$$\begin{bmatrix} \bar{c}_1 \\ \bar{c}_2 \end{bmatrix} = \mathbb{X}_j^{-1} \begin{bmatrix} \bar{c}_R \\ \bar{c}_C \end{bmatrix}, \tag{4.14}$$

where $\bar{c}_i$ is the bit representation of appropriate symbol.

Invertibility of the pair of mappings $\mathbb{X}_R, \mathbb{X}_C$ can be easily checked by row rank (in GF(2) sense) of a matrix that is formed by a vertical concatenation of $\mathbb{X}_R$ and $\mathbb{X}_C$. Notice that any FULL mapping forms invertible pair with arbitrary WPLNC mapping. Any FULL mapping, such as $\mathbb{X}_i^4$ in Eq.(4.12), is full rank on its own.

In a centralised network the relay mappings f_i can be tailored to fit the situation properly, but this is not the case in distributed control scenarios. When the relays perform selfish selection of f_i, with lack of knowledge of the other relay actions, the invertibility of the source data may be violated since the inverse mapping in Eq.(4.13) may not exist. The situation is even more complicated when the relay does not perfectly know the type of the opponent – whether it is friendly or malicious.

Incomplete information game

To analyse and study the problem of distributed selection of WPLNC mapping by the relays in scenarios with potential presence of malicious relays we describe the situation as an incomplete information non-cooperative game G of two players $\mathcal{R}_R, \mathcal{R}_C$. The game is the incomplete information game since the relays do not know the behaviour (friendly/malicious) of the other relay. Thus, the players do not know all the information relevant for their decision, a part of the information is private, revealed only to one of the nodes. A complete information game corresponds to the DLA algorithm.

By Harsanyi transform [118] the incomplete information game can be transformed into a Bayesian game which is solvable by standard game theoretic tools. All the nodes' uncertainties about the game (behaviour type of the other relay in our case) are transformed into the payoff uncertainties. In Static Bayesian games the first move is played by an "artificial" player – nature. The nature selects and assigns a type to each player. The type t_i of player i is revealed only to the player i, it is its private information. The nature's turn is a lottery but its prior probability is a common knowledge available to all players. After the type of each relay is revealed by the nature, the relays simultaneously choose their action. We assume static, i.e. single shot, version of the game. Based on the actions played and also on the players' types the payoffs are received by the players.

More formally the game of interest G is defined as follows. The *set of players* consist of $\mathcal{R}_R$ and $\mathcal{R}_C$. The *set of actions* of each player consist of several various WPLNC mappings f_i. To describe the problem generally we do not exactly specify the functions f_i now. Although the players' action sets can be much wider, the minimal action subset is as follows: two different minimal mappings $\mathrm{MIN}_{1,2}$, two different extended ones $\mathrm{EXT}_{1,2}$ and one full mapping. This set is capable of describing all the situations that can occur in specific instances of G – such as two minimal mappings that form an invertible pair, i.e. f_j^{-1} exists, a non-invertible pair of the minimal and the extended mapping, i.e. f_j^{-1} does not exist, etc. For the sake of simplicity we assume the minimal action set to be common to both players. We also assume that channel conditions are such that they allow the relays to use any mapping from their action set. For some channel conditions the performance (minimal distance, BER, etc.) of various mapping may differ, exhaustive discussion about the parametric channel performance of WPLNC can be found in [119, 120] and a possible solution for complete information game can be found in Section 4.2.

The *set of players' types* is $t_R \in \{\mathrm{F}, \mathrm{M}\}$ and $t_C \in \{\mathrm{F}\}$, where F stands for friendly and M for malicious. The type of $\mathcal{R}_R$ in particular realisation of the game G depends on the nature's turn – friendly version is chosen with $\Pr\{t_R = \mathrm{F}\} = p$ and malicious with $\Pr\{t_R = \mathrm{M}\} = 1 - p$. Notice that for the sake of simplicity we assume that only $\mathcal{R}_R$ has two possible incarnations, $\mathcal{R}_C$ is always friendly, i.e. $\Pr\{t_C = \mathrm{F}\} = 1$.

The players have *beliefs* about the opponent types, such as $p_R(t_C|t_R)$. In the analysed game it is assumed that player's belief is independent of the player's own type, e.g.

Table 4.3: Payoff matrix of two relay game G – Friendly vs. friendly.

F vs. f	MIN_1	MIN_2	EXT_1	EXT_2	FULL
MIN_1	P, P	$\mathbf{A}, \mathbf{A}$	A, B	P, P	A, C
MIN_2	$\mathbf{A}, \mathbf{A}$	P, P	P, P	A, B	A, C
EXT_1	B, A	P, P	B, B	P, P	B, C
EXT_2	P, P	B, A	P, P	B, B	B, C
FULL	C, A	C, A	C, B	C, B	C, C

$p_R(\text{F}|\text{F}) = p_R(\text{F}|\text{M}) = 1$. Finally there is a *payoff* function π_i of each player that assigns a reward to the player i. In contrast to complete information game the payoff function of the player i depends not only on the actions played by all the players but also on their types, i.e. $\pi_i(a_R, a_C; t_R, t_C)$ where a_R, a_C are actions of $\mathcal{R}_R, \mathcal{R}_C$ and t_R, t_C are their appropriate types.

In complete information games the terms action and strategy more or less coincide. The situation in Bayesian games is quite different. A strategy is a function that for every player type assigns an action from the set of actions. Thus, the strategy for player $\mathcal{R}_R$, with two possible types, is a doublet $s_R = (a_R(t_R = \text{F}), a_R(t_R = \text{M}))$, e.g. $s_R = (\text{MIN}_1, \text{MIN}_2)$ means that the relay $\mathcal{R}_R$ uses MIN_1 when it is the friendly relay while its malicious version uses MIN_2. Since there is only one type of $\mathcal{R}_C$ its strategy coincides with its action, e.g. $s_C = \text{EXT}_2$. We will use notation $s_i(t_i)$ to denote particular action used by player i of type t_i.

Since there are two types of the player $\mathcal{R}_R$ we can construct two payoff matrices describing the two game incarnations. The first one, when both relays are friendly, is shown in Tab.4.3. The second one, when the $\mathcal{R}_R$ is malicious is in Tab.4.4. In both matrices the column is determined by the action played by $\mathcal{R}_C$ and the row by the action of $\mathcal{R}_R$. The entries of the matrices have a form $\pi_R(a_R, a_C; t_R, t_C), \pi_C(a_R, a_C; t_R, t_C)$, the first number is a reward of $\mathcal{R}_R$ of type t_R when a pair of actions a_R, a_C is played. The second number gives a reward of $\mathcal{R}_C$ of type t_C. Notice that those two games are complete information games, it is the nature who selects which one of those games will be played by selecting the players' types.

For friendly players, see Tab.4.3, a payoff function is assumed, such that the payoff is determined only by the cardinality of the mapping (MIN, EXT, FULL) and their invertibility. The friendly relay ($\mathcal{R}_C$ always, $\mathcal{R}_R$ if its type t_R is friendly) is rewarded by A for using minimal mapping, B for the extended one and C for the full mapping. However when the friendly relays select a non-invertible pair of mappings, that do not allow the destinations to recover the source data, both relays have to pay a non-positive penalty P as a punishment. This situation is expressed in Tab.4.3 by (P, P) entries. We illustrate this on simple example when both relays select XOR mapping then the destination fails

Table 4.4: Payoff matrix of two relay game G – Malicious vs. friendly.

M vs. f	MIN_1	MIN_2	EXT_1	EXT_2	FULL
MIN_1	W, P	P, A	P, B	W, P	P, C
MIN_2	P, A	W, P	W, P	P, B	P, C
EXT_1	P, A	W, P	P, B	W, P	P, C
EXT_2	W, P	P, A	W, P	P, B	P, C
FULL	P, A	P, A	P, B	P, B	P, C

to recover c_A and c_B from its observations because it has received $c_A \oplus c_B$ twice, both relays are therefore penalised by P since they waste their energy and fail to make reliable $\mathcal{S}$ to $\mathcal{D}$ communication.

The payoff preference for friendly relays is given by

$$A > B > C > 0 \geq P, \tag{4.15}$$

thus the friendly relay always tries to minimise its output cardinality while the invertibility of WPLNC at the destination must be always guaranteed. This utility function is related to energy efficiency of the nodes. Given the fixed energy for transmission the lower the cardinality the lower the resulting BER. Moreover, any waste of energy due to non-invertible WPLNC mappings is punished and penalised.

The malicious relay ($\mathcal{R}_C$ never, $\mathcal{R}_R$ if its type t_R is malicious) is rewarded for the caused disorder, see Tab.4.4. It obtains a payoff W whenever it selects a WPLNC mapping that forms a non-invertible pair with the friendly relay. Under this condition the destination is unable to decode the source data and thus the friendly relay is transmitting and spending its resources in vain. When the malicious relay node selects a mapping that forms invertible pair with the friendly node it has failed in its task to cause harm to the network. In this case the malicious relay is in fact supporting the friendly relay and thus spends in vain its resources and it is penalised by P.

The payoff preference of the malicious relays is given by

$$W > 0 \geq P, \tag{4.16}$$

thus the malicious relay thus prefers to create the non-invertible pair. It can be easily seen that the full mapping is never optimal for such a relay since it is always invertible.

Nash equilibria analysis and discussion

In incomplete information games the solution concept of Nash equilibria is extended to so called Bayesian Nash equilibria (BNE) [118]. A strategy profile $s^* = (s_R^*, s_C^*)$ is a pure

strategy Bayesian Nash equilibrium if for each player i and each of its type t_i holds

$$s_i^*(t_i) = \arg\max_{a_i} \sum_{t_{-i}} p_i(t_{-i}|t_i)\pi_i(a_i, s_{-i}^*(t_{-i}); t_i, t_{-i}), \tag{4.17}$$

where subscript $_{-i}$ denotes the player other than i [121]. BNE strategy maximises the expected payoff of the player given the player's type and belief about the opponent. Similarly to the Nash equilibria (NE) in complete information games the BNE represents the point where no player has any incentive to alter its strategy. In contrast to NE the BNE depends not only on the strategies but also on players' types and their beliefs.

It can be easily shown that friendly vs. friendly incarnation of the game has multiple (two for assumed action set of the players) pure strategy NE, highlighted by bold font in Tab.4.3. Both equilibria use MIN WPLNC mappings and give optimal (in Pareto sense) outcomes to both players. This friendly relay game was exhaustively studied in [122] and led to DLA described in Deliverable D3.02 and implemented in WP5. In contrast it can be shown that malicious vs. friendly game incarnation has no pure strategy equilibria. Notice that for both versions of game some mixed strategies also exist.

Since friendly version of the game has NE that provides optimistic results from global network point of view – the information from the sources is successfully delivered to the destination while the relays' utilities are maximised – we focus on those particular equilibria and test if they are also BNE of incomplete information game where the maliciously behaving relay is possible.

Particularly we search for the conditions under which the strategies containing some MIN WPLNC mappings are BNE.

Proposition 4. *A pair of strategies $s_R = (MIN_2, MIN_1)$, $s_C = MIN_1$ is the BNE of the presented game under some constrains on the probability p of $\mathcal{R}_R$ being friendly and valuation of the utility functions π_i.*

Proof. To prove the proposition we have to simply show that s_R and s_C are mutual best responses and find out the conditions for them to be the BNE.

According to Eq.(4.17) the best response $s_C^{\mathrm{BR}}(t_C)$ of $\mathcal{R}_C$ on $s_R = (\mathrm{MIN}_2, \mathrm{MIN}_1)$ is

$$\begin{aligned} s_C^{\mathrm{BR}}(t_C) &= \arg\max_{a_C} \sum_{t_R} p_C(t_R|t_C)\pi_C(s_R(t_R), a_C; t_R, t_C) \\ &= \arg\max_{a_C} \{p\pi_C(\mathrm{MIN}_2, a_C; \mathrm{F}, \mathrm{F}) + \\ &+ (1-p)\pi_C(\mathrm{MIN}_1, a_C; \mathrm{M}, \mathrm{F})\} \end{aligned}$$

This gives the following expected payoffs for various actions a_C of $\mathcal{R}_C$:

$$a_C = \mathrm{MIN}_1 : \quad pA + P - pP \tag{4.18}$$
$$a_C = \mathrm{MIN}_2 : \quad pP + A - pA \tag{4.19}$$
$$a_C = \mathrm{EXT}_1 : \quad pP + B - pB \tag{4.20}$$
$$a_C = \mathrm{EXT}_2 : \quad pB + P - pP \tag{4.21}$$
$$a_C = \mathrm{FULL} : \quad C \tag{4.22}$$

For MIN_1 to be the best response to $s_R(t_R)$ the expected payoff in Eq.(4.18) must be higher than expected payoff from any other action a_C in Eq.(4.19 – 4.22). It can be shown that this is true for

$$p > \frac{1}{2} \quad \text{and} \quad p > \frac{C-P}{A-P}. \tag{4.23}$$

Conversely, $s_R^{\mathrm{BR}}(t_R)$ of $\mathcal{R}_R$ of type t_R on $s_C = \mathrm{MIN}_1$ is

$$s_R^{\mathrm{BR}}(t_R) = \arg\max_{a_R} \sum_{t_C} p_R(t_C|t_R)\pi_R(a_R, s_C(t_C); t_R, t_C),$$

which is for friendly type of $\mathcal{R}_R$

$$s_R^{\mathrm{BR}}(F) = \arg\max_{a_R} \pi_R(a_R, \mathrm{MIN}_1; \mathrm{F}, \mathrm{F}) = \mathrm{MIN}_2,$$

since the preference relation $A > B > C > 0 \geq P$.

Similarly, for the malicious type of $\mathcal{R}_R$ it is

$$s_R^{\mathrm{BR}}(M) = \arg\max_{a_R} \pi_R(a_R, \mathrm{MIN}_1; \mathrm{M}, \mathrm{F}) = \mathrm{MIN}_1 \text{ or } \mathrm{EXT}_2,$$

since the preference relation $W > 0 \geq P$.

Thus, we have shown that $s_R = (\mathrm{MIN}_2, \mathrm{MIN}_1)$, $s_C = \mathrm{MIN}_1$ are mutual best responses and thus form BNE of the presented incomplete information game. □

By similar reasoning and the game symmetry the following strategy pair can be also shown to be BNE under the same conditions

$$s_R = (\mathrm{MIN}_1, \mathrm{MIN}_2) \qquad s_C = \mathrm{MIN}_2$$

Under slightly varied conditions (details omitted here) also

$$s_R = (\mathrm{MIN}_1, \mathrm{EXT}_1) \qquad s_C = \mathrm{MIN}_2$$
$$s_R = (\mathrm{MIN}_2, \mathrm{EXT}_2) \qquad s_C = \mathrm{MIN}_1,$$

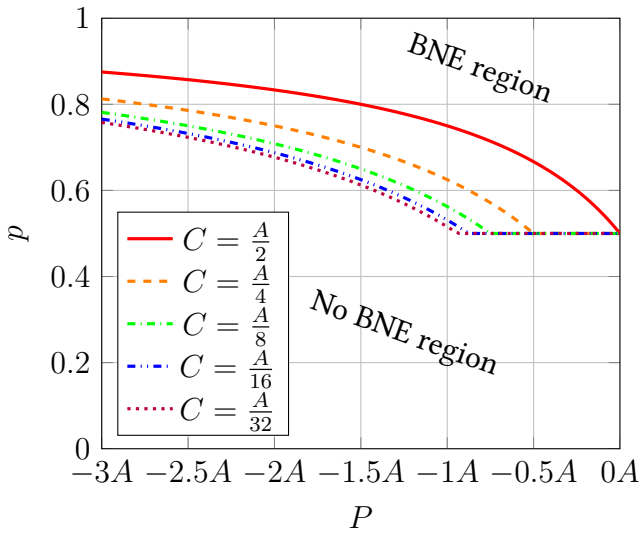


Figure 4.14: Regions where $s_R = (\text{MIN}_2, \text{MIN}_1), s_C = \text{MIN}_1$ is BNE based on p and utility valuation.

are the pure BNE. It is also possible that some other pure BNE, including both MIN and EXT WPLNC mappings, exist for completely different utility valuations. Also from the well-known game theoretic results a mixed strategy BNE is guaranteed to exist for any incomplete information game. Those results are not analysed here.

On the Figure 4.14 the region, defined by parameters in Eq.(4.23), where the strategy pair from Proposition 1 is the BNE is shown. The utility function valuations are expressed relatively to the value of A. The figure shows that given the probability p the lower the difference between the MIN and FULL mapping payoff A and C the lower penalty P must be paid by the nodes to have that particular BNE strategy. When the nodes are over-penalised and/or malicious relays are highly probable then there exits no studied strategy pair that is BNE.

4.3.2 Intentionally over-selfish relay

Here we focus on an example of relay network where a sequential selection of the WPLNC mappings takes place. We assume that the first relay may suffer from low battery level which may affect its selection of WPLNC mappings and that the second relay can deduce the battery state from observed actions and alter its behaviour to cope with the situation. For this scenario we discuss and evaluate an existence of game equilibria.

As highlighted in the previous part the WPLNC is quite sensitive to the selection of mapping functions and thus the maliciously behaving node may cause significant damage to the network simply by selecting improper WPLNC mapping. The malicious behaviour

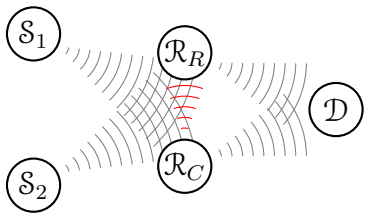


Figure 4.15: Two relay network with $\mathcal{R}_C$ to $\mathcal{R}_R$ signalling.

can be either intentional when evil node attacks the network (see above) or unintentional when the node running out of battery selects unexpected mappings to grab its last chance to deliver data. Our focus now will be laid on the latter case – unintentional malicious behaviour caused by emptying battery that forces the relay to be over-selfish and use the most energy saving WPLNC mappings.

It is also important to note that since we are focused on sensor networks, where the node are mostly simple battery powered devices that are not capable of any advanced channel coding technique, we investigate uncoded communication.

System model

A network of interest consists of two independent sources $\mathcal{S}_1, \mathcal{S}_2$, two relays $\mathcal{R}_R, \mathcal{R}_C$, where subscript R stands for row and C for column to distinguish the game players, and a destinations $\mathcal{D}$, see Figure 4.15. The important fact is that we assume that the relay $\mathcal{R}_C$ can suffer from low battery situation with probability p. When it is depleting its resources its main intent is to deliver as much as possible information to the destination in the most energy saving way. The other relay $\mathcal{R}_R$ has available observation of $\mathcal{R}_C$'s actions (shown in red in Figure 4.15) and can deduce his battery state and consequently support it. Since the $\mathcal{R}_C$ knows about its battery state we will call it a informed relay (player). The uninformed one – $\mathcal{R}_R$ – has only a chance to learn about $\mathcal{R}_C$ state from the received signal.

Signal processing related to the WPLNC is similar as in the malicious relay study please refer the system model described in Section 4.3.1 for details.

Sequential game

We describe the situation of differently informed relays, that sequentially select their WPLNC mappings, as a sequential incomplete information game. Our game of interest is defined as follows: A set of players is $\mathcal{R}_C$ and $\mathcal{R}_R$. Relay node $\mathcal{R}_C$ has two possible incarnations – types – it has either high or low battery level. Type of the relay $\mathcal{R}_C$ is $t_C \in \{H, L\}$, where H denotes high battery level and L low level. Relay $\mathcal{R}_R$ has only one type, it always has full battery. The type t_C is a private information of $\mathcal{R}_C$ and is

Table 4.5: Payoff matrix for high battery vs. high battery game.

	MIN1	MIN2	FULL
MIN1	P,P	2A,A	2A,B
MIN2	A,2A	P,P	A,B
FULL	B,2A	B,A	B,B

unknown to $\mathcal{R}_R$. A battery level is selected by nature prior to the beginning of the game. Low battery level is selected with probability p, this a priori probability is available to $\mathcal{R}_R$ too, but the actual battery level of $\mathcal{R}_C$ is not.

Each relay has a set of available actions, both sets contains various WPLNC mappings. To make the situation a bit easier we reduce the action set of $\mathcal{R}_R$ to two different minimal mappings MIN1 and MIN2 and a full mapping FULL. The action set of $\mathcal{R}_C$ depends on its type, high battery level type has the same action set as $\mathcal{R}_R$ since they are in fact the same. But low battery type uses minimal mappings only, i.e. its action set contains MIN1 and MIN2, this is because the minimal WPLNC mapping delivers data in the most energy saving way. For the sake of simplicity we do not assume EXT mappings in our example.

The relays are rewarded for the WPLNC mappings used. The payoff function is related to energy efficiency, since the MIN mappings compress the information more and thus save more energy. We define the payoff function as follows: At high battery level each node is more or less selfish, it tries to use the best (from its own point of view) possible map. Such a relay is awarded by $2A$ for using MIN1 mapping, by A for MIN2 and by B for full mapping, where $A > B > 0$. We assume without loss of generality that there are some minimal mappings that perform better (in terms of SER, etc.) in given channel conditions, that is why the rewards for various minimal mappings differ. We also assume that a minimal mapping is always better than a full mapping (from energy savings point of view). At low battery mode the relay $\mathcal{R}_C$ tries to deliver as much as possible information before its battery is depleted (ignoring any other performance measure, only energy efficiency matters). It is rewarded by $4A$ for any minimal mappings it uses, because of important energy savings. The relay $\mathcal{R}_R$, when facing low battery node, tries to help it and is rewarded by $4A$ for usage of full mapping since it helps most to $\mathcal{R}_C$ and by A for any invertible minimal mapping.

Whenever both relays select a pair of mapping that is non-invertible then the nodes waste their resources in vain and are penalised by $P \leq 0$ since they do not fulfil their task to enable $\mathcal{S}$ to $\mathcal{D}$ communication. Because wasting of resources is critical especially in low battery mode the relays are penalised by $2P$ when $\mathcal{R}_C$ is in low battery state. The payoff matrices of both game incarnations are given in Tabs. 4.5 and 4.6.

The uniformed player can form beliefs about the informed one based on observed actions and a priori type probability, e.g. after observing MIN1 mapping the $\mathcal{R}_R$ can create a

Table 4.6: Payoff matrix for high battery vs. low battery game.

