IMT Institute for Advanced Studies Lucca
Lucca, Italy

Ph.D. Program in Computer, Decision, and Systems Science

Ph.D. Thesis

Cooperative Techniques in Wireless Communications: a Game- and Information- Theoretic Analysis

Author:
Farshad Shams ....

Supervisor:
Prof. Marco Luise ....

December 2012

This thesis has been reviewed by the external examining committee:

1) Mérouane Debbah . . . .

Full professor at Supélec, Paris, France

2) Jianwei Huang . . . .

Assistant professor at the Chinese University of Hong Kong

3) Erik G. Larsson . . . .

Full professor at the Linköping University, Linköping, Sweden

Abstract

This dissertation consists of two parts. In the first part (Chapters 1-3) we study the notion of cooperative game theory and its applications to network engineering problems. The tools of cooperative game theory are shown to be advantageous for obtaining high-performance and good results in terms of fairness, and stability. In particular, the first part describes a theoretical framework for the design and analysis of resource allocation algorithms for wireless networks using OFDMA technology in uplink transmission. The resource allocation issue is modeled as a cooperative game in which every user terminal is assigned to a set of subcarriers, and then chooses its transmit power so as to achieve its demanded data rate exactly. The power distribution is obtained by a dynamic learning algorithm based upon Markov modeling. Simulation results show that the average number of operations of the proposed iterative algorithm are much lower than $K.N$, where N and K are the number of allocated subcarriers and of mobile terminals.

The second part (Chapters 4-6) deals with the information theoretic issues around multi hop communication where some nodes act as intermediate relays. We present the advantages and limitations of different relaying strategies. We first review point-to-point communication assisted by one relay, and derive the upper bounds on the capacity of memoryless AWGN relay channel. We bound the capacity region of the channel using, the max-flow min-cut (cutset) theorem and find achievable rate regions for the amplify and forward, decode and forward, and compress and forward protocols. Then, we aim at studying the capacity of AWGN channels assisted by multiple parallel relay channels. We are interested in two cooperation models involving point-to-point and multiple access (multiple transmitters and single receiver). We derive the capacities of AWGN channels for different relaying techniques. In addition, we study a reliable network wherein some relays perform decode and forward strategy and others relays as compress and forward. We show that the improvement of the overall data rate strongly depends on the combination of relays' positions and relaying strategy. Also, we will show that a multiple parallel relayed communication achieves its maximum capacity by either only one relay or all relays together. Adding a new relay/transmitter does not necessarily expand outer capacity region, and can even degrade the capacity of the network.

Contents

List of figures **ix**

List of tables **xiii**

List of acronyms **xv**

List of symbols and operators **xvii**

Introduction **1**

Motivation 1

Main contributions 6

Outline of the dissertation 8

I Cooperative Resource Allocation in OFDMA Wireless Communications **11**

1 A survey on resource allocation techniques in OFDM(A) **13**

1.1 OFDM(A) modulation and channel access 14

1.2 Problem formulation in OFDM 18

1.3 Taxonomy of problem types in OFDMA 21

1.4 Water-filling solution 25

1.5 Max-min fairness criterion 26

1.6 Weighted proportional fairness criterion 29

1.7 Utility maximization scheme 33

1.8 Cross-layer scheme 35
1.9 Game theory solution 36
1.10 Discussion 41

2 Cooperative game theory 45
2.1 Motivation 46
2.2 Preliminaries 48
2.3 The core solution 50
2.3.1 On core stability 56
2.4 Shapley value 57
2.5 The kernel and nucleolus 60
2.6 Cooperative Nash equilibria 63
2.7 Coordinated equilibrium 71
2.8 Dynamic learning 77
2.9 Discussion 81

3 A resource allocation cooperative game in OFDMA 83
3.1 Problem formulation 85
3.1.1 Subchannel assignment 85
3.1.2 Power allocation 86
3.2 The best-response algorithm 89
3.3 Numerical results 95
3.4 Discussion 105

II On the Capacity of Relay Communications 107

4 Basics of cooperative communications 109
4.1 Introduction 110
4.2 Relayed point-to-point communication 114
4.3 Cutset upper bound 118
4.4 Amplify and forward technique 122
4.5 Decode and forward technique 123

4.6 Compress and forward technique 127
4.7 Case study 132
4.8 Summary 135

5 Multiple parallel relayed point-to-point communication 137

5.1 System model 138
5.2 Cutset upper bound 140
5.3 Amplify and forward technique 145
5.4 Decode and forward technique 146
5.5 Compress and forward technique 149
5.6 Mixed DF & CF techniques 155
5.7 Case study 158
5.8 Summary 161

6 Multiple sources, parallel relays, one destination 163

6.1 System model 164
6.2 Cutset upper bound 167
6.3 Amplify and forward technique 174
6.4 Decode and forward technique 176
6.5 Compress and forward technique 181
6.6 Mixed DF & CF techniques 186
6.7 Case study 190
6.8 Summary 193

7 Summary and perspective 195

A Multivariate Gaussian distribution 199

B Some known results in information theory 201

B.1 Memoryless channels 201
B.1.1 Broadcast communication 206
B.1.2 Multiple-access communication 208
B.1.3 General multi node network 210

B.2 Wyner-Ziv source coding 212

Bibliography **215**

List of Figures

2.1 The network allocates power among three players according to their will to cooperate with each other. A selfish player receives zero, a pair of cooperative players receive 0.8 mW, and the network supply 1 mW to the grand coalition. 54

2.2 The network scenario of the forwarder's dilemma game. 64

2.3 The network scenario in the near-far effect game. 74

3.1 Block partitioning of the available bandwidth. 85

3.2 Shape of the utility as a function of the Shannon capacity. . . . 88

3.3 Achieved rates as functions of the iteration step. 97

3.4 Average normalized power expenditure as a function of K, with $W = 10\,\mathrm{MHz}$, $N = 1024$, and $R_k^\star = R^\star = 200\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ in the case of vacant-carrier assignment model. 98

3.5 Experimental average number of operations as a function of K, with $W = 10\,\mathrm{MHz}$, $N = 1024$, and $R_k^\star = R^\star = 200\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ in the case of vacant-carrier assignment model. 98

3.6 Average normalized power expenditure as a function of K, with $W = 10\,\mathrm{MHz}$, $N = 1024$, and $R_k^\star = R^\star = 200\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ in the case of best-carrier assignment model. 101

3.7 Experimental average number of operations as a function of K, with $W = 10\,\mathrm{MHz}$, $N = 1024$, and $R_k^\star = R^\star = 200\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ in the case of best-carrier assignment model. 101

3.8 Experimental average number of operations as a function of K, with $W = 10\,\mathrm{MHz}$, $N = 512$, and $R_k^\star = R^\star = 500\,\mathrm{kb/s}\ \forall k \in \mathcal{K}$ in the case of vacant-carrier assignment model. 102
3.9 Experimental average number of operations as a function of K, with $W = 20\,\mathrm{MHz}$, $N = 2048$, and $R_k^\star = R^\star = 2\,\mathrm{Mb/s}\ \forall k \in \mathcal{K}$ in the case of vacant-carrier assignment model. 103
3.10 Average number of assigned subcarriers as a function of the achieved rate, with $\overline{\Delta p}_{kn} = 600\,\mathrm{nW}$ in the case of vacant-carrier assignment model. 104

4.1 One source, one relay, one far destination network. 114
4.2 One source, one relay, one destination network scenario. 116
4.3 A single relay communication network scenario. 132
4.4 Rates for one relay with $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 1\,\mathrm{m}$, and $d_r = 0.1\,\mathrm{m}$. 133
4.5 Rates for one relay with $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 500\,\mathrm{m}$, and $d_r = 10\,\mathrm{m}$. 134

5.1 One source, multiple parallel relays, one destination cooperative communication network. 138
5.2 Point-to-point two relayed communication network scenario. . . 159
5.3 Rates for two relays with $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 1\,\mathrm{m}$, and $d_r = 0.1\,\mathrm{m}$. 160
5.4 Rates for one relay with $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 500\,\mathrm{m}$, and $d_r = 10\,\mathrm{m}$. 160

6.1 Multiple sources, multiple parallel relays, one destination communication network scenario. 164
6.2 Multiple access multiple relayed communication network scenario. 190
6.3 Overall rates of two transmitters relayed by two relays with $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 1\,\mathrm{m}$, $d_k = 0.05\,\mathrm{m}$, and $d_r = 0.1\,\mathrm{m}$. 192

6.4 Overall rates of two transmitters aided by two relays with $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 500\,\mathrm{m}$, $d_k = 5\,\mathrm{m}$, and $d_r = 10\,\mathrm{m}$. 192

B.1 The basic communication system. 202
B.2 Broadcast communication network. 207
B.3 Multiple access communication network. 209
B.4 General communication network. 211
B.5 Slepian-Wolf source coding scheme. 213
B.6 Wyner-Ziv source coding scheme. 213

List of Tables

2.1 The strategic form in the forwarder's dilemma game. In each cell, the first value is the payoff of player k_1, whereas the second is that of k_2. 64

2.2 Subcarrier allocation in OFDMA network game in strategic form. The three strategies for three players k_1, k_2 and k_3 are: Transmitting with the maximum power only on subcarrier number 1 (a_1), transmitting with the maximum power only on subcarrier number 2 (a_2) and equal division of the maximum power among both subcarriers (a_3). The player's payoff is the achieved channel capacity in kb/s. 70

2.3 The multiple access game in strategic form. The two moves for each player are: access (A) or wait (W). 72

2.4 The strategic form matrix of the multiple access game with preplay agreement. 73

2.5 Payoff matrix for the near-far effect game with power control and variable throughput. 75

2.6 The strategic form matrix of the near-far effect game with preplay agreement, and with $\kappa = \dfrac{\delta u}{(1-\eta)\,c}$. 77

List of Acronyms

1/2/3/4G	first/second/third/fourth generation
3GPP	third generation partnership project
ADSL	asymmetric digital subscriber lines
AF	amplify and forward
AP	access point
AWGN	additive white Gaussian noise
BER	bit error rate
BPSK	binary phase-shift keying
BS	base station
CDMA	code division multiple access
CF	compress and forward
DAB	digital audio broadcasting
DF	decode and forward
DFT	discrete Fourier transform
DVB-T	digital video broadcasting for terrestrial television
FEC	forward error correction
FFT	fast Fourier transform
GSM	global system for mobile communication
ICI	inter carrier interference
i.i.d.	independent and identically distributed
IP	Internet protocol

ISI	inter symbol interference
LP	linear programming
LTE	long term evolution
MAC	medium access control
MAI	multiple access interference
MC	multi-carrier
MC-CDMA	multi-carrier code division multiple access
MIMO	multiple-input multiple-output
MRC	maximal ratio combining
NE	Nash equilibrium
NTU	non transferable utility
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
OSI	open systems interconnection
PAPR	peak-to-average power ratio
QAM	quadrature adaptive modulation
QASK	quadrature amplitude shift keying
QoS	quality of service
QPSK	quadrature phase-shift keying
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
SW	Slepian-Wolf
TU	transferable utility
UMTS	universal mobile telecommunications system
UWB	ultra-wideband
WiMAX	worldwide interoperability for microwave access
WZ	Wyner-Ziv

List of symbols and operators

Throughout this thesis the following notational conventions are used:

absolute value	$\lvert\cdot\rvert$
cross product	$\times$
$\max\{x, 0\}$	$[x]^{+}$
characteristic vector of coalition $\mathcal{A}$	$\mathbf{1}_{\mathcal{A}}$
pure strategy	a
pure strategy space of player k	$\mathbf{A}_k$
joint pure strategy chosen by the player k	$\mathbf{a}_k \in \times_k \mathbf{A}_k$
set of jointly ϵ-typical	A_{ϵ}
generic coalition (subset) of players and transmitters	$\mathcal{A}$
complementary set of $\mathcal{A}$	$\mathcal{A}^C$
coalition at time step t	$\mathcal{A}^t$
block index of intended send symbols	b
total number of blocks	B
power expenditure cost	c
Shannon capacity formula: $\frac{1}{2}\log_2(1+\gamma)$	$C(\gamma)$
covariance between two random variables X and Y	$\mathbb{C}\mathrm{ov}\,(X\,,\,Y)$
information flow upper bound symbol	$\mathcal{C}$
generic index for a destination	d

determinant of a matrix Σ	$\det\{\Sigma\}$
number of destinations	D
set of destinations	$\mathcal{D}$
excess value of the coalition $\mathcal{A}$ with respect to the payoff distribution $\mathbf{u}$	$e\,(\mathcal{A}\,,\,\mathbf{u})$
vector of excess values of all coalitions arranged in non-icreasing order	$\mathbf{E}$
amplification factor at the relay r	E_r
expectation of a random variable X	$\mathbb{E}\,\{X\}$
generic relaying function	f_r
family of coalitions	$\mathcal{F}$
coalitional game	$\mathcal{G}$
real-valued channel response of the channel between terminal k and the base station over subcarrier n	h_{kn}
complex-valued channel response of the channel between terminal k on nth subcarrier	H_{kn}
amplitude of H_{kn}	$\|H_{kn}\|^2$
entropy of a random variable X	$\mathrm{H}\,(X)$
conditional entropy of X given Y	$\mathrm{H}\,(X\|Y)$
infimum under a given probability	$\inf_{p(.)}$
average mutual information between two sequence of random variables	$\mathrm{I}\,(X\,;\,Y)$
set of imputations in a game	$\mathcal{I}\,(\mathcal{K}\,,\,\boldsymbol{\nu})$
other feasible imputation	$\mathcal{J}$
generic index for a terminal/player	k
number of terminals/players	K
set of terminals/players	$\mathcal{K}$
generic index for carrier subblock	l
number of carrier subblocks	L

size of alphabet space	m
generic index for a subcarrier	n
number of carriers	N
set of carriers	$\mathcal{N}$
set of carriers of the lth subblock	$\mathcal{N}^{(l)}$
set of carriers assigned to transmitter k	$\mathcal{N}_k$
(real) Gaussian distribution with mean μ and variance σ^2	$\mathcal{N}(\mu\,,\,\sigma^2)$
probability	$\Pr\{\cdot\}$
probability density function of a channel	$p(x)$
channel conditional probability distribution of x given y	$p(x\|y)$
transmit power of terminal k over carrier n	p_{kn}
transmit power of terminal k over carrier n at time step t	p_{kn}^t
tentative transmit power of terminal k over carrier n	$\tilde{p}_{kn}$
previous tentative transmit power of terminal k over carrier n	$\hat{p}_{kn}$
maximum transmit power of terminal k over carrier n	$\overline{p}_{kn}$
maximum tentative transmit power of terminal k over carrier n	$\tilde{p}_{kn}^{\max}$
Q value of user k at time step t	Q_k^t
generic index for a relay node	r
number of relays	R
data rate (Shannon capacity) achieved by terminal k	R_k
data rate achieved by terminal k on the carrier n	R_{kn}
achieved data rate between two terminals k and m	$R_{(km)}$
target data rate of terminal k	$R_k^\star$
maximum demanded rate of a terminal user	$\overline{R}_k$
minimum demanded rate of a terminal user	$\underline{R}_k$
rate vector: $\mathbf{R} = [R_1, \ldots, R_k]$	$\mathbf{R}$

real numbers	$\mathbb{R}$
set of relay nodes	$\mathcal{R}$
maximum excess of player k against i	s_{ki}
supremum under a given probability	$\sup_{p(.)}$
core set of a game	$\mathcal{S}\left(\mathcal{K}\,,\, \boldsymbol{\nu}\right)$
generic time step	t
temperature function	T
generic subset of relays	$\mathcal{T}$
reward symbol	u
payoff (reward) of player k	u_k
payoff distribution across players	$\mathbf{u}$
variance of a random variable X	$\mathbb{V}\text{ar}\left(X\right)$
send symbol of the user k in the block index b	w_k^b
vector of $[w_1^b, \ldots, w_K^b]$	$\boldsymbol{w}_{\mathcal{K}}^b$
OFDM signal bandwidth	W
transmitter k message set	$\mathcal{W}_k$
generic symbols for channel distribution	$x\,,\, y$
joint of channel distributions x_i for all $i \in \mathcal{A}$	$\boldsymbol{x}_{\mathcal{A}}$
generic symbols for random variables	$X\,,\, Y$
X and Z are conditionally independent given Y	$X \leftrightarrow Y \leftrightarrow Z$
estimate for X	$\widehat{X}$
send signal from node i	X_i
joint of the random variables X_i for all $i \in \mathcal{A}$	$\boldsymbol{X}_{\mathcal{A}}$
Gaussian alphabet space at node i's encoder	$\mathcal{X}_i$
received signal at node i	Y_i
Gaussian alphabet space at node i's decoder	$\mathcal{Y}_i$
additive white Gaussian noise at the node i	Z_i
auxiliary alphabet space	$\mathcal{Z}$

generic positive constants	α, β
received SNR of terminal k over carrier n	γ_{kn}
carrier spacing	Δf
power step to update the tentative transmit power of terminal k over carrier n	$\Delta\tilde{p}_{kn}$
maximum power step to update the tentative transmit power of terminal k over carrier n	$\overline{\Delta p}_{kn}$
discount factor of the user k	δ_k
tolerance parameter for demanded data rate	ε
normalized average power expenditure of terminal k	ζ_k
stopping criterion of the iterative algorithm	Θ
probability of transmit power update	λ
balanced weight of coalition $\mathcal{A}$	$\mu_{\mathcal{A}}$
uniformly-distributed random variable	ξ_{kn}^t
correlation value between the outputs of source node k and a relay r	ρ_{kr}
correlation value between the outputs of two source nodes	σ_{km}
AWGN power	σ_w^2
determinant of the $k{*}k$ covariance matrix $[\mathbb{C}\mathrm{ov}\,(X_i\,,\,X_j)]\,,$ $i = 1, \ldots, k;\ j = 1, \ldots, k$	$\mathbf{\Sigma}[X_1X_2\cdots X_k]$
correlation value between the outputs of two relays	τ_{rj}
coalition utility function	ν
set of coalition utilities	$\boldsymbol{\nu}$
set of coalition utilities at time step t	$\boldsymbol{\nu}^t$
Shapley value of the player k with respect to coalition payoff ν	$\phi_k\,(\nu)$
weight coefficient of fairness of a terminal user	φ_k
ergodic set	Φ
set of disjoint coalitions	ψ

set of disjoint coalitions at time step t	ψ^t
set of all possible ψ	$\boldsymbol{\Psi}$
state of the Markov chain	ω
state of the Markov chain at time step t	ω^t
state space of the Markov chain	Ω

Introduction

Motivation

Wireless networks have come a long way in a remarkably short time. Evolution of wireless access technologies for cellular communications is about to the reach its fourth generation (4G). Looking past, wireless access technologies has followed different evolutionary paths towards a unified target: performance and efficiency in high mobility environment. Mobile and wireless communication systems have been successfully and extensively deployed since the 1980s. In the first generation (1G), different analogue systems were deployed mainly in the developed regions of the world to support voice telephony services for mobile subscribers. 2G introduced the global system for mobile communication (GSM), which was deployed internationally from 1991. In the beginning, the main objective of GSM was the support of voice telephony and international roaming with a single system across Europe. In parallel with the fast growth of 2G mobile communication, third generation (3G) systems were developed from about 1998 and deployed globally since 2002. 3G has quest for data at higher speeds to open the gates for good "mobile broadband" experience. It is considered broad because multiple types of services can be provided across the wide band at high speeds, and mobile broadband, on the other hand, launches these services on mobile platforms.

The great interest for higher data rates exhibited by users of mobile wireless services exhibits an exponential increase. Long term evolution (LTE) technology seeks to improve voice quality and expand broadband data services, to deliver high-definition video and audio and other on-demand and real-time services on a "anything-anywhere-anytime" basis. The Release-10 LTE-Advanced (LTE-A) is a major enhancement of the LTE standard developed by the 3rh Generation Partnership Project (3GPP), and LTE-A was ratified by the International Telecommunication Union (ITU) as an IMT-Advanced (4G) technology in November 2010. The 4G objective is to meet challenges presented by the ever increasing use of "smart" wireless devices that require significantly higher spectral resources than conventional cell phones. LTE-A addresses those challenges by targeting peak data rates 1 Gb/s with up to 100 MHz supported spectrum bandwidth.

In the short space of just a few years, wireless services have been transformed far more than in the previous decades. Mobile communication systems enable many new applications and allow more flexibility for users and thereby an improvement of quality of life and efficiency of business processes. From serving voice communication only, the advent of new high-definition and affordable/free-to-air entertainment services such as IPTV and web browsing on new powerful cell phones, has demanded unprecedented wireless quality of service (QoS) requirement. This justifies the need for wideband, high-capacity wireless communication technologies that use the available bandwidth efficiently and provide data rates close to channel capacity [1]. While the peak data rate continues to be important from a marketing perspective, the end-user experience is more closely associate with the achieved degree of uniformity of service provision. The trend of increasing demand for high quality of service at the user equipment, coupled with the shortage of wireless spectrum imply that future enhancements will be needed to augment cell edge data rates rather than simply peak rates. In comparison to wireline channel, the radio channel is time variant, and transmission should be adapted to channel quality as well as to service requirements. With user mobility, this imposes several challenges in providing effective communications. Multicarrier access techniques such as orthogonal frequency-division multiple access (OFDMA) [2] can be exploited to increase data rates in a multi-user environment, by dividing a frequency-selective broadband channel into a multitude of orthogonal narrowband flat-fading subchannels.

We are seeing that mobile broadband technologies are reaching a commonality in the air interface and networking architecture; they are converging towards an IP-based network architecture with OFDMA based technology. Although network evolution has not yet reached the point of true and full convergence, wireless access networks is being designed to support ubiquitous delivery of multimedia services via inter-networking. First worldwide debut of IP-OFDMA-based mobile broadband is with worldwide interoperability for microwave access (WiMAX) technology. The spread spectrum radio technology used in 3G systems, is abandoned in all 4G candidate systems and replaced by OFDMA transmission schemes, making it possible to transfer very high bit rates despite extensive multi-path radio propagation. Applying (multi-carrier) OFDMA make it even feasible to provide a full range of multiuser/multirate services, all with different QoS and performance requirements. Due to the concurrent access in multiuser environment, and the large degree of freedom when it comes to the allocation of a large number of subcarriers, an efficient radio resource management is of paramount importance for such systems. The goal of resource allocation can be either the maximization of the overall data rate or minimization of the total transmit power, at the base station or at the individual wireless terminal, in the downlink or uplink direction. Though significant progress has been made for the resource allocation in OFDMA, a "standard" and optimum approach is still not available. In fact, it is quite unlikely we would arrive at a generic approach for the resource allocation in wireless networks due to the inherent multi-faceted nature of the problem in question.

Regardless of being uplink or downlink communication, resource allocation in OFDMA consists of two sub-problems: subcarrier selection for each individual user, and transmit power over every subcarrier. An intelligent and scalable joint power and bandwidth allocation mechanism is crucial to ensure QoS to the consumer at a reasonable cost [2]. A well devised algorithm for subcarrier selection can significantly increase signal-to-interference-plus-noise ratio (SINR), that is necessary for throughput enhancement in dynamic access. Moreover, transmit power in wireless cellular networks is a key degree of freedom in the management of interference, energy, and connectivity. Energy efficiency is an area of increasing importance due to the rising cost of energy and enviromental concerns. "The 3G chipsets that are available to semiconductors work reasonably well except for power. They are real power hogs, so as you know, the handset battery life used to be 5-6 hours for GSM, but when we got to 3G they got cut in half. Most 3G phones have battery lives of 2-3 hours": said Steve Jobs[1] in the early days of universal mobile telecommunications system (UMTS). Whilst "integrated circuits will double in performance every 18 months" according to Moore's law, but batteries have not improved hardly at all. There are deep physical limits. The energy capacity is only doubling every 10 years, as battery industry foresees. This justifies the necessity of an energy efficiency resource allocation technique in modern wireless networks.

The limited resources and the increasing number of the users are inevitably accentuating the relevance of good management of the wireless network resources. In recent years, game theory has appeared as an expedient tool for network engineers to address competition among user terminals, and also wireless service providers. Besides, it might be possible that network equipment, e.g. wireless terminals and base-station, cooperate each other to access the resources, and through this cooperation effectively achieve a robust allocation strategy which promises significant benefits such as higher throughput, fairness, better QoS, and lower network interference. The cooperation upon which such networking concepts are based, may also be a key to realizing higher-order scalability. *Cooperative game theory* is a branch of game theory which is dedicated to model the various aspects of cooperative and interactive behavior. The requirement of applying cooperative game theory in wireless networks is to implement network equipment policies as a set of (social) strategies to be applied in a given scenario. In such environments, subcarrier selection policy, the amount of transmit power, time, space, network selection,

[1] Steve Jobs (1955-2011) is best known as the co-founder, chairman, and chief executive officer of Apple Inc.

security, etc. can be considered as set of strategies. It is obvious that such an implementation would be efficient if social rules (strategies) satisfy individual interests of end-users and network service provider.

Mobile broadband networks deployments span scenarios from very dense urban areas to remote rural areas, operating in a large range of carrier frequencies with different propagation characteristics and different coverage levels. With the ever growing demand of data applications, traditional cellular networks face the challenges of providing enhanced system capacity, extended cell coverage, and improved minimum throughput in a cost effective manner. The cost of the backhaul to the various network nodes has been identified as an essential cost component in the deployment of cellular networks, especially for future deployment of micro- and pico-cells for which the site acquisition cost is expected to be lower than that for macro-cells. Therefore, the ability to deploy network nodes not relying on a wired backhaul is an appealing option to reduce total network deployment cost and operating costs. *Co-operative communications* through wireless relay stations is a technology to enhance network throughput/coverage without incurring high site acquisition and backhaul costs. In order to provide economic coverage and higher capacity, new deployment schemes such as relay-based or multi hop systems are getting much attention in the research community.

LTE-A features relay system concepts, with the ultimate goal of designing a system that is drastically enhanced in both cell capacity and coverage. Relaying also provides higher deployment flexibility for operators by introducing low-power nodes without requiring additional wired backhaul. This new technique is already bringing significant performance benefits to Release LTE-A. These include increased capacity and spectral efficiency, improvement of throughput actually experienced by the user, fairness or throughput provision, reduction of cost per bit, and energy saving.

Main contributions

The main focus of this thesis is twofold: the applications of cooperative game theory to wireless communications and especially in OFDMA, and cooperative communications. Firstly, we review existing OFDMA resource allocation techniques in the literature. Then, we show that cooperative game theory is an appealing tool to tackle different problems in wireless engineering. The main focus of the first part is the application of cooperative games to resource management in the uplink of OFDMA infrastructure wireless networks. Due to the large number of subcarriers, there is a great freedom to resource assignment in a multiuser network. Low complexity, fairness from both end-user and wireless service provider viewpoint, and energy efficient approaches are the key requirements of the devised cooperative game-theoretic model.