	MIN1	MIN2
MIN1	2P,2P	2A,4A
MIN2	A,4A	2P,2P
FULL	4A,4A	4A,4A

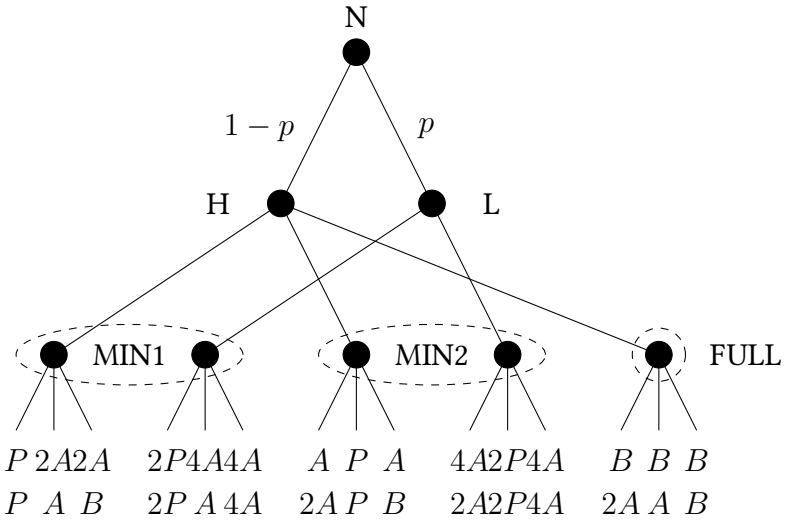


Figure 4.16: Game tree of sequential game.

belief $\mu(H|MIN1)$ of $\mathcal{R}_C$ being type H. The belief system is important when evaluating expected utilities and when searching for the game equilibria.

The game sequence and rules are depicted in Figure 4.16. The dashed regions show common information sets, e.g. the relay $\mathcal{R}_R$ is unsure about the type of $\mathcal{R}_C$ when it observes MIN1 mapping, it has only beliefs about the type, so both paths leading to MIN1 dashed region belong to the same information set. It can be also seen that FULL mapping is a singleton information set since it automatically means that $t_C = H$. Edges that represent relays' actions are from left to right MIN1, MIN2 and FULL. Rewards are shown at the leaves of the game tree, $\mathcal{R}_C$'s upper and $\mathcal{R}_R$'s lower.

Equilibria analysis and discussion

The proposed game is called a sequential incomplete information game. Also, due to presence of signalisation link, a signalling game. Widely accepted solution for this type of games is Perfect Bayesian Equilibrium (PBE) [121]. A PBE is a strategy profile $s^* = (a_C^*, a_R^*)$ and a belief system μ^* such that the strategies are sequentially rational given the

belief system and the belief system is consistent given the strategy profile. Sequential rationality means that the action taken in given information set is optimal given the belief and actions of the other players. Consistent belief systems means that the beliefs at least in the information sets on-the-equilibrium-path are given by Bayes' rule.

Since the relay $\mathcal{R}_C$ has two possible types its strategy is a doublet assigning an action to both types, e.g. $a_C = (a_C(H), a_C(L)) = (\text{MIN1}, \text{MIN2})$ means that $\mathcal{R}_C$ of type $t_C = H$ uses MIN1 while type $t_C = L$ plays MIN2. When the action played by different types is the same we call this a pooling strategy, otherwise it is a separating strategy.

In the proposed game there four separating:

$$\begin{aligned} a_C &= (a_C(H), a_C(L)) = (\text{MIN1}, \text{MIN2}) \\ a_C &= (a_C(H), a_C(L)) = (\text{MIN2}, \text{MIN1}) \\ a_C &= (a_C(H), a_C(L)) = (\text{MIN1}, \text{FULL}) \\ a_C &= (a_C(H), a_C(L)) = (\text{MIN2}, \text{FULL}), \end{aligned}$$

and two pooling strategies of $\mathcal{R}_C$:

$$\begin{aligned} a_C &= (a_C(H), a_C(L)) = (\text{MIN1}, \text{MIN1}) \\ a_C &= (a_C(H), a_C(L)) = (\text{MIN2}, \text{MIN2}) \end{aligned}$$

Since $\mathcal{R}_R$ has only one type its strategy is a singleton action a_R such as $a_R = \text{MIN1}$.

Proposition 5. *The only PBEs of the proposed games are connected with the pooling strategy* $a_C^* = (a_C^*(H), a_C^*(L)) = (\textit{MIN1}, \textit{MIN1})$. *Otherwise there is no PBE.*

Proof. To prove the Proposition 5 we have to show that there is a consistent belief system and sequentially rational strategies for both players. First of all define belief system: on the equilibrium path $\mathcal{R}_R$'s belief about $\mathcal{R}_C$ having high battery level after observing MIN1 is $\mu(H|\text{MIN1}) = 1-p$ which follows from Bayesian rule. Off the equilibrium path beliefs should be undefined such as $\mu(H|\text{MIN2})$ but due to singleton information set of FULL mapping the belief $\mu(H|\text{FULL})$ is simply 1. Thus, the belief system is consistent given a_C^*.

Having the consistent belief system we show the sequential rationality of the strategy. We first seek the best responses of $\mathcal{R}_R$ on observed signals from $\mathcal{R}_C$ having $\mathcal{R}_R$'s beliefs. The expected payoff of $\mathcal{R}_R$ obtained in response to $a_C = \text{MIN1}$ for all three actions of $\mathcal{R}_R$ is:

$$\begin{aligned} &\text{MIN1:} \quad (1-p)P + p2P = P - pP \\ &\text{MIN2:} \quad (1-p)A + pA = A \\ &\text{FULL:} \quad (1-p)B + p4A = B - pB + p4A \end{aligned}$$

Thus, the $\mathcal{R}_R$'s best response on action $a_C = \text{MIN1}$ is $a_R = \text{MIN2}$ if $p < \frac{A-B}{4A-B}$ or $a_R = \text{FULL}$ elsewhere, of course $p \in [0, 1]$.

The $\mathcal{R}_R$'s best response on action $a_C = \text{MIN2}$ is undefined since the belief $\mu(H|\text{MIN2})$ is not defined. Later we will refer to this potential best response as $a_R(\text{MIN2})$.

The $\mathcal{R}_R$'s best response on action $a_C = \text{FULL}$ is $a_R = \text{MIN1}$ simply because $2A > A > B$, see Figure 4.16.

Knowing the best response of $\mathcal{R}_R$, the strategy of $\mathcal{R}_C$ and belief system the assumed strategy is PBE if and only if $\mathcal{R}_C$ has no incentive to change its strategy, i.e. its expected payoff should be the highest among all possible alternative strategies. Let us start with the situation when the best response on $a_C = \text{MIN1}$ is $a_R = \text{MIN2}$. Then the expected payoff of $\mathcal{R}_C$ must be:

$$\pi_C(H, \text{MIN1}, \text{MIN2}) \geq \pi_C(H, \text{FULL}, \text{MIN1}) \tag{4.24}$$
$$\pi_C(H, \text{MIN1}, \text{MIN2}) \geq \pi_C(H, \text{MIN2}, a_R(\text{MIN2})) \tag{4.25}$$
$$\pi_C(L, \text{MIN1}, \text{MIN2}) \geq \pi_C(L, \text{MIN2}, a_R(\text{MIN2})), \tag{4.26}$$

where $\pi_C(t_C, a_C, a_R)$ is payoff of $\mathcal{R}_C$ of type t_C playing action a_C while $\mathcal{R}_R$ plays a_R. All equations are true since (4.24) is $2A \geq B$, (4.25) is $2A \geq A$ since A is maximal possible payoff of $\mathcal{R}_C$ given that strategy and (4.26) is $4A \geq 4A$ since $4A$ is maximal possible payoff of $\mathcal{R}_C$ given that strategy.

Secondly, assume that the best response on $a_C = \text{MIN1}$ is $a_R = \text{FULL}$. Then the expected payoff of $\mathcal{R}_C$ must be:

$$\pi_C(H, \text{MIN1}, \text{FULL}) \geq \pi_C(H, \text{FULL}, \text{MIN1}) \tag{4.27}$$
$$\pi_C(H, \text{MIN1}, \text{FULL}) \geq \pi_C(H, \text{MIN2}, a_R(\text{MIN2})) \tag{4.28}$$
$$\pi_C(L, \text{MIN1}, \text{FULL}) \geq \pi_C(L, \text{MIN2}, a_R(\text{MIN2})). \tag{4.29}$$

All equations are true since (4.27) is $2A \geq C$, (4.28) is $2A \geq A$ since A is maximal possible payoff of $\mathcal{R}_C$ given that strategy and (4.29) is $4A \geq 4A$ since $4A$ is maximal possible payoff of $\mathcal{R}_C$ given that strategy.

This holds true for arbitrary belief $\mu(H|\text{MIN2}) \in [0, 1]$. We have shown that there are two PBEs based on mutual relation between A, B and p – $s_1^* = (a_C^*(H), a_C^*(L), a_R^*) = (\text{MIN1}, \text{MIN1}, \text{MIN2})$ and $s_2^* = (a_C^*(H), a_C^*(L), a_R^*) = (\text{MIN1}, \text{MIN1}, \text{FULL})$ with belief system $\mu^*(H|\text{MIN1}) = 1 - p$, $\mu^*(H|\text{MIN2}) \in [0, 1]$ and $\mu^*(H|\text{FULL}) = 1$. s_1^* is PBE if $p < \frac{A-B}{4A-B}$ otherwise it is s_2^*.

By similar reasoning it can be shown that aforementioned pooling strategy of $\mathcal{R}_C$ is the only PBE of the proposed game. □

Figure 4.17 shows a surface that divides regions of existence of both PBE. Below this surface s_1^* is the PBE of the game, above it is s_2^*. The surface is given by probability of

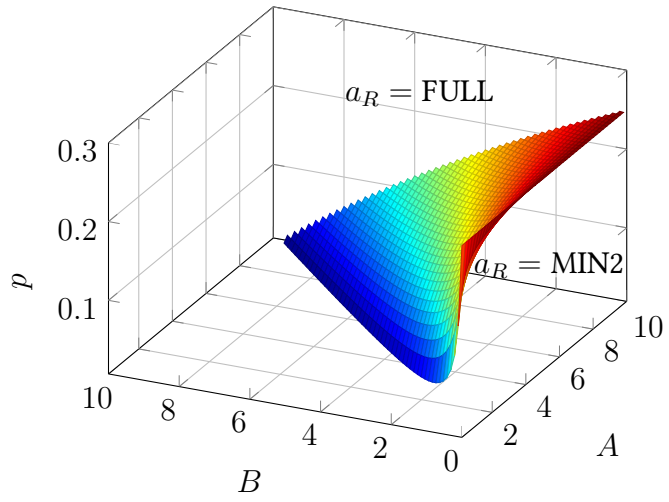


Figure 4.17: Regions of existence of individual PBEs.

low battery state $p = \frac{A-B}{4A-B}$ as a function of rewards A, B. It can be concluded that the higher the probability of depleted battery and the higher the reward B the relay $\mathcal{R}_R$ will prefer action FULL since it supports $\mathcal{R}_C$ and provides the higher payoff.

4.4 Self-selection of physical layer network coding parameters

4.4.1 Introduction

Compute-and-Forward (C&F) [123] is a relatively new relaying technique that relies on lattice codes through the linearity property of lattices. In the context of C&F relaying in fading channels, one challenging task is to find corresponding integer vectors in the relay nodes which is referred to as a network coding vector or an $\mathbf{a}$ vector in the literature, e.g. [123, 124]. There is intensive ongoing research on finding the best network coding vector ($\mathbf{a}$ vector), however, so far, most of the focus of the literature, including our works, is on obtaining an integer vector $\mathbf{a}$ that maximises the computation rate in the relay regardless of the corresponding $\mathbf{a}$ vectors computed in other relay nodes, i.e. local maximisation. Moreover, one necessary condition under which a destination node is capable of unambiguously decoding transmitted information by the source nodes is that the matrix (say $\mathbf{A}$ matrix[1]) obtained using the $\mathbf{a}$ vectors computed in the relay nodes must be non-singular. Using conventional methods for computing network coding vectors in the relays, although the rate is maximised locally, the overall computation rate of the network is zero if $\mathbf{A}$ matrix is singular, i.e. if $|\mathbf{A}| = 0$.

[1]The notions of $\mathbf{a}$ and $\mathbf{A}$ are introduced in [123] as integer vectors and matrices, respectively, and widely adopted in literature. We use the same notions in this chapter, too.

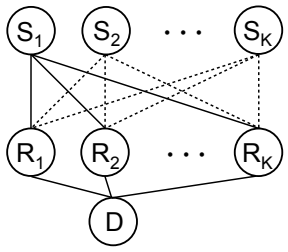


Figure 4.18: System model.

There are few chapter that directly address the problem of finding a proper $\mathbf{A}$ matrix that is not singular; in particular, [125–127] study a similar problem to the one that we address in this chapter, however, despite the ingenious methods proposed in [125] and [126], the problem is tackled assuming full coordination between the nodes that indeed imposes significant overhead signalling in practice. [127] proposes an algorithm that can be applied for a particular type of network with two source and two relay nodes; therefore, although it proposes a blind C&F relaying with no overhead signalling, the network topology is limited and cannot be applied to a network with an arbitrary number of source and relay nodes such as is of interest in this chapter.

Contribution In this chapter we study a multi-source multi-relay network and aim to compute network coding vector $\mathbf{a}$ in the relay nodes that tries to avoid singular $\mathbf{A}$. Based on two different assumptions, we propose two new algorithms in this chapter:

- Non-coordinated (blind) C&F where the relay nodes compute network coding vectors blindly, without knowledge of the network coding vectors used in the other relay nodes.
- Partially coordinated C&F where the relay nodes partially communicate to specify the order of transmission and use a network coding function that does not reduce the rank of the $\mathbf{A}$ matrix.

The chapter is organised as follows: In Subsection 4.4.2 system model is introduced and the rate description of C&F relaying is provided. In Subsection 4.4.3 two novel relaying strategies are proposed. In Subsection 4.4.4 numerical simulations are provided to validate the usefulness of the proposed methods and the chapter is finalised by some concluding remarks in Subsection 4.4.5.

4.4.2 System model

As shown by Figure 4.18, a cooperative network consisting of K source nodes, K relay nodes and one destination node is studied. The entire transmission from sources to the destination is divided into $K + 1$ time slots: in the first time slot all the source nodes

transmit their data to the destination using a shared interference channel (IC) that is referred to as the Multiple Access Channel (MAC) in the literature. In a second phase, that consists of K time slots, the relay nodes each compute an equation from the received superimposed signal and forward it to the destination node. The relay nodes exploit C&F and so in the second phase, relay nodes use orthogonal channels for transmission because the source node requires at least K equations to be capable of decoding all the messages transmitted from the source nodes. The transmissions from the source and relay nodes are summarised in the following:

Source Each source node selects a message $\mathbf{w}_l$ that is drawn from a set of M messages with equal probability. Every message is then mapped to a nested lattice codeword $\mathbf{x}_l$ and sent to the relay nodes in the first time slot.

Relay Since the relay nodes exploit C&F relying, each relay exploits lattice decoding and attempts to find a set of equations and the rates corresponding to each particular equation. Let us define the set $\mathbb{A}_r$ as a set of integer vectors defining possible network coding functions at relay r as follows

$$\mathbb{A}_r = \{\mathbf{a}_{r,1}, \mathbf{a}_{r,2}, \cdots, \mathbf{a}_{r,n}\}. \tag{4.30}$$

Each integer vector $\mathbf{a}_{r,i}$ results in a computation rate that is stored in set $\mathbb{R}_r$ as follows

$$\mathbb{R}_r = \{\mathcal{R}_{r,1}, \mathcal{R}_{r,2}, \cdots, \mathcal{R}_{r,n}\}. \tag{4.31}$$

It is assumed that $\mathcal{R}_{r,1} \geq \mathcal{R}_{r,2} \geq \cdots \geq \mathcal{R}_{r,n}$. The relay function will be discussed in further detail in the next subsections, however, note that it is proved in [123] that the computation rate in relay node r is obtained using following expression:

$$\mathcal{R}_r(\mathbf{h}, \mathbf{a}) = \log_2^+ \left(\frac{\gamma}{\gamma \parallel \alpha_{r,l}\mathbf{h}_r - \mathbf{a}_{r,l} \parallel^2 + \alpha_{r,l}^2} \right) \tag{4.32}$$

which depends on inflation coefficient $\alpha_{r,l}$, channel realisation $\mathbf{h}_r$ and choice of the integer vector $\mathbf{a}_{r,l}$, see [123] for detailed description of the parameters.

It is clear that choosing $\mathbf{a}_{r,1}$ is the best option if the intention is to maximise the computation rate locally in the relay nodes; this is indeed the main optimisation criterion in the original C&F chapter in [123]. However, in this chapter, we are interested in optimising the overall transmission rate of the network, defined as follows

$$\mathcal{R}(\mathbf{H}, \mathbf{A}) = \begin{cases} \min(\mathcal{R}_1, \cdots, \mathcal{R}_K), & \text{if } |\mathbf{A}| \neq 0 \\ 0, & \text{if } |\mathbf{A}| = 0 \end{cases} \tag{4.33}$$

where $\mathbf{H}$ is the channel realisation between the source and the relay nodes and $\mathbf{A}$ is the matrix whose columns are the $\mathbf{a}_r$ vectors exploited in the relay nodes as the network

coding vectors. We assume that each relay appends its chosen integer vector to the equation and transmits it to the destination; also we assume that the relays can overhear one another's transmissions.

4.4.3 Relay strategy

Upon reception of the source transmissions, each relay node r needs to choose an integer coefficient $\mathbf{a}_r$ and perform lattice decoding before forwarding an equation towards the destination. One can assume different criteria for computing $\mathbf{a}_r$ vectors as described below.

Non-coordinated (blind) compute-and-forward

Once the destination collects the relay transmissions, it will be capable of decoding the messages from the source nodes if the matrix $\mathbf{A}$ is a full rank matrix. In an attempt to reduce the occasions which result in non full rank $\mathbf{A}$, a blind C&F relay strategy is proposed in the following:

Proposition Instead of computing an equation that corresponds to the highest computation rate in the relay r, i.e. locally optimising rate, each relay computes a set of equations corresponding to different computation rates as described in (4.30) and (4.31). Moreover, a new parameter is defined as

$$\kappa_r \quad = \quad \{k_{r,1}, k_{r,2}, \cdots, k_{r,n}\} \tag{4.34}$$

which specifies the number of non-zero entries in $\mathbf{a}_r$ vectors. As an example, an integer vector $\mathbf{a}_{r,j} = [1, 0, 0]$ consists of information only from source 1, however an integer vector $\mathbf{a}_{r,j} = [1, 0, 1]$ consists of information from two source nodes, source 1 and source 3. We define $k_{r,j}$ as the number of non-zero entries in the $\mathbf{a}_{r,j}$ vector, i.e. $\mathrm{nnz}(\mathbf{a}_{r,j}) = k_{r,j}$. As a relaying strategy, instead of forwarding $\mathbf{a}_{r,1}$, we propose to transmit a function that includes information from, at least, m sources, i.e. $\mathbf{a}_r = \mathbf{a}_{r,j}$ where

$$\mathcal{R}_{r,j} = \max\{\mathbb{R}_r\}\,\mathbf{T} \text{ given } k_{r,j} \geq m. \tag{4.35}$$

This strategy is helpful, especially at low SNR where the integer vectors $\mathbf{a}_{r,j}$ usually have only one non-zero entry. Therefore, once the integer vector from other relays has a non-zero entry at the same position of $\mathbf{a}_{r,j}$, $|\mathbf{A}|$ becomes equal to zero, and hence setting the overall transmission rate of the network to zero. Whereas ensuring that at least m entries of the integer vectors $\mathbf{a}_{r,j}$ are non-zero, the probability of non-full rank $\mathbf{A}$ matrix decreases, and hence avoiding $\mathcal{R}(\mathbf{H}, \mathbf{A}) = 0$ due to $|\mathbf{A}| = 0$. In Subsection 4.4.4 computer simulations are provided to validate the benefits of the proposed algorithm.

Partially coordinated compute-and-forward

In Blind C&F algorithms, the relay nodes are indexed arbitrarily and so there is no rule to decide the order with which the relays transmit their equations. In other words, it is implicitly assumed that relay R_1 transmits first, and then the relay R_2 and etc. However, for a partially coordinated C&F algorithm as proposed in this subsection we define a parameter referred to as *rate-difference* as follows

$$d_r = \mathcal{R}_{r,1} - \mathcal{R}_{r,2}, \tag{4.36}$$

that is the rate difference between two largest rates in each relay. In the following, it will be proposed to give the priority for transmission to the relay nodes with larger d_r; for instance, in a two relay scenario, if $d_2 > d_1$, the relay R_2 transmits its computed equation first and then the relay R_1 transmits an equation.

Proposition Partially coordinated C&F protocol proposed in this subsection consists of two parts: (i) sorting relays and specifying the priority of the transmission and (ii) choosing the best equation in the relays, i.e. choosing proper integer vector $\mathbf{a}$, which simultaneously guarantees local optimisation of the computation rate as well as preserving the rank of the $\mathbf{A}$ matrix. The algorithm is described in the following:

- Upon reception, every relay computes a set of best equations, leading to largest rates and corresponding rates with which the relays calculate the rate difference and broadcast it. Since we assume the relays can overhear each other, each relay receives the rate difference of other relays and based on the rate differences, the relays are ordered for transmission as described earlier, i.e. the relays with larger rate-difference d_r get priority for transmission. At the end, a motivation for this is explained in *Remark*. For simplicity of notation, let us assume that the relay indices specify the order of transmission. In other words, we assume that $d_1 > d_2 > \cdots > d_n$ and so, R_1 is the first relay to transmit an equation, R_2 is the second relay and similarly, R_n is the last relay that transmits.

- Each relay appends the exploited integer vector to the frame and sends it to the destination. For instance, R_1 sends its integer vector $\mathbf{a}_1$ along with the equation; the R_2 overhears the $\mathbf{a}_1$ and exploits an integer vector $\mathbf{a}_2$ that does not reduce the rank of $[\mathbf{a}_1; \mathbf{a}_2]$ matrix. Relay R_3 overhears and decodes $\mathbf{a}_1$ and $\mathbf{a}_2$ from R_1 and R_2 transmissions and exploits a proper $\mathbf{a}_3$ that does not reduce the rank of $[\mathbf{a}_1; \mathbf{a}_2; \mathbf{a}_3]$. The transmission continues until all the relays transmit their corresponding data while ensuring that choosing an integer vector $\mathbf{a}_r$ does not lead to a non full-rank $\mathbf{A}$ matrix.