The main contributions of the first part of this thesis are as follows:

- In our fairness approach, each wireless terminal achieves its own data rate demand exactly;
- two different approaches to assign different subcarriers among wireless terminals are described. Our model allows each subcarrier to be shared by more wireless terminals;
- existence of the solution of the cooperative resource allocation game, the core set solution, is proved by means of the analytical tools of coalitional game theory;
- a dynamic learning algorithm for reaching one of the core set solutions of the power expenditure scheme is derived, and its convergence is demonstrated based on Markov modeling;
- the performed simulations show that the derived framework outperforms the results available in the literature in terms of complexity, power consumption, and utilization of the spectrum.

The second part of this dissertation is focused on the computation of the capacity of Gaussian channels when assisted by one or multiple parallel relays.

Two relays are said to be parallel if there is no cooperation among themselves, while both have direct-link from the source(s) and to the destination(s). We present the information-theoretic perspective, and outer range capacity of different relayed communication scenarios, using the max-flow min-cut (cutset) theorem, and also applying known relaying protocols: amplify and forward (AF), decode and forward (DF), and compress and forward (CF). We show that choosing the best cooperative approach can increase significantly the network channel capacity.

The contributions of the second part of this thesis are as follows:

- applying cutset theorem in a relay assisted communication between a transmitter and an out of coverage destination. We calculate the best transmit power at the transmitter and the relay to achieve the highest end-to-end capacity;

- the outer range capacity in a relayed point-to-point communication applying different relay strategies is calculated;

- we explore the capacity of different relaying techniques with two case studies;

- the outer range capacity in a multiple parallel relayed point-to-point communication applying different relay strategies is calculated. Our upper bound capacity is lower than that in the existing literature;

- the outer range capacity of a network communication consisting of multiple sources, multiple parallel relays, and one destination is calculated. The upper bound capacity is lower than that in the existing literature;

- we describe a new information-theoretic perspective for CF technique based on distributed Wyner-Ziv (WZ) source coding;

- we show that using relays can significantly increase upper bound capacity conditional upon the relays positions;

- we show networks wherein adding a new relay or transmitter degrades the overall data rate.

Outline of the dissertation

The remainder of this thesis is organized in two parts. The first part is devoted to introduce a resource allocation technique in OFDMA based on cooperative game and is structured as follows:

In Chapter 1, after a brief introduction of OFDM(A) technology, we review the existing techniques in the literature for radio resource allocation in OFDM and its multi-user version (OFDMA), where the "resources" can be bandwidth and transmit power. We then discuss the relevant features of each technique and we identify why a new fairness approach and resource allocation in OFDMA is needed.

In Chapter 2, we introduce the basic concepts of cooperative game theory and we try to extend it in order to apply it in wireless networks. To this end, we provide motivating examples for the application of cooperative game theory to network engineering problems, and we outline the trends in research into cooperative game theory applications to wireless networks.

In Chapter 3, we introduce the resource allocation problem for OFDMA-based wireless networks as a cooperative game. Firstly, we propose two different subcarrier allocation techniques. We then formulate the power control scheme using a fairness approach. Next, we prove the existence of the core solution(s) of the proposed game, and we describe an iterative algorithm to reach one of the core sets in a centralized fashion. We conclude this chapter with comparing of the simulation results to existing solutions in the literature. This chapter is a revised and extended form of the following published papers:

1. F. Shams, G. Bacci, M. Luise, "A coalitional game-inspired algorithm for resource allocation in orthogonal frequency division multiple access," in *Proc. European Wireless Conference,*, Lucca, Italy, Apr. 2010.

2. F. Shams, G. Bacci, M. Luise, "Low complexity resource allocation

for OFDMA based on coalitional game theory," in *Proc. Int. Workshop Game Theory in Communication Networks (GameComm)*, Paris, France, May 2006.

3. F. Shams, G. Bacci, M. Luise, "An OFDMA resource allocation algorithm based on coalitional games," *EURASIP J. Wireless Communications and Networking*, vol. 2011, no. 1, 2011:46, July 2011.

The second part of this thesis is focused on the study of cooperative communications, and it is structured as follows:

In Chapter 4, we summarize a background literature that contributes historical notes on cooperative communications. We then introduce the system model of peer-to-peer relayed communication. We detail the mathematical model of different relaying schemes (i.e., amplify and forward, decode and forward, and compress and forward) and introduce a capacity upper bound of each of them.

In Chapter 5, we treat the upper bound capacity of a peer-to-peer communication with multiple parallel relays wherein there is no cooperation among relays. We will show that in such a reliable network, the upper bound capacity is achieved by either only one relay or by all relays together.

In Chapter 6, we extend the previous chapter and investigate a network consists of multiple sources, multiple parallel relays, and one destination. We introduce the upper bound channel capacity of different relaying strategies.

Finally, Chapter 7 summarizes our conclusions for this thesis and discuss open issues and further areas for these research fields.

In principle all the theorems without references are firstly proven by the author of this dissertation.

Part I

Cooperative Resource Allocation in OFDMA Wireless Communications

Chapter 1

A survey on resource allocation techniques in OFDM(A)

Interest in orthogonal frequency-division multiplexing (OFDM), has grow steadily, as it appears to be the most efficient air-interface for wireless communications primarily due to its inherent resistance to frequency-selective multipath fading and the flexibility it offers in radio resource allocations. One of the crucial issues in OFDM transmission is the allocation of the power resources to the available subchannels.

This chapter reviews the existing resource allocation techniques in OFDM and its multiuser version, OFDMA. We start with historical notes in the following section. Then, we provide a description of resource allocation issue in OFDM in Sect. 1.2. In Sect. 1.3, we describe different fashions of allocation of radio resources in OFDMA. We continue with exploration of three classic power allocation solutions of water-filling, max-min fairness, and weighted proportional fairness in Sects. 1.4,1.5, and 1.6, respectively. Sects. 1.7 and 1.8 discuss about two important resource allocation issues in multi service traffic networks: utility maximization, and cross layer. Different solutions based on game theory are reviewed in Sect. 1.9. Finally, we summarize key features of the existing solutions in Sect. 1.10.

1.1 OFDM(A) modulation and channel access

The principle of transmitting data by dividing it into several interleaved bit streams, and using these to modulate several carriers, is a concept that helps reducing the detrimental effects of multi path fading in communication systems. It was proposed for the first time by Doelz *et al.* for the U.S. military HF communication applications in 1957 in the pioneering Collins Kineplex system [3]. The contribution of Doelz *et al.* led to a few orthogonal frequency division multiplexing (OFDM) schemes in the 60s, which were proposed by Saltzberg [4] and Chang [5]. In the late 1960s, the multicarrier concept was adopted in some military applications such as KATHRYN [6] and ANDEFT [7]. These systems involved a large hardware complexity since parallel data transmission was essentially implemented through a bank of oscillators, each tuned to a specific subcarrier. In 1970, the first patent was granted on OFDM [8].

OFDM is a parallel transmission scheme, where a high-rate serial data stream is split up into a set of low-rate sub-streams with generally equal bandwidth, each of which is modulated on a separate subcarrier (called also sub-channel or tone). Thereby, the bandwidth of the subcarriers becomes small compared with the coherence bandwidth of the channel; that is, the individual subcarriers experience flat fading, which allows for simple equalization. This implies that the symbol period of the sub-streams is made long compared to the delay spread of the time-dispersive radio channel. While each subcarrier is separately modulated by a data symbol, the overall modulation operation across all the sub-channels (multicarrier modulation) results in a frequency multiplexed signal.

OFDM is extremely effective in a time dispersive environment where signals can have many paths to reach their destinations, resulting in variable time delays. With classical modulations, these time delays cause one symbol to interfere with the next one (inter symbol interference) at high bit-rates. The major contribution to OFDM scheme came after realization of the results of Weinstein and Ebert [9]. They demonstrated that using discrete Fourier transforms (DFT) to perform the baseband modulation and demodulation considerably increases the efficiency of modulation and demodulation processing. All of the sinc (sinus cardinal) shaped sub-channel spectra exhibit zero crossings at all of the remaining subcarrier frequencies and the individual sub-channel spectra are orthogonal to each other. The orthogonality among different tones ensures that the subcarrier signals do not interfere with each other, when communicating over perfectly distortionless channels.

While the problems of inter symbol interference (ISI) is mitigated by the guard space between consecutive OFDM symbols and the raised-cosine filtering OFDM imposes, it is not eliminated. The adoption of OFDM is facilitated by the efficient implementation of fast Fourier transform (FFT) and inverse FFT (IFFT) algorithms in DSP chips. To attain perfect orthogonality between subcarriers in a time dispersive channel, Peled and Ruiz [10] introduced the notion of cyclic prefix (CP). The guard space is filled with a cyclic extension of each time domain OFDM symbol, in order to overcome the inter-OFDM symbol interference due to the channel memory. The CP performs the circular convolution by the channel under the assumption that the channel impulse response is shorter than the length of the CP, thus preserving the orthogonality of subcarriers. Using a CP extension, OFDM exhibits a high resilience against the ISI. The CP has made OFDM both practical and attractive to the radio link designer. Although addition of the CP causes power/spectrum efficiency loss, this deficiency was more than compensated by the ease of receiver implementation.

The idea of transmission over dispersive parallel sub-channels using OFDM scheme has been of enduring interest ever since. OFDM has been adopted by many European and American telecommunication standards. In the context of wired environment, OFDM was applied for high speed digital voice services, e.g. asymmetric digital subscriber lines (ADSL) [11], and in wireless communication, OFDM technique is the fundamental building block of the IEEE 802.16 standards and it has been considered as a solution to mitigate multi path harmful effect in broadband multimedia broadcasting, e.g. digital video broadcasting for terrestrial television (DVB-T), digital audio broadcasting (DAB), 3G mobile communication (3GPP-LTE). This wide interest in OFDM technique is due the following advantages:

- High spectral efficiency.
- Interference suppression capability through the use of the CP.
- Protection against narrowband interference and inter carrier interference (ICI).
- Efficient implementation using FFT.
- Flexible spectrum adaptation.
- Separated subcarrier modulation. Different constellations can be applied on individual subcarriers, which allows numerous resource allocation strategies.

Even though the concept of multicarrier transmission is simple in its basic principle, the design of practical OFDM systems is far from being a trivial task. Synchronization, channel estimation and radio resource management are only a few examples of the numerous challenges related to multicarrier technology. As a result of continuous efforts of many researchers, most of these challenging issues have been studied and several solutions are currently available in the open literature. Besides its significant advantages, OFDM suffers from the following disadvantages:

- High peak-to-average power ratio (PAPR), which requires high linear amplifiers and consequently high power consumption [12].
- Sensitivity to Doppler effect and carrier frequency offset [13].
- Sensitivity to phase noise and time and frequency synchronization problems. Synchronization of an OFDM signal requires finding the symbol arrival time, and carrier frequency offset [14].
- Loss in data rate due to the guard interval insertion.

OFDM is also good from the standpoint of multiple access opportunities. Compared to single carrier systems, OFDM is a versatile modulation and channel access scheme for multiple access systems in that it intrinsically facilitates both time-division multiple access and frequency-division (or subcarrier-division) multiple access. In a multi user scenario the available bandwidth must be shared among several users. Each user may experience different conditions in terms of path loss, and shadowing. Furthermore, each may have different requirements in terms of quality of service. An acceptable design of the network should therefore take into account the different user conditions while providing fairness, without a drastic reduction in the overall spectral efficiency. A higher spectral efficiency is actually the main goal of all radio interface design. In 1998 a combination of OFDM and frequency division multiple access (FDMA) called orthogonal frequency-division multiple access (OFDMA) was proposed by Sari *et al.* for cable TV (CATV) networks [15].

OFDMA is a promising multiple access scheme that has attracted interest for wireless MANs. OFDMA is based on OFDM and inherits its immunity to ISI and frequency selective fading. Furthermore, in OFDMA systems different modulation schemes can be employed for different users. For instance, each user, according to its distance to base station, can invoke different order of modulation scheme (either high- or low-order modulation) to increase its data rate. Its implementational flexibility, the low complexity equalizer required in the transceiver, as well as the attainable high performance make OFDMA a highly attractive candidate for high data rate communications over time-varying frequency selective multi user radio channels. Compared to classic FDMA, OFDMA presents higher spectral efficiency by avoiding the need for large guard bands between users' signals. The main advantages of OFDMA are the increased flexibility in resource management and the ability for dynamic channel assignment. OFDMA can exploit channel state information to provide users with the best subcarriers (in terms of channel condition between transmitter and receiver over different subcarriers) that are available, thereby leading to remarkable gains in terms of achievable data throughput. Thanks to its favorable features, OFDMA is widely recognized as the technique for fourth generation broadband wireless networks. An ongoing research activity is currently devoted to study MIMO-OFDMA as promising candidates for 4G wireless broadband systems.

Today, it can now be implemented using powerful integrated circuits optimized for performing DFT. Because of its increasingly widespread acceptance as the modulation scheme of wireless networks of the future, it attracts a lot of research attention, in areas like resource allocation, time-domain equalization, PAPR reduction, phase noise mitigation and pulse shaping. We will concentrate on resource allocation techniques which includes subcarriers selection and power allocation. In multiuser environment, a good resource allocation scheme leverages multiuser diversity and channel fading.

1.2 Problem formulation in OFDM

The research interest on resource allocation in multicarrier systems was encouraged by the successfully development of ADSL services in the nineties. This technology employs a digital multitone (DMT) modulation for high-speed wireline data transmissions. Due to crosstalk from adjacent copper twisted pairs, the ADSL channel is characterized by strongly frequency-selective noise. One of the attractive features of OFDM scheme is flexibility to allocate individual power and modulation on different subcarriers. Different criteria can be performed depending upon whether the network is trying to maximize the overall data rate under a total power constraint or to minimize the overall transmit power given a fixed data rate or bit error rate (BER). The optimum adaptation algorithm, called the water-filling criterion [16], tends to allocate more information bits onto the highest signal-to-noise ratios (SNRs) carriers. Note that the number of bits determines the constellation size as: 1 bit corresponds to binary phase-shift keying (BPSK) modulation, 2 bits to quadrature phase-shift keying (QPSK) modulation, 4 bits to 16-quadrature amplitude modulation (16-QAM), and so on. In some situations a number of subcarriers may even be left unassigned if the respective SNR is too low for reliable data transmission.

In the literature, the problem of efficiently allocating information bits on the available sub-channels and using the best modulation is referred to as either *bit loading*, *adaptive modulation* or *link adaptation*. In OFDM communication, the unique transmitter k spending power $p_{kn} \leq \overline{p}_{kn}$ over the nth subcarrier, achieves the number of bits R_{kn} that is calculated by the Shannon channel capacity formula as [17, Eq. 1]:

$$R_{kn} = \log_2 \left(1 + \frac{p_{kn} \left| H_{kn} \right|^2}{(\gamma_{margin} + \Gamma) \cdot \sigma_w^2} \right) \tag{1.1}$$

where $\left| H_{kn} \right|^2$ is the amplitude of the Gaussian complex frequency response (channel condition) of subcarrier n and σ_w^2 is the noise power spectral density on the subcarrier. The performance margin γ_{margin} is an additional amount of noise (in dB) that the system can tolerate, even if the noise level is increased by a factor of γ_{margin}. The parameter Γ, called as the SNR gap (also known as the normalized SNR), is used to evaluate the relative performance of a modulation scheme versus the theoretical capacity of the channel [18]. Therefore, by increasing the value of γ_{margin} we can improve the system robustness against noise, and hence have the new operating point of the constellations at a distance of $(\gamma_{margin} + \Gamma)$ dB from the Shannon limit. There are many theoretical works whose results show a sub-optimal adaptive bit loading. In the following we cite some of the pioneering and well-known bit loading algorithms in OFDM.

- Hughes-Hartogs in 1987 [19] designed a greedy algorithm for approximating the water-filling (e.g. see [16]) for twisted-pair channels over an additive white Gaussian noise (AWGN) channel with ISI. The goal of this discrete loading algorithm is minimization of the transmit power under a BER and data rate constraints for each tone. It accomplishes this end by successively assigning bits to carriers, each time choosing the carrier that requires the least incremental power, until the given target rate is reached. Bingham in [20] proposes to apply sinc functions for each individual spectra instead of using quadrature amplitude shift keying (QASK) in [19]. Applying this technique make separating signals at receiver easier using efficient FFT techniques. The complexity burden of the proposed algorithm is very expensive to be implemented in practice, specifically in high speed wireless networks.

- The principle of adaptive modulation and power over OFDM was recognized in 1989 by I. Kalet [21]. Kalet simulates a twisted pair OFDM modulation and power control in which for each subcarrier uses QAM to maximize the bit rate. Total transmit power division is based on water-filling solution under the assumption that all subcarriers have the same BER. Furthermore, multi carrier QAM performance is about 9 dB worse than the channel capacity, independent of the channel response.

- Chow *et al.* in [17] propose an iterative bit loading algorithm which offers a significant applicability advantages over the earlier Hughes-Hartogs algorithm [19] and water-filling method. The simulation results of ADSL service show a degradation of 1.3 dB SNR only. Even though the proposed algorithm is faster than that Hughes-Hartogs, it is not optimal from number of iterations and computational load points of view.

- A. Czylwik [22] in 1996 simulates an OFDM transmission system with time-variant channel functions measured with a wideband channel sounder composed of fixed carrier frequency antennas. The simulation results show that with the proposed subcarrier adaptive modulation the required signal power at BER $= 10^{-3}$ can be reduced by $5 \div 15$ dB depends upon SNR and propagation scenario. Different modulation formats of BPSK, QAM and no modulation can be selected such that the BER is minimized under a constant data rate constraint.

- Fischer *et al.* in 1996 [23] propose a bit loading algorithm to reduce the computational complexity of Hughes-Hartogs and Chow algorithms. This algorithm distributes bits and transmit power in order to maximize the SNR in each carrier. Van-der Perre *et al.* in [24] apply Fischer's algorithm to simulate the performance of OFDM-based high speed wireless LAN. The simulation results show that the proposed adaptive loading strategy improves the system performance considerably, as an SNR gain of 14 dB, under BER $= 10^{-3}$ constraint and using 4-QAM modulation.

All link adaptation studies have demonstrated that a performance improvement is attained in OFDM system applying individual power and data rate adjusting over each subcarrier in order to exploit channel frequency selectivity.

1.3 Taxonomy of problem types in OFDMA

A typical case of multiple access channel is the uplink of a cellular system. In general, in the OFDMA uplink scenario, each user will receive a channel assignment and power allocation from the base station (BS) that consists of a (usually exclusive) subset of subcarriers and power levels on each of them, that make up an OFDM symbol in the available frequency band. In an OFDMA network, the BS has to optimally allocate power and bits over different subcarriers based on instantaneous channel conditions of different active wireless terminals. The only requirement is that the fading rate is not too fast, as instantaneous resource allocation is hardly usable in the presence of rapidly-varying transmission channels of mobile terminals. Other difficulties with wireless terminals include interference effects and limited media resource such as bandwidth and transmission power. This makes the link adaptation task much more challenging than in single-user systems. It is clear that, in comparison to a point-to-point, single-user OFDM-based connection, a multi user OFDMA link adaptation is much more complicated and hardly scalable [25]. Now, we start to review the main concepts behind bit and power loading in OFDMA-based transmissions.

Generally, a resource allocation algorithm can either be *centralized* or *distributed*. In centralized schemes like [26, 27], the algorithm is run by a central unit (like the radio base station) that is aware of the channel conditions and the demands of all mobile terminals. In a distributed model (such as [28]), each mobile terminal tries to accomplish its own (minimum) QoS autonomously. In general, centralized techniques show better performance at the expense of a higher signaling between terminals and central unit, and lower scalability. In the context of distributed algorithms, several cross-layer approaches [29] (discussed in Sect. 1.8) to reduce the total power consumption and to support different services and traffic classes, mostly for the downlink of an OFDMA system. Maximizing the power efficiency in uplink OFDMA has also been tackled in [27, 30, 31] using different formulations for the joint resource allocation problem.

One of the main problems with multiuser OFDM is the large amount of feedback required from the users when applying centralized schemes. Since different users can be scheduled on different frequency subcarriers, users must feed back measurement information about *every* subcarrier. Consider a network with K active OFDM mobile terminals and N overall available subcarriers. The scheduler requires channel state information about the different users. The full channel state information consists of $K \cdot N$ complex numbers (the values of the channel frequency response at each subcarrier for every user). This feedback information represents a very large overhead if there are many users and subcarriers in the system. Cimini *et al.* [32] propose to group adjacent subcarriers into (so-called) clusters and feedback the information about the best cluster(s) in terms of channel quality. In [33, 34], it is shown that sending back only heavily quantized channel status information dramatically reduces the feedback needs without sacrificing the essential of the scheme performance. P. Svedman *et al.* [35] show that the cluster-size is an important design parameter to achieve a good downlink throughput. A suitable cluster-size depends upon the average channel delay spread of the users. Since it is compulsory to use the same cluster-size for different users, an arrangement that attains good performance for most expected channels is a suitable choice.

The problem of subcarrier and power assignment in OFDMA has been extensively considered in the literature during the last few years. The proposed solutions mainly fall into two different categories: *margin-adaptive* and *rate-adaptive* methods. The goal of margin adaptive schemes [36] is to minimize the total transmit power expenditure given a set of fixed user data rates and BER requirements. Algorithms based on the rate-adaptive criterion [37] aim on the contrary at achieving the maximum total sum (continuous) rate over all users subject to different QoS constraints, e.g. power expenditure. In addition, in some broadband systems e.g. code division multiple access (CDMA), ultra-wideband (UWB), and multi carrier CDMA (MC-CDMA)[1], a resource allocation strategy, the so-called *mean BER minimization*, was studied where the robustness of the system is enhanced by allocating bits and powers to subcarriers in such a way that the error rate of an entire symbol is minimized for a given target bit rate. This scheme is not a major interest of OFDMA systems, since in (almost) all OFDMA resource allocation techniques, as it will be seen in the next sections, each subcarrier is not permitted to be assigned to more users. This means that a well devised algorithm for total data rate maximization also results in minimization of users' BER.

The first resource allocation strategy that is formulated here is the minimization of the OFDM system power expenditure for a given target data rate. The margin-adaptive optimization problem is formulated as:

$$\begin{aligned} &\min_{\{\mathbf{p},\,\mathcal{N}\}} \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} p_{kn} \\ &\text{s.t.} \quad \sum_{n\in\mathcal{N}_k} R_{kn} \geq \underline{R}_k \qquad \forall\, k \in \mathcal{K} \\ &\text{and} \quad \sum_{n\in\mathcal{N}_k} p_{kn} \leq \overline{p}_k \qquad \forall\, k \in \mathcal{K} \\ &\text{and} \quad \mathcal{N}_k \cap \mathcal{N}_{m\neq k} = \emptyset \qquad \forall\, k,\, m \in \mathcal{K} \end{aligned} \tag{1.2}$$

wherein $k \in \mathcal{K} = [1, \ldots, K]$ denotes the index of the wireless terminal which transmits with power $\mathbf{p}_k = \left[p_{k1}, \ldots, p_{kn}, \ldots, p_{kN}\right]$ over subcarriers which are

[1]Unlike OFDMA, in the CDMA, UWB, and MC-CDMA systems, whole bandwidth is shared by all active wireless terminals.

represented by the set of $\mathcal{N} = [1, \ldots, n, \ldots, N]$. Let $\mathcal{N}_k \subset \mathcal{N}$ be the set of subcarriers assigned to user k, and R_{kn} the achieved channel capacity by user k over the nth subcarrier. The sets of assigned subcarriers are disjoint and this means that each subcarrier is not allowed to be shared by more than one terminal. Each user k wishes to attain its target rate $\underline{R}_k$, under the constraint of $\overline{p}_k$ as individual total transmit power. It is clear that for each terminal k and every $n \notin \mathcal{N}_k$, we have $p_{kn} = 0$, and accordingly $R_{kn} = 0$. The overall data rate of each user is obtained by the Shannon capacity formula as:

$$R_k = \sum_{n \in \mathcal{N}_k} R_{kn} = \sum_{n \in \mathcal{N}_k} \log_2 \left(1 + \frac{p_{kn} \left| H_{kn} \right|^2}{\sigma_w^2} \right) \tag{1.3}$$

wherein $\left| H_{kn} \right|^2$ denotes the amplitude of the Gaussian complex path gain of user k on subcarrier n. The parameter σ_w^2 is the variance of the AWGN zero mean Gaussian noise on each subcarrier. Two levels of decomposition are necessary to turn this NP-hard problem into the set of sub-problems of [38] subcarrier allocation and power control. In fact, the exclusive assignment of subcarriers to users is a way to reduce the complexity computation of optimization equation. On the other hand, as users are not allowed to share a common subcarrier, the allocation process boils down to a combinatorial optimization problem for which no optimal greedy solution exists. Kivanc *et al.* [39] develop a computationally inexpensive method for OFDMA resource assignment which achieves a comparable performance at reduced computational complexity. However, this approach did not provide a fair opportunity so that some users may be dominant in resource occupancy even when the minimum rate requirement is not satisfied for the others.

The most common optimization problem for OFDMA systems is the rate-adaptive or maximization of bit rate which is instead formulated as:

$$\begin{aligned}
&\max_{\{\mathbf{p},\,\mathcal{N}\}} \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} R_{kn} \\
&\text{s.t.} \quad \sum_{n\in\mathcal{N}_k} R_{kn} \geq \underline{R}_k \qquad \forall\, k \in \mathcal{K} \\
&\text{and} \quad \sum_{n\in\mathcal{N}_k} p_{kn} \leq \overline{p}_k \qquad \forall\, k \in \mathcal{K} \\
&\text{and} \;\; \mathcal{N}_k \cap \mathcal{N}_{m\neq k} = \emptyset \qquad \forall\, k,\, m \in \mathcal{K}
\end{aligned} \tag{1.4}$$

The objective of this problem is to distribute bits and power among different subcarrier in such a way that the overall data rate of the system is maximized. Most algorithms focus on the downlink scenario, with constraints on the total power transmitted by the radio base station. In the uplink scenario, the restrictions apply on an individual basis to each user terminal, and the simplest solution to maximize channel capacity of mobile devices under a power constraint is the water-filling (WF) criterion [40].

1.4 Water-filling solution

Cheng and Verdù in [41] pioneered to apply WF solution in an uplink OFDMA network scenario and derive capacity region and the optimal power allocation of individual users. In rate-adaptive optimization, the channel capacity is obtained by maximizing the right-hand-side of Eq. 1.3 with respect to $\sum_{n\in\mathcal{N}_k} p_{kn} \leq \overline{p}_k \; \forall\, k \in \mathcal{K}$, and assigned the channel gains $|H_{kn}|^2$, i.e.

$$\max_{\{\mathbf{p}\,,\mathcal{N}\}} \left\{ \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} \log_2 \left(1 + \frac{p_{kn}\,|H_{kn}|^2}{\sigma_w^2} \right) \right\} \tag{1.5}$$

Since the objective function in (1.5) is convex in the variables $\{\mathbf{p}_k\}$, the optimum power allocation under the convex constraints of overall transmit power can be found using Lagrangian methods. The optimal strategy to (1.4) is such that the user with the largest channel gain is first selected, i.e., on each subcarrier $n \in \mathcal{N}$ select the best terminal as:

$$k \longleftarrow \arg\max_{k \in \mathcal{K}} |H_{kn}|^2 \tag{1.6}$$

The resulting optimal power allocation for user k is given by:

$$p_{kn} = \left[\frac{1}{\lambda_k \cdot \ln 2} - \frac{\sigma_w^2}{|H_{kn}|^2} \right]^+ \tag{1.7}$$

where $[x]^+ = \max\{x, 0\}$, and the Lagrangian parameter λ_k ("water-level") is chosen such that the sum of the allocated powers satisfies the total power constraint $\overline{p}_k$.

$$\lambda_k = \frac{|\mathcal{N}_k|}{\ln 2} \cdot \left(\overline{p}_k + \sum_{n \in \mathcal{N}_k} \frac{\sigma_w^2}{|H_{kn}|^2} \right)^{-1} \tag{1.8}$$

WF is a greedy power allocation scheme with which channel capacity is increased when every subcarrier is assigned to the user with the best path gain, and the power is distributed according to the WF criterion of Eq. 1.7. However, the WF solution is highly unfair, since only users with the best channel gains receive an acceptable channel capacity, while users with bad channel conditions achieve very low data rates. More information-theoretic discussions on related topics can be found in [42]. To derive fair resource allocation schemes, we resort to other techniques, described in the following.