In order to better understand the algorithm, an example is provided in the following.

Example Assume a network with three source and three relay nodes, operating at SNR = 10 dB, with channel realisations between the source and the relay nodes as follows

$$\begin{aligned} \mathbf{h}_1 &= [0.85, 3.63, 1.91]^T \\ \mathbf{h}_2 &= [0.14, 13.7, 7.52]^T \\ \mathbf{h}_3 &= [2.37, 0.92, 4.51]^T . \end{aligned} \tag{4.37}$$

For each relay, one can compute a set of integer vectors $\mathbb{A}_r = \{\mathbf{a}_{r,1}, \mathbf{a}_{r,2}, \mathbf{a}_{r,3}, \cdots\}$ with which the rates of $\mathbb{R}_r = \{\mathcal{R}_{r,1}, \mathcal{R}_{r,2}, \mathcal{R}_{r,3}, \cdots\}$ can be achieved (note that we assume the entries of $\mathbb{R}_r$ are ordered in descending order). For instance for relay R_1, we have computed $\mathbb{A}_1$ and $\mathbb{R}_1$, with three entries, as follows

$$\mathbb{A}_1 = \{\mathbf{a}_{1,1}, \mathbf{a}_{1,2}, \mathbf{a}_{1,3}\} \quad \text{and} \quad \mathbb{R}_1 = \{\mathcal{R}_{1,1}, \mathcal{R}_{1,2}, \mathcal{R}_{1,3}\}$$

where

$$\mathbf{a}_{1,1} = [0, 0, 1],\ \mathbf{a}_{1,2} = [1, 2, 5], \quad \mathbf{a}_{1,3} = [0, 1, 2] \tag{4.38}$$

$$\mathcal{R}_{1,1} = 0.971, \quad \mathcal{R}_{1,2} = 0.943, \quad \mathcal{R}_{1,3} = 0.234. \tag{4.39}$$

Likewise, one can compute the entries of $\mathbb{A}_2$ and $\mathbb{A}_3$ as follows

$$\mathbf{a}_{2,1} = [0, 0, 1],\ \mathbf{a}_{2,2} = [2, -5, 5],\ \mathbf{a}_{2,3} = [1, -3, 3] \tag{4.40}$$

$$\mathbf{a}_{3,1} = [1, 0, 0],\ \mathbf{a}_{3,2} = [0, 0, 1], \quad \mathbf{a}_{3,3} = [2, 1, 0] \tag{4.41}$$

and $\mathbb{R}_2$ and $\mathbb{R}_3$ as follows

$$\mathbb{R}_2 = \{3.15, 1.77, 0.02\} \quad \text{and} \quad \mathbb{R}_3 = \{1.15, 1.01, 0.02.\} .$$

Consequently, the rate-difference d_r defined in (4.36) for the three relays can be defined as

$$d_1 = 0.028, \quad d_2 = 1.38 \ \text{ and } \ d_3 = 0.14. \tag{4.42}$$

Since $d_2 > d3 > d1$ in (4.42), we propose to order relay transmission based on the rate loss. In this example, second relay R_2 transmits as the first relay because the largest rate loss occurs in R_2; therefore, it selects the best $\mathbf{a}$ vector corresponding to largest rate, i.e. the second relay chooses $\mathbf{a}_2 = [0, 0, 1]$ that corresponds to $\mathcal{R}_2 = 3.15$. Along with the transmission of the equation based on $\mathbf{a}_2$, the relays transmit $\mathbf{a}$ vector, too. Upon reception of the $\mathbf{a}$ vector by the other relays, they decode it and store for future use. Now there are two more relays to transmit their equations, however, since $d_3 > d_1$, the third relay transmits first. The best option for third relay is to choose $\mathbf{a}_3 = [1, 0, 0]$ and note that this choice does not reduce the rank of $\mathbf{A}$ matrix. Relay R_3 sends its equation along with the chosen $\mathbf{a}_3$ that is overheard and decoded by relay R_1. The first relay is the last relay to send its equation, however, although the best option for relay R_1 is $\mathbf{a}_1 = [0, 0, 1]$, this choice reduces the rank of $\mathbf{A}$ matrix and sets $|\mathbf{A}| = 0$; therefore as it selects second integer vector from set $\mathbb{A}_1$, i.e. $\mathbf{a}_1 = [1, 2, 5]$. Note that although the first relay selects its second best $\mathbf{a}$ vector, it leads to insignificant rate loss because the corresponding rate-difference is low ($d_1 = 0.028$).

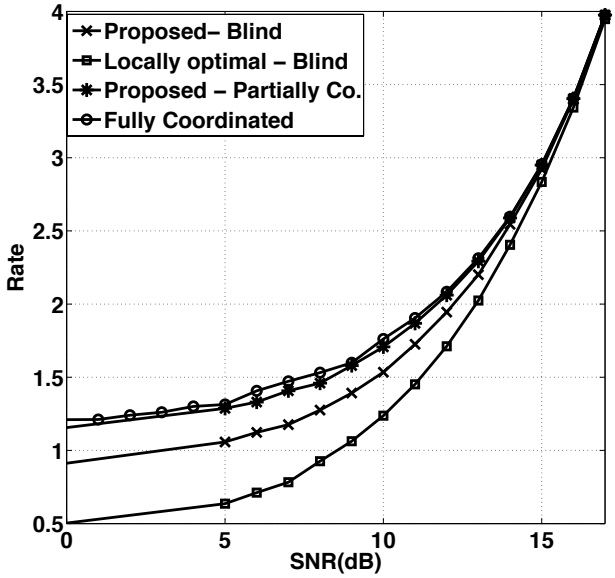


Figure 4.19: Computation rate: 3 users ($K = 3$).

Remark It was proposed, without a motivation, to sort relays based on rate-difference in descending order, i.e. the relay with larger rate-difference d_r gets priority for transmission. Note that in order to avoid singular $\mathbf{A}$, every relay decodes the integer vectors transmitted by the previous relay nodes and based on previous integer vectors, chooses an integer vector that does not reduce the rank of the $\mathbf{A}$ matrix. Hence, if an integer vector corresponding to the best rate in a relay lead to a singular $\mathbf{A}$, the relay selects another integer vector with lower rate but full rank $\mathbf{A}$. Therefore, if a relay with larger d_r is forced to choose its second integer vector, this will lead to large rate loss in the relay nodes locally, and so we propose to give priority for transmission for the relays with larger d_r.

Fully coordinated compute-and-forward

In fully coordinated C&F relaying, all the nodes know the parameters required for optimisation of the whole network, see [125]. This algorithm can result in significant overhead signalling, however, this is exploited in next subsection as a benchmark which validates the usefulness of the algorithms proposed in this chapter.

4.4.4 Numerical results

In this subsection numerical results for two relay networks with three and five source/relay nodes are provided, i.e. $K = 3$ and $K = 5$ in Figure 4.18. In Figure 4.18 we assume that the distance between any two neighbouring nodes is one meter and the path

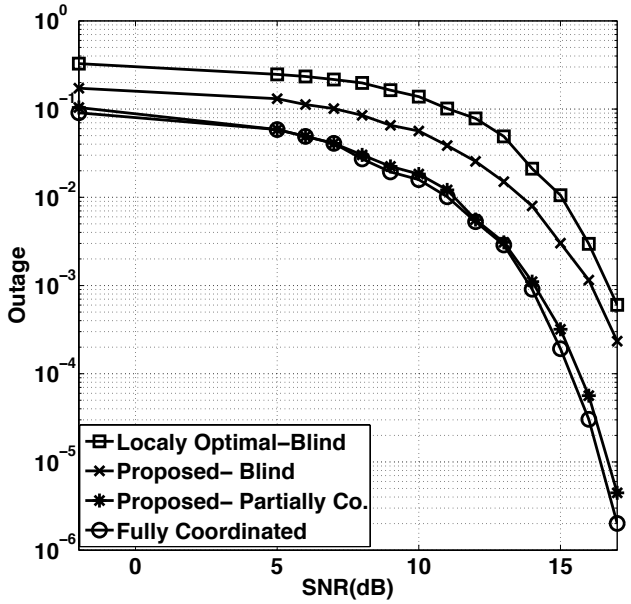


Figure 4.20: Outage rate: 3 users ($K = 3$) and threshold rate $\mathcal{R}_{th} = 1$.

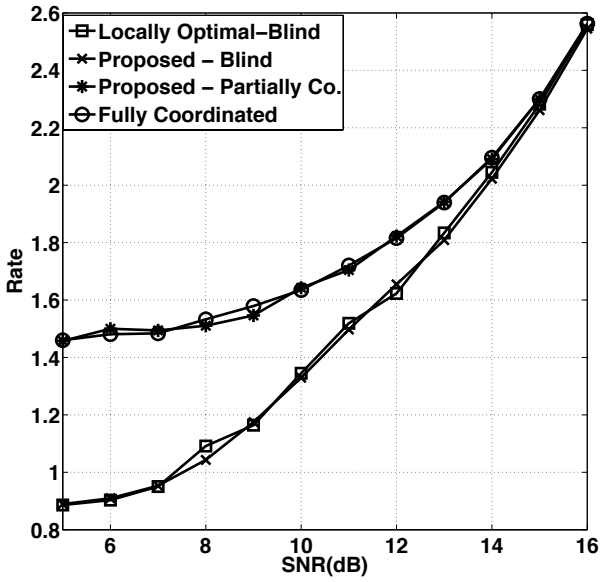


Figure 4.21: Computation rate: 5 users ($K = 5$).

loss coefficient is $\alpha = 3$. We assume block Rayleigh fading channels that are obtained through $h_{ij} = (\frac{d_{ij}}{d_0})^{-\alpha}\tilde{h}_{ij}$ where $\tilde{h}_{ij}$ represent fading realisation between S_i and R_j. d_0 is the largest distance between a source node and a relay node, e.g. S_1 and R_K.

Figure 4.19 illustrates the computation rate (defined in (4.33)) using the proposed blind C&F algorithm; each relay makes sure that network coding function includes data from at least two transmitters, i.e. $k_{r,j} \geq 2$ in (4.35). For comparison, the computation rate of

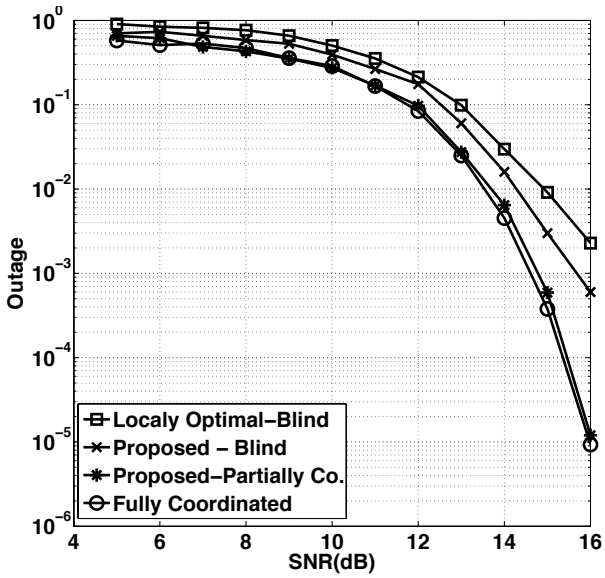


Figure 4.22: Outage rate: 5 users ($K = 5$) and threshold rate $\mathcal{R}_{th} = 1.5$.

the conventional blind C&F algorithm is also provided; it is clear that the proposed blind C&F achieves higher rates. Figure 4.20 illustrates the outage rate assuming threshold rate $\mathcal{R}_{th} = 1.5$. Clearly, the outage rate of the proposed blind algorithm is lower than that of the conventional blind algorithm; this validates the usefulness of the blind C&F algorithm proposed in this chapter. In Figures 4.19 and 4.20, the computation rate and outage are also shown for the partially coordinated C&F algorithm. It is clear that the proposed algorithm that is developed by exchanging a few parameters among the relays, i.e. rate difference and integer coefficients, approaches the fully coordinated C&F that requires significant signalling. Figures 4.21 and 4.22 illustrate computation rate and outage for a system with five source and relay nodes. The superior performance of the proposed algorithms is evident.

4.4.5 Conclusion

Compute-and-Forward (C&F) relaying in a multi-source multi-relay network is studied in this chapter and two novel algorithms are proposed. Assuming no coordination between the nodes, a blind C&F technique is developed. Another algorithm is proposed that requires the exchange of a few parameters between the nodes. This algorithm is called partially coordinated C&F and it is demonstrated to perform nearly as well as a fully coordinated C&F system.

4.5 Novel approach for computing network coding function in compute-and-forward

4.5.1 Introduction

Since it was proved that it is possible to achieve Shannon capacity using lattice codes [128–130] and lattice decoding [131] on a point-to-point communication system over a Gaussian channel, lattice encoding/decoding has received significant attention in research community as another class of capacity achieving codes in communication systems. In addition, a new relaying protocol referred to as Compute-and-Forward (C&F) was introduced based on the ideas developed in [123]. This relaying protocol exploits lattice codes in a scenario where multiple users transmit data in an interference channel, and thus extends its application to multihop wireless networks such as occur in Wireless Sensor Networks (WSN) and the "Internet of Things". Since it is able to handle signals from multiple sources in an optimum way without treating them as deleterious interference, it has the potential to greatly increase energy and spectrum efficiency in such networks. C&F is in fact a Physical Layer Network Coding[2] (PLNC) scheme that performs two main functions:

- Forward Error Correcting (FEC) encoding/decoding using lattice codes,
- Network coding using mod-lattice operation,

where for the FEC decoding using lattice decoder, the receiver aims at finding a "superimposed" lattice point that is corrupted by thermal noise and self-noise. Then, the network coding function uses the mod-lattice function. Note that the idea of PLNC has been investigated in the literature in the past, e.g. the XOR function is a very common network coding function [119]. However, exploiting the mod-lattice operation as the network coding function was first proposed in [123].

In the context of fading channels, one challenging task in C&F relaying, which might in practice limit the applicability of the scheme, is to find an integer vector which is referred to as the $\boldsymbol{a}$ vector, e.g. [123, 132]. There is intensive ongoing research on finding the best $\boldsymbol{a}$ vector. Note that finding $\boldsymbol{a}$ vector is in fact a well known problem referred to as a Shortest Vector Problem (SVP) in discrete mathematics and it is known to be an NP-hard problem. Owing to the complexity of solving this, there are several schemes with lower complexity to tackle the problem: e.g. the Fincke–Phost method was proposed in [125] as one of the methods to find the integer vector $\boldsymbol{a}$; however, since it is of exponential complexity, in higher dimensions the algorithm will be prohibitively complex and so not appealing from a practical point of view. A branch-and-bound algorithm for convex quadratic integer programming was proposed in [133] which is suitable for estimating the $\boldsymbol{a}$ vector and it was exploited in [134] as another method for calculating $\boldsymbol{a}$, however, the performance of the algorithm decreases as the number of users increases and also in

[2] In the literature, it is also called Wireless Network Coding (WNC).

the high SNR regime. One well-known approach to this problem is using lattice reduction, e.g. [135, 136], and notably the LLL (Lenstra–Lenstra–Lovasz) algorithm[137, 138]. This algorithm is widely adopted in C&F systems, however, it becomes also computationally very expensive at high SNR and when the number of the users increases due to its polynomial complexity. In order to manage the complexity of the LLL, a complexity-performance trade-off parameter referred to as δ has been offered in [137] where the performance of the algorithm is sacrificed in favour of providing manageable complexity. Therefore, the LLL algorithm in practice provides an estimate for the "optimal" $\boldsymbol{a}$ vector and does not guarantee to provide the optimum $\boldsymbol{a}$ with affordable complexity. Indeed, it will be observed in this chapter that for a large number of users the computation rate obtained using the LLL algorithm is outperformed by our proposed algorithm, depending on the SNR and the number of users, when the LLL algorithm is set to operate in an affordable complexity region. A Quadratic Programming (QP) approach was proposed in [139] with low complexity and giving computation rates close to the LLL algorithm. However our proposed algorithm outperforms this QP approach as well as the LLL algorithm. An exhaustive search approach was also introduced in [140], which provides an upper bound on the performance of these algorithms, but is more complex: our algorithm is extremely close in performance to this bound, except at very low SNR. Apart from the works cited here, there is a large body of work on the problem of finding the $\boldsymbol{a}$ vector and so for a complete survey, the interested reader is recommended to read [125, 133–140] and references therein.

Contribution

In this chapter, we propose a new approach to solve the problem of finding the integer vector $\boldsymbol{a}$. In this scheme, by relaxing $\boldsymbol{a}$ to be *non*-integer (say $\bar{\boldsymbol{a}} \in \mathbb{R}$), we first study the noise function and decompose it into two additive terms where one of which is a symmetric function around zero, almost[3] independent of the channel as well as the transmit power and the other is a convex function which has its minimum on a line. Therefore, in this chapter we look for an integer vector $\boldsymbol{a}$ in the proximity of this minimising line.

It will be revealed that the proposed algorithm on the one hand reduces the complexity of the system compared to existing algorithms because it only performs a set of rounding operations and on the other hand outperforms the other well known algorithms, e.g. LLL for $\delta \leq 0.95$, in terms of computation rate. This enables the practical C&F protocol to achieve higher computation rates using the proposed algorithm. The complexity-performance trade-off of the proposed algorithm is defined and compared with that of the existing algorithms through numerical simulations, where the superior performance of the proposed scheme is demonstrated.

[3]The symmetric part of noise depends only on the channel norm and so its symmetric property is maintained.

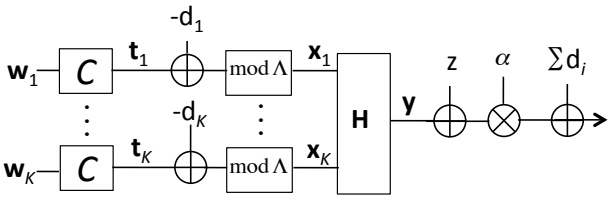


Figure 4.23: System model.

The chapter is organised as follows. In Subsection 4.5.2, the system model is explained and a summary of related literature is provided. In Subsection 4.5.3, we focus on partially minimising noise and develop an algorithm to find the integer $\boldsymbol{a}$. The performance of the system based on the proposed algorithm is verified in Subsection 4.5.4 using numerical simulations. We briefly discuss the complexity in Subsection 4.5.5 and finalise the chapter with some concluding remarks in Subsection 4.5.6.

Notation

Matrices are represented by boldface upper cases ($\boldsymbol{M}$). Row vectors are denoted by boldface lower cases ($\boldsymbol{h}$), and h_i indicates the i-th element of $\boldsymbol{h}$. The superscript $(\cdot)^T$ stands for transposition and we refer to the identity matrix by $\boldsymbol{I}$. Round, floor and ceiling values of real-valued x are indicated by $\lfloor x \rceil$, $\lfloor x \rfloor$ and $\lceil x \rceil$, respectively.

4.5.2 System Model and Preliminaries

The MAC phase of an interference channel based on C&F as shown in Figure 4.23 is studied in this chapter. A detailed description of the system can be found in [123], however, in order to keep the consistency of notation, the MAC phase and required formulations are repeated in this subsection.

We consider a MAC phase of a system, where n users send their messages $\boldsymbol{w}_i$ via a shared interference channel to a receiver node after encoding the messages using a common lattice code and applying the $[\,\cdot\,] \bmod \Lambda$ operation. The signal received at the destination is y, which upon reception is multiplied by a factor α and dither components are subtracted as $\bar{y} = \alpha y - \sum_{i=1}^{n} d_i$. After basic algebraic manipulation, $\bar{y}$ can be written as

$$\bar{y} = \sum_{i=1}^{n} a_i \boldsymbol{t}_i + \underbrace{\sum_{i=1}^{n} (\alpha h_i - a_i)\boldsymbol{x}_i}_{\text{self noise}} + \underbrace{\alpha^2 z}_{\text{thermal noise}} \tag{4.43}$$

where a_i is an integer value to be derived in the chapter, α is scaling factor, h_i is the channel coefficient, $\boldsymbol{x}_i$ a uniformly distributed vector inside the shaping lattice and z is

the normally distributed thermal noise[4]. For the two noise components distinguished in (4.43), the noise power corresponding to the "self noise" (P_{SN}) and the thermal noise (P_{G}) are given by

$$P_{\text{SN}} = \gamma \parallel \alpha \boldsymbol{h} - \boldsymbol{a} \parallel^2 \tag{4.44}$$

$$P_{\text{G}} = \alpha^2, \tag{4.45}$$

where it is assumed that the thermal noise variance $N_0 = 1$ throughout this chapter. It is proved in [123] that the computation rate of this system is

$$\mathcal{R}(\boldsymbol{h}, \boldsymbol{a}) = \log_2^+ \left(\frac{\gamma}{\gamma \parallel \alpha \boldsymbol{h} - \boldsymbol{a} \parallel^2 + \alpha^2} \right) \tag{4.46}$$

and the optimal value of α leading to the optimal computation rate derived in (4.46) is

$$\alpha = \frac{\gamma \boldsymbol{a} \boldsymbol{h}^T}{1 + \gamma \parallel \boldsymbol{h} \parallel^2}. \tag{4.47}$$

Note that although the optimal solution (4.47) for α is proposed in [123], the optimal $\boldsymbol{a}$ is not further investigated, and consequently, the optimal α was left as a function of $\boldsymbol{a}$.

In the following subsections, we propose a novel approach to handle the problem of finding the optimal $\boldsymbol{a}$ using simple mathematics based on partial derivative practice. Note that the proposed algorithm is optimal at high SNR and/or when the number of the users is large, i.e. when $\gamma \parallel \boldsymbol{h} \parallel^2 \gg 1$. Since we assume high SNR in this chapter, the α derived in (4.47) can be simplified to

$$\alpha_{\text{SO}} = \frac{\boldsymbol{a} \boldsymbol{h}^T}{\parallel \boldsymbol{h} \parallel^2} \tag{4.48}$$

when $\gamma ||h||^2 \gg 1$ and the subscript "SO" stands for suboptimal[5]. This provides an alternative α for high SNR regime. Moreover, not only is α_{SO} in (4.48) optimal at high SNR, it also minimises the self noise, regardless high or low SNR. In other words, (4.48) is also the solution for the following minimisation problem:

$$\begin{aligned} \alpha_{\text{SO}} &= \arg\min_{\boldsymbol{a} \in \mathbb{Z}} \left(P_{\text{SN}} \right) \\ &= \arg\min_{\boldsymbol{a} \in \mathbb{Z}} \left(\alpha^2 \parallel \boldsymbol{h} \parallel^2 - 2\alpha \boldsymbol{a} \boldsymbol{h}^T + \parallel \boldsymbol{a} \parallel^2 \right) \\ &= \frac{\boldsymbol{a} \boldsymbol{h}^T}{\parallel \boldsymbol{h} \parallel^2}, \end{aligned} \tag{4.49}$$

which is obtained by solving $\frac{\partial P_{\text{SN}}}{\partial \alpha} = 0$, which is the same as the value derived in (4.48) for the high SNR regime. Hence, choosing the suboptimal scaling factor α_{SO} guarantees (i) to minimise total noise at high SNR and (ii) to minimise the self noise in entire SNR region. It will be used in the following subsections for finding the integer vector $\boldsymbol{a}$.

[4] A complete description of the system parameters can be found in [123]

[5] Although α_{SO} derived in (4.48) is suboptimal at low SNR, it is optimal at high SNR when $\gamma \parallel \boldsymbol{h} \parallel^2 \gg 1$.

4.5.3 Finding Optimal a at High SNR

In this subsection, the integer vector $\boldsymbol{a}$ that is optimal at high SNR is calculated. The noise power (P_{tot}) can be written through (4.44), (4.45) and (4.48) as follows

$$\begin{aligned} P_{\text{tot}} &= P_{\text{SN}} + P_{\text{G}} \\ &= \gamma \parallel \alpha \boldsymbol{h} - \boldsymbol{a} \parallel^2 + \alpha^2 \\ &= \gamma \boldsymbol{a} \Big(\frac{\boldsymbol{h}^T \boldsymbol{h}}{\parallel \boldsymbol{h} \parallel^2} - \boldsymbol{I} \Big) \Big(\frac{\boldsymbol{h}^T \boldsymbol{h}}{\parallel \boldsymbol{h} \parallel^2} - \boldsymbol{I} \Big)^T \boldsymbol{a}^T + \boldsymbol{a} \frac{\boldsymbol{h}^T \boldsymbol{h}}{\parallel \boldsymbol{h} \parallel^4} \boldsymbol{a}^T \\ &= \gamma \boldsymbol{a} \Big(\boldsymbol{I} - \frac{\boldsymbol{h}^T \boldsymbol{h}}{\parallel \boldsymbol{h} \parallel^2} \Big) \boldsymbol{a}^T + \boldsymbol{a} \frac{\boldsymbol{h}^T \boldsymbol{h}}{\parallel \boldsymbol{h} \parallel^4} \boldsymbol{a}^T; \end{aligned} \tag{4.50}$$

then by adding and subtracting $\boldsymbol{I}$ from the second term in (4.50) and through basic algebraic manipulation, we have

$$P_{\text{tot}} = (\gamma - \frac{1}{\parallel \boldsymbol{h} \parallel^2}) \boldsymbol{a} \Big(\boldsymbol{I} - \frac{\boldsymbol{h}^T \boldsymbol{h}}{\parallel \boldsymbol{h} \parallel^2} \Big) \boldsymbol{a}^T + \frac{\parallel \boldsymbol{a} \parallel^2}{\parallel \boldsymbol{h} \parallel^2}. \tag{4.51}$$

Careful inspection of (4.51) reveals that the second term of the noise power, i.e. $\frac{\|\boldsymbol{a}\|^2}{\|\boldsymbol{h}\|^2}$, is a convex quadratic function which depends only on the channel's second norm and has its minimum point at $\mathbf{0}$; moreover, considering that we assume $\gamma \parallel \boldsymbol{h} \parallel^2 >> 1$ in this chapter, i.e. high SNR, it is clear that the second term in (4.51) is negligible when $\parallel \boldsymbol{a} \parallel^2$ is relatively small and so focusing on minimising the first term in (4.51) is justified.

In the following, we focus on the first term of (4.51) and find arbitrary $\boldsymbol{a}$ (say $\bar{\boldsymbol{a}} \in \mathbb{R}$) that minimises the first term in (4.51), i.e.

$$\begin{aligned} \bar{\boldsymbol{a}} &= \arg\min_{\bar{\boldsymbol{a}} \in \mathbb{R}} \Big\{ (\gamma - \frac{1}{\parallel \boldsymbol{h} \parallel^2}) \bar{\boldsymbol{a}} \Big(\boldsymbol{I} - \frac{\boldsymbol{h}^T \boldsymbol{h}}{\parallel \boldsymbol{h} \parallel^2} \Big) \bar{\boldsymbol{a}}^T \Big\} \\ &\triangleq \arg\min_{\bar{\boldsymbol{a}} \in \mathbb{R}} \Big\{ \bar{\boldsymbol{a}} \underbrace{\Big(\boldsymbol{I} - \frac{\boldsymbol{h}^T \boldsymbol{h}}{\parallel \boldsymbol{h} \parallel^2} \Big)}_{\boldsymbol{M}} \bar{\boldsymbol{a}}^T \Big\}, \end{aligned} \tag{4.52}$$

where in the second line the $(\gamma - \frac{1}{\|\boldsymbol{h}\|^2})$ coefficient is neglected because the minimisation is independent of this coefficient. The result of this minimisation problem is summarised in the following theorem.

Theorem 4. *Assuming $\boldsymbol{M} = \boldsymbol{I} - \frac{\boldsymbol{h}^T \boldsymbol{h}}{\|\boldsymbol{h}\|^2}$, the $\bar{\boldsymbol{a}}$ value minimising $\bar{\boldsymbol{a}} \boldsymbol{M} \bar{\boldsymbol{a}}^T$ is on a line as follows*

$$\begin{aligned} \bar{\boldsymbol{a}} &= \arg\min_{\bar{a} \in \mathbb{R}} \big\{ \bar{\boldsymbol{a}} \boldsymbol{M} \bar{\boldsymbol{a}}^T \big\} \\ &= \begin{cases} \bar{a}_1 = \frac{u_1}{u_k} \bar{a}_k \\ \vdots \\ \bar{a}_n = \frac{u_n}{u_k} \bar{a}_k \end{cases}, \text{ for every } \bar{a}_k \in \mathbb{R} \end{aligned} \tag{4.53}$$

where $\boldsymbol{u} = [u_1, u_2, \cdots, u_n]$ is a unitary vector corresponding to the zero eigenvalue of $\boldsymbol{M}$.

Proof. Regardless of the channel distribution, since $\boldsymbol{h}^T\boldsymbol{h}$ is a positive symmetric matrix, it is clear that $\boldsymbol{M}$ is also positive symmetric and can be written through eigenvalue decomposition (EVD) as

$$\boldsymbol{M} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T. \tag{4.54}$$

It is proved in Appendix 4.5.7 that $\boldsymbol{\Lambda}$ is a diagonal matrix where all the diagonal entries are one except only the last diagonal entry, which is zero, i.e.

$$\boldsymbol{\Lambda}_{n\times n} = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{bmatrix} \tag{4.55}$$

and $\boldsymbol{V}$ is a unitary matrix with its columns as the eigenvectors of the $\boldsymbol{M}$ matrix as follows

$$\boldsymbol{V} = [\boldsymbol{v}_1 \cdot \boldsymbol{v}_{n-1}\boldsymbol{v}_n] = \begin{bmatrix} v_{11} & \cdots & v_{1n-1} & v_{1n} \\ v_{21} & \cdots & v_{2n-1} & v_{2n} \\ \vdots & & \vdots & \vdots \\ v_{n1} & \cdots & v_{nn-1} & v_{nn} \end{bmatrix} \tag{4.56}$$

where $\boldsymbol{v}_n$ corresponds to the zero eigenvalue in (4.55). Considering that the last diagonal entry of the $\boldsymbol{\Lambda}$ matrix is zero, one can write the argument of (4.52) as

$$\mathcal{S} = \bar{\boldsymbol{a}}\boldsymbol{M}\bar{\boldsymbol{a}}^T \doteq \bar{\boldsymbol{a}}\boldsymbol{W}\boldsymbol{W}^T\bar{\boldsymbol{a}}^T \tag{4.57}$$

where $\boldsymbol{W}$ is a non-square matrix as

$$\boldsymbol{W}_{n\times n-1} = [\boldsymbol{v}_1 \cdots \boldsymbol{v}_{n-1}] \tag{4.58}$$

and note that $\boldsymbol{v}_n$ is omitted from $\boldsymbol{W}$. For simplicity of notation, let us define a new parameter $\boldsymbol{u} \doteq \boldsymbol{v}_n$, i.e.

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} v_{1n} \\ v_{2n} \\ \vdots \\ v_{nn} \end{bmatrix}. \tag{4.59}$$

It is proved in Appendix 4.5.8 that the expression in (4.57) can be written in the following form

$$\mathcal{S} = \bar{a}_1^2 + \bar{a}_2^2 + \cdots + \bar{a}_n^2 - (\bar{a}_1u_1 + \bar{a}_2u_2 + \cdots + \bar{a}_nu_n)^2 \tag{4.60}$$

where, interestingly, the u_i values are the entries of the $\boldsymbol{v}_n$ vector that were cancelled in (4.57). Through further algebraic manipulation, one can rewrite (4.60) as

$$\mathcal{S} = \sum_{i=1}^{n}(1-u_i^2)\bar{a}_i^2 - \prod_{j=i+1}^{n} 2\bar{a}_i\bar{a}_j u_i u_j. \tag{4.61}$$

Note that by deriving (4.61), the problem of minimising $\bar{\boldsymbol{a}}\boldsymbol{M}\bar{\boldsymbol{a}}^T$ in (4.52), which depends on $\boldsymbol{M}$ with n^2 entries has reduced to minimising equation (4.61) with n entries of the vector $\boldsymbol{u} = \boldsymbol{v}_n$. In order to get a better understanding about the optimisation of the expression in (4.61), we have derived it for $n = 2$ and 3 in the following. These examples will act as illustrations of the general case. For $n = 2$, we have

$$\mathcal{S} = (1-u_1^2)\bar{a}_1^2 + (1-u_2^2)\bar{a}_2^2 - 2u_1u_2\bar{a}_1\bar{a}_2 \tag{4.62}$$

and for $n = 3$, one can write

$$\begin{aligned}\mathcal{S} \;=\; & (1-u_1^2)\bar{a}_1^2 + (1-u_2^2)\bar{a}_2^2 + (1-u_3^2)\bar{a}_3^2 \\ & -2u_1u_2\bar{a}_1\bar{a}_2 - 2u_1u_3\bar{a}_1\bar{a}_3 - 2u_2u_3\bar{a}_2\bar{a}_3.\end{aligned} \tag{4.63}$$

Careful inspection of (4.61) reveals that $\mathcal{S}$ is a convex function because $\boldsymbol{u}$ is a unitary vector where $u_1^2+u_2^2+\cdots+u_n^2 = 1$, and so one can write $u_i < 1$; therefore the expression of $1-u_i^2 > 0$ for all i and so

$$\frac{\partial^2\mathcal{S}}{\partial\bar{a}_i^2} > 0 \tag{4.64}$$

for all i. Therefore, the $\mathcal{S}$ expression in (4.57), or equivalently, the $\bar{\boldsymbol{a}}\boldsymbol{M}\bar{\boldsymbol{a}}^T$ expression in (4.52) has a global minimum point that is *not* necessarily integer. We will find that $\mathcal{S}$ has its global minimum on a line (instead of single point) and this provides the possibility of choosing desired values for $\boldsymbol{a}$ instead of single minimum point.

In order to find the global minimum point, one should find the following set of partial derivative equations and set them equal to zero as follows

$$\begin{aligned}\frac{\partial\mathcal{S}}{\partial\bar{a}_1} &= 2(1-u_1^2)\bar{a}_1 - \sum_{i=1,i\neq1}^{n} 2u_1u_i\bar{a}_i = 0 \\ \frac{\partial\mathcal{S}}{\partial\bar{a}_2} &= 2(1-u_2^2)\bar{a}_2 - \sum_{i=1,i\neq2}^{n} 2u_2u_i\bar{a}_i = 0 \\ &\vdots \\ \frac{\partial\mathcal{S}}{\partial\bar{a}_n} &= 2(1-u_n^2)\bar{a}_n - \sum_{i=1,i\neq n}^{n} 2u_nu_i\bar{a}_i = 0.\end{aligned} \tag{4.65}$$

Note that we use $\bar{a}_i$ instead of a_i because the idea is to find the global minimum where the variables are not necessarily integer; therefore, we reserve a_i to represent integer

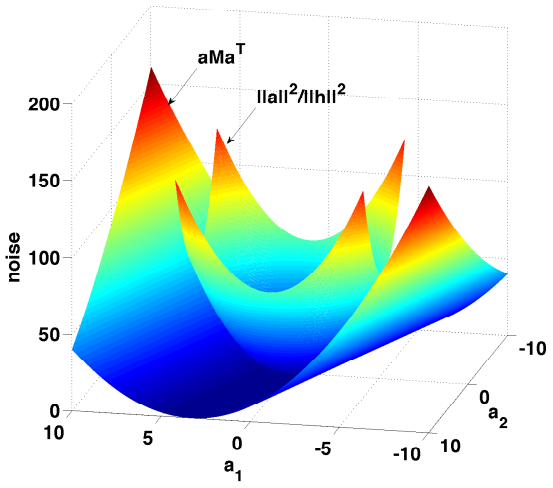


Figure 4.24: The noise power as a function of $\bar{a}_1$ and $\bar{a}_2$ decomposed into two parts in (4.52): $\boldsymbol{aMa}$ and $\frac{\|\boldsymbol{a}\|^2}{\|\boldsymbol{h}\|^2}$.

values and $\bar{a}_i$ for real values, i.e. $a_i \in \mathbb{Z}$ and $\bar{a}_i \in \mathbb{R}$. At this stage, we focus on solving the set of equations using examples; this will allow us to find simpler equations for the problem than the set of equations in (4.65).

Example 1, n = 2: Assuming $n = 2$, one should solve only two equations from (4.65) as

$$\frac{\partial \mathcal{S}}{\partial \bar{a}_1} = 2(1-u_1^2)\bar{a}_1 - 2u_1u_2\bar{a}_2 = 0 \tag{4.66}$$

$$\frac{\partial \mathcal{S}}{\partial \bar{a}_2} = 2(1-u_1^2)\bar{a}_2 - 2u_1u_2\bar{a}_1 = 0, \tag{4.67}$$

where assuming $u_1^2 + u_2^2 = 1$, one can solve (4.66) as follows

$$\frac{\partial \mathcal{S}}{\partial \bar{a}_1} = 0 \quad \text{if} \quad \bar{a}_1 = \frac{u_1}{u_2}\bar{a}_2; \tag{4.68}$$

then, substituting (4.68) in(4.67) reveals that $\frac{\partial \mathcal{S}}{\partial \bar{a}_2} = 0$ is always valid regardless of $\bar{a}_2$. Therefore, one can conclude that the $\mathcal{S}$ function has its global minimum on the line

$$\bar{a}_1 = \frac{u_1}{u_2}\bar{a}_2. \tag{4.69}$$

Figure 4.24 illustrates the $\mathcal{S}$ function and its global minimum line indicated by $\boldsymbol{aMa}^T$. The plot indicated by $\frac{\|\boldsymbol{a}\|^2}{\|\boldsymbol{h}\|^2}$ shows the symmetric function of (4.51). Note that at high SNR or with a large number of users the symmetric function becomes nearly flat and so the minimisation of the total noise reduces to minimisation of the $\mathcal{S}$ function.

Example 2, n = 3: Assuming $n = 2$, one should solve three equations from (4.65) as

$$\frac{\partial \mathcal{S}}{\partial \bar{a}_1} = 2(1-u_1^2)\bar{a}_1 - 2u_1u_2\bar{a}_2 - 2u_1u_3\bar{a}_3 = 0 \tag{4.70}$$

$$\frac{\partial \mathcal{S}}{\partial \bar{a}_2} = 2(1-u_2^2)\bar{a}_2 - 2u_1u_2\bar{a}_1 - 2u_2u_3\bar{a}_3 = 0 \tag{4.71}$$

$$\frac{\partial \mathcal{S}}{\partial \bar{a}_3} = 2(1-u_3^2)\bar{a}_3 - 2u_1u_3\bar{a}_1 - 2u_2u_3\bar{a}_2 = 0. \tag{4.72}$$

By solving (4.70), one can obtain

$$\bar{a}_1 = \frac{u_1u_2}{1-u_1^2}\bar{a}_2 + \frac{u_1u_3}{1-u_1^2}\bar{a}_3, \tag{4.73}$$

then, substituting $\bar{a}_1$ in (4.71) and solving it leads to

$$\bar{a}_2 = \frac{u_2}{u_3}\bar{a}_3; \tag{4.74}$$

by substituting (4.74) in (4.73), it is interesting to obtain the familiar equation

$$\bar{a}_1 = \frac{u_1}{u_3}\bar{a}_3. \tag{4.75}$$

Finally, having (4.74) and (4.75), one can prove that $\frac{\partial \mathcal{S}}{\partial \bar{a}_3}$ in (4.72) holds regardless of $\bar{a}_3$. Therefore, the global minimum of $\mathcal{S}$ occurs on the following line

$$\begin{cases} \bar{a}_1 = \frac{u_1}{u_3}\bar{a}_3 \\ \bar{a}_2 = \frac{u_2}{u_3}\bar{a}_3. \end{cases}, \qquad \bar{a}_3 \in \mathbb{R}. \tag{4.76}$$

For $n = 2$ and 3, it was proved that the global minimum of $\mathcal{S}$ occurs on a line. Moreover, we found the $\bar{a}_1$ and $\bar{a}_2$ as a function of $\bar{a}_3$, however, one can choose an arbitrary $\bar{a}_k$ and minimise the function $\mathcal{S}$ based on $\bar{a}_k$ instead of $\bar{a}_n$.

Further focusing on (4.65) reveals that for $n > 3$ the equations are similar to the case of $n = 3$ in (4.70)–(4.72) except the number of equations increases, and hence for arbitrary n one can obtain $\bar{a}_1$ by solving $\frac{\partial \mathcal{S}}{\partial \bar{a}_1} = 0$ in (4.65) as

$$\bar{a}_1 = \frac{1}{1-u_1^2}\left(\bar{a}_2u_1u_2 + \bar{a}_3u_1u_3 + \cdots + \bar{a}_nu_1u_n\right); \tag{4.77}$$

then, substituting (4.77) in (4.65) and solving $\frac{\partial \mathcal{S}}{\partial \bar{a}_2} = 0$ leads to

$$\bar{a}_2 = \frac{1}{1-u_1^2-u_2^2}\left(\bar{a}_3u_2u_3 + \bar{a}_4u_2u_4 + \cdots + \bar{a}_nu_2u_n\right). \tag{4.78}$$

By successively solving the partial derivative equations in (4.65), one will reach $\frac{\partial \mathcal{S}}{\partial \bar{a}_{n-1}} = 0$ which leads to

$$\bar{a}_{n-1} = \frac{1}{1-u_1^2-u_2^2-\cdots-u_{n-1}^2}\bar{a}_nu_{n-1}u_n. \tag{4.79}$$

Since $1 - u_1^2 - u_2^2 - \cdots - u_{n-1}^2 = u_n^2$, the expression in (4.79) will be simplified to

$$\bar{a}_{n-1} = \frac{u_{n-1}}{u_n}\bar{a}_n, \tag{4.80}$$

that is the same expression derived in (4.69) and (4.75) for particular scenarios of $n = 2$ and $n = 3$, respectively. Consequently, the global minimum of $\mathcal{S}$ function is

$$\begin{aligned} \bar{\boldsymbol{a}} &= \underset{\bar{\boldsymbol{a}}\in\mathbb{R}}{\arg\min}\left\{\bar{\boldsymbol{a}}\boldsymbol{M}\bar{\boldsymbol{a}}^T\right\} \\ &= \begin{cases} \bar{a}_1 = \frac{u_1}{u_k}\bar{a}_k \\ \bar{a}_2 = \frac{u_2}{u_k}\bar{a}_k \\ \quad\vdots \\ \bar{a}_n = \frac{u_n}{u_k}\bar{a}_k \end{cases} \quad \text{for every } \bar{a}_k \in \mathbb{R}. \end{aligned} \tag{4.81}$$

This completes proof of the theorem. □

Note that so far we assume arbitrary $\bar{\boldsymbol{a}} \in \mathbb{R}$, however, below we propose to quantise $\bar{\boldsymbol{a}}$ to obtain the $\boldsymbol{a}$ vector at the proximity of the minimum line.

Estimating optimal integer vector $\boldsymbol{a}$ from $\bar{\boldsymbol{a}}$

Now that we have found the global minimum of $\mathcal{S}$ function, one can expect that the global minimum of P_{tot} with integer $\boldsymbol{a}$ also occurs in the proximity of the global minimum of $\mathcal{S}$, i.e. in the proximity of the line derived in (4.81). In order to focus on the proposed algorithm, it is assumed that the channels are non-negative in this subsection, however, it will be revealed that the algorithm is consistent regardless of the channel sign. This assumption is relaxed in *Remark 1* at the end of the subsection.

Proposition: Since there is no constraint on the choice of $\bar{a}_k$ in (4.81), we propose to choose $\bar{a}_k$ to be integer, i.e. $a_k = \bar{a}_k$; then the corresponding values for $\bar{a}_i$, $i = 1, \cdots, n$ will be obtained from (4.81) as $\bar{a}_i = \frac{u_i}{u_k}a_k$. Then, the integer a_i can be determined by rounding $\bar{a}_i$, i.e. $a_i = \lfloor \bar{a}_i \rceil$. Through numerical simulations, it will be revealed that although $a_i = \lfloor \bar{a}_i \rceil$ results in superior performance compared to the existing algorithms at high SNR, the performance is not satisfactory at low SNR because in some cases $\lfloor \bar{a}_i \rceil = \lceil \bar{a}_i \rceil$ which results in enhancing the symmetric part of the noise $||\bar{\boldsymbol{a}}||^2$ in (4.51), and hence reducing the performance. Therefore, we propose deriving $a_i = \lfloor \bar{a}_i \rceil$ and $\lfloor \bar{a}_i \rfloor$ and choose the one leading to a better performance.

Although the proposed algorithm introduces a way to obtain integer values for a_i, it does not explain how to choose a_k, so that the total noise power in (4.50) is minimised. For this scenario, we propose to perform an exhaustive search over integer values for a_k and select the integer a_k that minimises P_{tot}. The algorithm is summarised in Table 4.7:

Table 4.7: Algorithm 1.