1.5 Max-min fairness criterion

In a multiuser OFDM network, one possible approach to overcome the unfairness of WF is described in [43]. This alternative formulation aims at the maximization of the minimum user's data rate. This enforced a notion of *max-min rate-maximization fairness*, so that the starvation of some users can be avoided.

Definition 1 *A feasible flow rate vector* $\mathbf{R} = [R_1, \ldots, R_k, \ldots, R_K]$ *is defined to be max-min fair if any rate* R_k *cannot be increased without decreasing some* $R_{m \neq k}$*, which is smaller than or equal to* R_k*.*

In the max-min power control, roughly speaking, the objective is to optimize the performance of the worst link amongst all users for fixed QoS-based power control approaches. The idea behind the max-min fair approach is to treat all users as fairly as possible by making all rates as large as possible [44]. The work of Rhee *et al.* in [43] is an extension of [45] which is a dual problem of minimizing total transmit power for given data rate requirements. The given problem is formulated as the following convex optimization problem [43]:

$$\begin{aligned} &\max_{\{\mathbf{p},\,\mathcal{N}\}} \min_{k\in\mathcal{K}} \sum_{n\in\mathcal{N}_k} R_{kn} \\ &\text{s.t.} \quad \sum_{n\in\mathcal{N}_k} p_{kn} \le \overline{p}_k \qquad \forall\, k \in \mathcal{K} \\ &\text{and} \quad \mathcal{N}_k \cap \mathcal{N}_{m\neq k} = \emptyset \qquad \forall\, k,\, m \in \mathcal{K} \end{aligned} \tag{1.9}$$

The Lagrangian relaxation algorithm proposed in [43, 45] approaches the solution by slowly increasing the power level for each user. In contrast to WF result, the rationale behind max-min fairness solution is to assign more power to users exhibiting poor channel conditions so that they can achieve a data rate comparable to that of other users with better channel quality. It is worthwhile to note that the max-min fair rate allocation is unique when the number of resources and flows, i.e., subcarriers and of wireless terminals, are both finite [44]. Unfortunately, due to the nonlinear nature of the integer problem, the algorithm proposed in [43,45] is computationally very expensive.

In the paper [43, Eq. 2], formulation (1.9) is also extended to:

$$\begin{aligned} &\max_{\{\mathbf{p},\,\mathcal{N}\}} \min_{k\in\mathcal{K}} \sum_{n\in\mathcal{N}_k} t_{kn}\, R_{kn} \\ &\text{s.t.} \quad \sum_{k=1}^{K} t_{kn} \le 1 \qquad \forall\, n \in \mathcal{N} \\ &\text{and} \quad \sum_{n\in\mathcal{N}_k} p_{kn} \le \overline{p}_k \qquad \forall\, k \in \mathcal{K} \end{aligned} \tag{1.10}$$

wherein the positive coefficient $t_{kn} \in [0,\, 1]$ introduces the percentage of time each subcarrier is used by a given user. With parameter t_{kn}, each sub-channel can be shared by different users in TDMA fashion. Clearly, the assumption behind this approach is that the users' channel responses do not change significantly over a timing interval. However, in the simulation results of [43], it is assumed that $K \ll N$ and no sub-channel is shared among users, i.e. t_{kn} is a binary value and $\sum_{k\in\mathcal{K}} t_{kn} = 1 \;\forall\, n \in \mathcal{N}$, or equivalently $\mathcal{N}_k \cap \mathcal{N}_{m\neq k} = \emptyset$. In addition, determining the best values for $t_{kn} \in (0, 1)$ and indicating a time-sharing allocation policy is not always feasible for $K > N$, as reported in [46]. Here we stick to the original problem formulation (1.9).

B. Bertsekas *et al.* in [47, p. 527] propose a simple iterative algorithm for computing a sub-optimal max-min fair rate vector which is simply extendable for an OFDMA network as:

1) *Zero Initialization:* Supposing $K \ll N$, the algorithm starts with an all zero data rate vector: $R_k = 0$ and $\mathcal{N}_k = \emptyset \quad \forall\, k \in \mathcal{K}$.

2) *Round-robin fashion at once:* Assign every user $k \in \mathcal{K}$ the subcarrier n whose channel gain $|H_{kn}|^2$ is the best for it, and uniformly increase the respective data rate as:

$$n \longleftarrow \arg\max_{n\in\mathcal{N}} |H_{kn}|^2;$$
$$\mathcal{N}_k = \mathcal{N}_k \cup \{n\}; \qquad \mathcal{N} \leftarrow \mathcal{N} \setminus \{n\};$$
$$R_k = R_{kn}^{1/N} = \log_2\left(1 + \frac{\frac{\overline{p}_k}{N} |H_{kn}|^2}{\sigma_w^2}\right)$$

At this point, every user $k \in \mathcal{K}$ is assigned to one subcarrier.

3) *Best user rate updating:* Find the user k with the smallest attained data rate, i.e. $k \leftarrow \arg\min_{k\in\mathcal{K}} R_k$, and then assign the subcarrier $n \in \mathcal{N}$ with the best channel condition $|H_{kn}|^2$, and update the respective data rate as:

$$k \leftarrow \arg\min_{k\in\mathcal{K}} R_k;$$
$$n \longleftarrow \arg\max_{n\in\mathcal{N}} |H_{kn}|^2;$$
$$\mathcal{N}_k = \mathcal{N}_k \cup \{n\}; \qquad \mathcal{N} \leftarrow \mathcal{N} \setminus \{n\};$$

$$R_k = R_k + R_{kn}^{1/N}$$

4) *Exit condition:* If there exists some unassigned subcarrier then go to step (3), else exit.

In addition, to achieve a max-min fairness data rate vector, F. Kelly [48] suggests the problem formulation as:

$$\max_{\{\mathbf{p},\mathcal{N}\}} \sum_{k=1}^{K} -\left(-\log_2\left(\frac{R_k}{\beta}\right)\right)^{\alpha} \tag{1.11}$$

wherein $\alpha > 1$ is a constant parameter and β is a positive number satisfying: $R_k < \beta \ll \infty \ \forall\ k \in \mathcal{K}$. Furthermore, [48] defines the following alternative criterion instead of choosing the best user with the smallest data rate in step (3):

$$k \longleftarrow \arg\max_{k\in\mathcal{K}} \left\{ \frac{\alpha}{R_k} \cdot \left(\log_2\left(\frac{\beta}{R_k}\right)\right)^{\alpha-1} \cdot R_{kn}^{1/N} \right\} \tag{1.12}$$

For $\alpha \to \infty$, the condition became:

$$k \longleftarrow \arg\min_{k\in\mathcal{K}} R_k \tag{1.13}$$

which coincides with the original strategy of max-min fairness to allocate a subcarrier to the user with the minimum achieved data rate.

Although in the max-min criterion is attained at the cost of a reduction of the overall throughput of the cell. Although in the max-min criterion all users are treated fairly, this solution cannot practically be used. This is because, in general, the number of allocated bits may not correspond to any practical modulation scheme [49]. Furthermore, the results show that under the max-min fair solution, some users may consume significantly more bandwidth than others [50].

1.6 Weighted proportional fairness criterion

Achieving traffic fairness and efficiency are two conflicting goals. Hence, the optimization of the radio resource utilization tends to penalize terminals with low SINR, independently of their traffic level performance. Max-min fairness scheme is inappropriate when different users have different priorities. Generally, the problem is how to balance between fairness and the utilization of resources. This led F. Kelly *et al.* to formulate in [51] the notion of *weighted proportional fairness*. Under a proportional maximization rate constraint, the rate of each user should adhere to a set of predetermined proportionality constants which make a concrete way of assigning priorities to the users.

$$R_1 : \cdots : R_k : \cdots : R_K = \varphi_1 : \cdots : \varphi_k : \cdots : \varphi_K \tag{1.14}$$

wherein φ_ks are proportion constants.

Definition 2 *A vector data rate* $\mathbf{R} = [R_1, \ldots, R_K]$ *is proportional fair if it is feasible and for any other feasible rate vector* $\mathbf{R}' = \left[R_1', \ldots, R_K'\right]$ *the aggregate of proportional change is non positive:*

$$\sum_{k=1}^{K} \varphi_k \frac{R_k' - R_k}{R_k} \leq 0. \tag{1.15}$$

This method is also useful for service level differentiation, which allows for flexible allocation mechanisms for different classes of users with separable constraints. The proportional fair objective of (1.15) is continuously differentiable, monotonically increasing and strictly concave, therewith admitting a convex optimization formulation. F. Kelly *et al.* [51] suggested an algorithm that converges to the proportionally fair rate vector. It is shown that considering the Kuhn–Tucker conditions for problem (1.4), proportional fair resource assignment is a maximizer of the sum of the logarithms of the long-run average data rates provided to the users. They show that to achieve proportional fairness rate allocation, the problem formulation should be:

$$\max_{\{\mathbf{p}, \mathcal{N}\}} \sum_{k=1}^{K} \varphi_k \log_2 (R_k) \tag{1.16}$$

over all feasible rate allocations. Thus, since a logarithm function is strictly concave, it may be derived that proportional fair rates are unique [52, Sec.

6.7]. The logarithmic utility function indicates that users with low average rates benefit more in utility from being scheduled than users with high average rates. The iterative algorithm for computing proportionally max-min fair rate vectors is like that for max-min fairness, except for choosing of the best user in step (3) that follows the following criterion instead:

$$k \longleftarrow \arg\max_{k \in \mathcal{K}} \frac{R_{kn}^{1/N}}{R_k} \tag{1.17}$$

wherein $R_{kn}^{1/N}$ is the same formula in step (2) in the max-min fairness algorithm. This is basically the maximum throughput scheduling, but weighted by the inverse of the achieved data rate. If a user has been given channel access recently and therefore has a high past throughput, it will need a very high channel quality $R_{kn}^{1/N}$ to be scheduled. This gives a compromise between multiuser diversity and fairness.

The time-update of the historical data rate R_k can be done in different ways. A low complexity update equation that also has low memory requirements is defined in [52, Sec. 6.7] which keeps track of the average throughput R_k of each user in an exponentially weighted time-window of length t_c as:

$$\begin{cases} R_k = \left(1 - \frac{1}{t_c}\right) R_k + \frac{1}{t_c} R_{kn}^{1/N} & m = k; \\ R_k = \left(1 - \frac{1}{t_c}\right) R_k & m \neq k. \end{cases} \tag{1.18}$$

wherein user k is the best preferred user for the next updating round, and m is the selected user for the current round. The update Eq. 1.18 is an exponentially weighted filter that includes all historical rates in the average rate. With a very large time-scale t_c, the rate updating (1.18), is equivalent to maximization problem of the Eq. 1.16 [52, Sec. 6.7].

For OFDMA, some other instantaneous sum-rate maximization methods with proportional rate constraints have been studied previously in [53–55]. The main emphasis of these papers, in terms of formulation, is the maximization of the data rates with instantaneous proportional rate constraints, exclusive subcarrier assignment, and given limited total transmit power. These works propose integer programming solution methods with time complexity (the number of time steps in the iterative algorithm) of $\mathcal{O}(NK\log_2 N)$ or higher. The notion of weighted proportional fairness has been extended by J. Mo *et al.* in [56]. They observed some particular TCP network traffics with which the total throughput of weighted proportional fairness is not optimum in terms of spectral efficiency, however it outperforms max-min fair solution. Let α be a non negative constant and $\boldsymbol{\varphi} = [\varphi_1, \ldots, \varphi_K]$ be a positive weight vector. A vector of $\mathbf{R}$ is said to be $(\boldsymbol{\varphi}, \alpha)$ proportionally fair if:

Definition 3 *A vector data rate* $\mathbf{R} = [R_1, \ldots, R_K]$ *is* $(\boldsymbol{\varphi}, \alpha)$ *proportional fair if it is feasible and for any other feasible rate vector* $\mathbf{R}' = \left[R_1', \ldots, R_K'\right]$*:*

$$\sum_{k=1}^{K} \varphi_k \frac{R_k' - R_k}{{R_k}^{\alpha}} \leq 0. \tag{1.19}$$

Obviously, if $\alpha = 1$, Def. 3 reduces to the weighted proportional fairness definition (2) and as $\alpha \to \infty$ it approaches to that of the max-min rate vector [56, Lemma 3]. In other words, this generalization includes arbitrarily close approximation of max-min fairness. Unfortunately, the challenge of choosing the best value of α makes this framework (almost) impractical. Further examination clarifies that $(\boldsymbol{\varphi}, \alpha > 1)$ proportional fairness maximizes [56, Lemma 2]:

$$\sum_{k=1}^{K} \varphi_k (1-\alpha)^{-1} {R_k}^{1-\alpha} \tag{1.20}$$

over all feasible data rate vectors.

Mathematically, Eq. 1.19 is a twice continuously differentiable and strictly concave function. Algorithms for computing $(\boldsymbol{\varphi}, \alpha)$ proportionally fair rates have been developed in [56]. Here, each transmitter adapts its window size based on the total delay. The main drawback of the proportionally fair rate allocation is that utility (maximization) functions are commonly assumed as concave. Lee *et al.* [57] showed that if the above mentioned algorithms developed for concave utility functions are applied to non concave utility functions, the system can be unstable and can cause excessive congestion in the network. Since the rate adaptive functions of some real-time applications are not concave [58], those applications cannot be dealt with in this system. For example, a multimedia communication user corresponds to a S-shaped[2] utility function [59] because his contentment is at a maximum if the allowed data rate is larger than the encoding rate, and is at a minimum if the allowed rate is smaller than the encoding rate.

1.7 Utility maximization scheme

Max-min fairness and weighted proportional fairness consider a same QoS requirements among network users with a strictly concave rate adaptive function. In some systems, e.g. real-time applications, the rate maximization functions are not concave. Furthermore, the above mentioned contexts are not able to formulated real-time constraints requirement, e.g. delay. In the seminal work [60], Cao *et al.* overcome these disadvantages of the above mentioned schemes. They introduce *utility maximization* (application layer performance) whose aim is providing individual QoS requirements for each user with a non necessarily concave function for rate maximization. A utility function is used to mathematically describe the QoS characteristics of an application. Utility maximization guarantees the application specific demand which can be characterized by bandwidth, delay and delay jitter or time spent to complete data deliveries. The drawback in [60] is a high delay in the communication network among users. Cao *et al.* in [61] extend their previous paper to overcome the limitation of the delay and propose a control-theoretic utility max-min flow control algorithm which solves the problems of [60], and showed that the algorithm converges to a utility max-min fair rate vector by using Dewey and Jury's stability criterion [62].

The utility maximization fairness scheme can be formulated in different ways according to the goal of the system. For instance, power control for optimal

[2] An increasing function is S-shaped if there is a point below which the function is convex, and above which the function is concave.

SNR assignment is formulated as:

$$\begin{aligned}
&\max_{\{\mathbf{p},\,\mathcal{N}\}} \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} U_k(\gamma_{kn}) \\
&\text{and} \quad \sum_{n\in\mathcal{N}_k} p_{kn} \leq \overline{p}_k \qquad \forall\, k \in \mathcal{K} \\
&\text{and} \quad \mathcal{N}_k \cap \mathcal{N}_{m\neq k} = \emptyset \qquad \forall\, k,\, m \in \mathcal{K}
\end{aligned} \tag{1.21}$$

wherein γ_{kn} denotes the SNR on the nth carrier of the user k at the BS. $U_k(.)$ is an individual maximization function of the user k. The maximization function can be represented as a greedy function for each user:

$$\max_{\{\mathbf{p}_k,\,\mathcal{N}_k\}} \sum_{n\in\mathcal{N}_k} U_k(\gamma_{kn}) \qquad \forall\, k \in \mathcal{K} \tag{1.22}$$

C. Zhou *et al.* [63] introduce a new scheduling and resource allocation problem in an OFDMA system wherein an approach based on utility functions results in a discrete optimization problem with non-differentiable non-convex objective with minimum data rate constraint. The idea is to transform the discrete problem in a suitable weighted max-min fairness problem which is easy to implement. Authors of [64] present a general utility-based framework for joint uplink/downlink optimization where user overall satisfaction is modeled by two different utility functions; one for the uplink and another for the downlink direction. The resource allocation is formulated as a maximization problem with an objective based on the sessions' utility functions and allocation probabilities as scheduling constraints that are solved via dual optimization techniques.

Even though the utility maximization approach has made advances in dealing with congestion control and resource allocation, it also exhibits a serious limitations. As already mentioned, there exists a tradeoff between average throughput and fairness in the system. Sometimes there exists a conflict between the QoS balance and the utility maximization. If users select utility functions based on their real QoS requirements, then the optimal achieved data rate may result in a totally unfair resource allocation within the network.

Applying advanced optimization methods of geometric programming [65], and majorization theory [66] may achieve an admissible tradeoff between fairness and overall throughput [67, 68].

1.8 Cross-layer scheme

So far, we talked about the design of an OFDM(A) system based on classic link-level approach. The wireless link-level primarily addresses two challenges that arise from the physical medium: channel fading and multiple access interference (MAI). Advances in link design for wireless channels have led to different modulation and channel coding schemes that provide increased robustness to MAI and multi path and thereby, enhance link-level or radio band capacity. While OFDM(A) provides a powerful physical layer engine for broadband communications, applying it without thorough application level considerations may lead to disappointing results. In high speed data networks, traffic is highly diverse with distinct QoS parameters; e.g. channel conditions may vary dramatically over a short time scale. However, the traditional decoupled layer design cannot meet such requirements. For instance, if the MAC layer does not interact with upper layers, it cannot obtain information regarding the type of service and the associated QoS parameters. As a consequence, MAC has no ability to adjust itself to the changing characteristics of traffic.

An OFDMA radio allocation module can be designed to be both channel-aware and application-aware through *cross-layer* interactions [69]. Cross-layer solutions break the traditional layered paradigm of communication since they rely on the concept of joint optimization across multiple layers. The cross-layer approach allows different layers to be grouped and/or assumes the existence of protocols that work with more than one layer, thus optimizing the protocol stack. With cross-layer techniques, decision making can be more accurate, bringing forth several benefits to the performance of the proposals that use such technique.

In the context of cross-layer design, joint scheduling-routing-flow control algorithms have been proposed and shown to achieve utility maximization approach while guaranteeing network stability. Sometimes the differences between cross-layer and utility maximization fairness schemes blur away since cross-layer schemes may require to improve their performance applying a non concave utility function, and utility maximization functions may consist of the parameters of different layers. Cross-layer multiuser techniques maximize the rate delivered on the radio channel, guarantee a fair allocation of resources among users belonging to the same traffic class, consider the dynamics of traffic sources by limiting the delay of data packets in the queues, and help to maximize QoS at the application layer. The common idea behind cross-layer schemes is to use properly maintained packet queues to make dynamic decisions about new packet transmission as well as rate allocation. Some pioneering works in the field of resource allocation in OFDMA using cross-layer design have appeared in [70–72].

Papers [29, 73, 74] propose some feasible solutions to resource allocation problems in downlink multiuser OFDM system with the goal not only to maximize system throughput under QoS guarantee, but also to reduce computational complexity. These works use a number of simplifications to find out the optimal frequency and power distribution for a given set of Lagrangian relaxations. The results show that the users with very low SINR achieve a good performance. However, in the case where the channel conditions and QoS requests vary significantly between successive frames, a new set of multiplier values must be found in each frame. The necessity of getting best values for Lagrangian multipliers may reveal an impractical issue.

1.9 Game theory solution

In the utility maximization and cross-layer schemes, different utility functions apply for different users. Sometimes the interests of wireless terminals are not aligned and they compete for the scarce wireless resources, bandwidth and power. Each user's interest could be also in conflict with others. In this situation, the wireless terminals can decide to behave in altruistic or egoistic manner. In both cases, the related problems can be formulated applying *game theory*. Depending on the interaction rules, there exist various types of games. For instance, if the users are allowed to exchange their proper interests and information before the game in order to form coalitions and coordinate their actions, the game is said to be *cooperative* (*coalitional game theory*). If coordination is not present, the game is said to be *non-cooperative*. The players act according to their *strategies*. The strategy of a player can be a single move or a set of moves during the game. For games in wireless communication , each transmitter represent a player whose strategy space covers the choices of modulation level, coding rate, transmit power, transmission frequency, etc. Another factor that identify different types of games is the number of time that users interact. If they play the game over multiple rounds the game is said to a *repeated game*. Contexts where the users only interact once are referred to as *static games* [75, 76].

In a multiuser OFDM network, there are multiple interacting users which use a fraction of the whole bandwidth and they must also decide the amount of transmit power on each subcarrier taking into account that the decisions and interests of each mobile terminal that affect the others. The resource allocation problem in OFDMA can be analyzed within the framework of game theory. In network resource allocation one of the challenges is to achieve a *Pareto optimal* rate vector. A rate allocation vector is Pareto optimal when there exists no other rate allocation that leads to higher performance for some users without degrading the performance of some other users.

Non-cooperative game theory has been vastly applied to wireless communication problems, and much progress has been made on distributed power control in Gaussian interference channels. In [77], Wu *et al.* investigate a joint power and exclusive subcarrier assignment scheme in single cell uplink OFDMA systems based on non-cooperative game theory. The aim of the utility function is to maximize the rate-sum capacity with the minimum power. They prove the existence and uniqueness of the Nash equilibrium

(NE) point. Instead, Yu *et al.* [28] apply a different convex utility functions in the same scenario whose objective is maximization of the power efficiency. In the utility function, a (transmit) power pricing factor (multiplier) [78] is introduced to overcome the near-far effect reaching a (nearly) Pareto optimal NE point. The fairness of both approaches [28, 77] is experimentally showed among small number of users.

Kwon *et al.* in [79] aim to maximize the weighted sum rate of the users with less transmit power in a multi-cell scenario. This objective, together with power and rate constraints defines the non-cooperative game. The system performance takes into account the intra-cell co-channel interference in an uplink OFDMA network. The simulation results show that the performance of the proposed algorithm strictly depends on the power price coefficient which represents the cost imposed on each base station for the co-channel interference generated by it as well as its power consumption.

Z. Han *et al.* in [27] analyze the previous mentioned works aimed to maximization of data rate under a bounded transmit power constraint. They show that the pure non-cooperative game may have in some undesirable NE points with low system and individual performance. The authors suggest to introduce a centralized "virtual referee" whose role is to prevent users having high co-channel interference to share one subcarrier or to reduce the infeasible required transmission rates. Even though the results significantly outperforms WF solution in terms of reducing transmit power and increasing data rate, the proposed algorithm suffers from high computational complexity.

The problem of resource allocation is extended in [80, 81] to a multi cell OFDMA system. The authors of [80] devise a non-cooperative potential game [82] aimed at maximizing the users' energy efficiency. Results clearly show that the proposed game brings performance improvements in terms of goodput (error-free delivery) for each unit of energy. In [81] the same purpose is accomplished by a centralized subcarrier allocation procedure and a distributed non-cooperative power control game. The simulation results in a realistic multi cell network scenario show that the proposed algorithm achieves an acceptable performance and computational complexity burden.

Recently, several other methods which use various heuristics based on cooperative (coalitional) game theory [75, 76] has been proposed to address the problem of fair resource allocation for OFDMA systems using either centralized or distributed algorithms. Nash bargaining solution (NBS) [75] is the most refired technique applied to wireless resource allocation problems in a multiuser OFDM network. The NBS proves the existence and uniqueness of NE point of the following convex utility function:

$$\begin{aligned}
&\max_{\{\mathbf{p},\,\mathcal{N}\}} \prod_{k=1}^{K} (R_k - \underline{R}_k) \\
&\text{s.t.} \quad R_k \geq \underline{R}_k \qquad \forall\, k \in \mathcal{K} \\
&\text{and} \sum_{n \in \mathcal{N}_k} p_{kn} \leq \overline{p}_k \qquad \forall\, k \in \mathcal{K} \\
&\text{and} \quad \mathcal{N}_k \cap \mathcal{N}_{m \neq k} = \emptyset \qquad \forall\, k,\, m \in \mathcal{K}
\end{aligned} \tag{1.23}$$

wherein $R_k = \sum_{n \in \mathcal{N}_k} R_{kn}$. In other words, the goal is to maximize the product of the excesses of the transmitter's rates over their own minimum demands. The NBS guarantees each user to achieve its own demand, thus providing an individual rationality to the resource allocation. The important result of applying NBS is that the final rate allocation vector is Pareto optimal. When $\underline{R}_k = 0 \ \forall\ k \in \mathcal{K}$, taking into consideration the strictly concave increasing property of a logarithm function, we can transform the utility maximization of NBS into the same problem with the following objective function:

$$\max_{\{\mathbf{p},\,\mathcal{N}\}} \sum_{k=1}^{K} \log_2 (R_k) \tag{1.24}$$

Clearly, when $\underline{R}_k = 0$, NBS fairness scheme is the same as the weighted proportional fairness at $\varphi_k = 1$.

Z. Han *et al.* in [26] introduce a distributed algorithm for an OFDMA uplink based on the NBS and the Hungarian method [83] to maximize the overall system rate under individual power and rate constraints. The underlying idea is that once the minimum demands are provided for all users, the rest of the resources are allocated proportionally to different users according to their own conditions. The proposed algorithm shows a complexity $\mathcal{O}\left(K^2 N \log_2 N + K^4\right)$, without considering the expensive computational load to solve the (convex) equations of the NBS. In [84], Lee *et al.* resolve two subproblems of exclusive subcarrier assignment and power control in an OFDMA network aiming at the maximization of NBS fairness. The simulation results show an overall end-to-end rate between the nodes comparable to [26].