Algorithm 1. Finding integer vector $\boldsymbol{a}$
Initialise: Set m to be a positive integer
$[\boldsymbol{V}, \boldsymbol{\Lambda}] = \text{eig}(\boldsymbol{M})$, descending order
$\boldsymbol{u} = \boldsymbol{v}_n$
for $a_k = 1$ to m
Step 1) $a_i = \lfloor \frac{u_i}{u_k} a_k \rceil$ and $\lfloor \frac{u_i}{u_k} a_k \rfloor$
Step 2) Calculate total noise power P_{tot}
Step 3) Save P_{tot} and $\boldsymbol{a}$
end
Select $\boldsymbol{a}$ corresponding to minimum P_{tot}

Note that the exhaustive search here is over only one variable a_k and so it does not introduce much complexity to the proposed algorithm. Moreover, from [123] it is well known that

$$\| \boldsymbol{a} \|^2 < 1 + \gamma \| \boldsymbol{h} \|^2, \tag{4.82}$$

and so the search for appropriate a_k, i.e. the value of m in Table 4.7, is limited by (4.82). It will be demonstrated through numerical simulations that this exhaustive search over a_k is performed over fifty values at high SNR and as few as five values at low SNR and that is enough to achieve better performance than the LLL algorithm which operates with a high complexity, e.g. complexity-performance trade-off $\delta = 0.95$.

Further simplification of the problem

The main complexity of the proposed algorithm in Table 4.7 depends on performing eigenvalue decomposition which has a complexity of order $\mathcal{O}(n^3)$. In the following theorem, we provide another solution which does not require the eigenvalue decomposition in the proposed algorithm.

Theorem 5. *Assuming* $\boldsymbol{M} = \boldsymbol{I} - \frac{\boldsymbol{h}^T \boldsymbol{h}}{\|\boldsymbol{h}\|^2}$ *and* $\boldsymbol{u}$ *the eigenvector corresponding to the non-zero (unit) eigenvalue of* $\frac{\boldsymbol{h}^T \boldsymbol{h}}{\|\boldsymbol{h}\|^2}$ *(as defined in* (4.59)*), the following ratios hold:*

$$\frac{u_i}{u_k} = \frac{h_i}{h_k}, \text{ for } i = 1, \cdots, n. \tag{4.83}$$

where u_i *values are the entries of the unit vector* $\boldsymbol{u}$*.*

Proof. In order to prove $\frac{u_i}{u_k} = \frac{h_i}{h_k}$, we use the fact that it has been proved in Appendix 4.5.7 that $\boldsymbol{V} = \boldsymbol{Q}$ where $\boldsymbol{V}$ is the eigenvector matrix of $\boldsymbol{M}$ as defined in (4.54) and $\boldsymbol{Q}$ is the

eigenvector matrix of $\frac{\boldsymbol{h}^T\boldsymbol{h}}{\|\boldsymbol{h}\|^2}$ that is defined in (4.89) in Appendix 4.5.7. Consequently, we have $\boldsymbol{v}_n = \boldsymbol{q}_n$ and due to the definition $\boldsymbol{u} = \boldsymbol{v}_n$ in (4.59), one can also write $\boldsymbol{u} = \boldsymbol{q}_n$.

As described in (4.89)-(4.90) in Appendix 4.5.7, the eigenvalue corresponding to $\boldsymbol{u} = \boldsymbol{q}_n$ is equal to one, and so using the definition of the eigenvalue, one can write

$$\frac{\boldsymbol{h}^T\boldsymbol{h}}{\| \boldsymbol{h} \|^2}\boldsymbol{u} = 1\boldsymbol{u}. \tag{4.84}$$

Then, considering that

$$\boldsymbol{h}^T\boldsymbol{h} = \begin{bmatrix} h_1^2 & h_1h_2 & \cdots & h_1h_n \\ h_1h_2 & h_2^2 & \cdots & h_2h_n \\ \vdots & \vdots & & \vdots \\ h_1h_n & h_2h_n & \cdots & h_n^2 \end{bmatrix}, \tag{4.85}$$

by substituting (4.85) in (4.84) and doing simple algebraic manipulation, (4.84) can be written in the form of a set of equations as follows

$$\begin{cases} h_1u_1 + h_2u_2 + \cdots + h_nu_n = \frac{\|\boldsymbol{h}\|^2}{h_1}u_1 \\ h_1u_1 + h_2u_2 + \cdots + h_nu_n = \frac{\|\boldsymbol{h}\|^2}{h_2}u_2 \\ \vdots \\ h_1u_1 + h_2u_2 + \cdots + h_nu_n = \frac{\|\boldsymbol{h}\|^2}{h_n}u_n \end{cases}. \tag{4.86}$$

Since the left-hand side of the equations are equal, so are the right-hand sides, and hence it can be easily concluded that

$$\frac{u_i}{u_k} = \frac{h_i}{h_k}, \text{ for } i = 1, \cdots, n, \tag{4.87}$$

and hence the theorem is proved. □

Since it is proved that $\frac{u_i}{u_k} = \frac{h_i}{h_k}$, the algorithm in Table 4.7 does not require the eigenvalue decomposition and the $\frac{u_i}{u_k}$ can be replaced with the ratio of the channel coefficients $\frac{h_i}{h_k}$. The simplified algorithm is summarised in Table 4.8 where the eigenvalue decomposition is no longer required, and hence the complexity is further reduced to a set of rounding operations.

Remark 1: In the proposed algorithm in Table 4.8, the a_i values are determined using a_k, h_k and h_i through $\frac{h_i}{h_k}a_k$ where a_k is defined to be positive. Therefore, one can determine the absolute value of the components of $\boldsymbol{a}$ assuming that all the channels are non-negative; then, the sign of a_i can be determined by the sign of $\frac{h_i}{h_k}$, i.e. $\text{sign}(a_k) = \text{sign}(\frac{h_i}{h_k})$.

Remark 2: The algorithm in this chapter is developed for a real scenario where both the integer vector $\boldsymbol{a}$ and the lattice codes $\boldsymbol{t}_i$ are assumed to be real, however, the results

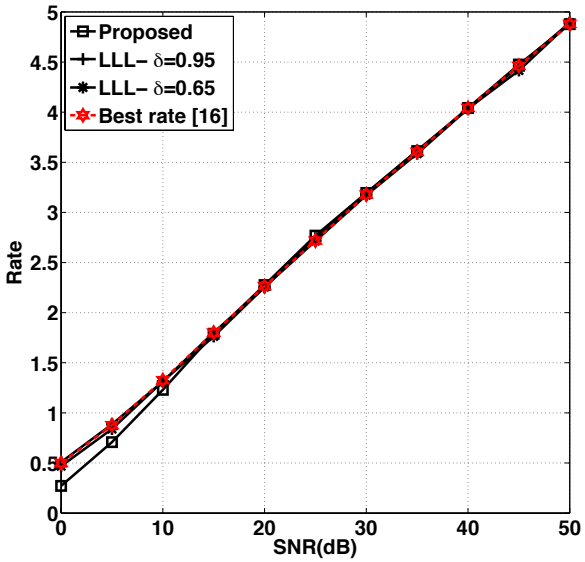


Figure 4.25: Computation rate assuming two transmitter ($n = 2$). The proposed algorithm compared with the LLL algorithm when $\delta = 0.65$ and 0.95 and the method proposed in [140].

can readily be extended to a complex scenario. A thorough discussion about a complex C&F using Gaussian integers can be found in [123].

Remark 3: In the algorithms proposed in Table 4.7 and 4.8, we propose to search for the minimum total noise over m integer values for a_k (and accordingly, $a_1, \cdots a_n$) assuming that m is restricted by (4.82). However, in the next subsection, it is demonstrated through numerical simulations that one can obtain an excellent estimate of the optimal integer vector $\boldsymbol{a}$ with relatively small m. The numerical simulations in the next subsection are obtained by assuming $m = 5$ at low/medium SNR and $m = 50$ at high SNR.

4.5.4 Numerical Results

In this subsection numerical results are provided in order to verify the performance of the system according to the proposed algorithm in Table 4.8.

We assume various numbers of transmitters and assume the real channels to be normally distributed with zero mean and unit variance, i.e. $h_i \in \mathcal{N}(0, 1)$. The performance of the proposed algorithm is compared with that of the well-known LLL algorithm for different complexity-performance trade-off (δ) values. Through simulations it will be noticed that the proposed algorithm performs as well as the LLL algorithm when complexity-performance trade-off factor is set to *almost* its maximum value ($\delta \to 1$) which results in

Table 4.8: Algorithm 2.

Algorithm 2. Finding integer vector $\boldsymbol{a}$
Initialise: Set m to be a positive integer
for $a_k = 1$ to m
Step 1) $a_i = \lfloor \frac{h_i}{h_k} a_k \rceil$ and $\lfloor \frac{h_i}{h_k} a_k \rfloor$
Step 2) Calculate total noise power P_{tot}
Step 3) Save P_{tot} and $\boldsymbol{a}$
end
Select $\boldsymbol{a}$ corresponding to minimum P_{tot}

significant complexity[6]. Apart from the LLL algorithm, the computation rates using the algorithm proposed in [140] is also provided which finds the best equation corresponding to highest rate, and hence is an invaluable benchmark for comparison. Moreover, it will be observed, through simulations, that the rate obtained exploiting the proposed algorithm is identical to that obtained using the method in [140], and hence validating the usefulness of the proposed algorithm.

Assuming two transmitters, Figure 4.25 illustrates the computation rate for the proposed algorithm in comparison with that of the LLL algorithm with complexity-performance trade-off coefficient values $\delta = 0.95$ and 0.65. Note that for a system with only two transmitters the computation rate is almost equal at high SNR regardless of the algorithm applied. Moreover, although the proposed algorithm is designed for high SNR region, its performance loss at low SNR is negligible. Figure 4.26 depicts the computation rate of another system with five transmitting nodes; clearly the proposed algorithm shows the best computation rate where the performance of the LLL algorithm is evaluated with different complexity-performance trade-off values. Note that with $\delta = 0.95$, although the LLL algorithm performs nearly as well as the proposed algorithm, the complexity is higher because $\delta \approx 1$ and so the complexity of the LLL algorithm is nearly equal to its maximum complexity. It is worth to mention that as the number of users increases, the complexity increases exponentially, which is difficult to manage in a practical scenario. Indeed the complexity of the system becomes crucial when the number of users is large, otherwise with only a few users, one even can find the optimal $\boldsymbol{a}$ even using exhaustive search with manageable complexity.

Figures 4.27 and 4.28 illustrate the performance of two other systems with ten and fifteen users, respectively. With a large number of users, the superiority of the proposed algorithm becomes evident, e.g. for the system with fifteen users at high SNR, the proposed algorithm outperforms the LLL algorithms even when it is set to $\delta = 0.95$).

[6]Note that LLL algorithm is optimal when $\delta \to 1$, and so achieving/approaching the performance of LLL algorithm using our proposed algorithm validates its precision.

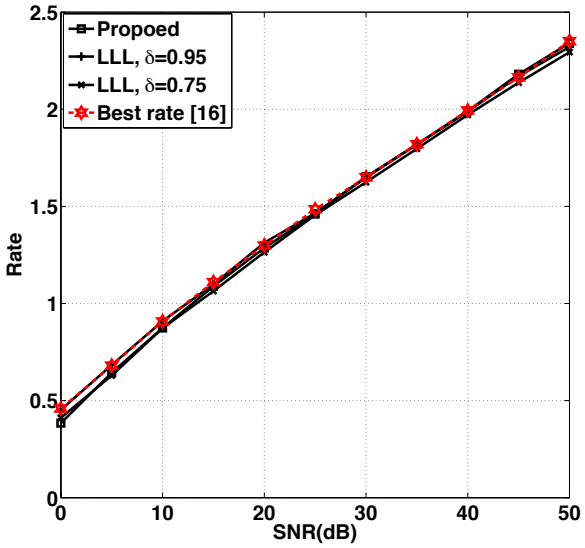


Figure 4.26: Computation rate assuming five transmitter ($n = 5$). The proposed algorithm compared with the LLL algorithm when $\delta = 0.65, \cdots, 0.95$ and the method proposed in [140].

In Figure 4.29 and 4.30, computer simulations for evaluating the complexity-performance trade-off are provided. The figures illustrate computation rate vs. m for various values of SNR. Note that by increasing m, as expected, the computation rate as well as the complexity of the system increase. Clearly the complexity increases by m since the number of iterations increases; however, as observed in Figure 4.29 and 4.30, it is worth mentioning that the computation rate achieves its largest values relatively quickly with small m, ensuring that the complexity of the system remains low.

4.5.5 Complexity analysis

In the following the complexity of the proposed algorithm described in Table 4.8 is studied and its computation complexity in terms of floating-point operations is evaluated[7]. We calculate only the number of operations in "Step 1" and "Step 2"; the total number of operations is m times the sum of operations in "Step 1" and "Step 2":

- *Step 1:* There are two operations to be performed in this step: multiplication and rounding/flooring. Since $i = 1, \cdots, n$, there are n multiplying operations due to $\frac{h_i}{h_k} a_k$ and n rounding/flooring operations, and hence overall $2n$ operations are performed in this step. Note that the values of $\frac{h_i}{h_k}$ are calculated once and so the complexity introduced by them are negligible compared to the rest of the algorithm.

[7]The complexity due to saving a variable is neglected.

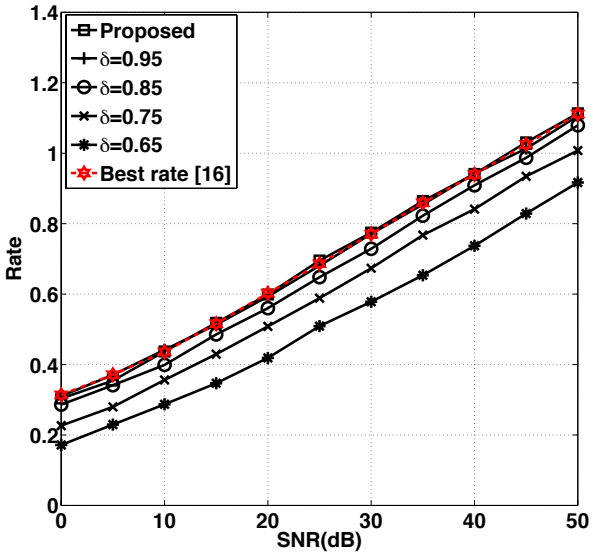


Figure 4.27: Computation rate assuming ten transmitter ($n = 10$). The proposed algorithm compared with the LLL algorithm when $\delta = 0.65, \cdots, 0.95$ and the method proposed in [140].

- *Step 2:* In this step the total noise is calculated based on the vectors $\boldsymbol{a} = \lfloor \bar{\boldsymbol{a}} \rceil$ and $\lfloor \bar{\boldsymbol{a}} \rfloor$ derived in the previous step. Using (4.51) and (4.60), P_{tot} can be implemented in hardware by performing $8n$ operations. For $\boldsymbol{a} = \lfloor \bar{\boldsymbol{a}} \rfloor$, implementing (4.60) requires n squaring (a_i^2), n multiplication ($a_i u_i$) and $2n$ summations which results in $4n$ operations in total; the same number of operation are required for $\boldsymbol{a} = \lfloor \bar{\boldsymbol{a}} \rceil$ and so almost $8n$ operations are required to calculate P_{tot}.

Consequently, for performing the two steps in the proposed algorithm $10n$ operations are required to be performed and so it does not introduce much of complexity to the system because the order of the complexity is n. Note that the complexity of the algorithm proposed in [140] is of the order $\mathcal{O}(n^2\sqrt{1 + P\|\|\boldsymbol{h}\|\|^2})$, and hence the proposed algorithm in this chapter is significantly less complex than the algorithms in [140].

The complexity analysis of LLL was studied in [137] and it is well known that for integer bases, the LLL algorithm has complexity bound $\mathcal{O}(n^4)$. For real-valued bases where the entries follow uniform distribution, the complexity of the LLL algorithm is studied in[141] and it is found that the number of iterations complexity is of the order $\mathcal{O}(n^2 \log n)$. Assuming that the entries of the basis vectors follow the i.i.d. standard normal distribution, the complexity of the LLL algorithm is bounded by $\mathcal{O}(n^3 \log n)$ [142]. Note that the LLL algorithm in this chapter is applied to the matrix $\boldsymbol{M} = \boldsymbol{I} - \frac{\boldsymbol{h}^T \boldsymbol{h}}{\|\boldsymbol{h}\|^2}$ where $\boldsymbol{h}$ is normally distributed; nevertheless, to the best of the authors, knowledge, the distribution of the matrix $\boldsymbol{M}$ as well as the complexity bounds of the LLL algorithm for this problem are unknown. Therefore, we are not aware of any complexity analysis of

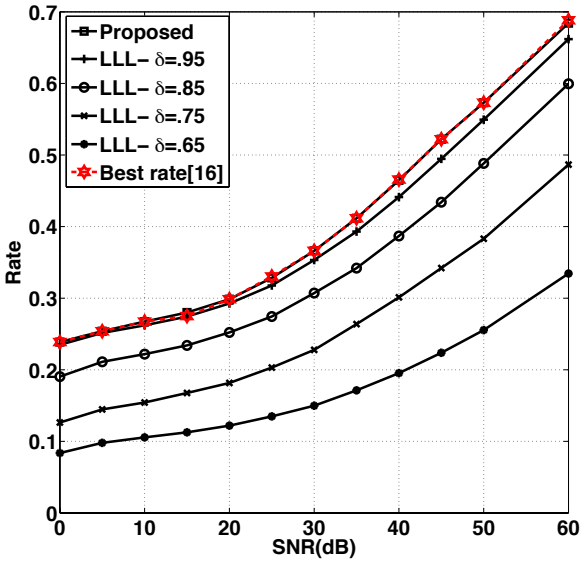


Figure 4.28: Computation rate assuming fifteen transmitter ($n = 15$). The proposed algorithm compared with the LLL algorithm when $\delta = 0.65, \cdots, 0.95$ and the method proposed in [140].

the LLL algorithm for the problem addressed here and so a fair comparison cannot be provided. Nevertheless, clearly the proposed algorithm has lower complexity, especially for small values of m, compared to the complexity of the LLL algorithm addressed in [137, 141, 142].

4.5.6 Conclusion

In this chapter, a novel approach was proposed to compute the integer vector $\boldsymbol{a}$ for compute-and-forward (C&F) relaying. We first decompose the total noise power into two functions, one of which is symmetric around zero. Then, we focus on finding the global minimum of the second function and prove that it is minimised on a line. Exploiting this finding, we proposed a new algorithm to determine the integer vector in the context of C&F relaying protocol. As the proposed algorithm requires eigenvalue decomposition, we have proposed an alternative algorithm which does not require this decomposition and operates only based on channel parameters. The performance and complexity of the proposed algorithm was compared with existing algorithms and its superiority was demonstrated through numerical simulations. This result significantly improves the practical applicability of C&F, and will help it realise its potential in improving the efficiency of wireless multihop networks for many applications.

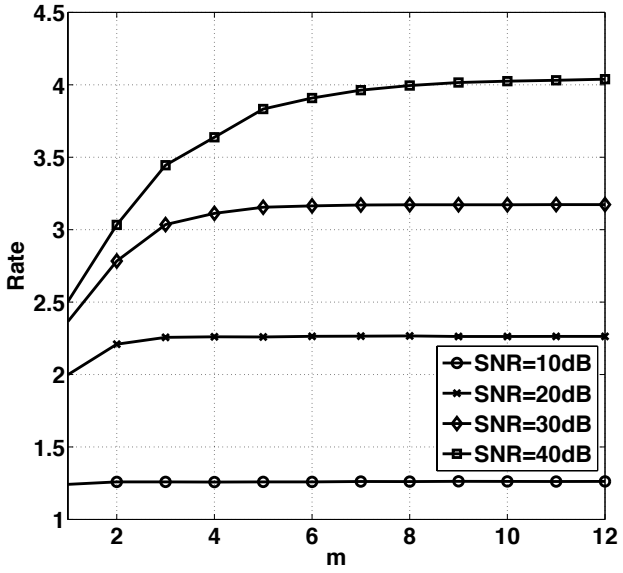


Figure 4.29: Computation rate vs. iteration m (complexity), assuming $n = 2$.

4.5.7 Proof of (4.54) and (4.55)

Rewriting $\boldsymbol{M}$ form (4.52), we have

$$\boldsymbol{M} = \boldsymbol{I} - \frac{\boldsymbol{h}^T\boldsymbol{h}}{\| \boldsymbol{h} \|^2} \tag{4.88}$$

where the eigenvalue decomposition of $\frac{\boldsymbol{h}^T\boldsymbol{h}}{\|\boldsymbol{h}\|^2}$ can be written as

$$\frac{\boldsymbol{h}^T\boldsymbol{h}}{\| \boldsymbol{h} \|^2} = \boldsymbol{Q}\boldsymbol{L}\boldsymbol{Q}^T \tag{4.89}$$

with $\boldsymbol{L}$ a diagonal matrix where only the last entry is one and the rest of the values are zero, i.e.

$$\boldsymbol{L} = \begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & 1 \end{bmatrix}. \tag{4.90}$$

$\boldsymbol{Q}$ is a unitary matrix corresponding to eigenvalues in $\boldsymbol{L}$. Considering that $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}$, one can write (4.88) as follows

$$\begin{aligned} \boldsymbol{M} &= \boldsymbol{I} - \frac{\boldsymbol{h}^T\boldsymbol{h}}{\| \boldsymbol{h} \|^2} \\ &= \boldsymbol{Q}\boldsymbol{I}\boldsymbol{Q}^T - \boldsymbol{Q}\boldsymbol{L}\boldsymbol{Q}^T \\ &= \boldsymbol{Q}(\underbrace{\boldsymbol{I} - \boldsymbol{L}}_{\boldsymbol{\Lambda}})\boldsymbol{Q}^T \end{aligned} \tag{4.91}$$

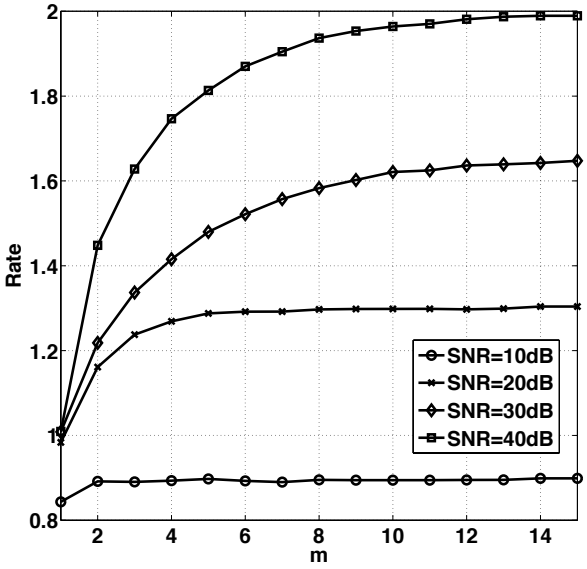


Figure 4.30: Computation rate vs. iteration m (complexity), assuming $n = 5$.

and so $\mathbf{\Lambda}$ in (4.55) is proved. Moreover, it is easy to see that $\boldsymbol{Q} = \boldsymbol{V}$ where $\boldsymbol{V}$ was defined as the eigenvector matrix of $\boldsymbol{M}$ in (4.54).

4.5.8 Proof of (4.60)

In this subsection, we simplify the objective function of $\mathcal{S} = \boldsymbol{a}\boldsymbol{M}\boldsymbol{a}^T$ in (4.52) according to (4.61). In order to simplify the proof and avoid confusion due to the parameters, we assume a system with three users where $n = 3$; note that the proof can easily be extended to arbitrary n using a similar method.

From (4.57), we have

$$\begin{aligned} \boldsymbol{a}\boldsymbol{W} &= [a_1 a_2 a_3] \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \\ v_{31} & v_{32} \end{bmatrix} \\ &= [a_1 v_{11} + a_2 v_{21} + a_3 v_{31} \ , \ a_1 v_{12} + a_2 v_{22} + a_3 v_{32}] \, , \end{aligned} \tag{4.92}$$

and so $\mathcal{S} = \boldsymbol{a}\boldsymbol{W}\boldsymbol{W}^T\boldsymbol{a}^T$ in (4.57) is

$$\begin{aligned} \mathcal{S} &= \\ & (a_1 v_{11} + a_2 v_{21} + a_3 v_{31})^2 + (a_1 v_{12} + a_2 v_{22} + a_3 v_{32})^2 \, , \end{aligned} \tag{4.93}$$

which by means of some algebraic manipulation can be written as

$$\mathcal{S} = (v_{11}^2 + v_{12}^2)a_1^2 + (v_{21}^2 + v_{22}^2)a_2^2 + (v_{31}^2 + v_{32}^2)a_3^2 \tag{4.94a}$$
$$+2(v_{11}v_{21} + v_{12}v_{22})a_1a_2 \tag{4.94b}$$
$$+2(v_{11}v_{31} + v_{12}v_{32})a_1a_3 \tag{4.94c}$$
$$+2(v_{21}v_{31} + v_{22}v_{32})a_2a_3. \tag{4.94d}$$

Considering that $\boldsymbol{V}$ is a unitary matrix ($\boldsymbol{V}\boldsymbol{V}^T = \boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}$), we have

$$\| \boldsymbol{v}_i \|^2 = 1 \tag{4.95a}$$
$$\boldsymbol{v}_i\boldsymbol{v}_j^T = 0 \qquad \text{for } i \neq j, \tag{4.95b}$$

and hence in (4.94a) we can write

$$\begin{aligned} v_{11}^2 + v_{12}^2 &= 1 - v_{13}^2 \\ v_{21}^2 + v_{22}^2 &= 1 - v_{23}^2 \\ v_{31}^2 + v_{32}^2 &= 1 - v_{33}^2 \end{aligned} \tag{4.96}$$

and for (4.94b),(4.94c) and (4.94d), one can write

$$\begin{aligned} v_{11}v_{21} + v_{12}v_{22} &= -v_{13}v_{23} \\ v_{11}v_{31} + v_{12}v_{32} &= -v_{13}v_{33} \\ v_{21}v_{31} + v_{22}v_{32} &= -v_{23}v_{33} \end{aligned} \tag{4.97}$$

and so by substituting (4.97) and (4.96) in (4.94) we have

$$\begin{aligned} \mathcal{S} \;=\; & (1 - v_{13}^2)a_1^2 + (1 - v_{13}^2)a_2^2 + (1 - v_{33}^2)a_3^2 \\ & -2v_{13}v_{23}a_1a_2 - 2v_{13}v_{33}a_1a_3 - 2v_{23}v_{33}a_2a_3. \end{aligned} \tag{4.98}$$

where by applying the simplified notation of (4.59) the proof is completed.

4.6 Design for adaptive physical layer network coding over cooperative relaying

In the next generation cellular networks, the architecture of network MIMO plays a key role on capacity improvement. A well-known example is the coordinated multipoint (CoMP) [143] technique included in LTE-A, where the access points (AP) of a cluster perform cooperative detection of multiple mobile terminals (MT). All MTs may share the same radio resources, and then be served by the corresponding APs where the multiuser interference can be effectively mitigated. The drawback of CoMP that has been left unsolved is that it increases the backhaul loads on the uplink, especially for wireless backhaul link (the link between AP and hub base station (HBS)). The number of bits that

the AP transmits over backhaul link depend on the quantisation scheme, co-channel interference and modulations, which normally result in increased backhaul loads for acceptable performance.

Physical layer network coding (PNC) [144][145] is a scheme in which the AP attempts to infer and forward combinations of multiple sources over an algebraic field, given the simultaneously arrived noisy, faded and superimposed signals. The initial work for PNC was implemented in a two-way relay channel (TWRC) which easily double the network throughput without routine operation. An important property of PNC is that the relay decodes the multiple sources information to a linear function over the algebraic field. This greatly reduces the data flows from the relay.

Although, in theory, PNC is well suitable for the cooperative relaying, especially it will effectively manage the backhaul load, there are a few technical limitations that remains unsolved.

1. Coefficients of each relay should be selected such that all source symbols can be derived at the destination.

2. Engineering applicable – thus, PNC must operate over the binary system so that the forward error correction codes (FEC) and conventional modulation methods can be readily used.

Compute-and-forward (C&F) [123] generalises PNC of TWRC to multiuser relay networks by utilising structured nested lattice codes. This should have been considered as a good candidate. However, it is based on construction A or D which operates over a finite field and the coset size of the quotient lattices is typically not binary-based.

Here we propose an adaptive PNC (APNC) technique, and the main contributions are:

1. We propose an APNC scheme.

2. We propose an offline algorithm which finds the coefficient matrices for each relay, such that 1). the composite full matrix guarantees all source symbols to be decoded at destination; 2) the matrices at each relay can resolve all singular fade states; 3) the number of coefficient matrices at each relay is minimised. 4) randomly selected coefficient matrix from each relay forms a full rank matrix at high probability.

3. The whole scheme operates over binary system.

4. We compare the performance of the proposed algorithm with the optimum global online search.

4.6.1 Design criteria

System model

The system model is illustrated in Figure 4.31. The first stage link between APs and MTs is referred to as the access link, where u mobile terminals transmit symbols to n access points. Each AP receives data from all MTs and then infers and forwards a linear combination (which is the network coded symbols (NCS) in this section) of the entire messages over a finite field or ring. We assume MTs and APs are all equipped with a single antenna. The second stage link between APs and HBS is referred to as the backhaul link, where n APs send the NCSs to HBS via a lossless but bandwidth-limited bit-pipe. Access link is modelled as wireless link and this is indeed the case in 5G. The backhaul link may be deployed on wireless or wireline. The techniques presented here is in particular suitable for wireless backhaul which is normally more cost-effective.

Each MT employs a 2^m-ary digital modulation scheme where m is the modulation order. Let $\mathscr{M} : \mathbb{F}_2^m \longrightarrow \Omega$ denote a one-to-one mapping function, where Ω is the set of all possible complex constellation points. Hence, the messages $\mathbf{w}_\ell \in \mathbb{F}_2^m$, $\ell = 1, 2, \cdots, u$, of the ℓ^{th} MT can be mapped to the complex symbol $s_\ell = \mathscr{M}(\mathbf{w}_\ell)$, where $\mathbf{w}_\ell = [w_\ell^{(1)}, \cdots, w_\ell^{(m)}]$ is an m-tuple with each element $w_\ell^{(i)} \in \mathbb{F}_2$.

The link between all MTs and the j^{th} ($j = 1, 2, \cdots, n$) AP forms a multiple access channel (MAC), where the j^{th} AP observes the noisy, faded and superimposed signals at the a certain time slot is:

$$y_j = \sum_{\ell=1}^{u} h_{j,\ell} s_\ell + z_j \tag{4.99}$$

where z_j is the outcome of a complex Gaussian random variable Z_j with zero mean and variance σ^2 per dimension, and $h_{j,\ell}$ represents the channel fading coefficient between the wireless link of the ℓ^{th} MT and the j^{th} AP, which is the outcome of a random variable with Rayleigh distribution.

Design of PNC

PNC is a technique that compresses all source messages at relay using some algebraic approaches based on the simultaneously received signals, which makes network throughput greatly improved and the cardinality of relay outputs considerably eliminated. The original PNC is proposed and designed in a TWRC based on BPSK. Although only BPSK is used, the idea of PNC motivates many research outcomes thereafter, e.g. C&F and lattice network coding [132]. However, for practical applications, lattice-based PNC is not a good candidate as it does not operates over the binary systems, and zero-point is always here which reduces the power efficient.

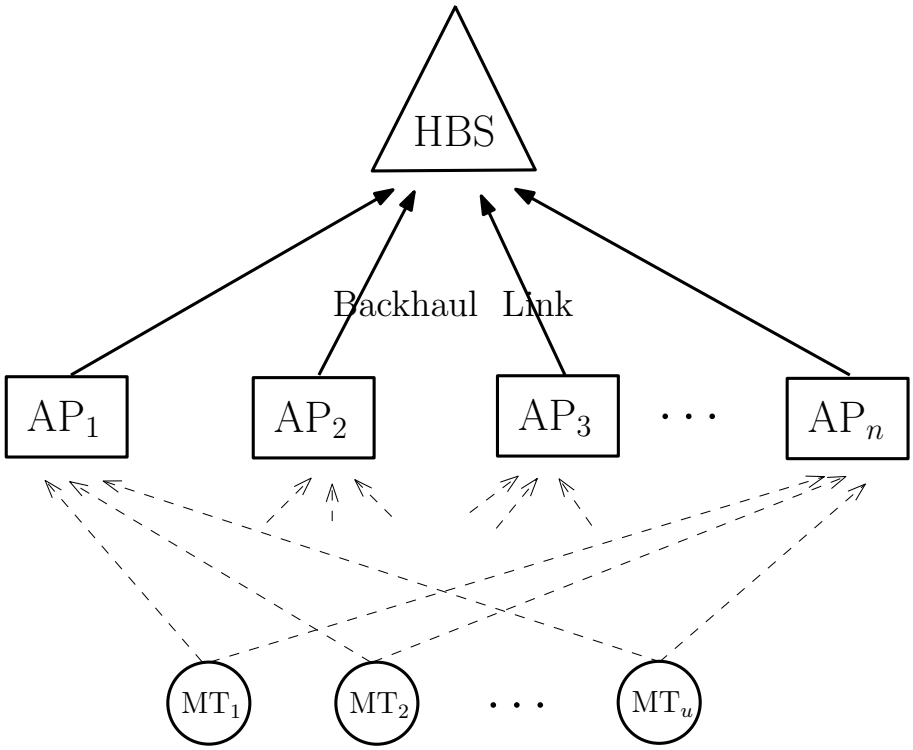


Figure 4.31: The uplink system diagram.

Our attention focuses back on the conventional 2^m-ary digital modulation. When QPSK or higher order modulations are used, PNC has to solve the so-called singular fading problem which is typically unavoidable at the MAC phase. Failure to resolve the singular fade states will result in severe performance degradation. Toshiaki *et. al.* [119] proposes a scheme, namely the denoise-and-forward, which employs a non-linear 5QAM PNC mapping to mitigate all singular fade states, and gives good performance. Other researches on this issue including the design of linear functions over the integer finite field or ring, e.g. linear PNC (LPNC) which can only be optimised for the q-ary PNC mapping where q is a prime in $\mathbb{Z}^+$. All these approaches, however, do not operate over the binary systems, and hence cannot be readily applied in the current mobile communications networks. The three main restrictions for PNC to be engineering applicable:

1. PNC decoding must operate over $\mathbb{F}_2$; thus, the NCS needs to be binary-based.

2. The PNC function must be well designed such that all singular fade states can be resolved.

3. Ensure that HBS can unambiguously decoding all source messages based on the forwarded NCSs from multiple APs.

4. The computation load in finding good PNC mapping is normally unaffordable.

The design method proposed for APNC relaxes all three aforementioned constraints, as explained below in detail.

Engineering applicable PNC function We are primarily concerned with the MAC phase between the u MTs and the j^{th} AP in the design of the PNC function for AP_j. Instead of using traditional approach in PNC where the linear combinations are performed on symbol level. We consider to design the PNC function directly in the message domain; thus, the AP will decode the linear network coded function (LNCF) over $\mathbb{F}_2$ among the bit level.

Definition 1: The bit-level LNCF of the j^{th} AP for u MTs is defined as:

$$\mathscr{N}_j : (\mathbf{M}_j, \mathbf{w}) \longrightarrow \mathbf{x}_j \tag{4.100}$$

and mathematically can be expressed as:

$$\mathbf{x}_j = \mathscr{N}_j(\mathbf{M}_j, \mathbf{w}) = \mathbf{M}_j \star \mathbf{w} \tag{4.101}$$

where $\mathbf{w} \triangleq [\mathbf{w}_1, \cdots, \mathbf{w}_u]^T$ denotes the joint message vector with $\mathbf{w} \in \mathbb{F}_2^{mu}$, $\mathbf{M}_j$ denotes the element of a vector space over $\mathbb{F}_2^{t^{(j)} \times mu}$, $t^{(j)} \geq m$, and $\star$ denotes the matrix multiplication over $\mathbb{F}_2$. $\mathbf{x}_j \in \mathbb{F}_2^{t^{(j)}}$ consists of all $t^{(j)}$ linear network coded bits (LNCB):

$$\mathbf{x}_j = [x_j^{(1)}, x_j^{(2)}, \cdots, x_j^{(t)}]^T \tag{4.102}$$

and is called the network coded vector (NCV). □

It is obvious that each LNCB is indeed a linear combination of all source bits over $\mathbb{F}_2$, thus,

$$x_j^i = \mathbf{M}_j^{(i,1)} \boxdot w_1^{(1)} \boxplus \cdots \boxplus \mathbf{M}_j^{(i,um)} \boxdot w_u^{(m)} \tag{4.103}$$

where $\boxplus$ and $\boxdot$ denote the addition and multiplication operations over $\mathbb{F}_2$, and $\mathbf{M}_j^{(i,1)}$ denote the entry at the i^{th} row and the 1^{st} column of $\mathbf{M}_j$.

Definition 2: Let $\mathbf{s} \triangleq [s_1, \cdots, s_u]$, we define

$$\mathbf{s}_{j,\triangle} \triangleq \sum_{\ell=1}^{u} h_{j,\ell} s_\ell, \quad \forall \mathbf{s} \in \Omega^u \tag{4.104}$$

the superimposed constellation set of the j^{th} AP at a given channel coefficient vector $\mathbf{h}_j \triangleq [h_{j,1}, \cdots, h_{j,u}]$. $s_{j,\triangle}^{(\tau)}$ denotes a particular constellation point in $\mathbf{s}_{j,\triangle}$, where $\tau = 1, 2, \cdots, 2^{mu}$.

Theorem 6. *For the MAC link between u MTs and the j^{th} AP, there exists a surjective function:*

$$\Theta : \mathbf{s}_{j,\triangle} \longrightarrow \mathbf{x}_j \tag{4.105}$$

when $t^{(j)} < mu$.

Proof: Since $\mathscr{M}$ is a bijective function, we have the following relationship:

$$\mathbf{x}_j \overset{\mathscr{N}_j}{\Leftarrow} \mathbf{w} \underset{\mathscr{M}^{-1}}{\overset{\mathscr{M}}{\Longleftrightarrow}} \mathbf{s} \tag{4.106}$$

where $\Leftarrow$ and $\Longleftrightarrow$ represent surjective and bijective relationships, respectively. Following (4.104), for each element of $\mathbf{s}$, there exists a superimposed constellation point $s_{j,\triangle}$ at a given channel coefficient vector $\mathbf{h}_j$, and this proves the theorem. ■

We call Θ the PNC function which maps the superimposed constellation point to the NCV, and plays the key role in PNC decoding. The PNC decoding performs estimation of the possible NCV outcomes $\mathbf{x}_j$ for the j^{th} relay, based on the received signals y. Let $\mathbf{X}_j$ denote the vector-based random variable with its realisation $\mathbf{x}_j$. The *a posteriori* probability of the event $\mathbf{X}_j = \mathbf{x}_j$ conditioned on the MAC outputs $Y_j = y_j$ is:

$$\begin{aligned}
&\Pr(\mathbf{X}_j = \mathbf{x}_j | y_j, \mathbf{h}_j) \\
=&\frac{\Pr(Y_j | \mathbf{X}_j = \mathbf{x}_j, \mathbf{h}_j)\Pr(\mathbf{X}_j = \mathbf{x}_j)}{\Pr(Y_j = y_j)} \\
=&\frac{\sum\limits_{\forall \mathbf{w}: \mathscr{N}_j(\mathbf{w}, \mathbf{M}_j) = \mathbf{x}_j} \Pr(Y_j | \mathbf{w}, \mathbf{h}_j)\Pr(\mathbf{w})}{\Pr(Y_j = y_j)} \\
=&\frac{\sum\limits_{\forall \mathbf{s}: \Theta(\mathbf{s}_{j,\triangle}) = \mathbf{x}_j} \Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle})\Pr(\mathbf{S} = \mathbf{s})}{\Pr(Y_j = y_j)}
\end{aligned} \tag{4.107}$$

The conditional probability density function $\Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle})$ is given by:

$$\Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|y_j - \mathbf{s}_{j,\triangle}|^2}{2\sigma^2}\right) \tag{4.108}$$

The *a posteriori* L-value $L_{\mathbf{x}_j}$ for the event $\mathbf{X}_j = \mathbf{x}_j$ is:

$$L_{\mathbf{x}_j} = \log\left(\frac{\sum\limits_{\forall \mathbf{s}: \Theta(\mathbf{s}_{j,\triangle}) = \mathbf{x}_j} \Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle})\Pr(\mathbf{S} = \mathbf{s})}{\sum\limits_{\forall \mathbf{s}: \Theta(\mathbf{s}_{j,\triangle}) = \mathbf{0}} \Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle})\Pr(\mathbf{S} = \mathbf{s})}\right) \tag{4.109}$$

where $\mathbf{0}$ is a length-$t^{(j)}$ all-zero vector over $\mathbb{F}_2^{t^{(j)}}$.

Resolving the singular fading We have set up the PNC mapping approach based on the binary systems, which establishes the fundamental PNC system structure available for practical engineering application, e.g. 5G. The next upcoming problem lies in how to solve the singular fading of the multiple access channels. In this subsection we demonstrate that the PNC function Θ_j proposed above is capable of resolving all singular fade states with a simple design approach. We first define the singular fade states as follows:

Definition 3: The singular fade state at the j^{th} AP is defined as the channel fading coefficients $\mathbf{h}_j$ which makes $s_{j,\triangle}^{(\tau)} = s_{j,\triangle}^{(\tau')}$ when $\tau \neq \tau'$.

In other words, at a given channel coefficients $\mathbf{h}_j$, if there are two or more elements in the set $\mathbf{s}_{j,\triangle}$ are the same, $\mathbf{h}_j$ is one singular fade state (SFS). Normally, SFS is unavoidable at MAC, and multiuser detection is in principle infeasible if the j^{th} AP expects to decode all source messages. PNC provides a good solution to overcome SFS when the coincident superimposed constellation points are well labelled by the same NCV $\mathbf{x}_j$, which helps HBS to recover all source messages finally.

Definition 4: The minimum distance between the NCVs are defined as:

$$d_{\min} = \min_{\forall \mathbf{s}_{j,\triangle}: \Theta(s_{j,\triangle}^{(\tau)}) \neq \Theta(s_{j,\triangle}^{(\tau')})} |s_{j,\triangle}^{(\tau)} - s_{j,\triangle}^{(\tau')}|^2 \tag{4.110}$$

Thus, if all superimposed constellation points that are labelled by the same NCV are placed in a certain cluster, the $d_{\min}$ is the minimum distance of the inter-clusters.

Definition 5: The singular PNC function of the j^{th} AP is defined as the surjective function Θ_j which results in:

$$d_{\min} = 0 \tag{4.111}$$

We call Θ_j the singular PNC mapping when the superimposed constellation points that are located at the same complex plane are labelled by more than one NCVs.

Theorem 7. *The singular PNC function cannot resolve singular fading if* $d_{\min} = 0$.

Proof: When $d_{\min} = 0$, the posterior probability of some outcomes of $\mathbf{X}_j$ will be very similar (in terms of (4.107)). This definitely introduces the ambiguities in estimating the real NCV, $\mathbf{x}_j$, especially when the received complex signal is close to a superimposed constellation point that is labelled by more than one NCVs. Hence, the singular PNC function is in principle not capable of decoding the NCV reliably. ■

Normally m is the minimum dimension of NCV $\mathbf{x}_j$ at the j^{th} relay. When the number of source increases (a large MAC), the singular fading problem becomes more severe. However, by simply increasing the dimension t of NCV (thus, increasing the number of rows of $\mathbf{M}_j$), there definitely exists non-singular PNC function which is capable of resolving a kind of SFS.

Remark 10. *We can obtain non-singular PNC function* Θ_j *for the* j^{th} *AP if the cardinality* $t^{(j)}$ *of the PNC decoding outcomes is determined in terms of the following criterion:*

$$t^{(j)} = \arg\min_{m \leq t^{(j)} < mu} \{d_{\min} - d_\alpha \geq 0\} \tag{4.112}$$

where $d_\alpha > 0$ *is a distance threshold.*

Remark 10 reveals the second design criterion for PNC function Θ_j over a u-MT and 2^m-ary digital modulation MAC, which guarantees the reliable PNC decoding with the minimum possible cardinality expansion.