One main drawback of applying NBS in resource allocation problems is that this scheme guarantees minimum requirements of the users, but it does not impose any upper bound constraint. In fact, the achieved data rate may be much higher than the initial demands and this is an unsatisfactory from wireless network provider viewpoint. One of the most prominent alternatives to the NBS is the Raiff-Kalai-Smorodinsky bargaining solution (RBS), defined by Raiffa [85] and characterized by Kalai and Smorodinsky [86]. RBS requires that a user's payoff data rate should be proportional not only to his minimal rate but also to his maximal rate. Whereas NBS takes into account the individuals gain, RBS emphasizes the importance of one's gain and others' losses. For an OFDMA resource allocation problem, the RBS bargaining outcome is the solution to:

$$\begin{aligned}
&\max_{\{\mathbf{p},\,\mathcal{N}\}} \prod_{k=1}^{K} \left(R_k - \underline{R}_k + \frac{1}{K-1} \left(\sum_{k \neq m \in \mathcal{K}} \overline{R}_m - R_k \right) \right) \\
&\text{s.t.} \quad \underline{R}_k \leq R_k \leq \overline{R}_k \qquad \forall\, k \in \mathcal{K} \\
&\text{and} \quad \sum_{n \in \mathcal{N}_k} p_{kn} \leq \overline{p}_k \qquad \forall\, k \in \mathcal{K} \\
&\text{and} \quad \mathcal{N}_k \cap \mathcal{N}_{m \neq k} = \emptyset \qquad \forall\, k,\, m \in \mathcal{K}
\end{aligned} \tag{1.25}$$

wherein $\overline{R}_k$ denotes the upper bound of the transmission rate of the each user. When applying RBS, if the channel quality of a terminal improves, it will get a better capacity without any reduction to that of the other users (individual monotonicity). The existence and uniqueness of RBS were shown, but for more than two players a Pareto optimal NE point is not always attained as

Roth stated in [87]. The utility maximization of RBS is transformed into the same problem with the following objective function:

$$\max_{\{\mathbf{p},\,\mathcal{N}\}} \sum_{k=1}^{K} \log_2 \left(\frac{R_k - \underline{R}_k}{\overline{R}_k - \underline{R}_k} \right) \tag{1.26}$$

RBS is a point at which each individual's gain is proportional to its maximum gain. When $\underline{R}_k = 0 \; \forall \, k \in \mathcal{K}$ and $R_1 : \cdots : R_K = \overline{R}_1 : \cdots : \overline{R}_K$, RBS achieves the same result of max-min fairness criterion.

In [88], Chee *et al.* propose a centralized algorithm for the OFDMA downlink scenario based on RBS. RBS guarantees data rate achieved by each user to be bounded to a minimum and a maximal rate. The results show a good performance only when the gap between the maximum and the minimum rate is (very) large. Even though the subcarriers are assigned in an exclusive manner, the computational complexity of this algorithm is $\mathcal{O}\left(KN + K^2\right)$.

Auction methods are another cooperative game scheme which has recently drawn attention in the resource allocation research literature. In [89], Noh proposes a distributed and iterative auction-based algorithm in the OFDMA uplink scenario with incomplete information. The time complexity of the algorithm is experimentally revealed as $\mathcal{O}\left(KN \log_2 K\right)$. However, the simulation parameters are not realistic (three users and subcarriers), and it is thus hard to estimate the computational complexity when using real-world network parameters. For multi cellular scenario, Yang *et al.* [90] propose an auction solution to OFDMA resource allocation in an uplink direction. The proposed distributed (among base stations) algorithm converges to suboptimal weighted sum rate result. The number of the users and subcarriers in the simulation are small and it is not clear how the computational complexity considerations extend to a realistic network.

1.10 Discussion

It is a matter of controversy whether the OFDMA resource allocation techniques in the literature are actually usable in the practice. All the mentioned schemes, which represent, to the author's knowledge, the most relevant algorithms for OFDMA resource allocation with cooperative game theory, exhibit a good trade-off between overall system rate and fairness. The fairness schemes in the solutions based on cooperative game theory are extended approaches of that in classic solutions of: max-min fairness, and weighted proportional fairness schemes. Unfortunately, they also present a number of common problems:

1) In almost all algorithms the utility function is restricted to either be convex or strictly concave;

2) Most algorithms are based on non-linear programming, which is computationally intensive and hardly scalable when considering thousands of subcarriers and tens of users. Thus, they are not suitable for a cost-effective real-time implementation by network designers;

3) Although the resource apportionment turns out to be fair from the users point of view, the achieved QoS may be much larger than demanded. This implies a waste of network resources from a network service provider perspective, which is often overlooked by previous works;

4) To reduce the computational complexity, each subcarrier is allocated to mobile terminals in an exclusive manner, although this may limit the number of concurrent connections in the uplink channel;

5) To reduce the computational complexity, the power constraint is usually defined as the overall energy consumption of each user over all subcarriers rather than individual limitation on each subcarrier, and this may result in impractical spectral power distribution.

After reviewing cooperative game theory in Chapter 2, in Chapter 3 we will introduce an algorithm based on cooperative games to overcome most of the above mentioned disadvantages of the existing schemes. We aim at designing a low-complexity algorithm that achieves each users QoS requirement in terms of target transmit rates, with the best utilization of the network resources, so as to satisfy both the users and the network service provider.

Chapter 2

Cooperative game theory

Game theory is the study of decision making in an interactive environment. *Cooperative games* fulfill the promise of group efficient solutions to problems involving strategic actions. Formulation of optimal player behavior is a fundamental element in this theory. This chapter comprises a self-instructive didactic means indicating how cooperative game theory tools can provide a framework to tackle different network engineering problems.

This chapter is divided into nine sections. After a brief motivation in the following section, Sect. 2.2 provides an introductory discussion of cooperative game theory. We systematically study fundamental definitions and conditions of cooperative games: superadditivity and convexity. Then, Sect. 2.3 and the sub-section inside discuss the core set solution as the most known solution for payoff distribution. Sect. 2.4 is devoted to a study of a strong payoff distribution, the so-called Shapley value. In Sect. 2.5 we present a systematic study of two others reward division called the *kernel* and *nucleolus*. Then, in Sect. 2.6, we extend the concept of Nash equilibria in cooperative games. Sect. 2.7 is an investigation of the concept of coordinated equilibria where players of game are admitted to pre-communicate among themselves at once. Finally, Sect. 2.8 helps a reader to understand the basic concepts and importance of dynamic learning in cooperative games. Every sections contain some motivation examples that are expedient to understand how different communication networks problems can be modeled as cooperative game. We conclude this chapter in Sect. 2.9.

2.1 Motivation

The increase of the number of wireless services, combined with demand for high definition multimedia communications, have made the radio resources, and particularly the spectrum and power, a very precious and scarce resource, not because of their unavailability but because they are used inefficiently. For licensed spectrum, the measurements by Shared Spectrum Company [91] shows that the maximal usage of the spectrum is a low percentage of the whole licensed. While the number of users and the spectrum usage steadily increase, the amount of spectrum is still considered a limited resource. Beside, to differentiate between the true signal and background noise is complex for a radio equipment. Generally, this complex process enforces terminals to transmit strong version of signals, that wastes energy of a transmitter.

The modern wireless entities, i.e. wireless terminals and base stations, have considerable capacities to execute dynamic processes. This capability encourages wireless service providers to consider wireless entities as autonomous agents which could cooperate and negotiate with each other to achieve an efficient resource allocation in different situations. Cooperation among wireless terminals is usually intended to achieve a fair radio resource allocation. Cooperation between base stations can be devised to mitigate interference, and promote soft handover where channel gain is varying rapidly which is a challenge in LTE [92].

Game theory is the most prominent tool to analyze interaction issue in social sciences wherein often cooperation amongst autonomous agents is essential for successful task completion. In many settings, groups of competing agents are simultaneously concerned of both individual and overall benefits. In the game theory literature, this branch is known as cooperative (coalitional) game [75, 76]. The players, as the main decision making entities in the game, are considered to negotiate with each other to determine a binding agreement among them. If we assume that all users act rationally and we know what the behavior of the users are, it is possible to determine the overall performance of a system since the actions of one user becomes part of the circumstances for another user. Thus, we are interested in individual performance and overall system performance under a specific set of rules. To fully develop the different possibilities within a game for cooperation among players we have to address which groups the players can achieve collectively. Indeed, if a player assesses that within a certain group it does not receive what it is able to get by itself, then it might decide to abandon the cooperation and pursue an alternative allocation by itself. Cooperative game theory offers the opportunity to extend and expand the treatment of the players in traditional non-cooperative games, especially where selfish players compete over a set of resources.

In the last few years, cooperative game theory has been successfully applied to communications and networking. The current literature is mainly focused on applying cooperative games in various applications such as distributed/centralized radio resource allocation [88, 93, 94], power control [95, 96], spectrum sharing in cognitive radio [97,98], cooperative automatic repeat request (ARQ) mechanism [99], cooperative routing [100], and cooperative communications [101, 102]. These problems in wireless networks can be modeled as a cooperative game since it is highly likely that each wireless user can obtain a better utility value by forming groups and controlling resources cooperatively rather than individually. It has been shown that cooperation can result in an enhanced QoS in terms of throughput expansion, bit error rate reduction, or energy saving [103].

Cooperation can be realized at various layers of the network. At the physical layer, different separate antennas can constitute a cluster and then cooperate with each other to exploit multiple-input multiple-output (MIMO) gains. At the MAC sublayer, some wireless terminals can cooperate with each other to share a common wireless medium in an efficient manner and consequently mitigate the interference hazard. There is also the possibility of cooperation of physical and application layers among individual terminals to adapt channel and source codings in multimedia communications. The altruistic decision of cooperation with others network entities may result in an improvement on overall network performance, and concurrently achieve an egoistic interest of self improvement.

2.2 Preliminaries

Game theory deals with the study, through mathematical models, of conflict situations in which two or more rational players make decisions that will influence each other's welfare. The theory of cooperative games [75, 76] also assumes that binding agreements may be established among the players in the course of the conflict situation. In transferable utility (TU) games, agreement may be reached by any subset of the players, and the gain obtained from this agreement is a real number and it is transferable among these players. In non-transferable utility (NTU) games, agreement may be reached by any subset of the players, but the gain may be non-transferable. The main focus of this dissertation will be on the study of TU games.

A TU game is a pair $\mathcal{G} = (\mathcal{K}, \nu)$, where $\mathcal{K} = [1, \ldots, K]$ denotes the set of players and ν the *coalition (characteristic) function* which is interpreted as the maximum outcome (a real number) to each coalition (subset of $\mathcal{K}$) whose players can jointly produce. An NTU game is a pair $\mathcal{G} = (\mathcal{K}, V)$ where V is a mapping which for each coalition $\mathcal{A}$, defines a characteristic set, $V(\mathcal{A})$, satisfying:

1. $V(\mathcal{A})$ is non-empty and closed subset of $\mathbb{R}^{|\mathcal{A}|}$,
2. For each $k \in \mathcal{A}$ there is a $V_k \in \mathbb{R}$ such that $V(\{k\}) = (-\infty\,, V_k]$,
3. $V(\mathcal{A})$ is comprehensive, i.e. for all $\mathbf{u} \in V(\mathcal{A})$ and for all $\mathbf{u}' \in \mathbb{R}^{|\mathcal{A}|}$, if $\mathbf{u}'[k] \leq \mathbf{u}[k]\ \forall\, k \in \mathcal{A}$ then $\mathbf{u}' \in V(\mathcal{A})$,
4. The set $V(\mathcal{A}) \bigcap \left\{\mathbf{u}' \in \mathbb{R}^{|\mathcal{A}|} \mid \mathbf{u}'[k] \geq V_k \ \ \forall\, k \in \mathcal{A}\right\}$ is bounded.

The characteristic set, $V(\mathcal{A})$, is interpreted as the set of achievable outcomes the players in $\mathcal{A}$ can guarantee themselves without cooperating with the players in $\mathcal{K} \backslash \mathcal{A}$. In particular, an NTU-game $\mathcal{G} = (\mathcal{K}, V)$ is called a TU game when the characteristic set for each coalition $\mathcal{A}$, takes the form:

$$V(\mathcal{A}) = \left\{\mathbf{u} \in \mathbb{R}^{|\mathcal{A}|} : \sum_{k \in \mathcal{A}} u_k \leq \nu(\mathcal{A})\right\} \tag{2.1}$$

where $\mathbf{u} = \left[u_1, \ldots, u_{|\mathcal{A}|}\right] \in \mathbb{R}^{|\mathcal{A}|}$ and u_k is the payoff of player k in $\mathcal{A}$ and $\nu : 2^{\mathcal{K}} \longrightarrow \mathbb{R}$. If $\mathcal{A}$ is a coalition (subset) of $\mathcal{K}$ formed in $\mathcal{G}$, then its members get an overall payoff $\nu(\mathcal{A})$, zero for the empty set. Each coalition can be represented as a pure strategy in non-cooperative game theory.

An important property of interest in characteristic form TU games is *superadditivity*, which, if present, implies that the value of the union of any two disjoint coalitions is at least as big as the sum of their values.

Definition 4 *A TU game $\mathcal{G}$ is superadditive if*

$$\nu(\mathcal{A}_i \cup \mathcal{A}_j) \geq \nu(\mathcal{A}_i) + \nu(\mathcal{A}_j) \quad \forall\, \mathcal{A}_i, \mathcal{A}_j \subset \mathcal{K} \;\; s.t. \;\; \mathcal{A}_i \cap \mathcal{A}_j = \varnothing \qquad (2.2)$$

In a superadditive TU game there are positive synergies and the players prefer to join each other rather than act alone. Under superadditivity condition, the players are willing to form the *grand coalition* (the set $\mathcal{K}$).

Convex, or alternatively *supermodular* coalitional games were introduced by L. Shapley [104]. They model cooperative situations where the marginal contribution of a player to a coalition increases as the coalition becomes larger.

Definition 5 *A TU game $\mathcal{G}$ is convex or supermodular if for all $k \in \mathcal{K}$:*

$$\nu(\mathcal{A}_i \cup \{k\}) - \nu(\mathcal{A}_i) \;\leq\; \nu(\mathcal{A}_j \cup \{k\}) - \nu(\mathcal{A}_j) \quad \forall\, \mathcal{A}_i \subseteq \mathcal{A}_j \subset \mathcal{K}\backslash\{k\} \quad (2.3)$$

Equivalently:

Definition 6 *A TU game $\mathcal{G}$ is convex or supermodular if:*

$$\nu(\mathcal{A}_i) + \nu(\mathcal{A}_j) \;\leq\; \nu(\mathcal{A}_i \cap \mathcal{A}_j) + \nu(\mathcal{A}_i \cup \mathcal{A}_j) \quad \forall\, \mathcal{A}_i, \mathcal{A}_j \subseteq \mathcal{K} \qquad (2.4)$$

Convexity means that there are increasing returns to scale. Note that a convex game is superadditive. To better understand the importance of convexity approach in network probems, we verify the convexity condition in a K-user channel access game. The payoff of each coalition of players (transmitters) is defined as the outer MAC capacity region. [105, Lemma 1] shows that in a multiple access channel scenario, the inequality (2.4) is not met. This means the game is not convex, and thus adding a new player does not give benefit to others transmitters.

2.3 The core solution

A central question in a coalitional game is how to divide the extra earnings (or cost savings) among the members of the formed coalition. In a TU game, an allocation is a function $\mathbf{u}$ from $\mathcal{K}$ to $\mathbb{R}$ that specifies for each player $k \in \mathcal{K}$ the payoff $u_k \in \mathbb{R}$ that this player can expect when it cooperates with the other players. The payoff of each player can show the cost borne by the player, the power of influence, and so on depending on the problem setting.

Definition 7 *Let $\mathcal{K}$ be the set of K players of the superadditive TU game $\mathcal{G}$, and let ν be the payoff of the game. The set of all "imputations" of $\mathcal{G}$ is the set:*

$$\mathcal{I}(\mathcal{K}, \boldsymbol{\nu}) = \left\{ \mathbf{u} \in \mathbb{R}^K : \left| \begin{array}{ll} i) & \sum_{k \in \mathcal{K}} u_k = \nu(\mathcal{K}) \\ ii) & u_k \geq \nu(\{k\}) \quad \forall\, k \in \mathcal{K} \end{array} \right. \right\} \tag{2.5}$$

where $\mathbf{u} = [u_1, \ldots, u_k, \ldots, u_K] \in \mathbb{R}^K$ is the imputation vector of the players. The former condition is called the "feasibility", and the latter "individually rational" condition.

The *core* concept was introduced in [106] and is the most attractive and natural way to define a payoff distribution: if a payoff distribution is in the core, no agent has any incentive to be in a different coalition. The core of a TU game is the subset of all imputations $\mathbf{u} \in \mathcal{I}(\mathcal{K}, \boldsymbol{\nu})$ that no other imputation *directly dominates*, that is $\nexists\, \mathbf{u}' \in \mathcal{I}(\mathcal{K}, \boldsymbol{\nu})$ s.t. $u'_k > u_k$ $\forall\, k \in \mathcal{K}$. As can be seen, for cooperative games as well as non-cooperative games, the notion of dominance is essentially equivalent; the payoffs under the various situations are compared and one situation dominates the others if these payoffs are higher. The core actually presents a condition stronger than Nash equilibrium in non-cooperative game: no group of agents should be able to profitably deviate from a configuration in the core. Equivalently, no set of players can benefit from forming a new coalition, which corresponds to the group rationality assumption.

In an NTU game $\mathcal{G} = (\mathcal{K}, V)$, the core apportionment is defined as:

Definition 8 *Let $\mathcal{K}$ be the set of K players of the superadditive NTU-game $\mathcal{G}$, and let V be the payoff of the game. The core of $\mathcal{G}$ is the set*

$$\mathcal{S}\left(\mathcal{K}\,,\,\mathbf{V}\right)=\left\{\,\mathbf{u}\in V\left(\mathcal{K}\right):\ \ \forall\,\mathbf{u}'\in V\left(\mathcal{A}\right)\ \ \exists\,k\in\mathcal{A}\ \ s.t.\ \ u_k\geq u'_k\,\right\} \tag{2.6}$$

where $\mathbf{u}$ is the payoff distribution across players, and $u_k \in \mathbf{u}$ if and only if no coalition can improve upon u_k.

In a TU game $\mathcal{G} = (\mathcal{K}, \nu)$, the core apportionment is defined as follows:

Definition 9 *Let $\mathcal{K}$ be the set of K players of the superadditive TU game $\mathcal{G}$, and let ν be the payoff of the game. The core of $\mathcal{G}$ is the set*

$$\mathcal{S}\left(\mathcal{K}\,,\,\boldsymbol{\nu}\right)=\left\{\ \mathbf{u}\in\mathbb{R}^K:\ \left|\ \begin{array}{ll} i) & \displaystyle\sum_{k\in\mathcal{K}}u_k=\nu\left(\mathcal{K}\right) \\ ii) & \displaystyle\sum_{k\in\mathcal{A}}u_k\geq\nu\left(\mathcal{A}\right)\quad\forall\,\mathcal{A}\subset\mathcal{K} \end{array}\right.\right\} \tag{2.7}$$

where $\mathbf{u} = [u_1, \ldots, u_k, \ldots, u_K] \in \mathbb{R}^K$ is the payoff distribution across players, and $u_k \in \mathbf{u}$ if and only if no coalition can improve upon u_k. The second condition is called "non-blocking" condition.

The core consists of the set of allocations that can be blocked by any coalition of agents. If for some set of agents $\mathcal{A}$, the non-blocking condition does not hold, then the agents in $\mathcal{A}$ have an incentive to collectively deviate from the coalition structure and to divide $\nu\left(\mathcal{A}\right)$ among themselves. In general, the core of a given TU game $(\mathcal{K}, \nu)$ is found by linear programming (LP) as:

$$\min_{\mathbf{u}\in\mathbb{R}^K}\sum_{k\in\mathcal{K}}u_k\ ;\qquad \text{s.t.}\ \sum_{k\in\mathcal{A}}u_k\geq\nu\left(\mathcal{A}\right)\quad\forall\,\mathcal{A}\subseteq\mathcal{K} \tag{2.8}$$

Madiman in [107] introduces some intuitive applications of core solution to information theory contexts e.g. source coding and multiple access channel, and summarize some of its limitations in multi user scenarios. Li *et al.* in [108] show that the cooperation among wireless nodes and core apportionment can increase spectrum efficiency in a TDMA cooperative communication. In [109], Niyato *et al.* applies the core solution in a cooperative among different wireless access networks to offer a stable and efficient bandwidth allocation.

Indeed, there is a number of realistic application scenarios, in which the emergence of the grand coalition is either not guaranteed, or might be perceivably harmful, or is plainly impossible [110]. For a non-superadditive coalitional game, the coalition formation process does not lead the players to form the grand coalition. In this case, Def. 9 does not apply. Let us redefine the core set in a general (not necessarily superadditive) coalitional formation TU game. Let $\psi = [\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m]$ denote a partition of the set $\mathcal{K}$ wherein $\mathcal{A}_i \cap \mathcal{A}_j = \varnothing$ for $i \neq j$, $\bigcup_{i=1}^{m} \mathcal{A}_i = \mathcal{K}$ and $\mathcal{A}_i \neq \varnothing$ for $i = 1, \ldots, m$, and let $\mathbf{\Psi}$ denote the set of all possible partitions ψ. Let us also define $\mathcal{F} = [\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n]$, such that $\bigcup_{i=1}^{n} \mathcal{A}_i = \mathcal{K}$ and $\mathcal{A}_i \neq \varnothing$ for $i = 1, \ldots, n$, as a family of (not necessarily disjoint) coalitions.

Definition 10 *A "core apportionment"* $\mathbf{u} \in \mathbb{R}^K$ *is a payoff distribution with the following property:*

$$\mathcal{S}(\mathcal{K}, \boldsymbol{\nu}) = \left\{ \mathbf{u} \in \mathbb{R}^K : \left| \begin{array}{ll} i) & \sum\limits_{k \in \mathcal{K}} u_k = \max\limits_{\psi \in \mathbf{\Psi}} \sum\limits_{\mathcal{A} \in \psi} \nu(\mathcal{A}) \\ ii) & \sum\limits_{k \in \mathcal{A}} u_k \geq \nu(\mathcal{A}) \quad \forall \mathcal{A} \subset \mathcal{K} \end{array} \right. \right\} \tag{2.9}$$

Note that, if $\mathcal{G}$ *is superadditive, then* $\max\limits_{\psi \in \mathbf{\Psi}} \sum\limits_{\mathcal{A} \in \psi} \nu(\mathcal{A}) = \nu(\mathcal{K})$.

The core allocation set can be found through linear programming and its existence, in general, depends upon the feasibility of (2.8). Unfortunately, the core is a strong notion, and there exist many games where it is empty. We can study the non-emptiness of the core without explicitly solving the core equation. The following notation helps to simplify the dual of (2.8).

Definition 11 *A superadditive TU game* $\mathcal{G}$ *for a family* $\mathcal{F}$ *of coalitions is "totally balanced" if, for any* $\mathcal{A} \in \mathcal{F}$, *the inequality*

$$\sum_{\mathcal{A} \in \mathcal{F}} \mu_{\mathcal{A}} \cdot \nu(\mathcal{A}) \leq \nu(\mathcal{K}) \tag{2.10}$$

holds, where $\mu_{\mathcal{A}}$ *is a collection of numbers in* $[0,1]$ *(balanced collection of weights) such that*

$$\sum_{\mathcal{A}\in\mathcal{F}} \mu_{\mathcal{A}} \cdot \mathbf{1}_{\mathcal{A}} = \mathbf{1}_{\mathcal{K}} \tag{2.11}$$

with $\mathbf{1}_{\mathcal{A}} \in \mathbb{R}^{K}$ *denoting the characteristic vector whose elements are*

$$(\mathbf{1}_{\mathcal{A}})[i] = \begin{cases} 1, & i \in \mathcal{A} \\ 0, & \textit{otherwise} \end{cases} \tag{2.12}$$

The following pathbreaking result in the theory of TU games was independently gave by Bondareva [111] and L. Shapley [112].

Lemma 1 ([75]) *A totally balanced TU game has a non-empty core set.*

Where forming the grand coalition is not guaranteed, the following notation is applied.

Definition 12 *A (not necessarily superadditive) TU game* $\mathcal{G}$ *for a family* $\mathcal{F}$ *of coalitions is totally balanced if, for every balanced collection of weights* $\mu_{\mathcal{A}}$*, and for any* $\mathcal{A} \in \mathcal{F}$*,*

$$\sum_{\mathcal{A}\in\mathcal{F}} \mu_{\mathcal{A}} \cdot \nu(\mathcal{A}) \leq \max_{\psi\in\boldsymbol{\Psi}} \sum_{\mathcal{A}\in\psi} \nu(\mathcal{A}) \tag{2.13}$$

So, if a TU game is totally balanced, then the core is non empty and therefore it is a convenient solution concept on the class of totally balanced TU games. There is an interesting relation between convex and balanced games.

Lemma 2 ([76]) *A convex game is totally balanced, but the converse is not necessarily true.*

The other key feature of cooperative convex games is

Lemma 3 (L. Shapley [104]) *The core set of a convex game is unique.*

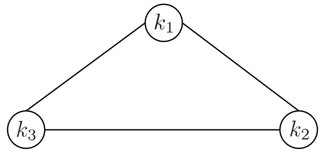

Fig. 2.1: *The network allocates power among three players according to their will to cooperate with each other. A selfish player receives zero, a pair of cooperative players receive* $0.8\,\mathrm{mW}$*, and the network supply* $1\,\mathrm{mW}$ *to the grand coalition.*

Now, we illustrate an intuitive example of *power distribution* based on core set solution. This example is an extended form of the example established by [113, Ch. 12]. The network sketched in Fig. 2.1 wishes to allocate power among three players $\mathcal{K} = \{k_1, k_2, k_3\}$, according to their will to cooperate with each other. A power of $1\,\mathrm{mW}$ is provided to the network if three players decide to cooperate, or equivalently if the grand coalition will form. If only one player refuses to cooperate, a power of $0.8\,\mathrm{mW}$ will be assigned to the pair of cooperating nodes. The coalition game of Fig. 2.1 is defined by:

$$\nu(\mathcal{A}) = \begin{cases} 0 & \text{if } |\mathcal{A}| = 1; \\ 0.8 & \text{if } |\mathcal{A}| = 2; \\ 1 & \text{if } |\mathcal{A}| = 3. \end{cases} \tag{2.14}$$

The players of each coalition will cooperate with each other. The player of a singleton coalition will be isolated.

Each player receives a positive payoff if it decides to cooperate, whereas all players receive zero if no agreement is bound. To divide the total payoff (power) in some appropriate way, we rest on the core set definition. It is straightforward to show that the cooperative TU game defined by (2.14) is superadditive. From Eqs. 2.3 and 2.4, it is easy to show that TU game (2.14) is not convex (supermodular). To check whether the core set of TU game (2.14) is empty or not, we resort to the balanced solution. TU game (2.14) is not balanced even though assigning the balanced weights as $\mu_{\mathcal{A}} = 1$ for singleton coalitions, and $\mu_{\mathcal{A}} = 0$ otherwise, inequality (2.10) holds. By using the fact that there exists other balanced collection of weights in which $\mu_{\mathcal{A}} = \frac{1}{2}$ for $|\mathcal{A}| = 2$, and $\mu_{\mathcal{A}} = 0$ otherwise, the game is not balanced, and its core set may be empty. Note that, this result does not mean that the core set of the game is *surely* empty.