The theoretic minimum cardinality for NCV, $\mathbf{x}_j$, is m. However, when the number of MTs increases, the singular fading becomes more serious, and hence $t^{(j)}$ must increase to ensure non-singular PNC functions can be found in terms of a kind of SFS.

Although this means that the cardinality of the NCV increases, we will explain later that the increase of this cardinality has no effect on the overall backhaul load.

Algebraic work for the unambiguous decodability We have set up two design guidelines of the engineering-applicable PNC approach for the uplink system. The next criterion is that the destination can guarantee all source messages to be unambiguously decoded.

Theorem 8. *Assume $\mathbf{M} = M_{n\times n}(R)$, where the coefficients are from a commutative ring R. Source messages are drawn from a subset of R. All source messages can be unambiguously decoded at the destination iff the determinant of the transfer matrix is a unit in R,*

$$\det(\mathbf{M}) = \mathcal{U}(R) \tag{4.113}$$

Proof: We first prove that (4.113) gives the sufficient and necessary conditions that make $\mathbf{B}$ invertible. Suppose $\mathbf{B}$ is invertible: then, there exists a matrix $\mathbf{C} \in M_{n\times n}(R)$ such that $\mathbf{BC} = \mathbf{CB} = \mathbf{I}_n$. This implies $1 = \det(\mathbf{I}_n) = \det(\mathbf{BC}) = \det(\mathbf{B})\det(\mathbf{C})$; According to the definition of a unit, we say $\det(\mathbf{B}) \in U(R)$.

We know $\mathbf{B}\cdot\mathrm{adj}(\mathbf{B}) = \mathrm{adj}(\mathbf{B})\cdot\mathbf{B} = \det(\mathbf{B})\mathbf{I}_n$. If $\det(\mathbf{B}) \in U(R)$, we have

$$\mathbf{B}\cdot(\det(\mathbf{B})^{-1}\mathrm{adj}(\mathbf{B})) = (\det(\mathbf{B})^{-1}\mathrm{adj}(\mathbf{B}))\mathbf{B} \tag{4.114}$$

$$= \det(\mathbf{B})^{-1}\det(\mathbf{B}) = \mathrm{I}_n \tag{4.115}$$

Hence, $\mathbf{C} = (\det(\mathbf{B})^{-1}\mathrm{adj}(\mathbf{B}))$ is the inverse of $\mathbf{B}$ since $\mathbf{BC} = \mathbf{CB} = \mathbf{I}_n$.

If $\mathbf{B}$ is invertible, then its inverse $\mathbf{B}^{-1}$ is uniquely determined. Assuming $\mathbf{B}$ has two inverses, say, $\mathbf{C}$ and $\mathbf{C}'$. Then

$$\mathbf{B}\cdot\mathbf{C} = \mathbf{C}\cdot\mathbf{B} = \mathbf{I}_n \tag{4.116}$$

$$\mathbf{B}\cdot\mathbf{C}' = \mathbf{C}'\cdot\mathbf{B} = \mathbf{I}_n \tag{4.117}$$

and hence we have

$$\mathbf{C} = \mathbf{C}\cdot\mathbf{I}_n = \mathbf{C}\cdot\mathbf{B}\cdot\mathbf{C}' = \mathbf{I}_n\cdot\mathbf{C}' = \mathbf{C}' \tag{4.118}$$

It proves the uniqueness of the invertible matrix $\mathbf{B}$ over R.

Assume $\mathbf{s} \neq \mathbf{s}'$, $\mathbf{B}\cdot\mathbf{s} = \mathbf{F}$, $\mathbf{B}\cdot\mathbf{s}' = \mathbf{F}'$, and $\mathbf{F} = \mathbf{F}'$. This means

$$\mathbf{s} = \mathbf{B}^{-1}\cdot\mathbf{F} = \mathbf{B}^{-1}\cdot\mathbf{F}' = \mathbf{s}' \tag{4.119}$$

This contradicts $\mathbf{s} \neq \mathbf{s}'$. Hence, it ensures unambiguously decodability:

$$\mathbf{B}\cdot\mathbf{s} \neq \mathbf{B}\cdot\mathbf{s}',\ \forall \mathbf{s} \neq \mathbf{s}' \tag{4.120}$$

■

Definition 4: The ideal in R generated by all $\nu \times \nu$ minors of $M_{m\times n}(R)$ is denoted by $I_\nu(M_{m\times n}(R))$, where $\nu = 1, 2, \cdots, r = \min\{m, n\}$.

A $\nu\times\nu$ minor of $M_{m\times n}(R)$ is the determinant of a $\nu\times\nu$ matrix obtained by deleting $m-\nu$ rows and $n-\nu$ columns. Hence, there are $\binom{m}{\nu}\binom{n}{\nu}$ minors of size $\nu \times \nu$. $I_\nu(M_{m\times n}(R))$ is the ideal of R generated by all these minors.

Theorem 9. *The destination is able to unambiguously decode ι source messages if:*

1. $\iota \geq \max\{\nu \mid \mathrm{Ann}_R(I_\nu(\mathbf{A}^j)) = \langle 0\rangle\}, \forall j = 1, 2, \cdots, K.$
2. $\mathbf{A}^j = \arg\max_{\mathbf{A}^j}\left\{I\left(\overrightarrow{Y}; \overrightarrow{F}^j\right)\right\}$

where $\langle x\rangle$ denotes the ideal generated by x and K is the number of layers (here we consider $K = 1$).

Condition 1 can be proved as follows. According to Laplace's theorem, every $(\nu+1)\times(\nu+1)$ minor of $M_{m\times n}(R)$ must lie in $I_\nu(M_{m\times n}(R))$. This suggests an ascending chain of ideals in R:

$$\langle 0\rangle = I_{r+1}(\mathbf{A}^j) \subseteq I_r(\mathbf{A}^j) \subseteq \cdots \subseteq I_1(\mathbf{A}^j) \subseteq I_0(\mathbf{A}^j) = R \tag{4.121}$$

Computing the annihilator of each ideal in (4.121) produces another ascending chain of ideals,

$$\begin{aligned}\langle 0\rangle = \mathrm{Ann}_R(R) \subseteq \mathrm{Ann}_R(I_1(\mathbf{A}^j)) \subseteq \cdots \subseteq \mathrm{Ann}_R(I_r(\mathbf{A}^j))\\ \subseteq \mathrm{Ann}_R(\langle 0\rangle) = R\end{aligned} \tag{4.122}$$

It is obvious that:

$$\begin{aligned}&\mathrm{Ann}_R(I_t(\mathbf{A}^j)) \neq \langle 0\rangle\\ \Rightarrow&\mathrm{Ann}_R(I_k(\mathbf{A}^j)) \neq \langle 0\rangle,\ \ \forall t \leq k.\end{aligned} \tag{4.123}$$

The maximum value of ν which satisfies $\mathrm{Ann}_R(I_\nu(\mathbf{A}^j)) = \langle 0\rangle$ guarantees that $I_t(\mathbf{A}^j) \in R, \forall t < \nu$. Hence, we define the rank of $\mathbf{A}^j$ as $\mathrm{rk}(\mathbf{A}^j) = \max\{\nu \mid \mathrm{Ann}_R(I_\nu(\mathbf{A}^j)) = \langle 0\rangle\}$. Suppose $\mathbf{A}^t \in M_{m\times p}(R)$ and $\mathbf{A}^k \in M_{p\times n}(R)$, then $\mathrm{rk}(\mathbf{A}^t\mathbf{A}^k) \leq \min\{\mathrm{rk}(\mathbf{A}^t), \mathrm{rk}(\mathbf{A}^k)\}$, and we can easily prove that $0 \leq \mathrm{rk}(M_{m\times n}(R)) \leq \min\{m, n\}$. Thus, in order to guarantee there are at least ι unambiguous linear equations available at the destination, $\mathrm{rk}(\mathbf{A}^j)$ must be at least $\iota, \forall j = 1, 2, \cdots, K$.

The special case of condition 1 is that the entry of the coefficient matrix $\mathbf{A}^j \in M_{m\times n}(F)$ is from a finite field $F \in \mathbb{F}$. Then, condition 1 may be changed to be "the maximum number of linearly independent rows (or columns)"since $\mathrm{Ann}_R(I_\nu(\mathbf{A}^j)) = \langle 0\rangle$ if and only if $I_\nu(\mathbf{A}^j) \neq 0$. In other words, the largest ν such that the $\nu\times\nu$ minor of $\mathbf{A}^j$ is a non-zero divisor represents how many reliable linear combinations the j^{th} layer may produce. Hence, condition 1 is a strict definition which ensures unambiguous decodability of the ι sources. Condition 3 ensures that the selected coefficient matrix maximises the mutual information of the particular layer, giving finally the maximum overall throughput.

Adaptive mapping design and offline algorithm

Finding an appropriate coefficient matrix $\mathbf{M}_j$ for the j^{th} AP such that the aforementioned three points are resolved are extremely computation-consuming, especially when the number of MTs increases. We propose to pre-compute a group of L' candidate coefficient matrices for each AP and perform APNC decoding online. This will greatly reduce the computational load of each relay in the real time decoding when L' is small. With the pre-computed L' candidate coefficient matrix, each relay is capable of adapting multiple access channels, and especially is capable of resolving all possible singular fade states that will occur at that relay. It is also important that the obtained candidate at each relay will have high probability to form a composite matrix which satisfies Theorem 8 and Theorem 9.

The basic idea of the proposed algorithm is to remove all possible images corresponding to the same type of singular fade states and then find for each relay the minimum number of candidate matrices (here we denote this number by L') such that the composite full matrix satisfies the two theorems. Note that each relay will have different candidates, otherwise it will be difficult to form full matrix which satisfies the two theorems. The proposed offline algorithm is detailed in Algorithm 1.

Algorithm 1 Off-line algorithm for candidate matrices.

for $i = 1 : L$ **do** ▷ each singular fade state
  $h = \mathcal{S}(i)$ ▷ h is a $1 \times m$ vector.
  for $j = 1 : K$ **do** ▷ each binary matrix
    $[\xi, T_\xi] = N(\mathcal{M}(j))$
    $\xi_{\mathrm{f}} \leftarrow \mathscr{F}(T_\xi, h)$ ▷ $\mathscr{F}(\cdot)$ produces all faded NCSs.
    $d_{\min} \leftarrow \mathscr{D}(\xi_{\mathrm{f}})$ ▷ $\mathscr{D}(\cdot)$ calculates the minimum distance of all NCSs.
    $\mathcal{Q}_{\mathrm{d}} \leftarrow \mathcal{Q}_{\mathrm{d}} \cup d_{\min}$ ▷ store all $d_{\min}$ in $\mathcal{Q}_{\mathrm{d}}$.
  end for
  $[\boldsymbol{\beta}(i), \boldsymbol{\alpha}(i)] \leftarrow \mathscr{C}(\mathcal{Q}_{\mathrm{d}})$ ▷ $\mathscr{C}(\cdot)$ sorts $\mathcal{Q}_{\mathrm{d}}$ in descending order stored in $\boldsymbol{\beta}(i)$ and outputs the rearranged index vector $\boldsymbol{\alpha}(i)$.
end for
$\mathcal{S}' \leftarrow \mathscr{J}(\mathcal{S}, \boldsymbol{\alpha})$ ▷ delete all mirror singular fade states and $\mathcal{S}'$ has L' singular states, $L' < L$.
$\boldsymbol{\alpha} \leftarrow \boldsymbol{\alpha} \setminus \boldsymbol{\alpha}(\boldsymbol{\beta} = 0)$ ▷ delete the index element of $\boldsymbol{\beta} = 0$.
$\boldsymbol{\alpha}' \leftarrow \boldsymbol{\alpha}(i|\mathcal{S}')$ ▷ $\boldsymbol{\alpha}'$ corresponds to only $\mathcal{S}'$.
for $l_{L'} = 1 : L'$ **do**
  $\mathcal{S}^{\dagger}_{L'-1} \leftarrow \mathcal{S}' \setminus \mathcal{S}'(l_{L'})$
  $\theta_{L'-1} \leftarrow \mathscr{E}(\mathcal{S}^{\dagger}_{L'-1})$ ▷ Index set of $\mathcal{S}'$ excluding the l^{th} element.
  for $l_{L'-1} = \theta_{L'-1}$ **do**
    $\mathcal{S}^{\dagger}_{L'-2} \leftarrow \mathcal{S}^{\dagger}_{L'-1} \setminus \mathcal{S}^{\dagger}_{L'-1}(l_{L'-1})$
    $\theta_{L'-2} \leftarrow \mathscr{E}(\mathcal{S}^{\dagger}_{L'-2})$
    $\vdots$
    for $l_{L'-n+1} = \theta_{L'-n+1}$ **do**
      for $i_1 = 1 : K$ **do**
        $\vdots$
        for $i_n = 1 : K$ **do**

$$\mathbf{M} = \begin{bmatrix} \mathcal{M}[\boldsymbol{\alpha}(l_{L'}, i_1)] \\ \vdots \\ \mathcal{M}[\boldsymbol{\alpha}(l_{L'-n+1}, i_n)] \end{bmatrix}$$

          $\delta \leftarrow \det(\mathbf{M})_{|\mathbb{F}_2}$ ▷ determinant over $\mathbb{F}_2$.
          if $\delta = 1$ **then**
            $\mathscr{R} \leftarrow \mathscr{R} \cup (l_{L'} \cdots l_{L'-n+1}; i_1 \cdots i_n)$
            $\mathbf{G} \leftarrow \mathbf{G} \cup \mathbf{M}$ ▷ $\mathbf{M} \leftrightarrow \mathbf{G}_{A(k)}$ in $\mathbf{G}$ has unique address $A(k) = (l^{(k)}_{L'} \cdots l^{(k)}_{L'-n+1}; i^{(k)}_1 \cdots i^{(k)}_n)$, $k = 1, \cdots \frac{L'!}{(L'-n)!}$.
            return (21)
          end if
        end for
      end for
    end for
  end for
end for

$$[\mathbf{G}_{A(k_1)} \cdots \mathbf{G}_{A(k_n)}] \leftarrow \mathscr{X}(\mathbf{G})$$

▷ find n $\mathbf{M}$ from $\mathbf{G}$ satisfying bijection relations $(l^{(k_e)}_{L'} \cdots l^{(k_e)}_{L'-n+1}) \Leftrightarrow \mathcal{S}'$ for $k_e = k_1 \cdots k_n$.
for $i = 1 : n$ **do**
  $\mathcal{Q}_i \leftarrow [\mathbf{G}^i_{A(k_1)} \cdots \mathbf{G}^i_{A(k_n)}]$ ▷ $\mathbf{G}^i_{A(k_i)} = \mathcal{M}[\boldsymbol{\alpha}(l^{(k_i)}_{L'-i+1}, i^{(k_i)}_i)]$
end for
Output: n stacks $\mathcal{Q}_i$ with each including L' binary matrices.

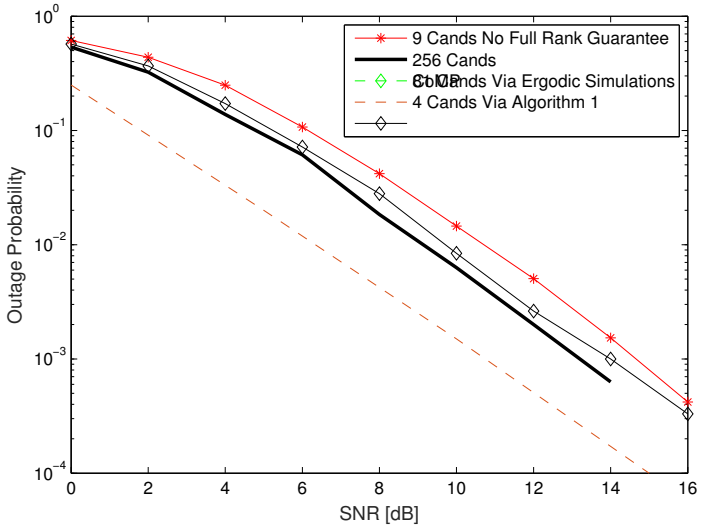


Figure 4.32: Comparison of the outage probability for the proposed algorithm.

4.6.2 Simulations

We consider here $u = 2$ and $n = 2$. The destination (HBS) aims to decode the two source data by solving some linearly independent equations which are transmitted from the two relays. Note that we assume the backhaul link is noise-free bit-pipe. Each mobile terminal employs QPSK digital modulation. As explained in Subsection 4.6.1, the j^{th} relay needs to choose appropriate binary matrix $\mathbf{M}_j \in \mathbb{F}_2^{2\times 4}$ such that the minimum distance $d_{\min}$ between the inter-clusters labelled by the NCV is larger than a threshold.

The Algorithm 1 carries out offline computation and finally assigns each relay four candidate matrices $\mathbf{M}_j$, $j = 1, 2$. This means that each relay is capable of resolving all possible singular fade states based on one of the assigned four candidate matrices. This hugely reduces the real-time computational complexity since normally we need to choose one out of 256 candidates online. The two groups of candidate matrices are in principle capable of constructing 16 different full matrices (4×4) and we expect that most of them are full rank because this will reduce the number of online search and provide more degrees of freedom of resolving singular fade states at each relay. The two groups form 10 full rank matrices. This is the largest number we can obtain when only 4 candidates are chosen for each relay.

Simulation results are depicted in Figure 4.32. The dashed line is the bound for ComP. The black curve gives the outage probability when each relay online searches all 256 candidate matrices. It should be viewed as a benchmark to compare with the proposed scheme as it is not a practically feasible method due to the unmanageable detection complexity. The outage probability based on the output of Algorithm 1 (thus, 4 candidates per relay) nearly approximates the black curve, with approximately 0.5 dB performance loss.

However, the online computational load is dramatically relaxed. A small performance loss is worthwhile in hugely reducing the detection complexity for practical systems. Note also that the two outage curves are based on the fact that the resultant composite full matrix is full rank, and hence the HBS is capable of unambiguously decoding all source bits. We also carry out an experiment where 9 half matrices are selected (which corresponds to 9 singular fade states), but without considering unambiguous decodability. Note that in this case, we only randomly select a half matrix which is capable of resolving the corresponding singular fade state. The outage performance is shown in Figure 4.32. It is observed that there is more than 1 dB loss at 10^{-3}. This implies that even if full rank condition is not considered, the candidate matrices must be carefully designed. This also suggests that Algorithm 1 performs well. Without loss of generality, Algorithm 1 can be applied in more general case.

4.6.3 Conclusions

We have proposed the design guideline for engineering-applicable physical layer network coding with multiple relay cooperation. The design criterion proposed guarantees that: (i) the whole system operates over binary system, (ii) each relay can choose non-singluar PNC function to overcome all the singular fade states, and (iii) the destination can unambiguously recover all source messages while the overall back-haul load remains at the lowest level regardless of how many sources or which digital mapping is used. We develop an offline search algorithm which finds a small number of best efficient coefficient matrices. This greatly reduces the on-line computation complexity with negligible performance loss.

5 Conclusions

This final deliverable focused on advanced scenarios of cloud network processing where algorithms have been designed and assessed by taking into account the limits, specifications and practical requirements as feedback from the experimental activity carried out in WP5.

The report presented a comprehensive overview of all the algorithms for distributed channel/system state identification, relay coding and cloud self-organisation that have been successfully implemented into the HW/SLS demonstrators. This activity was carried out in close interaction with WP5. Furthermore, gaining from the outcomes of the close cooperation with the demostrating activities in WP5, new advanced cloud processing methods have been taylored for a feasibility in perspective, paving the road for cloud processing evolution in future scenarios.

In perspective, the distributed synchronisation has been calibrated and adapted to track a common TO and CFO in dense network. DIWINE has shown for the first time the practical feasibility of distributed synchronisation in self-coordinated wireless networks, far beyond the early biology-inspired algorithms. On the same line, consensus based methods for distributed inference (validated in D3.03 for interference estimation, resource allocation, localisation) proved to be an alternative to centralised algorithms attaining the same performance with inter-node signalling.

DIWINE has also shown ability of self-optimisation of WPLNC mapping selection across the cloud network. Local relay functions are selected in distributed way such that the global goal – reliable source to destination communication – is fulfilled and simultaneously the local utility function – related for example to energy efficiency – is optimised. The cloud network is also able to recover and self-heal from relays' failures. All of this newly added distributed features do not unbearably increase the communication overhead. The information is provided by existing pilot signals transmitted along data packets. Close cooperation with WP5 showed fruitfulness of the proposed algorithms in real scenarios.

A mapping scheme and a mapping matrix selection algorithm have been developed which are suitable for conventional binary approaches to modulation and coding, which allows both that each relay can choose a matrix to overcome singular fade states and that the overall mapping matrix is full rank, so that the destination can unambiguously decode source messages. The algorithm comprises an offline search algorithm which finds a small number of best coefficient matrices corresponding to the singular fade states, and hence greatly reduces the complexity of the online search for best coefficient matrices for

a particular channel state, with negligible performance loss compared to the exhaustive search. This mapping approach and algorithm has been implemented in the SLS.

Bibliography

[1] D. Halls, W. Thompson, T. Hyenek, P. Procházka, and T. Uřičář, "Measurement and performance evaluation of SMN demonstrator," DIWINE European Project, Tech. Rep. D5.42, 2016.

[2] S. Galimberti, S. Savazzi, L. Ascorti, G. Soatti, M. B. Nicoli, and D. Halls, "Measurement and performance evaluation of CIMC demonstrator," DIWINE European Project, Tech. Rep. D5.52, 2016.

[3] D. Boxiade, J. S. Vardakas, R. Torne, H. Barah, Y. Wang, C. Chen, T. Peng, A. G. Burr, A. Parichehreh, and D. Halls, "SLS and performance evaluation of advanced (non-HW demonstration) scenarios," DIWINE European Project, Tech. Rep. D5.33, 2016.

[4] U. Spagnolini, M. B. Nicoli, A. Pascale, G. Soatti, V. Forutan, A. Parichehreh, M. A. Alvarez Villanueva, S. Savazzi, J. Sýkora, T. Hynek, E. A. Jorswieck, A. G. Burr, and M. Molu Mortazawi, "Cloud network processing for core demonstration scenarios," DIWINE European Project, Tech. Rep. D3.02, 2014.

[5] M. A. Alvarez Villanueva, B. Azari, and U. Spagnolini, "Time and frequency self-synchronization in dense cooperative networks," in *Proceedings of the IEEE Asilomar Conference on Signals, Systems, and Computers (ACSSC)*, Nov. 2014.