Now, we heuristically find a core apportionment studying various possible networks. When there is no cooperation among players, the players are not provided with any power. That is, $\mathcal{F} = [\{k_1\}, \{k_2\}, \{k_3\}]$ with payoff distribution:

$$u_{k_1} = u_{k_2} = u_{k_3} = 0$$

If only one player decides to stay alone, the payoff 0.8 is equally divided between the two cooperative players and the isolated player gets zero. That is, for instance, $\mathcal{F} = [\{k_1, k_2\}, \{k_3\}]$ with payoff distribution:

$$\begin{cases} u_{k_3} = 0 \\ u_{k_1} = u_{k_2} = 0.4 \end{cases}$$

Now, we suppose a player, for example, k_2 decides to cooperate with both k_1 and k_3, but the two players k_1 and k_3 do not bind an agreement to mutually cooperate. It is reasonable to suppose that the player k_2 can act as a relay between k_1 and k_3 and it must be provided more power. That is $\mathcal{F} = [\{k_1, k_2\}, \{k_2, k_3\}]$ with payoff distribution:

$$\begin{cases} u_{k_1} = u_{k_3} = 0.2 \\ u_{k_2} = 0.6 \end{cases}$$

Finally, in the complete network each player receives the same payoff. That is $\mathcal{F} = [\{k_1, k_2, k_3\}]$ with payoff distribution:

$$u_{k_1} = u_{k_2} = u_{k_3} = 1/3$$

As can be easily seen, the above argument satisfies feasibility and non-blocking conditions of the core set apportionment in Def. 9. It is worthwhile to note that the core set definition does not imply an even division of the whole payoff across players. Thus, it is clear that this game consists of multiple core sets.

2.3.1 On core stability

The goal of the network Fig. 2.1 is to allocate power among players in order to stimulate all of them to cooperate. Obviously, each player tries to get the highest possible payoff. Let us predict the behavior of the players after having known the definition of the game. Suppose that the players k_1 and k_2 find an opportunity to meet each other. Obviously, they quickly take advantage to cooperate and achieve payoff distribution $\mathbf{u} = [\,0.4\,,\,0.4\,,0\,]$. Then, it is profitable for player k_1 to invite player k_3 to join and therefore, improving its own payoff from 0.4 to 0.6 and that of player k_3 from zero to 0.2. On the other hand, this new agreement causes a decreasing payoff of player k_2 from 0.4 to 0.2, and now the players k_2 and k_3 have an incentive to cooperate and increase their proper payoff from 0.2 to $1/3$. Note that this agreement makes the player k_1's payoff dcrease from 0.6 to $1/3$. The unfavorable decision of player k_2 would tempt player k_1 to retaliate. A negotiation between k_1 and k_3 to release cooperation with k_2 results increasing their payoffs and boiling down k_2's payoff to zero. The result of above argument concerns: The network is sustained by only one pair cooperation under the threat of: "If you cooperate with the third player, then I will do the same".[1] It is fairly clear that the players would seek to cooperate only as pairs for the purpose of negotiation, and not cooperate in the grand coalition framework, even though the game is superadditive. This is due to fact of being superadditive but not balanced. The pairs can be changed as time goes on. In fact, the core apportionment suffers the lack of "farsighted" (i.e., long-term) stability.

A coalition structure based on core set, is not adequately farsighted to avoid the elusiveness of negotiation structure. At first sight, the core appears to be an extremely myopic notion, requiring the stability of a proposed allocation to deviations or blocks by coalitions, but not examining the stability of the deviations themselves. In general, the stability requirement is that the outcome be immune to deviations of a certain sort by coalitions. To provide the formal definition of farsighted stability, we need some additional notation.

---

[1] *Two is cooperation, three is a crowd.*

Definition 13 *For* $\mathbf{u}, \mathbf{u}' \in \mathcal{I}(\mathcal{K}, \boldsymbol{\nu})$, $\mathbf{u}$ *"indirectly dominates"* $\mathbf{u}'$, *which is denoted by* $\mathbf{u}' \ll \mathbf{u}$, *if there exist a finite sequence of imputations* $\mathbf{u}' = \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m = \mathbf{u}$ *and a finite sequence of nonempty coalitions* $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m$ *such that for each* $j = 1, 2, \ldots, m-1$*: i) by the deviation of* $\mathcal{A}_j$*, the imputation of* $\mathbf{u}_j$ *is replaced to* $\mathbf{u}_{j+1}$*, and ii)* $\mathbf{u}_j[k] < \mathbf{u}[k]$ *for all* $k \in \mathcal{A}_j$*.*

Condition $i)$ says that each coalition $\mathcal{A}_j$ has the power to replace imputation $\mathbf{u}_j$ by imputation $\mathbf{u}_{j+1}$, and the condition $ii)$ says that each player in $\mathcal{A}_j$ strictly prefers imputation $\mathbf{u}$ to imputation $\mathbf{u}_j$. It is clear that the indirect dominance relation contains the direct dominance relation.

Definition 14 *Let* $\mathcal{G} = (\mathcal{K}, \nu)$ *be a TU game. A subset* $\mathcal{J}$ *of* $\mathcal{I}(\mathcal{K}, \boldsymbol{\nu})$ *is a "farsighted stable" set if: i) for all* $\mathbf{u}, \mathbf{u}' \in \mathcal{J}$*, neither* $\mathbf{u} \ll \mathbf{u}'$ *nor* $\mathbf{u}' \ll \mathbf{u}$*, and ii) for all* $\mathbf{u}' \in \mathcal{I}(\mathcal{K}, \boldsymbol{\nu}) \backslash \mathcal{J}$ *there exists* $\mathbf{u} \in \mathcal{J}$ *such that* $\mathbf{u}' \ll \mathbf{u}$*. Conditions i) and ii) are called "internal stability" and "external stability", respectively.*

By internal stability, there is no imputation in $\mathcal{J}$ that is dominated by another imputation in $\mathcal{J}$. By external stability, an imputation outside a stable set $\mathcal{J}$ is unlikely to be attained. Let us introduce three other different payoff distribution concepts which capture foresight of the players.

2.4 Shapley value

The Shapley value is an alternative solution for the payoff distribution in TU games. The Shapley value has long been a central solution concept in cooperative game theory. It was introduced by L. S. Shapley in the seminal paper [114] and it was seen as a reasonable way of distributing the gains of cooperation, in a fair and unique way, among the players in the game. In the Shapley solution, those who contribute more to the groups that include them are paid more. Let us denote $\phi_k(\nu)$ as the Shapley value of player k in the TU game defined by ν. The surprising result due to Shapley is the following theorem.

Theorem 1 *There is a unique single-valued solution to TU games satisfying efficiency, symmetry, additivity and dummy. It is the well-known Shapley value, the function that assigns to each player k the payoff:*

$$\phi_k(\nu) = \sum_{\substack{\forall \mathcal{A} \subseteq \mathcal{K} \\ s.t.\ k \in \mathcal{A}}} \frac{(|\mathcal{A}| - 1)! \cdot (K - |\mathcal{A}|)!}{K!} \big(\nu(\mathcal{A}) - \nu(\mathcal{A} \backslash \{k\}) \big) \tag{2.15}$$

The expression $\nu(\mathcal{A}) - \nu(\mathcal{A}\backslash\{k\})$ is the marginal payoff of player k to the coalition $\mathcal{A}$. The Shapley value can be interpreted as the expected marginal contribution made by a player to the value of a coalition, where the distribution of coalitions is such that any ordering of the players is equally likely. That makes the Shapley value exponentially hard to compute. Shapley characterized such value as the unique solution that satisfies the following four axioms:

1. *Efficiency*: The payoffs must add up to $\nu(\mathcal{K})$, which means that all the grand coalition surplus is allocated. That is:

   $$\sum_{k \in \mathcal{K}} \phi_k(\nu) = \nu(\mathcal{K})$$

   In the absence of superadditivity, instead we use: $\max_{\psi \in \mathbf{\Psi}} \sum_{\mathcal{A} \in \psi} \nu(\mathcal{A})$.

2. *Symmetry*: This axiom requires that the names of the players play no role in determining the value. If two players are substitutes because they contribute the same to each coalition, the solution should treat them equally. That is:

   $$\nu(\mathcal{A} \cup \{k\}) = \nu(\mathcal{A} \cup \{i\}) \implies \phi_k(\nu) = \phi_i(\nu).$$

3. *Additivity*: The solution to the sum of two TU games must be the sum of what it awards to each of the two games. That is:

   $$\phi_k(\nu + \omega) = \phi_k(\nu) + \phi_k(\omega) \qquad \forall\, k \in \mathcal{K}.$$

4. *Dummy player*: The player k is dummy (null) if $\nu(\mathcal{A} \cup \{k\}) = \nu(\mathcal{A})$ for all $\mathcal{A}$ not containing k. If a player k is dummy, the solution should pay it nothing; i.e. $\phi_k(\nu) = 0$.

The Shapley value is a feasible allocation, but need not be individually rational. Whenever the TU game is superadditive, the Shapley value is feasible and individually rational, but need not be in the core and hence can be directly dominated by another imputation. Reference [104] shows that the Shapley value of a supermodular TU-game is a core imputation, that is, the Shapley value is not dominated. For a superadditive TU game The Shapley value is an internal and external stable imputation, and for NTU games, it is formulated in [115, 116]. To make an example, let us calculate the Shapley value of the players in the power distribution game of Fig. 2.1:

$$\nu = \begin{cases} 0 & \{k_1\}, \{k_2\}, \{k_3\}; \\ 0.8 & \{k_1, k_2\}, \{k_1, k_3\}, \{k_2, k_3\}; \\ 1 & \{k_1, k_2, k_3\}. \end{cases} \Longrightarrow$$

$$\phi_{k_1}(\nu) = \phi_{k_2}(\nu) = \phi_{k_3}(\nu) =$$
$$0 + \frac{1!\,.\,1!}{3!}(0.8 - 0) + \frac{1!\,.\,1!}{3!}(0.8 - 0) + \frac{2!\,.\,0!}{3!}(1 - 0.8) = 1/3.$$

Young in [117] defines an equivalent definition for Shapley value. He withdraws the additivity axiom, and instead, adds an axiom of marginality.

1. *Marginality*: If the marginal contribution to coalitions of a player in two games is the same, then the the award of the player must be the same. That is, if:

   $$\nu(\mathcal{A}_i) - \nu(\mathcal{A}_i \backslash \{k\}) = \omega(\mathcal{A}_j) - \omega(\mathcal{A}_j \backslash \{k\}) \quad \forall\, \mathcal{A}_i \in \nu \text{ and } \forall\, \mathcal{A}_j \in \omega,$$

   then $\phi_k(\nu) = \phi_k(\omega)$.

Marginality is an idea with a strong tradition in economic theory. In Young's definition, marginality is assumed and additivity is dropped. Young in [117] shows that the Shapley value is unique.

Theorem 2 (Young [117]) *There exists a unique single-valued solution to TU games satisfying efficiency, symmetry and marginality, and this solution is the Shapley value.*

In the network engineering literature, S. Kim in [118] proposes an energy efficient routing protocol based on the Shapley value. The concept of Shapley value is used by Khouzani *et al.* [119] to achieve a fair aggregate cost of link sharing, among primary and secondray users in a cognitive network. Using the Shapley value, a suitable network resource sharing among multimedia users is fairly achievable, as Park *et al.* propose in [120].

2.5 The kernel and nucleolus

Let $\mathcal{G} = (\mathcal{K}, \nu)$ be a coalitional game with transferable payoff. The *excess* of the coalition $\mathcal{A}$ with respect to the payoff vector $\mathbf{u} \in \mathbb{R}^K$ is defined as

$$e\,(\mathcal{A}\,,\,\mathbf{u}) = \nu\,(\mathcal{A})\; - \sum_{k\in\mathcal{A}} u_k \tag{2.16}$$

A positive excess can be interpreted as an incentive for a coalition to generate more utility. Using the excess notion, the core apportionment in a TU game can be redefined as:

$$\left\{\,\mathbf{u} \in \mathbb{R}^K :\;\; e\,(\mathcal{K}\,,\,\mathbf{u}) = 0 \quad, \text{and} \quad e\,(\mathcal{A}\,,\,\mathbf{u}) \le 0 \quad \forall\,\mathcal{A} \subset \mathcal{K}\,\right\} \tag{2.17}$$

The *maximum excess* of player k against i is defined as

$$s_{ki}\,(\mathbf{u}) = \max\left\{\,e\,(\mathcal{A}\,,\,\mathbf{u}) \mid \mathcal{A} \subset \mathcal{K},\;\; k \in \mathcal{A},\;\; i \in \mathcal{K}\backslash\mathcal{A}\,\right\} \tag{2.18}$$

If player k departs from $\mathbf{u}$, the most it can hope to gain (the least to lose) without the consent of player i is the amount of maximum excess. Extensions of the excess for NTU games is formalized in [121].

As defined by Osborne and Rubinstein [75, Ch. 14], a coalition $\mathcal{A}_i$ is an *objection* of k against i to $\mathbf{u}$, if $\mathcal{A}_i$ includes k but not i and $u_i > \nu\,(\{i\})$. Equivalently, $\mathcal{A}_i$ is a coalition that contains k, excludes i and which gains too little. A coalition $\mathcal{A}_j$ is a *counter-objection* to the objection $\mathcal{A}_i$ of k against i, if $\mathcal{A}_j$ includes i but not k and $e(\mathcal{A}_j, \mathbf{u}) \geq e(\mathcal{A}_i, \mathbf{u})$. Equivalently, $\mathcal{A}_j$ is a coalition that contains i and excludes k and that gains even less. Objections and counter-objections are exchanged between members of the same coalition in $\mathcal{A}_i$.

The idea captured by the *kernel* is that if at a non empty imputation $\mathbf{u}$ the maximum excess of player k against any other player i is less than the maximum excess of player i against the player k, then player k should get less. Of course, the players cannot get less than their individual worths if $\mathbf{u}$ is an imputation. The definition of the kernel follows:

Definition 15 *The kernel is the set of all imputations* $\mathbf{u}$ *with the property that for every objection* $\mathcal{A}_i$ *of any player* k *against any other player* i *to* $\mathbf{u}$ *there is a counter-objection of* i *to* $\mathcal{A}_i$ *such that:*

a) $s_{ki}(\mathbf{u}) = s_{ik}(\mathbf{u})$; *or*

b) $s_{ki}(\mathbf{u}) < s_{ik}(\mathbf{u})$ *and* $u_k = \nu(\{k\})$; *or*

c) $s_{ki}(\mathbf{u}) > s_{ik}(\mathbf{u})$ *and* $u_i = \nu(\{i\})$.

The kernel is the set of imputations $\mathbf{u}$ such that for any coalition $\mathcal{A}_i$, for each objection $\mathcal{A}_j$ of a user $k \in \mathcal{A}_i$ over any other member $i \in \mathcal{A}_i$, there is a counter-objection of i to $\mathcal{A}_j$. The kernel is contained in the (nonempty) core in any assignment game ν [122, Th. 1]. In Fig. 2.1, the unique kernel element is the equal split $\mathbf{u} = [1/3, 1/3, 1/3]$, otherwise for the single player coalition objection of the player with the minimum payoff, there is no any counter-objection.

The last type of a stable imputation we will study is the *nucleolus*. With the nucleolus no confusion regarding the player set can arise. The basic motivation behind the nucleolus is that one can provide an allocation that minimizes the excess of the coalitions in a given cooperative game $\mathcal{G} = (\mathcal{K}, \nu)$. For a TU game $\mathcal{G} = (\mathcal{K}, \nu)$ and the payoff vector $\mathbf{u} \in \mathbb{R}^K$, let us denote $\mathbf{E}(\mathbf{u}) = [\cdots \geq e(\mathcal{A}, \mathbf{u}) \geq \cdots : \ \emptyset \neq \mathcal{A} \neq \mathcal{K}]$ as a $2^K - 2$ dimensional vector whose components are the values of the excess function for all $\mathcal{A} \subset \mathcal{K}$, arranged in a non-increasing order. The nucleolus of a game is the imputation which minimizes the excess with respect to the lexicographic order [2] over the set of imputations. The nucleolus of $\mathcal{G}$ with respect to $\mathcal{I}(\mathcal{K}, \boldsymbol{\nu})$ is given by:

$$\{\mathbf{u} \in \mathcal{I}(\mathcal{K}, \boldsymbol{\nu}) \mid \mathbf{E}(\mathbf{u}) \preceq_{lex} \mathbf{E}(\mathbf{u}') \quad \forall\, \mathbf{u}' \in \mathcal{I}(\mathcal{K}, \boldsymbol{\nu})\} \tag{2.19}$$

The definition of the nucleolus of a cooperative game in characteristic function form entails comparisons between vectors of exponential length. Thus, if one attempts to compute the nucleolus by simply following its definition, it would take an exponential time. In the network engineering literature, Han and Poor in [123] apply the Shapley value, excess and nucleolus solutions to study a possible cooperative transmission among intermediate nodes to help relay the information of wireless users.

This defining property makes the nucleolus appealing as a fair single-valued solution. It is easy to see that, whenever the core of a game is nonempty, the nucleolus lies in it [76]. Moreover, the nucleolus always belongs to the kernel and satisfies the symmetry and dummy axioms of Shapley: dummy players receive zero payoffs. If a null player is removed from the game, the payoff allocation of the remaining players is uninfluenced by its departure. Because of these desirable properties, the nucleolus solution has found a lot of applications in cost sharing and resource allocation as Maschler in [124] reports. However, the nucleolus possesses certain features that makes it less agreeable. The original definition treats the excesses of any two coalitions as equally important, regardless of coalition sizes and coalition composition. Some unappealing features of utility distribution, derived with the nucleolus are listed in [117]. For instance, the nucleolus lacks many monotonicity properties. That is, if a game changes so that some player's contribution to all coalitions increases then the player's allocation should not decrease. Monotonicity states that as the underlying data of game change, the utility must change in a parallel fashion.

[2] The lexicographic order between two vectors $\mathbf{u}$ and $\mathbf{u}'$ is defined by $\mathbf{u} \preceq_{lex} \mathbf{u}'$, if there exists an index k such that $\mathbf{u}[l] = \mathbf{u}'[l]$ for all $l < k$, and $\mathbf{u}[k] < \mathbf{u}'[k]$.

2.6 Cooperative Nash equilibria

Coalitional games aim at identifying the best coalitions of the agents and a fair distribution of the payoff among the agents. The classic core solution is an extension of the Nash equilibrium, since the coalitions bind agreements of agents with each other and earns a vector value rather than a real number. In [125, Sec. 7.6] it is shown that the core set of an underlying coalitional game, if it exists, asymptotically coincides with the set of Nash equilibria of the repeated game, in the long run. The result of the Nash equilibrium is not always a satisfactory outcome for an external observer (e.g., prisoner's dilemma game). R. Aumann[3] in [126] and Bernheim *et al.* [127] introduce a stronger notion of Nash equilibria based on coalitional game theory. First, let us review the definition of the Nash equilibrium where each pure strategy in a static game is presented as a coalition in a cooperative game. Thus, each player belongs to only one coalition.

Definition 16 *A pure strategy (coalition) combination* $\psi = [\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m]$ *wherein* $\mathcal{A}_i \bigcap \mathcal{A}_{j \neq i} = \emptyset$, $\bigcup_{i=1}^{m} \mathcal{A}_i = \mathcal{K}$ *and a payoff distribution* $\mathbf{u} = [u_1, \ldots, u_K]$ *is a pure Nash equilibrium if there does not exist a player* $k \in \mathcal{K}$ *whose unilateral deviation to a different coalition (pure strategy) yields a new distribution* $\mathbf{u}' = [u'_1, \ldots, u'_K]$ *such that* $u'_k > u_k$.

In other words, in a Nash equilibrium no agent is motivated to deviate from its coalition (strategy) given that the others do not deviate. As an example, we study the forwarder's dilemma game [128] presented in Fig. 2.2. This game is intended to represent a basic wireless relay operation between two different wireless terminals. These two agents, represented by players k_1 and k_2, are supposed to operate a direct link that enables them to communicate without intermediaries. Each players wants to send a packet to its destination, d_1 and d_2 respectively, in each time step using the other player as a forwarder. We assume that each forwarding has a energy cost $0 < c \ll 1$. If player

[3] Robert Aumann has received in 2005 the Nobel prize in economy for his contributions to game theory, together with Thomas Schelling.

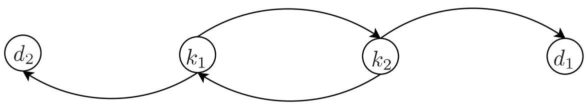

Fig. 2.2: *The network scenario of the forwarder's dilemma game.*

		k_2: F	k_2: D
k_1	F	$1-c\,,\,1-c$	$-c\,,\,1$
k_1	D	$1\,,\,-c$	$0\,,\,0$

$$\mathbf{u} = [\,u_{k_1}\,,\,u_{k_2}\,]$$

Tab. 2.1: *The strategic form in the forwarder's dilemma game. In each cell, the first value is the payoff of player k_1, whereas the second is that of k_2.*

k_1 forwards (F) the player's k_2 packet, player k_2 gets a reward 1 and vice versa. Each player's utility is its reward minus the cost. Each player is allured to drop (D) the received packet for saving energy. The strategic form of this game is depicted in Tab. 2.1. In the cooperative representation of the forwarder's dilemma game there are two coalitions $\psi = [\mathcal{A}_F\,,\,\mathcal{A}_D]$ and each player in $\mathcal{K} = \{k_1\,,\,k_2\}$ must choose one coalition. For instance, $\psi = [\mathcal{A}_F = \{k_1\,,\,k_2\}\,,\,\mathcal{A}_D = \emptyset]$ is equivalent to the strategy profile (F, F) and $\psi = [\mathcal{A}_F = \{k_2\}, \mathcal{A}_D = \{k_1\}]$ corresponds to the strategy profile (D, F), and so on.

Unilateral deviation of player k_1 from $\psi = [\mathcal{A}_F = \{k_1, k_2\}\,,\,\mathcal{A}_D = \emptyset]$ to $\psi = [\mathcal{A}_F = \{k_2\}, \mathcal{A}_D = \{k_1\}]$ increases its own payoff, and therefore the pure strategy profile (F, F) is not a Nash equilibrium point. The same applies to the departure of player k_2 from $\psi = [\mathcal{A}_F = \{k_1, k_2\}\,,\,\mathcal{A}_D = \emptyset]$ to the pure strategy $\psi = [\mathcal{A}_F = \{k_1\}\,,\,\mathcal{A}_D = \{k_2\}]$. We can easily check the different combinations of $\psi = [\mathcal{A}_F = \{k_1\},\, \mathcal{A}_D = \{k_2\}]$, $\psi = [\mathcal{A}_F = \{k_2\},\, \mathcal{A}_D = \{k_1\}]$, and finally $\psi = [\mathcal{A}_F = \emptyset,\, \mathcal{A}_D = \{k_1,\, k_2\}]$. The unilateral move of user k_1 (resp. k_2) from the strategy profile $\psi = [\mathcal{A}_F = \emptyset,\, \mathcal{A}_D = \{k_1,\, k_2\}]$ to $\psi = [\mathcal{A}_F = \{k_1\},\, \mathcal{A}_D = \{k_2\}]$ (resp. to $\psi = [\mathcal{A}_F = \{k_2\},\, \mathcal{A}_D = \{k_1\}]$), does not yield any benefit. This game has a unique Nash equilibrium at the pure joint strategy $\psi = [\mathcal{A}_F = \emptyset, \mathcal{A}_D = \{k_1, k_2\}]$ with unsatisfactory payoff distribution $\mathbf{u} = [\,0\,,\,0\,]$. At the Nash equilibrium point either players choose the "competetive" and "egoistic" strategy D.

In many games, there are opportunities for joint deviations that are mutually beneficial for a subset of players. This led Aumann [126] to propose the idea of *strong Nash equilibrium* which ensures a more restrictive stability than the conventional Nash equilibrium. Strong Nash equilibrium reflects the unprofitability of coalition deviations. It is a strategy profile that is stable against deviations not only by single players but by all coalitions of players. A strong equilibrium is defined as a strategic profile for which no subset of players has a joint deviation that strictly benefits all of them, while all other players (in the subset) are expected to maintain their equilibrium strategies.

Definition 17 *A strategy (coalition) combination $\psi = [\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m]$ where $\mathcal{A}_i \bigcap \mathcal{A}_{j\neq i} = \emptyset$ and $\bigcup_{i=1}^{m} \mathcal{A}_i = \mathcal{K}$ with payoff distribution $\mathbf{u} = [u_1, \ldots, u_K]$ is a strong Nash equilibrium if there do not exist a coalition $\mathcal{A}_i \in \psi$ whose deviation yields a new distribution $\mathbf{u}' = [u'_1, \ldots, u'_K]$ such that $u'_k \geq u_k \;\; \forall\, k \in \mathcal{A}_i$ and $\exists\, k \in \mathcal{A}_i$ such that $u'_k > u_k$.*

This definition of strong equilibrium is actually slightly different from those of [126] and [127]. Def. 17 allows a coalition to deviate from a strategy profile that strictly increases the payoffs of some of its members without decreasing those of the other members, whereas the original definition allows only deviations that strictly increase the payoffs of all members of a deviating coalition. We note that if a game implements a strategy for strong equilibrium, it does not necessarily implement it for Nash equilibrium. Both interpretations of strong Nash equilibrium are prominent in the literature, and in most games the two definitions lead to the same sets of strong Nash equilibria; however, the one that we use here is slightly more appealing in the context of network-formation games (see, e.g., [129]). Network formation games involve a number of independent players that interact with each other in order to form a suited graph that connects them.

Now, we restudy the forwarder's dilemma game and try to find strong Nash equilibria profile. We will show that the game possesses strong Nash equilibria which are not equivalent to the Nash equilibrium. We pick different coalition combination and test whether there exist any coalition whose deviation satisfies its own members or not.

1. $\psi = [\mathcal{A}_F = \{k_1\}\, , \, \mathcal{A}_D = \{k_2\}]$ is not strong Nash equilibrium because the deviation of $\mathcal{A}_F$ increases its member's payoff.

2. $\psi = [\mathcal{A}_F = \{k_2\}\, , \, \mathcal{A}_D = \{k_1\}]$ is not strong Nash equilibrium because the deviation of $\mathcal{A}_F$ renders its member's payoff higher.

3. $\psi = [\mathcal{A}_F = \emptyset\, , \, \mathcal{A}_D = \{k_1\, , \, k_2\}]$ is not strong Nash equilibrium because the deviation of both players from $\mathcal{A}_D$ to $\mathcal{A}_F$ increases payoff distribution.

4. $\psi = [\mathcal{A}_F = \{k_1\, , \, k_2\}\, , \, \mathcal{A}_D = \emptyset]$ is strong Nash equilibrium because the departure of one or both players from $\mathcal{A}_F$ to $\mathcal{A}_D$ decreses at least one player's payoff.

The unique strong Nash equilibrium is the strategy profile $(F\, , \, F)$ which corresponds to coalition set of $\psi = [\mathcal{A}_F = \{k_1\, , \, k_2\}\, , \, \mathcal{A}_D = \emptyset]$, since no deviation can better off the payoff distribution vector $\mathbf{u} = [\, 1 - c\, , \, 1 - c\,]$. In fact, at the strong Nash equilibrium, both players choose the "cooperative" and "altruistic" strategy of F in spite of the energy transmission cost.