[6] O. Simeone and U. Spagnolini, "Distributed time synchronization in wireless sensor networks with coupled discrete-time oscillators," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 57054, pp. 1–13, Sep. 2007.

[7] M. M. U. Gul, L. Sungeun, and M. Xiaoli, "Robust synchronization for OFDM employing Zadoff-Chu sequence," in *Proceedings of the Conference on Information Sciences and Systems (CISS)*, Mar. 2012.

[8] M. A. Alvarez Villanueva and U. Spagnolini, "Half-duplex scheduling in distributed synchronization," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2015.

[9] *Universal Software Radio Peripheral (USRP) N210*, Ettus Research, Sep. 2012. [Online]. Available: http://www.ettus.com/product/details/UN210-KIT

[10] M. A. Alvarez Villanueva, W. H. Thompson, and U. Spagnolini, "Distributed time and frequency synchronization: USRP hardware implementation," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2015.

[11] N. Varanese, U. Spagnolini, and Y. Bar-Ness, "Distributed frequency-locked loops for wireless networks," *IEEE Transactions on Communications*, vol. 59, no. 12, pp. 3440–3451, Dec. 2011.

[12] U. Spagnolini, M. B. Nicoli, G. Soatti, J. Sýkora, T. Uřičář, E. A. Jorswieck, A. Wolf, M. Mittelbach, Z. K. M. Ho, P.-H. Lin, Y. Wang, and A. G. Burr, "Fundamental limits for core demonstration scenarios," DIWINE European Project, Tech. Rep. D2.31, 2014.

[13] G. Soatti, M. B. Nicoli, S. Savazzi, and U. Spagnolini, "Distributed consensus-based algorithms for network calibration and localization," *IEEE Transactions on Signal and Information Processing over Networks*, 2016, submitted.

[14] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[15] P. Castiglione, S. Savazzi, M. B. Nicoli, and T. Zemen, "Partner selection in indoor-to-outdoor cooperative networks: An experimental study," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 77–85, Aug. 2013.

[16] H. Mark, "Dense cooperative wireless cloud network (DIWINE)," in *Proceedings of the Future Network and Mobile Summit (FuNeMS)*, Jul. 2013.

[17] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[18] M. B. Nicoli, G. Soatti, and S. Savazzi, "Distributed estimation of macroscopic channel parameters in dense cooperative wireless networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2014.

[19] A. Bolognino and U. Spagnolini, "Consensus based distributed estimation with local-accuracy exchange in dense wireless systems," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2014.

[20] J. E. Dennis Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Non-linear Equations*. Prentice Hall, 1983.

[21] G. Soatti, M. B. Nicoli, S. Savazzi, and U. Spagnolini, "Distributed sensing of interference pattern in dense cooperative wireless networks," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2015.

[22] J. M. Winter, I. Müller, G. Soatti, S. Savazzi, M. B. Nicoli, L. Bus Becker, J. C. Netto, and C. E. Pereira, "Wireless coexistence and spectrum sensing in Industrial Internet of Things: An experimental study," *International Journal of Distributed Sensor Networks*, vol. 2015, no. ID 627083, pp. 1–12, Oct. 2015.

[23] F. Dressler and O. B. Akan, "Bio-inspired networking: From theory to practice," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 176–183, Nov. 2010.

[24] J. Degesys, I. Rose, A. Patel, and R. Nagpal, "DESYNC: Self-organizing desynchronization and TDMA on wireless sensor networks," in *Proceedings of the International Symposium on Information Processing in Sensor Networks (IPSN)*, Apr. 2007.

[25] S. Mishra, A. Sahai, and R. Brodersen, "Cooperative sensing among cognitive radios," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2006.

[26] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, Nov. 2005.

[27] S. Atapattu, C. Tellambura, and H. Jiang, "Energy detection based cooperative spectrum sensing in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1232–1241, Apr. 2011.

[28] Z. Li, F. R. Yu, and M. Huang, "A distributed consensus-based cooperative spectrum-sensing scheme in cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 383–393, Jan. 2010.

[29] W. Zhang, Z. Wang, Y. Guo, H. Liu, Y. Chen, and J. Mitola, "Distributed cooperative spectrum sensing based on weighted average consensus," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2011.

[30] A. D. Whalen, *Detection of signal noise.* Academic Press, 1971.

[31] U. Spagnolini, "Cancellation of polarized impulsive noise using an azimuth-dependent conditional mean estimator," *IEEE Transactions on Signal Processing*, vol. 1, no. 46, pp. 3333–3344, Dec. 1998.

[32] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872–1905, Oct. 2011.

[33] S. Pawar, A. S. Avestimehr, and D. N. C. Tse, "Diversity-multiplexing tradeoff of the half-duplex relay channel," in *Proceedings of the Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2008.

[34] "Dynamic QMF for half-duplex relay networks," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Jul. 2012.

[35] R. Kolte, A. Özgür, and S. Diggavi, "When are dynamic relaying strategies necessary in half-duplex wireless networks?" *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1720–1738, Apr. 2015.

[36] S.-N. Hong and G. Caire, "Compute-and-forward strategies for cooperative distributed antenna systems," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5227–5243, Sep. 2013.

[37] D. Halperin, B. Greenstein, A. Sheth, and D. Wetherall, "Demystifying 802.11 n power consumption," in *Proceedings of the International Conference on Power Aware Computing and Systems (HotPower)*, Oct. 2010.

[38] S. Tsai and A. Soong, "Effective-SNR mapping for modeling frame error rates in multiple-state channels," Third Generation Partnership Project 2, Tech. Rep. 3GPP2-C30-20030429-010, 2003.

[39] L. Wan, S. Tsai, and M. Almgren, "A fading-insensitive performance metric for a unified link quality model," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2006.

[40] M.-A. Badiu, M. Varga, and V. Bota, "Link performance prediction methods for cooperative relaying in wireless networks," in *Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Sep. 2010.

[41] A. El Gamal, "Capacity theorems for relay channels," in *Proceedings of the MSRI Workshop on Mathematics of Relaying and Cooperation in Communication Networks*, Apr. 2006.

[42] V. Nagpal, I.-H. Wang, M. Jorgovanović, D. N. C. Tse, and B. Nikolić, "Quantize-map-and-forward relaying: Coding and system design," in *Proceedings of the Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2010.

[43] Z. Yang, S. Li, H. Feng, T. Honold, and G. Yu, "Cross-layer iterative decoding of irregular LDPC codes using cyclic redundancy check codes," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2009.

[44] V. Nagpal, I. Wang, M. Jorgovanović, D. N. C. Tse, and B. Nikolić, "Coding and system design for quantize-map-and-forward relaying," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 8, pp. 1423–1435, Aug. 2013.

[45] P. Sadeghi, R. A. Kennedy, P. B. Rapajić, and R. Shams, "Finite-state Markov modeling of fading channels-a survey of principles and applications," *IEEE Signal Processing Magazine*, vol. 25, no. 5, pp. 57–80, Sep. 2008.

[46] J. P. Coon and M. Sandell, "Combined bulk and per-tone transmit antenna selection in OFDM systems," *IEEE Communications Letters*, vol. 14, no. 5, pp. 426–428, May 2010.

[47] M. Sandell and J. P. Coon, "Performance of combined bulk and per-tone antenna selection precoding in coded OFDM systems," *IEEE Transactions on Communications*, vol. 60, no. 3, pp. 655–660, Mar. 2012.

[48] Y. Li, W. Wang, and F.-C. Zheng, "Combined bulk and per-tone relay selection in cooperative OFDM systems," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Aug. 2012.

[49] S. Kashyap and N. B. Mehta, "Joint antenna selection and frequency-domain scheduling in OFDMA systems with imperfect estimates from dual pilot training scheme," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3473–3483, Jul. 2013.

[50] T. Korakis, M. Knox, E. Erkip, and S. Panwar, "Cooperative network implementation using open-source platforms," *IEEE Communications Magazine*, vol. 47, no. 2, pp. 134–141, Feb. 2009.

[51] C. Politis, T. Oda, S. Dixit, A. Schieder, H.-Y. Lach, M. I. Smirnov, S. Uskela, and R. Tafazolli, "Cooperative networks for the future wireless world," *IEEE Communications Magazine*, vol. 42, no. 9, pp. 70–79, Sep. 2004.

[52] A. Chandra, C. Bose, and M. K. Bose, "Wireless relays for next generation broadband networks," *IEEE Potentials*, vol. 30, no. 2, pp. 39–43, Mar. 2011.

[53] P. Guo, Y. Bai, Z. Ma, S. Wu, and S. Dang, "Relay technology for multi-carrier systems: A research overview," in *Proceedings of the International Conference on Computer, Communication, Control and Information Technology (C3IT)*, Feb. 2015.

[54] H. Zhang and R. U. Nabar, "Transmit antenna selection in MIMO-OFDM systems: Bulk versus per-tone selection," in *Proceedings of the IEEE International Conference on Communications (ICC)*, May 2008.

[55] S. Dang, J. P. Coon, and D. E. Simmons, "Combined bulk/per-tone relay selection in two-hop OFDM systems," in *Proceedings of the IEEE Vehicular Technology Conference (VTC 2016-Spring)*, May 2016.

[56] ——, "Combined bulk and per-tone relay selection in super dense wireless networks," in *Proceedings of the IEEE International Conference on Communications (ICC) – Workshop on Advanced PHY and MAC Techniques for Super Dense Wireless Networks*, Jun. 2015.

[57] H. A. David and H. N. Nagaraja, *Order Statistics*. Wiley, 2004.

[58] M. O. Hasna and M.-S. Alouini, "A performance study of dual-hop transmissions with fixed gain relays," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Apr. 2003.

[59] D. E. Simmons and J. P. Coon, "Two-way OFDM-based nonlinear amplify-and-forward relay systems," *IEEE Transactions on Vehicular Technology*, 2016, to appear.

[60] G. Huang, Y. Wang, and J. P. Coon, "Performance of multihop decode-and-forward and amplify-and-forward relay networks with channel estimation," in *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PacRim)*, Aug. 2011.

[61] B. V. Nguyen, R. O. Afolabi, and K. Kim, "Dependence of outage probability of cooperative systems with single relay selection on channel correlation," *IEEE Communications Letters*, vol. 17, no. 11, pp. 2060–2063, Nov. 2013.

[62] Y. Chen and C. Tellambura, "Distribution functions of selection combiner output in equally correlated Rayleigh, Rician, and Nakagami-m fading channels," *IEEE Transactions on Communications*, vol. 52, no. 11, pp. 1948–1956, Nov. 2004.

[63] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill Education, 2007.

[64] L. Dai, B. Gui, and L. J. Cimini, "Selective relaying in OFDM multihop cooperative networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2007.

[65] S. Dang, J. P. Coon, and G. Chen, "An equivalence principle for OFDM-based combined bulk/per-subcarrier relay selection over equally spatially correlated channels," *IEEE Transactions on Vehicular Technology*, 2016, submitted.

[66] S. Dang, D. E. Simmons, and J. P. Coon, "Comparison of multicarrier relay selection schemes in super dense networks," in *Proceedings of the IEEE International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, Sep. 2015.

[67] S. Dang, J. P. Coon, and D. E. Simmons, "Combined bulk/per-subcarrier relay selection in two-hop OFDM systems," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, May 2016.

[68] Y. Liang, H. V. Poor, and S. Shamai (Shitz), *Foundations and Trends in Communications and Information Theory: Information Theoretic Security*. now Publishers, 2009.

[69] A. D. Wyner, "The wiretap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[70] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[71] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[72] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multi-antenna wiretap channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[73] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas – I: The MISOME wiretap channel," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[74] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[75] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.

[76] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[77] Z. Li, R. Yates, and W. Trappe, "Achieving secret communication for fast Rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2792–2799, Sep. 2010.

[78] S.-C. Lin and P.-H. Lin, "On ergodic secrecy capacity of multiple input wiretap channel with statistical CSIT," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 2, pp. 414–419, Feb. 2013.

[79] M. R. Bloch and J. N. Laneman, "Exploiting partial channel state information for secrecy over wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1840–1849, Sep. 2013.

[80] Y. Liang, L. Lai, H. V. Poor, and S. Shamai (Shitz), "A broadcast approach for fading wiretap channels," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 842–858, Feb. 2014.

[81] S.-C. Lin and C.-L. Lin, "On secrecy capacity of fast fading MIMOME wiretap channels with statistical CSIT," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 3293–3306, Jun. 2014.

[82] P. Mukherjee and S. Ulukus, "Fading wiretap channel with no CSI anywhere," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Jul. 2013.

[83] J. Li and A. Petropulu, "On ergodic secrecy rate for Gaussian MISO wiretap channels," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1176–1187, Apr. 2011.

[84] P.-H. Lin, S.-H. Lai, S.-C. Lin, and H.-J. Su, "On optimal artificial-noise assisted secure beamforming for the fading eavesdropper channel," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1728–1740, Sep. 2013.

[85] M. Shaked and J. G. Shanthikumar, *Stochastic Orders*. Springer, 2007.

[86] S. M. Ross, *Stochastic process*. John Wiley & Sons, 1996.

[87] P.-H. Lin and E. A. Jorswieck, "On the fast fading Gaussian wiretap channel with statistical channel state information at transmitter," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 1, pp. 46–58, Jan. 2016.

[88] T. Tao and V. Vu, "Random matrices: The distribution of the smallest singular values," *Geometric and Functional Analysis*, vol. 20, no. 1, pp. 260–297, Jun. 2010.

[89] G. A. F. Seber, *A matrix handbook for statisticians*. Wiley, 2007.

[90] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Application*. Academic Press, 1979.

[91] S. Jin, M. R. McKay, X. Gao, and I. B. Collings, "MIMO multichannel beamforming: SER and outage using new eigenvalue distributions of complex noncentral Wishart matrices," *IEEE Transactions on Communications*, vol. 56, no. 3, pp. 424–434, Mar. 2008.

[92] S. Loyka and C. D. Charalambous, "On optimal signaling over secure MIMO channels," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Jul. 2012.

[93] J. Li and A. Petropulu, "Optimality of beamforming and closed form secrecy capacity of mimo wiretap channels with two transmit antennas," in *Proceedings of the IEEE International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Jun. 2013.

[94] S. A. A. Fakoorian and A. L. Swindlehurst, "Full rank solutions for the MIMO Gaussian wiretap channel with an average power constraint," *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2620–2631, May 2013.

[95] A. Gupta and D. Nagar, *Matrix variate distributions*. Chapman & Hall/CRC, 2000.

[96] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: The received signal density in closed form," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1473–1484, Jun. 2002.

[97] R. Love, "Physical layer procedures for evolved universal terrestrial radio access (EUTRA)," Third Generation Partnership Project (3GPP), Tech. Rep. 3GPP-TS-36.213-V11.0.0, 2012.

[98] X. Lin, J. G. Andrews, and A. Ghosh, "Spectrum sharing for device-to-device communication in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 6727–6740, Sep. 2014.

[99] C.-H. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2752–2763, Aug. 2011.

[100] Y. Pei and Y.-C. Liang, "Resource allocation for device-to-device communications overlaying two-way cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3611–3621, Jun. 2013.

[101] D. Zheng, C. He, L. Jiang, J. Ding, Q. Zhang, and Q. Xi, "QoS-based resource allocation for multi-D2D communications in heterogeneous networks," in *Proceedings of the IEEE International Conference on Communication Workshop (ICCW)*, Jun. 2015.

[102] Z. Liu, T. Peng, H. Chen, and W. Wang, "Optimal D2D user allocation over multi-bands under heterogeneous networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2012.

[103] M. A. Alvarez Villanueva, G. Soatti, M. B. Nicoli, and U. Spagnolini, "Device-to-device resource scheduling by distributed interference coordination," in *Proceedings of the IEEE International Conference on Communications (ICC)*, May 2016, submitted.

[104] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 86–92, May 2014.

[105] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference coordination and cancellation for 4G networks," *IEEE Communications Magazine*, vol. 47, no. 4, pp. 74–81, Apr. 2009.

[106] K. Andersson, S. Al Mahmud Mostafa, and R. Ul-Islam, "Mobile VoIP user experience in LTE," in *Proceedings of the IEEE Conference on Local Computer Networks (LCN)*, Oct. 2011.

[107] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.

[108] D. Halls, P. Procházka, T. Hyenek, M. A. Alvarez Villanueva, V. Rampa, J. Sýkora, U. Spagnolini, E. A. Jorswieck, K. Ramantas, D. Boxiade, R. Torne, and P.-H. Lin, "SMN platform description and high-level routines," DIWINE European Project, Tech. Rep. D5.41, 2015.

[109] J. Sýkora, P. Procházka, T. Uřičář, T. Hynek, M. Hekrdla, D. Fang, A. G. Burr, and J. Yuan, "Terminal node processing for core demonstration scenarios," DIWINE European Project, Tech. Rep. D4.02, 2014.

[110] T. Uřičář and J. Sýkora, "Non-uniform 2-slot constellations for bidirectional relaying in fading channels," *IEEE Communications Letters*, vol. 15, no. 8, pp. 795–797, Aug. 2011.

[111] A. B. MacKenzie and L. A. DaSilva, *Game Theory for Wireless Engineers*. Morgan & Claypool, 2006.

[112] T. Hynek, D. Halls, and J. Sýkora, "SER-based utility optimization for wireless physical layer network coding mapping selection in relay networks," *Wireless Newtorks*, 2016, submitted.

[113] W. Wang, M. Chatterjee, and K. Kwiat, "Coexistence with malicious nodes: A game theoretic approach," in *Proceedings of the International Conference on Game Theory for Networks (GameNets)*, May 2009.

[114] S. Dehnie, H. T. Sencar, and N. Memon, "Detecting malicious behavior in cooperative diversity," in *Proceedings of the Conference on Information Sciences and Systems (CISS)*, Mar. 2007.

[115] M.-H. Chen, S.-C. Lin, and Y.-W. P. Hong, "A game theoretic approach for the cooperative network with the presence of malicious relays," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2011.

[116] T. Hynek and J. Sýkora, "Wireless physical layer network coding in potential presence of malicious relays – incomplete information game approach," *Electronics Letters*, vol. 51, no. 16, pp. 1292–1294, 2015.

[117] R. F. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[118] J. C. Harsanyi, "Games with incomplete information played by "Bayesian" players I–III, Part I. The basic model," *Management Science*, vol. 14, no. 3, pp. 159–182, Nov. 1967.

[119] T. Koike-Akino, P. Popovski, and V. Tarokh, "Optimized constellations for two-way wireless relaying with physical network coding," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 5, pp. 773–787, Jun. 2009.

[120] T. Uřičář and J. Sýkora, "Design criteria for hierarchical exclusive code with parameter-invariant decision regions for wireless 2-way relay channel," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, no. 921427, pp. 1–13, Jul. 2010.

[121] D. Fudenberg and J. Tirole, *Game Theory*. MIT Press, 1991.

[122] T. Hynek and J. Sýkora, "Non-cooperative broadcast game for distributed decision map selection of relay wireless network coding processing," in *Proceedings of the IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2013.

[123] B. Nazer and M. Gastpar, "Compute-and-Forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[124] M. Molu Mortazawi, K. Cumnan, Y. Wang, and A. G. Burr, "A novel simple approach for estimating the best compute-and-forward equation at high SNR," *IEEE Transactions on Communications*, 2016, to appear.

[125] L. Wei and W. Chen, "Compute-and-forward network coding design over multi-source multi-relay channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3348–3357, Sep. 2012.

[126] Z. Chen, P. Fan, and K. Ben Letaief, "Compute-and-forward: Optimization over multisource-multirelay networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 1806–1818, May 2015.

[127] A. Barreal, J. Pääkkönen, D. Karpuk, C. Hollanti, and O. Tirkkonen, "A low-complexity message recovery method for compute-and-forward relaying," *CoRR*, vol. abs/1504.03182, 2015. [Online]. Available: http://arxiv.org/abs/1504.03182

[128] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 273–278, Jan. 1998.

[129] R. de Buda, "Some optimal codes have structure," *IEEE Journal on Selected Areas in Communications*, vol. 7, no. 6, pp. 893–899, Aug. 1989.

[130] T. Linder, C. Schlegal, and K. Zeger, "Corrected proof of de Buda's theorem," *IEEE Transactions on Information Theory*, vol. 39, pp. 1735–1737, Sep. 1993.

[131] U. Erez and R. Zamir, "Achieving 1/2 log (1+SNR) on the AWGN channel with lattice encoding and decoding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[132] C. Feng, D. Silva, and F. R. Kschischang, "An algebraic approach to physical-layer network coding," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7576–7596, Nov. 2013.

[133] C. Buchheim, A. Caprara, and A. Lodi, "An effective branch-and-bound algorithm for convex quadratic integer programming," *Mathematical programming*, vol. 135, no. 1, pp. 369–395, Oct. 2012.

[134] J. Richter, C. Scheunert, and E. A. Jorswieck, "An efficient branch-and-bound algorithm for compute-and-forward," in *Proceedings of the IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2012.

[135] D. Wubben, D. Seethaler, J. Jalden, and G. Matz, "Lattice reduction," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 70–91, May 2011.

[136] Y. H. Gan, C. Ling, and W. H. Mow, "Complex lattice reduction algorithm for low-complexity full-diversity MIMO detection," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2701–2710, Jul. 2009.

[137] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Annalen*, vol. 261, no. 4, pp. 515–534, 1982.

[138] N. Gama and P. Q. Nguyen, "Finding short lattice vectors within Mordell's inequality," in *Proceedings of the ACM symposium on Theory of computing*, May 2008.

[139] B. Zhou and W. H. Mow, "A quadratic programming relaxation approach to compute-and-forward network coding design," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Jun. 2014.

[140] S. Sahraei and M. Gastpar, "Compute-and-forward: Finding the best equation," in *Proceedings of the Allerton Conference on Communication, Control, and Computing (Allerton)*, Oct. 2014.

[141] H. Daudé and B. Vallée, "An upper bound on the average number of iterations of the LLL algorithm," *Theoretical Computer Science*, vol. 123, no. 1, pp. 95–115, Jan. 1994.

[142] C. Ling and N. Howgrave-Graham, "Effective LLL reduction for lattice decoding," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Jun. 2007.

[143] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. P. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[144] S. Zhang, S.-C. Liew, and P. P. K. Lam, "Hot topic: Physical-layer network coding," in *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, Sep. 2006.

[145] P. Popovski and H. Yomo, "Bi-directional amplification of throughput in a wireless multi-hop network," in *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, May 2006.