In network problems, Zhong *et al.* show that using strong Nash equilibria context makes possible a collusion-resistant routing in non-cooperative wireless ad hoc networks [130]. Altman *et al.* in [131] examine a dynamic random access game with orthogonal power constraints in which the probability of transmission of a terminal in each slot depends on the amount of energy left prior to that slot. They show the existence of a strong Nash equilibrium point.

Conventional Nash equilibrium is concerned with the possibilities of only one step deviation by any player. The notion of strong Nash equilibrium requires an agreement not be subject to an improving (one step) deviation by any coalition of players given that all others coalitions be inert. This notion is stronger than Nash equilibrium, but it is not resistant to further deviation by sub-coalitions (the subsets of a coalition). Recognizing this problem, Bernheim *et al.* [127] introduced the notion of *coalition-proof Nash equilibrium*, which requires only that an agreement be immune to improving deviations which are *self-enforcing*. The definition of a self-enforcing deviation is recursive.

Definition 18 *For a singleton coalition, a deviation is self-enforcing if it maximizes the player's payoff. For a coalition of more than one player, a deviation is self-enforcing if: i) it is profitable for all its members, and ii) if there is no further self-enforcing and improving deviation available to a proper sub-coalition of players.*

Generally, a deviation by a coalition is self-enforcing if no sub-coalition has an incentive to initiate a new deviation. In the forwarder's dilemma game, the Nash equilibria is upset by a deviation of the coalition of both players k_1 and k_2. At the pure strategy Nash equilibrium where each player choose strategy D, they each obtain a payoff of 0. By jointly deviating (both choosing F instead) k_1 and k_2 each earn a payoff $1 - c$. This deviation is not self-enforcing even thought the movement to the pure strategy $\psi = [\mathcal{A}_F = \{k_1 , k_2\} , \mathcal{A}_D = \emptyset]$ is profitable for both players. At strong Nash pure strategy (F , F), the player k_1 tempts to move to strategy (D , F) to get more payoff, and player k_2 to that (F , D). Thus, the strong Nash equilibrium is not immune against self-enforceability.

This notion of self-enforceability provides a useful means of distinguishing coalitional deviations that are viable from those that are not resistant to further deviations. With the concept of self-enforceability, our notion of coalition-proofness is easily formulated.

Definition 19 *In a one player game, a strategy is a coalition-proof Nash equilibrium if it maximizes the player utility. In a game with more than one player, a combination strategy is coalition-proof Nash equilibrium, if no sub-coalition has a self-enforcing deviation that makes all its members better off.*

This solution concept requires that there is no sub-coalition that can make a mutually beneficial deviation (keeping the strategies of non-members fixed) in a way that the deviation itself is stable according to the same criterion. In the forwarder's dilemma game, the strong Nash equilibrium profile $(F\,,\,F)$ is not equivalent to coalition-proof Nash equilibrium. This is due to the fact that, the deviation of $\{k_1\} \subset \mathcal{A}_F = \{k_1\,,\,k_2\}$ to the strategy $(D\,,\,F)$ increases payoff of k_1. In this game there does not exist any coalition-proof Nash equilibrium, due to the fact that all pure strategies have at least one self-enforcing deviation.

Bernheim *et al.* [127] note that for 2-person games the set of coalition-proof equilibria coincides with the set of Nash equilibria that are not Pareto dominated by any other Nash equilibrium. However in n-person games ($K \geq 3$) the equilibrium concepts are independent. At coalition-proof Nash equilibrium, the deviations are restricted to be stable themselves against further deviations by sub-coalitions. Moldovanu in [132] discusses the situations of a 3-player game, wherein coalition-proof Nash equilibrium is equivalent to the core set. The conditions under which the set of coalition-proof Nash equilibria coincides with the set of strong Nash equilibria, are formulated by H. Konishi *et al.* in [133].

In the network engineering literature, Félegyházi *et al.* in [134] apply the concept of coalition-proof Nash equilibria to achieve a stable and fair channel allocation solution in a competitive multi-radio multi-channel wireless cognitive network. Gao *et al.* investigate multi-radio multi-channel allocation in multi-hop ad-hoc networks [135]. To better understand the concepts of self-enforceability and coalition-proof Nash equilibrium, let us introduce an intuitive *subcarrier allocation game in an OFDMA network.* Let us focus on three wireless transmitters $\mathcal{K} = \{k_1, k_2, k_3\}$ and an OFDMA base-station with two subcarriers $\mathcal{N} = \{1, 2\}$. Every subcarrier $n \in \mathcal{N}$ has a frequency spacing Δf. Each user $k \in \mathcal{K}$ experiences a Gaussian complex-valued channel gain $|H_{kn}|^2$ on the nth subcarrier to the base station. We assume that each subcarrier can be shared among more than one transmitter. The payoff of each player (transmitter) is defined as the achieved Shannon channel capacity. Each user $k \in \mathcal{K}$ is allowed to either spend a certain power $\overline{p}_k$ on only one choosen subcarrier, or equally divide it among both subcarriers. In the pure strategy a_1, player k transmits with the maximum power $\overline{p}_k$ on subcarrier $n = 1$ and does not transmit any information on subcarrier $n = 2$. The strategy a_2 is contrary to a_1, i.e. exclusively transmitting on subcarrier $n = 2$ with maximum power. Finally strategy a_3 equally divides its power on two subcarriers and exploits transmitting on both tones. The terminal k achieves a channel capacity:

$$R_k = \sum_{n \in \mathcal{N}} R_{kn} \tag{2.20}$$

where R_{kn} is the Shannon capacity achieved by user k on the nth subcarrier:

$$R_{kn} = \Delta f \cdot \log_2 \left(1 + \frac{|H_{kn}|^2 \, p_{kn}}{\sum_{k \neq i \in \mathcal{K}} |H_{in}|^2 \, p_{in} + \sigma_w^2} \right) \tag{2.21}$$

wherin p_{kn} represents the power allocated by terminal k over the nth subcarrier and where the interference term $\sum_{k \neq i \in \mathcal{K}} |H_{in}|^2 \, p_{in}$ is appriximated with a Gaussian random variable of equal mean and variance. Chooisng the strategy a_1 means selecting $p_{k1} = \overline{p}_k$ and $p_{k2} = 0$. For the strategy a_2, $p_{k1} = 0$ and $p_{k2} = \overline{p}_k$, and for strategy a_3, $p_{k1} = p_{k2} = \frac{\overline{p}_k}{2}$. The parameter σ_w^2 is the power of the additive white Gaussian noise (AWGN). Note that, in an OFDMA system, there is no interference between adjacent subcarriers. Hence, R_{kn} considers only intra-subcarrier noise, that occurs when the same subcarrier is shared by more terminals.

Tab. 2.2 reports the simulation results obtained after 100 random realizations of a network with terminals distributed at a distance between 3m and 50m

	k_3 (a_1)			k_3 (a_2)			k_3 (a_3)		
	k_2			k_2			k_2		
	a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3
k_1 a_1	8, 5, 6	11, 11, 7	9, 11, 6	12, 7, 10	15, 10, 10	11, 10, 9	9, 5, 10	12, 9, 10	9, 10, 9
k_1 a_2	12, 8, 6	8, 7, 10	9, 11, 8	11, 11, 8	7, 6, 6	8, 11, 7	13, 9, 11	8, 6, 11	8, 10, 10
k_1 a_3	15, 6, 6	14, 8, 8	13, 10, 7	14, 8, 9	15, 6, 7	13, 9, 7	14, 6, 10	14, 7, 10	12, 9, 10

$\mathbf{u} = [\, u_{k_1} \, , \, u_{k_2} \, , \, u_{k_3}]$

Tab. 2.2: *Subcarrier allocation in OFDMA network game in strategic form. The three strategies for three players k_1, k_2 and k_3 are: Transmitting with the maximum power only on subcarrier number* 1 (a_1)*, transmitting with the maximum power only on subcarrier number* 2 (a_2) *and equal division of the maximum power among both subcarriers* (a_3)*. The player's payoff is the achieved channel capacity in* kb/s.

from the base-station. In the pure strategy matrix form of Tab. 2.2, player k_1 chooses the row, player k_2 chooses the column, and player k_3 chooses the matrix. Each payoff reports the (rounded) value of the achieved Shannon channel capacity in kb/s. We consider the following parameters for our simulations: the maximum power of each terminal k is $\overline{p}_k = 10\,\text{mW}$; the power of the ambient AWGN noise on each subcarrier is $\sigma_w^2 = 100\,\text{pW}$, and finally the carrier spacing is $\Delta f = \frac{10}{1024}$ MHz.[4] The path coefficients $|H_{kn}|^2$, corresponding to the frequency response of the multipath wireless channel, are computed using the 24-tap ITU modified vehicular-B channel model adopted by the IEEE 802.16m standard [136].

It is easy to show that the (pure) Nash equilibrium strategies of Tab. 2.2 are $(a_3\,, a_3\,, a_3)$ equivalent to $\psi = [\mathcal{A}_{a_1} = \emptyset\,, \mathcal{A}_{a_2} = \emptyset\,, \mathcal{A}_{a_3} = \mathcal{K}\,]$ and $(a_1\,, a_2\,, a_2)$ to $\psi = [\mathcal{A}_{a_1} = \{\,k_1\,\}\,,\ \ \mathcal{A}_{a_2} = \{\,k_2\,, k_3\,\}\,, \mathcal{A}_{a_3} = \emptyset]$. The Nash equilibrium strategy $(a_3\,, a_3\,, a_3)$ is neither coalition-proof nor strong. With deviation of the coalition $\mathcal{A}_{a_3}$ to the strategy profile $(a_2\,, a_1\,, a_3)$ all players profit more with payoff distribution [13, 9, 11]. This change is no longer valid since, there exists a self-enforceability for player k_1 to transit to the strategy profile

[4] This is the carrier spacing of each subcarrier at a base station with 10 MHz bandwidth and 1024 subcarriers.

(a_3, a_1, a_3). This transition is not favorable for players k_2 and k_3. The player k_2 is tempted to transit to the Nash equilibrium point to earn a higher payoff. Whereas, the Nash equilibrium strategy profile (a_1, a_2, a_2) with payoff vector $[15, 10, 10]$ is a strong and coalition-proof Nash equilibrium. This is due to the fact that, in $\psi = [\mathcal{A}_{a_1} = \{k_1\}, \mathcal{A}_{a_2} = \{k_2, k_3\}, \mathcal{A}_{a_3} = \emptyset]$ there is no deviation and self-enforceability that can improve the payoff distribution. As can be seen, all players prefer to stay at the coalition-proof Nash equilibrium rather than the pure Nash equilibrium strategy (a_3, a_3, a_3). Note that, a strong or coalition-proof Nash equilibrium does not necessarily coincide with a Nash equilibrium strategy profile, and the result of Tab. 2.2 is an exception.

In general, the existence of a pure cooperative or non-cooperative Nash equilibrium for subcarrier allocation game in OFDMA network is not guaranteed. Given different parameters approaches to quite different channel capacities and this may results a matrix form without any type of Nash equilibrium. There even might exist a Nash equilibrium which is Pareto-dominated by another strategy profile. This shows that in OFDMA networks, an appropriate resource allocation technique is needed.

2.7 Coordinated equilibrium

The most common solution concept in (non-cooperative) game theory, Nash equilibrium, assumes that players take mixed actions independently of each other. Cooperative games allow players to coordinate each other to find out possible equilibria and (joint) optimizations that the players can perform on their own. Unlike evolutionary games [75, Ch. 3], in coordinated games the interaction between players is implemented once, among all players by a central authority, to increase their throughput. The notion of *correlated equilibrium* was introduced by R. Aumann [137]. Correlated equilibria are defined in a context where there is an intermediator who sends random (private or public) signals to the players. An intermediator needs not have any intelligence or knowledge of the game. These signals allow players to coordinate their actions,

		k_2: W	k_2: A
k_1	W	$0\,,0$	$0\,,1-c$
	A	$1-c\,,0$	$-c\,,-c$

$$\mathbf{u} = [\,u_{k_1}\,,\,u_{k_2}\,]$$

Tab. 2.3: *The multiple access game in strategic form. The two moves for each player are: access (A) or wait (W).*

and, in particular, to perform joint randomization over strategies. "Correlated strategies are familiar from cooperative game theory, but their applications in non-cooperative games are less understood", says R. Auman [137]. This is because, the players of a coordination game, are not totally isolated and without a communication between them, achieving to coordinated strategy profile is not possible.

Let us start with an intuitive example. Consider the multiple access game [128, Table III] described in Tab. 2.3. The players k_1 and k_2 wish to send some packets to their receivers sharing a common resource, i.e., the wireless medium. They are in the sight of each other and accordingly, they interferer if transmitting at the same time. The users have two possible pure strategies: access (A) and wait (W). In this game two identical transmitters must simultaneously decide whether to access to channel or wait. The transmission of each packet has an energy cost of $0 < c \ll 1$. Each player earns a payoff 1 if it succeeds to transmit its packet without collision with the other. Waiting does not bring neither cost nor reward for the player. Each player's utility is its reward minus the cost. This game has three Nash equilibria: $(A\,,\,W)$, $(W\,,\,A)$ and a mixed strategy Nash equilibrium where each player transmits with the probability $1-c$ [128, Sec. 2.3, 2.4]. The utilities of Nash equilibria strategies are: $(1-c\,,\,0)$, $(0\,,\,1-c)$ and $(0\,,\,0)$, respectively. It is clear that the

		k_2: W	k_2: A
k_1	W	$\frac{10+32c}{64}$	$\frac{11-16c}{32}$
	A	$\frac{11-16c}{32}$	$\frac{10+32c}{64}$

$$\mathbf{u} = [\, u_{k_1} = u_{k_2} \,]$$

Tab. 2.4: *The strategic form matrix of the multiple access game with preplay agreement.*

mixed strategy is not resistant to an improving deviation. In the following, we give the possibility of preplay communication to achieve a stable Nash equilibria.

In the game with "cheap conversation", each player simultaneously and publicly announces whether it decides to access or wait. Following the announcements each player makes its choice. Suppose the players agree to participate to the game binding the following agreement: each player announces A with probability $\frac{3}{4}$. If the profile of announcements is either $(A\,,\,W)$ or $(W\,,\,A)$, then each player plays its own announcement. Otherwise, each player plays A with probability $\frac{1}{2}$. Note that no further communication is possible. The use of joint deviation requires the unanimity of all members of the deviating coalition. A player agrees to be a part of a joint deviation if given its own information the deviation is profitable. Thus, if a joint deviation is used, it is common knowledge that each deviator believes that deviation is profitable. This tradeoff results in an expected payoff for each player of $\frac{11-16c}{32} > 0$, while in the mixed Nash equilibrium of the original game each player has an expected payoff of 0. In this coordinated Nash equilibrium of the game, the players effectively play the *correlated strategy* [137,138] (of the original game) given in Tab. 2.4, in order to face a higher utility in strategy profiles $(A\,,\,W)$ and $(W\,,\,A)$. It is important to note that, this joint probability distribution is replacemen

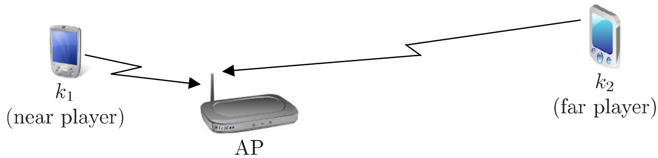

Fig. 2.3: *The network scenario in the near-far effect game.*

not the product of its marginal distributions and therefore cannot be achieved from a mixed strategy profile of the game without correlation among players.

As can be seen, the proposed correlated deviation from the mixed strategy equilibrium makes both players better off. Note that the players are allowed to bind an agreement only on the space of feasible outcomes. In the correlated multiple access game the outcome is feasible since the correlated results are in the range between the smallest and highest possible payoff. In fact, the set of correlated equilibria contains those equilibria from which no coalition has a self-enforcing deviation making all members better off.

Let us describe a more complicated correlated equilibrium. We study the near-far effect game established by G. Bacci *et al.* in [139, Fig. 6]. The basic idea of near-far effect game scheme is depicted in Fig. 2.3. Two wireless terminals k_1 and k_2, are placed close to and far from a certain access point (AP), respectively, in a code division multiple access (CDMA) network with high SINR regime. The strategy of each player is either to transmit with the maximum power $\overline{p}$, or with a weakened level $\eta\overline{p}$, where $0 < \eta < 1$. Due to the interference at the AP, the throughput (the amount of delivered information) of each player depends on the strategies chosen by both players. Transmitting with a higher power increases the BER, and this results decreasing the throughput. Each player is rewarded u if it successfully delivers its packet and a reduced δu, if it delivers a corrupted version of the packet, where $0 < \eta < \delta < 1$. If the near player k_1 decides to transmit with the power $\overline{p}$, the farther player k_2 will not be able to deliver any information to the AP.

This results in no benefit for k_2 and causes a power consumption cost equal

		k_2: $\eta\overline{p}$	k_2: $\overline{p}$
k_1	$\eta\overline{p}$	$\delta u - \eta c\,,\, -\eta c$	$\delta u - \eta c\,,\, \delta u - c$
k_1	$\overline{p}$	$u - c\,,\, -\eta c$	$u - c\,,\, -c$

$$\mathbf{u} = [\, u_{k_1}\,,\, u_{k_2}\,]$$

Tab. 2.5: *Payoff matrix for the near-far effect game with power control and variable throughput.*

to $-\eta c$ if k_2 chooses strategy $\eta\overline{p}$ and $-c$ otherwise, where $c \ll u$. Obviously, transmitting with power $\overline{p}$ for k_1, results in a complete information delivery. This concerns a payoff equal to reward minus power consumption cost, i.e. $u - c$, irrespective of the k_2 strategy. The packets of player k_2 are successfully delivered if it chooses the maximum power $\overline{p}$, and player k_1 that reduced $\eta\overline{p}$. On the other hand, if both players decide to transmit with reduced power $\eta\overline{p}$, the near player takes the payoff $\delta u - \eta c > 0$, whilst the farther player k_2 will not successfully deliver any packet and suffers only a power cost $-\eta c$.

The payoff matrix of the near-far effect game is depicted in Tab. 2.5. As can be seen, the unique pure strategy of this game is represented by the strategy $(\overline{p}\,,\,\eta\overline{p})$ with benefits $u - c$ and $-\eta c$ for k_1 and k_2, respectively. This means that, at the Nash equilibrium point, the farther player is not able to send any information. On the other hand, the Pareto optimal solution of the game are the strategies $(\overline{p}\,,\,\eta\overline{p})$ and $(\eta\overline{p}\,,\,\overline{p})$. This is an unsatisfactory outcome for the far player k_2, while the near player k_1 takes the highest possible payoff. Now, let us find the mixed strategy of the game. We denote α_1 the probability with which the near player k_1 decides to transmit with the maximum power $\overline{p}$ and α_2 the same probability for the far player k_2. The payoffs of the players k_1

and k_2 are represented by:

$$u_{k_1} = \alpha_1\Big(\,(1-\delta)\,u \;-\; (1-\eta)\,c\Big) + (\delta u - \eta c) \tag{2.22a}$$

$$u_{k_2} = \alpha_2\Big(\,(1-\alpha_1)\,\delta u \;-\; (1-\eta)\,c\Big) - \eta c \tag{2.22b}$$

Both players want to maximize their own payoff. As can be seen, u_{k_1} takes its maximum value $u - c$ with $\alpha_1 = 1$. On the other hand, with $\alpha_1 = 1$, the far player k_2 earns a negative payoff whatever $\alpha_2 \in [0, 1]$. Instead, with $\alpha_1 = 0$ the near player k_1 gains $\delta u - \eta c$, and the player k_2 setting up $\alpha_2 = 1$ achieves the payoff of $\delta u - c$. Thus, the best values for α_1 and α_2 are 0 and 1, respectively. The conclusion is that the mixed strategy is equivalent to the pure strategy $(\eta\overline{p}\,,\,\overline{p})$ with payoff $\mathbf{u} = [\,\delta u - \eta c\,,\,\delta u - c\,]$. In this game there is no (totally) mixed strategy and that is equal to the one of the pure Pareto optimal points.

The near player earns the highest possible payoff at the Nash equilibrium, hence, it does not leave this strategy profile. The highest possible payoff for the far player is on the contrary $\delta u - c$. We show that an appropriate agreement among players can satisfy both of them at correlated equilibrium. Players k_1 and k_2 can guarantee an expected payoff of $\mathbf{u} = [\,u - c\,,\,\delta u - c\,]$ by playing the correlated strategy profile:

$$\frac{\delta u}{(1-\eta)\,c}\cdot(\overline{p}\,,\,\eta\overline{p}) \;+\; \left(1 - \frac{\delta u}{(1-\eta)\,c}\right)\cdot(\overline{p}\,,\,\overline{p}) \tag{2.23}$$

This is a plausible end, since both players earn their own highest possible payoff. The correlated strategy (2.23) is derived from the fact that, picking any real number κ in the expression $\kappa\cdot(\overline{p}\,,\,\eta\overline{p}) + (1-\kappa)\cdot(\overline{p}\,,\,\overline{p})$ is indifferent for the near player k_1, since it gets its own highest possible payoff, $u - c$ as well. To satisfy the far player k_2, it is enough to solve the following equation for u_{k_2}:

$$\kappa\cdot(\overline{p}\,,\,\eta\overline{p}) + (1-\kappa)\cdot(\overline{p}\,,\,\overline{p}) = [\,u - c\,,\,\delta u - c\,] \tag{2.24}$$

Supposing $\kappa = \dfrac{\delta u}{(1-\eta)\,c} < 1$, the correlated strategy (2.23) means that the near player always transmits at its highest power level $\overline{p}$, and the far player

		k_2: $\eta\overline{p}$	k_2: $\overline{p}$
k_1	$\eta\overline{p}$	$0\,,\ 0$	$0\,,\ 0$
	$\overline{p}$	$\kappa\,,\ \kappa\dfrac{(\delta u - c)}{(u-c)}$	$1-\kappa\,,\,(1-\kappa)\,\dfrac{(\delta u - c)}{(u-c)}$

$\mathbf{u} = [\,u_{k_1}\,,\,u_{k_2}\,]$ normalized to 1

Tab. 2.6: *The strategic form matrix of the near-far effect game with preplay agreement, and with* $\kappa = \dfrac{\delta u}{(1-\eta)\,c}$.

transmits at that reduced $\eta\overline{p}$ with probability $\dfrac{\delta u}{(1-\eta)\,c}$, and the maximum power $\overline{p}$ otherwise. Actually, the near and far players effectively play the matrix form game of Tab. 2.6.

Bonneau *et al.* in [140] show that the coordination among mobile users can significantly increase the performance of access to a common channel in ALOHA setting. A coordination mechanism is also considered by Bonneau *et al.* in [141] to achieve the optimal power allocation in a wireless network wherein each terminal knows only its own channel state. The concept of correlated equilibrium is also introduced in a multi-user interference channel context in [142]. Different types of coordination is deeply discussed and widely used in [138].

2.8 Dynamic learning

Until now, we have realized that the Nash equilibrium suffers from the lack of farsighted stability, i.e., the relative results can be unsatisfactory and because of this any player can have incentive to improve its outcome by moving to another strategy. The existence of the strong and coalition-proof Nash equilibrium is not guaranteed and even if so, when the number of pure strategies is large, finding such solutions is very complicated. The challenge of finding a profitable accord among players is persistent in coordinated equilibria solution. In this section, the main question we seek an answer to is: *How can the players be led to a stable joint pure strategy gaining an acceptable payoff?* This question is important, even if multiple equilibrium points with the same payoff have been identified, since each player may autonomously decide to stay in a different strategy.

Dynamic learning [143] has been widely used in order to get rid of the anarchy derived from the conflicts between selfish decisions. Learning is a joint adaptive process for agents to converge and to get the best final response. The agents either have a common interest like a team work, or each agent has its own greedy goal. Generally, there are three learning process types: *individual learning*, *joint-action learning* and *stochastic learning*. In individual learning process, the independent agents cannot observe one another's actions; i.e. for each players the opponents are passive agents. Instead, during joint-action learning process, the notion of the "optimality" is improved by adding the observation of other concurrent learners to accomplish a stable optimal solution. The stochastic learning framework, having Markovian property and a stochastic inter-state transition rule, enables each player to observe the opponents' actions history.

In the network engineering literature, Schaar *et al.* in [144] introduce a stochastic learning process among autonomous wireless agents for the optimization of dynamic spectrum access given QoS of multimedia applications. A reconfigurable multi-hop wireless network is studied by Shiang *et al.* [145] wherein a decentralized stochastic learning process optimizes the transmission decisions of nodes aimed at supporting mission-critical applications. In [146], Lin *et al.* propose a reinforcement learning among agents of a multi-hop wireless network based on Markov decision process. Each terminal autonomously adjust transmission power in order to maximize the network utility, in a dynamic delay-sensitive environment.

Here, we study a well-known individual reinforcement learning task, namely the so-called *Q-Learning* [147]. We assume a set of players $\mathcal{K}$, and each player k has a finite set of individual actions $\mathbf{A}_k$. Each agent k individually chooses a pure joint action (strategy) to be performed $\mathbf{a}_k = (a_1, \ldots, a_K) \in \mathbf{A}_1 \times \ldots \times \mathbf{A}_K$ from the available joint strategy space. Q-learning enables the individual learners to achieve optimal coordination from repeated trials. Q-learning introduces a certain value Q as the immediate reward obtained after having moved to the new strategy. Each player individually updates a Q value for each of its actions. In each time step and after having selected the new joint action $\mathbf{a}_k$, the values of Q_k^t is individually updated. In particular, the value of $Q_k^{t+1}(\mathbf{a}_k)$ estimates the utility of performing the joint strategy $\mathbf{a}_k$ for user k. In the seminal paper of Watkins *et al.* [147], the Q value is updated by the following recursion:

$$Q_k^{t+1}(\mathbf{a}_k) \longleftarrow \left(1 - f_k^{t+1}\right) \cdot Q_k^t(\mathbf{a}_k) + f_k^{t+1} \cdot \left(u_k + \delta_k \cdot Q_k^t(\mathbf{a}_k)\right) \tag{2.25}$$

where $\delta_k \in (0, 1)$ is a discount factor and u_k is a reward of the joint action $\mathbf{a}_k$ for the respective player, and f_k is a function of t which is related to "learning rate". Watkins *et al.* showed that given bounded rewards, learning rate $0 \leq f_k^t < 1$, and

$$\sum_{t=1}^{\infty} f_k^t = \infty\,, \quad \text{and} \quad \sum_{t=1}^{\infty} \left(f_k^t\right)^2 < \infty \qquad \forall k \in \mathcal{K} \tag{2.26}$$

all Q_k values updating (2.25) converge a common joint pure strategy with probability one. The reward u_k is defined by a learning policy and it is not necessarily equal to the payoff defined by the game. The learning policy is greedy with respect to the Q value, i.e. the particular action $\mathbf{a}_k$ will be selected in long-run if it makes Q value better off. Q-learning is guaranteed to converge to an optimal and stable joint strategy regardless of the action selection policy. Q-learning is not applicable where the strategy space is continuous or the number of strategies is not finite. Claus *et al.* [148] establish a simplified version of the Q recursion (2.25) which updates the Q value by the following recursion:

$$Q_k^{t+1}(\mathbf{a}_k) \longleftarrow Q_k^t(\mathbf{a}_k) + \delta_k \cdot \left(u_k - Q_k^t(\mathbf{a}_k)\right) \tag{2.27}$$

For the sake of simplicity, we apply the Q recursion (2.27). In a multi learners scenario, a major challenge of Q-learning is strategy selection. When the number of strategies and players are large, the number of time step to achieve an optimal joint action exponentially increases. It is fairly clear that the best manner is to start with "exploration" of different strategies and then focus on "exploitation" of the strategies with the best value of Q. Kaelbling *et al.* in [149] recall *Boltzmann function* as an efficient strategy selection to strike a balance between exploration and exploitation. Boltzmann functions define a probability distribution among different joint actions. At each time step $t+1$, every player will individually select the joint strategy $\mathbf{a}_k$ with the probability $p(\mathbf{a}_k)$:

$$p(\mathbf{a}_k) = \frac{e^{Q_k^t(\mathbf{a}_k)\,/\,T}}{\sum\limits_{\forall\, \mathbf{a}_i \in \underset{\forall k \in \mathcal{K}}{\times} \mathbf{A}_k} e^{Q_k^t(\mathbf{a}_i)\,/\,T}} \tag{2.28}$$

The T is a function which provides a randomness component to control exploration and exploitation of the actions. Practically, the "temperature function" T is a decreasing function over time to decrease the exploration and increase exploitation. High values of T yields a small $p(\mathbf{a}_k)$ value and this encourages exploration, whereas a low T makes $Q(\mathbf{a}_k)$ more important, and this encourages exploitation. At time $t = 0$, each player randomly chooses a strategy and assign a random number to its own Q value. At time step t, after having been updated funtion T, each concurrent agents' experience consists of a sequence of stages [148]:

1) Computing $p(\mathbf{a}_k)$ for all $\mathbf{a}_k \in \underset{\forall k \in \mathcal{K}}{\times} \mathbf{A}_k$,

2) Generating a random number ξ_k^t uniformly distributed in $[0, 1]$, and then choosing the best joint strategy $\mathbf{a}_k$, i.e. the highest $p(\mathbf{a}_k)$ such that $\xi_k^t \geq p(\mathbf{a}_k)$. If $\xi_k^t < p(\mathbf{a}_k)$ for all $\mathbf{a}_k \in \underset{\forall k \in \mathcal{K}}{\times} \mathbf{A}_k$, then the learner randomly picks a strategy,

3) Updating the Q_k^t value according to (2.27). If Q_k^t grows, then the learner

moves to selected joint strategy $\mathbf{a}_k$, otherwise it stays in the current joint action and do not update Q.

Despite the individual best strategy selection of the learners, this process reach a common stable joint strategy such that all players stay there forever, i.e. no player deviates from the (common) achieved joint strategy.

2.9 Discussion

This chapter has provided a unified reference for network engineers investigating the applicability of cooperative game theory to practical problems. Different approaches such as core solution, Shapley value, kernel and nucleolus, were shown to provide a strong foundation for finding possible and stable resource/cost sharing arrangements. The results confirm the apparent analogy between the definition of Nash equilibrium in non-cooperative and cooperative game theory: both strong and coalition-proof Nash equilibria reflect on unprofitability of coalition deviations rather than an individual player deviation. In a network wherein informational exchange is possible, either through a central controller or among players themselves, the concept of coordinated equilibrium arises. The results of intuitive examples show a significantly improvement in coordinated equilibrium when compared with non-cooperative schemes. When the number of agents or strategies is large, the ability of jointly reach a consensus through environmental learning guarantees convergence to the best joint action.

Chapter 3

A resource allocation cooperative game in OFDMA

Following what discussed in Chapter 1, various attempts have been made to standardize a certain protocol for resource allocation in OFDMA, but all have fallen into disuse largely because of their over-complexity, and unfairness from the network service provider point of view. We also showed that cooperative game theory is a suitable tool to face resource allocation problems, especially when altruism and fairness play crucial roles.

The focus of the first part of this thesis is to introduce a scheme for resource allocation in OFDMA based on cooperative games, wherein applicability and fairness are target criterions. This chapter investigates a fair adaptive resource management criterion for the uplink of an OFDMA network populated by mobile users with constraints in terms of target data rates. We aim at fulfilling each users QoS requirement in terms of target transmit rates *exactly* with the best utilization of the network resources, so as to satisfy both the users and the wireless service provider. We also aim at designing a low-complexity algorithm that allows a centralized solution for the joint power and bandwidth allocation for OFDMA uplink channels to be achieved in a few steps using typical network parameters. In our approach, we allow every subcarrier to be possibly *shared* among more than one user, and we add a constraint on the maximum number of used subcarriers per terminal. This is achieved by dividing the available bandwidth into a number of disjoint blocks of consecutive subcarriers, and forcing each terminal to use at most one subcarrier per block. The motivation of this is twofold: we wish to i) increase the signal-to-interference-plus-noise ratio (SINR) on the used subcarriers, which also simplifies channel estimation; and ii) exploit frequency diversity across carriers used by one user to increase the performance of forward error correction (FEC) techniques.

In Sect. 3.1 we propose two methods to allot subcarriers to mobile terminals, in a possibly shared assignment. Next, the inherent optimization problem is tackled with the analytical tools of cooperative game theory aiming at accomplishment of data rate demanded exactly. The definition of the players, coalitions and utility function are formally discussed, and we prove the existence of the core set solution by means of the analytical tools of cooperative game theory. To accomplish data rate demanded we propose a utility function which is neither convex nor concave. Sect. 3.2 proposes a dynamic learning algorithm based on Markov modeling to achieve optimum transmit power over each subcarrier at each individual wireless terminal. Simulation results in Sect. 3.3 show that the average number of operations of the proposed algorithm is much lower than $K \cdot N$, where N and K are the number of subcarriers and users. We also show that the transmit power is comparable to the remarkable existing power effective literature results. Low-complexity, efficient use of available spectrum, and low power consumption bring promise to usability of the proposed scheme in each time slot at physical layer in the 4th generation (4G) of cellular networks.

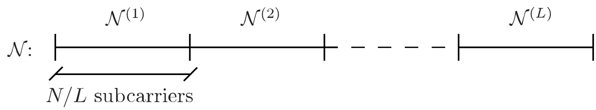

Fig. 3.1: *Block partitioning of the available bandwidth.*

3.1 Problem formulation

Let us consider the uplink of a single-cell infrastructure OFDMA system with total bandwidth W, subdivided in N subcarriers with frequency spacing $\Delta f = W/N$. The cell is populated by K mobile terminals, each terminal $k \in \mathcal{K} = [1, \ldots, K]$ experiencing a complex-valued channel gain H_{kn} on the nth subcarrier to the base station and having a data rate requirement $R_k^\star$ (in bit/s). We assume that fulfilling such constraints simultaneously by all terminals is feasible. To exploit frequency diversity, the subcarriers set $\mathcal{N} = [1, \ldots, N]$ is grouped in L blocks of N/L contiguous subcarriers $\mathcal{N}^{(l)} = \left[\frac{N}{L}(l-1)+1, \ldots, \frac{N}{L}l\right] \subset \mathcal{N}$, with $1 \leq l \leq L$, as shown in Fig. 3.1. Each terminal is allowed to take at most one subcarrier per each subblock.

Our resource allocation strategy consists in finding a vector of transmit powers $\mathbf{p}_k$, where $\mathbf{p}_k = \left[p_{k1}, \ldots, p_{kN}\right]$, with p_{kn} representing the power allocated by terminal k over its nth subcarrier, that allows the QoS constraint $R_k^\star$ to be satisfied. We decouple the problem into the subsequential resolution of subchannel assignment and (subsequent) power allocation.

3.1.1 Subchannel assignment

We describe here two different options to perform this function:

Best-carrier assignment

For every subblock $\mathcal{N}^{(l)}$, every terminal $k \in \mathcal{K}$ is assigned its best subcarrier $n_k^{(l)} = \arg\max_{n \in \mathcal{N}^{(l)}} |H_{kn}|^2$. The probability of assigning the same subcarrier to multiple mobile terminals is non-null.

Vacant-carrier assignment

In a sequential manner, for every subblock $\mathcal{N}^{(l)}$, every terminal $k \in \mathcal{K}$ is assigned its best subcarrier $n_k^{(l)} = \arg\max_{n \in \mathcal{N}^{(l)}} |H_{kn}|^2$. But, if $k \leq N/L$, we would like to ensure exclusive use of each subcarrier $n \in \mathcal{N}^{(l)}$ to better exploit the available bandwidth W (i.e., to reduce the multiple access interference). So, if $n_k^{(l)}$ has been already assigned to some other terminal $\ell < k$, then terminal k is assigned the nearest vacant (unassigned) subcarrier to $n_k^{(l)}$ within the channel coherence bandwidth. Clearly, this is not considered if $k > N/L$, so that terminal k is assigned its best subcarrier in the subblock anyway. Note that the ordering of $\mathcal{K}$ has a negligible impact on system performance when N is sufficiently high and, as usual, $N \gg K$ (e.g., 2048 subcarriers in LTE).

Both assignment strategies can be easily extended to the case in which each terminal is allowed to have a different number of assigned subcarriers (different L for each mobile terminal), based on its own data rate requirement $R_k^{\star}$, without any change in the strategy that we describe below. For the sake of simplicity, we consider the same L for all terminals.

3.1.2 Power allocation

To derive a stable solution to the power allocation subproblem, we consider it as a coalitional game, in which each subchannel $n_k^{(l)} \in \mathcal{N}$ is identified as a player in the game. To model the coalitional game, we build K coalitions $\psi = [\mathcal{A}_1, \ldots, \mathcal{A}_K]$, to be assigned to the K terminals. Each coalition $\mathcal{A}_k$, $k \in \mathcal{K}$, contains the L players $n_k^{(l)}$: $\mathcal{A}_k = \left[n_k^{(1)}, \ldots, n_k^{(L)}\right]$. Note that i) the members of each coalition are fixed, since one player cannot move from one coalition to another; and ii) since a subcarrier $n \in \mathcal{N}$ can be shared among multiple users, there exist virtual copies of it belonging to different coalitions. For the sake of notation, we will identify with a generic $n \in \mathcal{A}_k$ any of the subcarriers assigned to terminal k. The strategy of each player $n \in \mathcal{A}_k$ is represented by the optimal power expenditure $p_{kn} \leq \overline{p}_{kn}$. Note that i) if $n \notin \mathcal{A}_k$, $p_{kn} = 0$; and ii) if $n \in \mathcal{A}_k$, we can also have $p_{kn} = 0$, which means that the kth terminal does not transmit on the nth subcarrier, and it thus bears an actual number of active subcarriers $L'_k < L$.

The system under investigation aims at fulfilling the QoS requirement of every terminal k in terms of target rate $R_k^\star$. For simplicity, we estimate the achieved data rate as the Shannon capacity R_k of terminal k, that can be approached by using suitable channel coding techniques [150]:

$$R_k = \sum_{n \in \mathcal{N}} R_{kn} \tag{3.1}$$

where R_{kn} is the Shannon capacity achieved by terminal k on subcarrier n:

$$R_{kn} = \Delta f \cdot \log_2 \left(1 + \frac{|H_{kn}|^2 p_{kn}}{\sum_{m \neq k} |H_{mn}|^2 p_{mn} + \sigma_w^2} \right). \tag{3.2}$$

Clearly, $R_{kn} = 0$ if $n \notin \mathcal{A}_k$, since $p_{kn} = 0$. If $n \in \mathcal{A}_k$, R_{kn} depends on the received SINR at the base station on subcarrier n, which is a function of the strategy (i.e., the transmit power) chosen by player n (i.e., one of the L subcarriers assigned to the kth terminal), of the transmit power of other terminals on the same subcarrier (if $n \notin \mathcal{A}_m$, $p_{mn} = 0$), of the corresponding channel gains, and of the power of the additive white Gaussian noise (AWGN) σ_w^2. Note that, in an OFDMA system, there is no interference between adjacent subcarriers. Hence, R_{kn} considers only intra-subcarrier noise, that occurs when the same subcarrier is shared by more terminals. Each player $n \in \mathcal{A}_k$ causes interference only to its virtual copies, i.e. to the players of other coalitions such that $n_m^{(l')} = n \in \mathcal{A}_m$, with $m \neq k$ and for any l', $1 \leq l' \leq L$.

The network service provider are satisfied at most when each mobile terminal k achieves its own data rate requirement *exactly*: $R_k = R_k^\star$. In view of this goal, we can force all players in each coalition $\mathcal{A}_k$ to select their strategies (i.e., the power allocation for terminal k over the available bandwidth W) so as to maximize a utility function for the kth coalition $\mathcal{A}_k$, defined as:

$$\nu(\mathcal{A}_k) = \frac{1}{|R_k / R_k^\star - 1|} - \beta \cdot \mathrm{u}\left(1 - R_k / R_k^\star\right) \tag{3.3}$$

where $\mathrm{u}(\cdot)$ is the unit step function, with $\mathrm{u}(y) = 1$ if $y \geq 0$ and $\mathrm{u}(y) = 0$ otherwise (see Fig. 3.2).

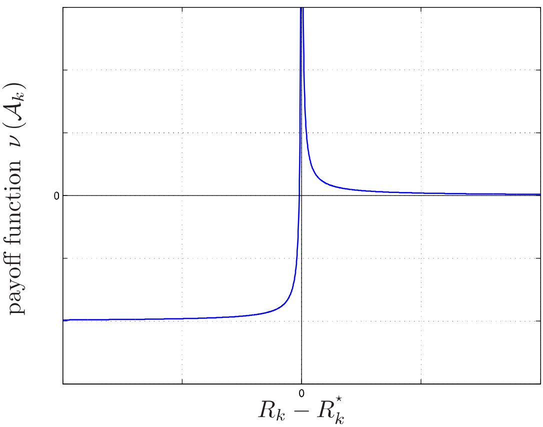

Fig. 3.2: *Shape of the utility as a function of the Shannon capacity.*

If $R_k = R_k^\star$, $\mathcal{A}_k$ earns the highest possible payoff $\nu(\mathcal{A}_k) = +\infty$. If $R_k > R_k^\star$, $\mathcal{A}_k$ gets a positive payoff, whereas it obtains a negative payoff if $R_k < R_k^\star$. The factor β is a positive constant (much) greater than zero that ensures $\nu(\mathcal{A}_k)$ to be negative when $R_k < R_k^\star$. This is expedient to let the players distinguish a capacity R_k that is lower/upper than $R_k^\star$ only by knowing their own coalition's payoff. Note that, in practice, $+\infty$ can be represented by the largest countable number available (e.g., $2^{64} - 1$) in a given computational platform.

The payoff of each coalition is a real number and, in our formulation, the most important parameter is the gain of each coalition, whereas the outcome of each player does not matter at all. Therefore, this game is a transferable utility (TU) one [75, 76]. The specific shape of our utility function (3.3) is actually immaterial, and was chosen to ensure fast convergence of the iterative algorithm that will be introduced later on. We could have considered any utility function which increases as its argument moves from $\pm\infty$ to 0, just to make sure that, for any $R_k \neq R_k^\star$, each coalition has an incentive to move towards $R_k = R_k^\star$.

To provide further insight into the problem, we investigate now some properties of the proposed game $\mathcal{G}$. As a first step, we note that the players in $\mathcal{G} = (\mathcal{K} = \bigcup_{k\in\mathcal{K}} \mathcal{A}_k, \nu)$ with the utility function (3.3), do *not* tend to form the grand coalition. This is because every player $n \in \mathcal{A}_k$ can not leave its coalition $\mathcal{A}_k$: the members of each coalition are fixed and do not change during the game. This may appear inappropriate to the notion of a coalitional game. However, our assumption is fairly common in economic problems like the study of a bargaining game between two corporations when each corporation has its own business branches [151]. In this case the members (branches) of each coalition (corporation) are fixed.

A relevant result for our game is the following:

Theorem 3 *The core of the game* $\mathcal{G} = (\mathcal{K} = \bigcup_{k\in\mathcal{K}} \mathcal{A}_k, \nu)$ *with utility function (3.3) is not empty.*

Proof The number of coalitions and the number of players in each coalition are both fixed. Since each player belongs just to one coalition, the unique balanced collection of weights $(\mu_{\mathcal{A}})_{\mathcal{A}\in\psi}$ is $\mu_{\mathcal{A}} = 1 \;\; \forall\, \mathcal{A} \in \psi$. To conclude the proof, we must verify that $\sum_{\mathcal{A}\in\psi} \nu(\mathcal{A}) \leq \max_{\psi\in\boldsymbol{\Psi}} \sum_{\mathcal{A}\in\psi} \nu(\mathcal{A})$. Since the target rates of all terminals are assumed to be feasible, then every coalition expects R_k to approach $R_k^{\star}$. Therefore, every coalition is allowed to earn the highest possible payoff. ■

In the following section, we will show how the fundamental properties of our game lead to a practical allocation algorithm.

3.2 The best-response algorithm

We are interested in answering questions like: *How do the players set their proper transmit powers?* Dynamic learning models provide a framework for analyzing the way the players may set their proper strategies. A player adopts a certain power amount if and only if this matches its coalition's interests, and this goal can be achieved through a best-response iterative algorithm [152] based on Markov modeling [153]. Each player takes its own decisions individually, myopically, and concurrently with the others, so as to lead its own coalition's payoff toward $+\infty$ ($R_k = R_k^\star$). At each (discrete) time step of the algorithm, the (autonomous) players simultaneously adjust their transmit powers based on a model to increase the payoff of their own coalitions. Although this leads to interference when virtual copies of the same subcarriers simultaneously change their powers, we show that this dynamic myopic procedure guarantees the maximum payoff to each coalition.

The process starts up at time step $t = 0$ with an arbitrary assignment of the transmit powers $p_{kn}^{t=0}$ to all $K \cdot L$ players in the game (that are grouped in K coalitions with players $n \in \mathcal{A}_k$ with $n = n_k^{(l)}$, $1 \leq l \leq L$). At the generic time step t, our system is in the state $\omega^t = (\psi^t, \boldsymbol{\nu}^t)$, where ψ^t is the set $[\mathcal{A}_1^t, \ldots, \mathcal{A}_K^t]$, and $\boldsymbol{\nu}^t = [\nu(\mathcal{A}_1^t), \ldots, \nu(\mathcal{A}_K^t)] \in \mathbb{R}^K$ contains the payoffs of the coalitions in ψ^t. The evolution of the Markov chain is then dictated by the strategy of the game. The strategy of each player $n \in \mathcal{A}_k$ is to find the best power amount p_{kn}^t that leads to an increase in the payoff $\nu(\mathcal{A}_k^t)$ of its own coalition $\mathcal{A}_k$. In practice, player $n \in \mathcal{A}_k$ decides whether to change its power allocation, making its coalition better off, or to keep transmitting at the same power level (e.g., when its coalition's payoff is infinite). The following pseudocode shows how each player $n \in \mathcal{A}_k$ takes its decision at time step t:

```
if ν(A_k^t) = +∞, then p_kn^{t+1} = p_kn^t, exit;
else //setting correct power range
    if ν(A_k^t) ≤ 0, then p̃_kn = p_kn^t, p̃_kn^max = p̄_kn;
                     else p̃_kn = 0,    p̃_kn^max = p_kn^t;
repeat
    p̂_kn = p̃_kn; //saving tentative power
    compute  ν(Ã_k); //tentative payoff
    Δp̃_kn = unif[0, Δp̄_kn]; //random power step
    p̃_kn = p̃_kn + Δp̃_kn; //tentative power
 until (ν(Ã_k) > ν(A_k^t)) or (p̃_kn > p̃_kn^max)
 if (ν(Ã_k) > ν(A_k^t)), then p_kn^{t+1} = p̂_kn; //accept
                         else p_kn^{t+1} = p_kn^t; //discard
```

In this algorithm, $\nu(\tilde{\mathcal{A}}_k)$ is the "trial" value of the current payoff of the coalition when the tentative power $\tilde{p}_{kn}$ is adopted: it is computed with $p_{mn} = p_{mn}^t$ for all $n \in \mathcal{N}$ and for any $m \neq k$, and $p_{kn} = \tilde{p}_{kn}$. At each step of the update process, the power step $\Delta\tilde{p}_{kn}$ is the particular outcome (value) of a random variable uniformly distributed between 0 and $\overline{\Delta p}_{kn}$, with $\overline{\Delta p}_{kn} \ll \overline{p}_{kn}$. As better detailed in Sect. 3.3, optimal values for $\overline{\Delta p}_{kn}$ can be found in order to minimize the algorithm computational load, based on experimental results. If $\nu(\mathcal{A}_k^t) \leq 0$, then $R_k < R_k^\star$, and the best strategy for player $n \in \mathcal{A}_k$ is to increase its current transmit power so as to increase its coalition's payoff. As a result of the random power stepping, the tentative power is a random number in the interval $[p_{kn}^t, \overline{p}_{kn}]$. Player $n \in \mathcal{A}_k$ accepts this value if and only if the coalition payoff $\nu(\mathcal{A}_k^t)$ increases, otherwise it ends up transmitting at its previous value. If $0 < \nu(\mathcal{A}_k^t) < \infty$, player $n \in \mathcal{A}_k$'s best strategy is on the contrary to decrease p_{kn}^t, and thus the tentative (random) transmit power belongs to the interval $[0, p_{kn}^t]$. At the end of each time step t, the base station computes the payoff $\nu(\mathcal{A}_k), \forall k \in \mathcal{K}$ with updated power amounts. As shown in the pseudocode, a uniformly distributed random power stepping is adopted to increase the probability of picking the best adjustment value, and thus both to reduce the convergence time of the algorithm and to possibly minimize the overall power consumption. As is apparent, the convergence speed of the algorithms depends not only on the parameters of the network, but also on the choice of the maximum update step $\overline{\Delta p}_{kn}$.

As already stated, two copies $n \in \mathcal{A}_k$ and $n \in \mathcal{A}_m$ (the virtual copies of the same subcarrier n) may happen to wish to adjust their transmit powers in a conflicting (and thus incompatible) way. If we assume that each player just follows the decision rules listed in the pseudocode above, then the probability of conflicting decisions will be high. To reduce the occurrence of this event, we modify our algorithm by requesting each player *not* to update its transmit power at every step of the game with a probability $\lambda \in [0, 1]$. At each time step t, every player $n \in \mathcal{A}_k$ selects a random number ξ_{kn}^t uniformly distributed in $[0, 1]$. If $\xi_{kn}^t > \lambda$, then the player applies the algorithm and (possibly) update p_{kn}^{t+1}, otherwise $p_{kn}^{t+1} = p_{kn}^t$ (i.e., during time step t, it skips the update process, and the value of p^t_{kn} is kept). If λ is close to 1, then the probability of conflicting decisions tends to 0, but the algorithm will have a large convergence time, since the probability of updates is low. In addition to the conflicts described above, another potentially disruptive condition may arise between different subcarriers belonging to the same coalition: if both (myopic) players simultaneously increase their powers $p^t_{kn} > 0$ and $p^t_{kn'} > 0$, it may occur that $R_k > R^\star_k$. To optimize the update mechanism and to cope with both negative kinds of events, we could consider a variable and adaptive threshold λ^t_{kn} for each virtual copy of the same subcarrier (each player). However, to reduce the complexity of the algorithm, we assume $\lambda^t_{kn} = \lambda > 0$ for all the players (i.e., virtual copies of the subcarriers). As better detailed in Sect. 3.3, the optimal value of λ must be selected as a suited trade-off. Note that the value of λ is common knowledge among the players at every step of the algorithm. Nevertheless, interference between concurrent, conflicting decisions may prevent the coalitions from achieving the expected payoff. If all coalitions earn less than the previous time step, all players assign the previous power amount for the next time step. There may exist network configurations in which the iterative algorithm is not guaranteed to converge. To account for these situations, we place a maximum number of operations Θ, beyond which the algorithm is stopped, and the sum of the users demands is thus labeled as unfeasible.

We show now that our proposed algorithm reaches a stable state at which no player attempts to change its own transmit power. Moreover we show that the stable state corresponds to the core apportionment of the game. We model the evolution of the algorithm as the output of a finite-state Markov chain with state space $\Omega = \{\omega = (\psi, \boldsymbol{\nu}) | \psi \in \boldsymbol{\Psi}, \boldsymbol{\nu} \in \mathbb{R}^K\}$. For all time steps t, $\psi^t = \psi$ belongs to the subset of all possible disjoint coalitions $\boldsymbol{\Psi}$ with exactly L members, and remains fixed for the whole duration of the algorithm. The time evolution of the algorithm as a Markov chain is due to the time variability of $\boldsymbol{\nu}^t$, which depends on the power levels p^t_{kn} chosen by the players in the coalitions collected by ψ^t. We use this notation for the sake of convenience, to emphasize that $\boldsymbol{\nu}^t$ is directly connected to ψ^t.

The Markov process asymptotically tends towards a stable coalition structure state, where no player has any incentive to change its power. In other words, all coalitions get their maximum payoffs. Our algorithm guarantees that, when $t \to \infty$, this Markov chain tends towards a singleton steady state with probability 1.

Definition 20 ([153]) *A set $\Phi \subset \Omega$ is an* ergodic set *if, for any $\omega \in \Phi$ and $\omega' \notin \Phi$, the probability of reaching the state ω' starting from ω is zero. Once the Markov chain falls into a state belonging to an ergodic set, it never leaves that set, and it wavers between the states in that ergodic set from then on. The probability of reaching any state in the ergodic set is strictly positive.*

Lemma 4 ([153]) *In any finite Markov chain, no matter which state the process starts from, the probability of ending up into an ergodic set tends to* 1 *as time tends to infinity.*

Definition 21 ([153]) *Singleton ergodic sets are called* absorbing states.

If Φ is an absorbing state and $\omega \in \Phi$, the probability of ending up into state ω when beginning from ω is one. In fact, absorbing states individually represent points of equilibrium.

Lemma 5 *The state $\omega = (\psi, \boldsymbol{\nu})$ is an absorbing state of the best-response process if and only if*

$$\nu\left(\mathcal{A}_k\right) = +\infty \quad \forall \mathcal{A}_k \in \psi \tag{3.4}$$

Proof This condition ensures that no player has any incentive to change its power amount. If this condition is met, then no coalition can get a higher payoff by deviating from state $\omega = (\psi, \boldsymbol{\nu})$. Since all the target rates are feasible, this condition is also necessary. ■

Theorem 4 *The best-response process has at least one absorbing state.*

Proof Since the best-response algorithm is a Markov process, Lemma 4 ensures that the best-response process reaches an ergodic set Φ. To conclude the proof, it is enough to show that Φ is singleton. Suppose that the number of states in the ergodic set is $|\Phi| > 1$. Then all players revise their strategies without conflicting decisions with a non-null probability. As a consequence, the Markov process moves to a new state, in which all coalitions' payoff are higher than those achieved in the previous state. This means that the probability of going back to the previous state is null, which contradicts the notion of an ergodic set. ■

Note that Theorem 4 does not ensure the uniqueness of the ergodic set in the best-response process. There may exist some different combinations of the power allocation for the players to reach to a steady state. It means that the game possesses multiple equilibria. The major finding of Theorem 4 is that, according to the way the players adjust their strategies, the best-response process leads to one of the steady states, in which no player has any incentive to revise its power allocation.

Theorem 5 *The set of payoffs associated to an absorbing state of the best-response process coincides with the set of core allocation:*

i) if $\omega = (\psi, \boldsymbol{\nu})$ is an absorbing state, then $\boldsymbol{\nu}$ is a core allocation.

ii) if $\boldsymbol{\nu}$ is a core allocation, then all $\omega = (\psi, \boldsymbol{\nu})$ are absorbing states.

Proof Part i) Suppose $\omega = (\psi, \boldsymbol{\nu})$ is an absorbing state but $\boldsymbol{\nu}$ is not a core allocation. In this case, there exist some coalitions that can obtain a higher payoff. This is contradictory, since the game reaches an absorbing state when every coalition gets the maximum payoff.

Part ii) If $\boldsymbol{\nu}$ is a core allocation, then no coalition can earn by letting its member change their powers. This implies that the state will not move to a new state, and thus the current state is absorbing. ■

Coalitional games aim at identifying the best coalitions of the agents and a fair distribution of the payoff among the agents. Interestingly, in this game the absorbing state coincides with one of the Nash equilibria [75] of the game. Suppose there are $K = 2$ mobiles connected to a base station with $N = 1$ subcarrier only. In this case, the $K = K \cdot N = 2$ copies of the subcarrier, each constituting a coalition, are engaged in a 2×2 game. Every player has two strategies: either $p_k = 0$ or $p_k = \overline{p}_k$. It is straightforward to verify that, in this game, a mixed (vs. pure) Nash equilibrium exists which satisfies the stability of the static game. With due attention to the notation, we can extend this result to a general case.

Theorem 6 *The set of absorbing states in the best-response process and the set of Nash equilibria of the static game are asymptotically (in the long run) equivalent.*

Proof Let us consider the coalitions in the best-response process as players in a static game. Lemma 4 ensures that this process reaches an ergodic set in the long run. According to Theorem 4, this set is singleton, and thus its member is an absorbing state. Hence, no coalition (i.e., no player in the static game) has any incentive to revise its strategy. In static games, this is the definition of a Nash equilibrium. ■

We can now conclude that the absorbing state is an extension of the Nash equilibrium, since the coalitions bind agreements with each other as economic agents and earn a vector value rather than a real number. Once the coalitions reach the absorbing state, their payoff is the highest possible ($+\infty$), and no coalition is willing to revise its current strategy. In general, as follows from Theorem 6, the Nash equilibrium of the game is Pareto-optimal (efficient), since no other strategy can achieve a payoff greater than $+\infty$.

3.3 Numerical results

In this section, we evaluate the performance of the best-response algorithm presented in Sect. 3.2. We consider some cases with different numbers of mobile terminals, target data rates, and subcarriers, showing that our suggested scheme reaches a steady state after a few steps only. To increase the convergence speed of the algorithm, we introduce a tolerance parameter ε in our utility function, such that, if $\left|R_k/R_k^\star - 1\right| < \varepsilon$, then we assume that the payoff is $+\infty$. We can possibly set an asymmetric range $[\varepsilon_1, \varepsilon_2]$ such that $\varepsilon_1 \leq (R_k/R_k^\star - 1) \leq \varepsilon_2$, so as to favor solutions with $R_k > R_k^\star$.

We consider the following parameters for our simulations: the maximum power of each terminal k on each subcarrier n is $\overline{p}_{kn} = \overline{p} = 3\,\mu\text{W}$; the power of the ambient AWGN noise on each subcarrier is $\sigma_w^2 = 100\,\text{nW}$, and the constant number in (3.3) is $\beta = 5000$. We also set $\Theta = 10K \cdot N$ as the stopping criterion of the iterative algorithm, where K and N depend on the network parameters of the simulation. The path coefficients $|H_{kn}|^2$, corresponding to the frequency response of the multipath wireless channel at the carrier frequency $n\Delta f$, are computed using the 24-tap ITU modified vehicular-B channel model adopted by the IEEE 802.16m standard [136]. To account for the large-scale path loss, we assumed the terminals to be uniformly distributed between 3 and 100 m. Based on numerical optimizations, the parameter λ that reduces the probability of conflicting decisions among members of different coalitions for different number of terminals, subcarriers, and signal bandwidth, is $\lambda = 0.97$. The initial power allocation is $p_{kn} = 0\ \forall k \in \mathcal{K}$ and $\forall n \in \mathcal{N}$. This experimentally provides the minimal power consumption at the steady state, and in most cases the minimum number of steps of the algorithm.

Fig. 3.3 reports the behavior of the achievable rate R_k as a function of the time step t in a network with $K = 10$ terminals, $N = 1024$ subcarriers, and bandwidth $W = 10\,\text{MHz}$ using the vacant-carrier assignment scheme. The target rates, reported in Fig. 3.3 with solid markers on the right axis, are assigned randomly to each terminal using a uniform distribution in the range $[100, 250]$ kb/s. Further parameters are: tolerance $\varepsilon_1 = 0, \varepsilon_2 = 0.01$, power update step $\overline{\Delta p}_{kn} = \overline{p}_{kn}/25 = 120\,\text{nW}$, and number of subblocks $L = 32$. Numerical results show the convergence of R_k to the respective target rates $R_k^\star$ after 31 steps of the best-response algorithm.

In the remainder of this section, we will evaluate by simulation the aver-

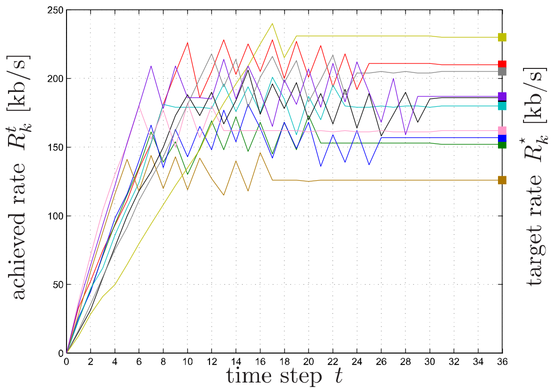

Fig. 3.3: *Achieved rates as functions of the iteration step.*

age performance of our proposed algorithm in terms of power expenditure and computational burden using realistic system parameters and extensive simulation campaigns. Note that we are not able to compare our technique with the joint resource allocation techniques available in the literature and reviewed in Chapter 1.9, mainly due to the unfeasible algorithmic complexity of the implementation of the latter when using tens of terminals, hundreds of subcarriers, and high data rates (on the order of Mb/s). As a consequence, in the following we will compare our measured results with the theoretical performance provided by the literature.

Figs. 3.4 and 3.5 report the simulation results obtained after 500 random realizations of a network with $R_k^\star = R^\star = 200\,\text{kb/s}\ \forall k \in \mathcal{K}$, $N = 1024$, $W = 10\,\text{MHz}$, and $\varepsilon_1 = 0, \varepsilon_2 = 0.04$ again with the vacant-carrier assignment strategy. Solid lines represent the case $\overline{\Delta p}_{kn} = \overline{p}_{kn}/5 = 600\,\text{nW}$, whereas dashed lines depict the case $\overline{\Delta p}_{kn} = \overline{p}_{kn}/25 = 120\,\text{nW}$. Circles, squares, upper triangles and lower triangles correspond to $L = \{8, 16, 32, 64\}$, respectively.

Fig. 3.4 shows the average normalized power expenditure ζ_k at the steady state as a function of K, computed by averaging $\zeta_k = \frac{1}{N}\sum_{n\in\mathcal{N}} \frac{p_{kn}}{\overline{p}_{kn}}$ over all terminals. This serves as a measure for the average total power consumption

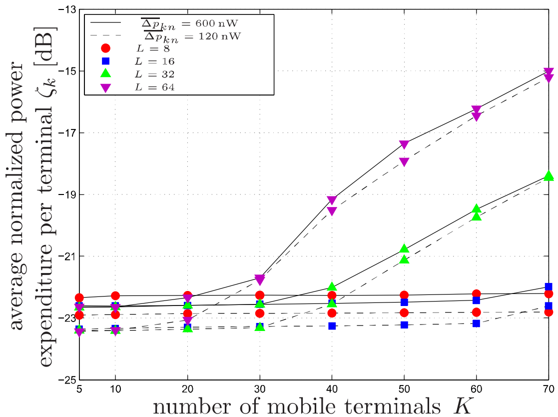

Fig. 3.4: *Average normalized power expenditure as a function of* K*, with* $W = 10\,\mathrm{MHz}$*,* $N = 1024$*, and* $R_k^\star = R^\star = 200\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ *in the case of vacant-carrier assignment model.*

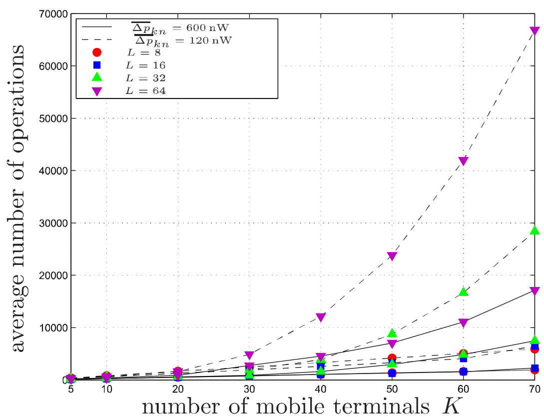

Fig. 3.5: *Experimental average number of operations as a function of* K*, with* $W = 10\,\mathrm{MHz}$*,* $N = 1024$*, and* $R_k^\star = R^\star = 200\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ *in the case of vacant-carrier assignment model.*

normalized to the maximum power expenditure available to each terminal. As can be noticed, ζ_k increases for $K \geq N/L$, since the number of shared subcarriers increases and the terminals must spend more power to overcome the intra-subcarrier noise. Interestingly, the power expenditure of the proposed centralized algorithm shows higher efficiency than the distributed and cross-layer schemes available in the literature (e.g., see [27, 30, 31, 154]). For instance, when considering 500 random realizations of a system with bandwidth $W = 10\,\text{MHz}$ and $N = 1024$ subcarriers, and using the vacant-carrier assignment model, we find that, in the case of a total sum-rate demand of 20 Mb/s (i.e., with a spectral efficiency of 2 b/s/Hz) and $R_k^\star = R^\star = 200\,\text{kb/s}$ (i.e., $K = 100$ terminals), the maximum power consumption per user is $31\,\mu\text{W}$ and the average power consumption of the system is 0.53 mW. In the multicell scenario of [27], the average power expenditure for each cell is 8 mW when the achievable data rate is 40 Mb/s. When considering the cross-layer algorithm proposed in [154], the average power expenditure per mobile terminal is 0.4 W with maximal spectral efficiency of 2 b/s/Hz, whereas the average power expenditure per mobile terminal required by the energy-efficient techniques proposed in [31] is 0.4 and 1.2 W when the achieved data rate is equal to 40 and 140 kb/s, respectively.

Fig. 3.5 shows the computational complexity of our algorithm expressed in terms of the average number of operations per terminal required to reach the steady state as a function of the number of terminals K with the vacant-carrier assignment model. The number of operations is measured experimentally by counting the number of steps required by the subchannel assignment plus the total number of trials required to update the transmit power according to the best-response algorithm. As can be seen, the complexity increases as L increases. This can be justified since increasing L increases the number of players $K \cdot L$, which yields an increase in the number of conflicting decisions. Note that the proposed algorithm is able to provide a spectral efficiency higher than 1 b/s/Hz, which occurs, for instance, when we assume more than $K = 50$ users with rates $R_k^\star = 200\,\text{kb/s}$ over a bandwidth $W = 10\,\text{MHz}$ in the proposed scenario, with a linear computational burden at the base station using appropriate values for the parameters. In this particular example, a good tradeoff between performance and complexity is $L = \{8, 16\}$ and $\overline{\Delta p}_{kn} = 600\,\text{nW}$. Using these values, the number of operations of the proposed algorithm is experimentally *lower* than the product $K \cdot N$, and so considerably *lower* than the complexity of the schemes available in the literature (e.g., see [26, 88, 89]). Our experiments with different data rate demands show that a smaller data rate reduces also the number of operations significantly. To further reduce the number of operations, we can also increase the tolerance parameters (e.g., with $\varepsilon_2 = 0.1$, we experience a complexity reduction on the order of $20 \div 30\%$). Note also that the spectral efficiency achieved by the proposed fair resource allocation method, while showing a linear computational burden, is comparable with that provided by sum-rate maximizing algorithms (e.g., see [155]). In the practice, a reasonable value for the maximum spectral efficiency achieved by the network in the region of linear complexity in all simulated scenarios (not reported here for the sake of brevity) is slightly lower than 2 b/s/Hz. For higher spectral efficiencies, no parameter selections can achieve the optimal resource allocation with linear complexity, and the number of operations appears to increase exponentially with the number of mobile terminals. However, note that the solutions can be found in most cases.

Figs. 3.6 and 3.7 depict the simulation results of a network with $R_k^\star = R^\star = 200\,\text{kb/s}\ \forall k \in \mathcal{K}$, $N = 1024$, $W = 10\,\text{MHz}$, and $\varepsilon_1 = 0, \varepsilon_2 = 0.04$ using the best-carrier assignment model. Solid lines represent the case $\overline{\Delta p}_{kn} = \overline{p}_{kn}/5 = 600\,\text{nW}$, whereas dashed lines depict the case $\overline{\Delta p}_{kn} = \overline{p}_{kn}/25 = 120\,\text{nW}$. Squares, upper triangles and lower triangles correspond to $L = \{16, 32, 64\}$, respectively. Fig. 3.6 shows the average normalized power expenditure ζ_k at the steady state as a function of K. As can be seen, the average power expenditure using the best-carrier assignment model is lower than with the vacant-carrier assignment, since the terminals having better channel conditions spend less power.

A drawback of the best-carrier assignment is an increased complexity of the algorithm. Fig. 3.7 shows the average number of operations per terminal

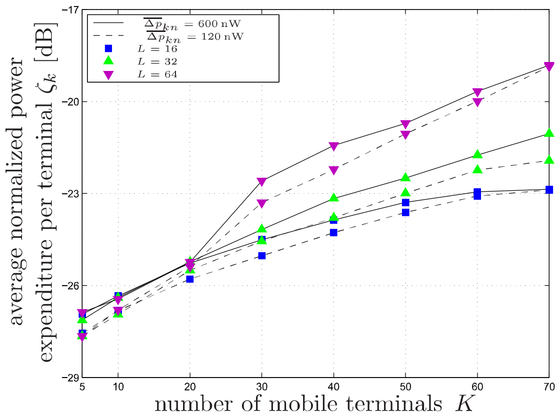

Fig. 3.6: *Average normalized power expenditure as a function of K, with $W = 10\,\mathrm{MHz}$, $N = 1024$, and $R_k^\star = R^\star = 200\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ in the case of best-carrier assignment model.*

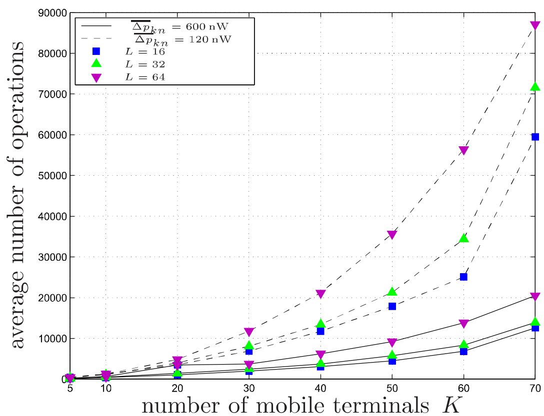

Fig. 3.7: *Experimental average number of operations as a function of K, with $W = 10\,\mathrm{MHz}$, $N = 1024$, and $R_k^\star = R^\star = 200\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ in the case of best-carrier assignment model.*

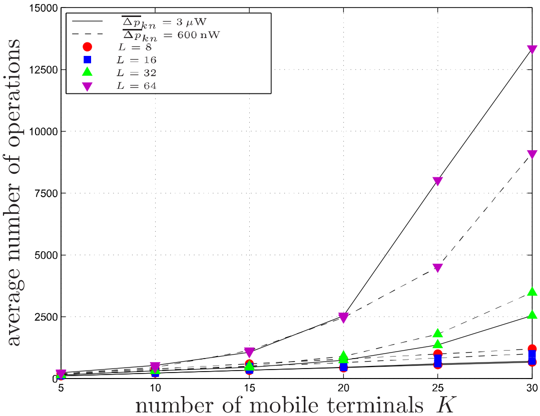

Fig. 3.8: *Experimental average number of operations as a function of K, with $W = 10\,\mathrm{MHz}$, $N = 512$, and $R_k^\star = R^\star = 500\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$ in the case of vacant-carrier assignment model.*

required to reach the steady state as a function of the number of terminals K. As can be seen, the best-carrier assignment model has a computational complexity higher than vacant-carrier assignment model, since the number of shared subcarriers in the best-carrier assignment model is larger than in the vacant-carrier assignment, which increases the probability of interference between simultaneous decisions in the best-reply algorithm. Note that, using the best-carrier assignment model, the case $L = 8$ appears to be computationally expensive.

Fig. 3.8 shows the average number of operations per terminal in the case of a network with parameters $R_k^\star = R^\star = 500\,\mathrm{kb/s}$ $\forall k \in \mathcal{K}$, $N = 512$, $W = 10\,\mathrm{MHz}$, and $\varepsilon_1 = 0, \varepsilon_2 = 0.04$ using vacant-carrier assignment model. Solid and dashed lines represents the cases $\overline{\Delta p}_{kn} = 3\,\mu\mathrm{W}$ and $\overline{\Delta p}_{kn} = 600\,\mathrm{nW}$, respectively, whereas circles, squares, upper triangles and lower triangles depict $L = \{8, 16, 32, 64\}$, respectively. Even in this case, with more severe requirements in terms of target data rates, the number of operations is shown to be *lower* than $K \cdot N$, again using spectral efficiencies higher than 1 b/s/Hz.

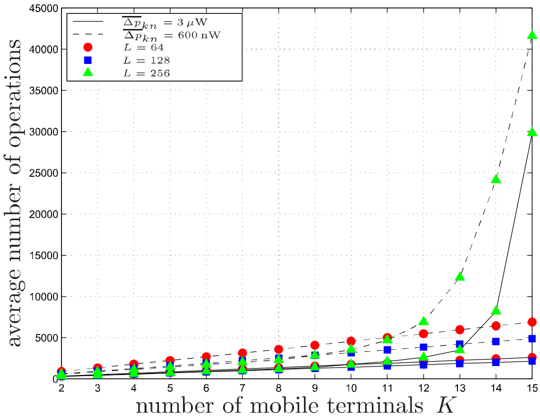

Fig. 3.9: *Experimental average number of operations as a function of* K*, with* $W = 20\,\mathrm{MHz}$*,* $N = 2048$*, and* $R_k^\star = R^\star = 2\,\mathrm{Mb/s}$ $\forall k \in \mathcal{K}$ *in the case of vacant-carrier assignment model.*

Finally, Fig. 3.9 shows the average number of operations per terminal in the case of a network with parameters $W = 20\,\mathrm{MHz}$, $N = 2048$, $R_k^\star = 2\,\mathrm{Mb/s}$, $\varepsilon_1 = 0$, and $\varepsilon_2 = 0.04$ with vacant-carrier assignment model. Solid and dashed lines represents the cases $\overline{\Delta p}_{kn} = 3\,\mu\mathrm{W}$ and $\overline{\Delta p}_{kn} = 600\,\mathrm{nW}$, respectively, whereas circles, squares and upper triangles depict $L = \{64, 128, 256\}$, respectively. The complexity is again *lower* than $K \cdot N$.

As can be seen in Fig. 3.5, 3.7, 3.8, and 3.9, due to the random behavior of the proposed algorithm, there is a strict relation between the average number of operations, the network parameters, and the algorithm parameters (including the channel assignment model). Depending on the parameter selection, we see different shapes (linear or exponential behavior) for the average number of operations. Thus, estimating the analytical complexity function for the best-response algorithm is hard to do. However, for *all* tested scenarios (not reported here for the sake of brevity), there exist properly tuned values (such as $L, \overline{\Delta p}_{kn}$) that provide an average number of operations for the proposed

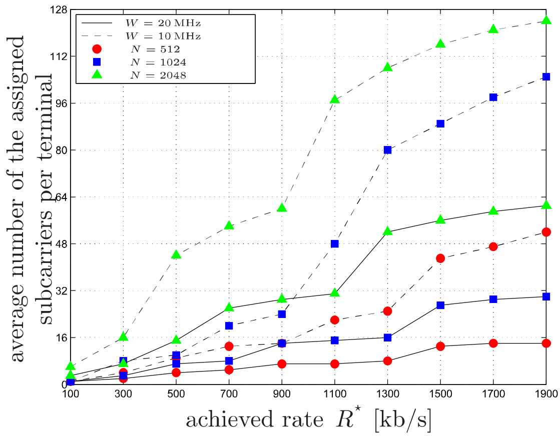

Fig. 3.10: *Average number of assigned subcarriers as a function of the achieved rate, with* $\overline{\Delta p}_{kn} = 600\,\mathrm{nW}$ *in the case of vacant-carrier assignment model.*

algorithm that are *lower* than the product $K \cdot N$, even with high data rate demands like in the cases of Figs. 3.8 and 3.9. The parameter that most impacts on the number of operations is L. Our experiments show that, for the optimal parameter selection (i.e., when the number of operations scales linearly with N and K), the average number of used subcarriers per terminal (i.e., those which bear $p_{kn} > 0$) is approximately $L/2$ when the vacant-carrier model is adopted. This rule-of-thumb can be used as a design criterion for the proposed algorithm. Let us consider Fig. 3.10, that reports the average number of assigned subcarriers to each mobile terminal as a function of the achieved rate $R^{\star}$, in the linear complexity regime and using $\overline{\Delta p}_{kn} = 600\,\mathrm{nW}$. Dashed and solid lines depict the cases $W = \{10, 20\}\,\mathrm{MHz}$, respectively, whereas circles, squares and upper triangles represent $N = \{512, 1024, 2048\}$, respectively. For instance, when $W = 20\,\mathrm{MHz}$, $N = 512$, and $R^{\star} = 500\,\mathrm{kb/s}$, the average number of used subcarriers is 4. If we look back at Fig. 3.8, we can verify that the linear complexity can be achieved using $L = 8$. Note that the number of assigned subcarriers in the case of $W = 10\,\mathrm{MHz}$ is higher than in the case $W = 20\,\text{MHz}$, since the subcarrier spacing is halved.

3.4 Discussion

This chapter described a computationally inexpensive centralized algorithm based on coalitional game theory to address the issue of fair optimal resource allocation (in terms of subcarrier assignment and power control) for the uplink of an infrastructure OFDMA wireless network. The scheme derived here is designed to meet the required data rates exactly, thus ensuring a fair performance apportionment to both users and service providers, with the best utilization of the network resources (minimum power expenditure and good spectral efficiency). The proposed algorithm can be analyzed as a Markov model that converges to an absorbing state with unitary probability in the long run. Our criterion also allows us to tradeoff system performance and computational burden of the algorithm, based on the number of subblocks used to apportion the available bandwidth and the data rate requirements of the terminals. Simulations show that the target rates are achieved with a low complexity procedure, even in the case of populated networks and stringent QoS requirements. The (greedy) best-carrier assignment rule results into a higher complexity but a lower power expenditure compared to the case with full use of the available subcarriers. The presented coalition-based strategy appears to be a good tradeoff between computational complexity and power efficiency in comparison with the schemes available in the literature, and achieves a spectral efficiency larger than 1 b/s/Hz.

Part II

On the Capacity of Relay Communications

Chapter 4

Basics of cooperative communications

Cooperative communication enables single-antenna mobiles in a multi-user environment to share their antennas and generate a virtual multiple-antenna transmitter, and consequently exploit some of the benefits of multiple-input multiple-output (MIMO) systems. Cooperative transmission can increase the data rate, save transmission power, and extend the coverage range of the network. As a result, it is considered to be a key-technique in the development of a robust and efficient communication system [156]. The first idea of cooperative transmission can be traced back to the proposal of the relay channel model, which consists of one source, one destination and one relay.

In the following, we will provide a brief literature overview of relay communication. In Sect. 4.2, we study relay assisted communication between a transmitter and a far destination which is out of its transmission range. Then, we extend the first scenario to a network wherein the direct-link between the transmitter and destination is indeed available. We next present an upper bound on the cutset capacity of relay communications in Sect. 4.3. The next three sections are devoted to the three well-known relaying strategies: the amplify and forward (AF), decode and forward (DF), and compress and forward (CF). We formally define these relaying protocols and calculate the upper bound capacity for each of them. We illustrate our results in Sect. 4.7, and then conclude in Sect. 4.8.

4.1 Introduction

In wired networks, using multiple wired connections that follow diverse paths, can significantly increase the end-to-end capacity of a link and establish a stable connection. A wireless network is traditionally viewed as a set of nodes trying to communicate with each other. Contrary to the inherent point-to-point wireless connection, wireless channel is by its nature of broadcast type and this causes multipath problem in a communication between two wireless devices. In wireless communication, cooperative diversity can exploit the broadcast nature of wireless transmission in order to increase the channel capacity of a link. The key idea is to have users cooperate in transmitting their messages to the destination, instead of operating independently and competing with each other for channel resources, as happens in conventional networks. The third party which acts as a relay node, receives the information from the source and deliver it to a destination via a channel that is independent from the (direct) source-destination link, the chances for a successful transmission would be better, thus improving the overall performance. The problems of reliable information transfer are in most cases intertwined with the problems of allocating the sparse resources available for use. Multiuser techniques try to optimize whole systems, by combining the multiple-access and information transmission aspects.

The three-terminal relay channel was introduced in 1971 by Van-der Meulen [157] and was initially investigated in the context of information theory. In his seminal paper, Van-der Meulen introduced upper and lower bounds for the capacity of a relay channel. Meanwhile, Sato [158] also looked at the relay channel in the framework of the ALOHA protocol. But, the booming interest in cooperative communications was initiated by the paper of Cover and El Gamal [159]. They evaluated the improvement of upper bound channel capacity of a point-to-point connection, placing one assistant node as relay and applying two different coding strategies of block Markov (*decode and forward*) and side information encoding (*compress and forward*). They also mixed these two strategies in [159, Th. 7]. Another relaying strategy proposed in the literature is *amplify and forward* [160] which augments the received signal without generating a new code at the relay.

M. R. Aref in his Ph.D. dissertation [161] formulated the max-flow min-cut (cutset) upper bound for the case of point-to-point connection through multiple relays. He also established the capacity of a degraded[1] relay network and in particular a cascade of degraded relay channels. Then, El Gamal and Aref in [162] established the capacity of the semi-deterministic relay channels where the received signal at the relay is a deterministic function of the emitted signals at the source and relay nodes. Zhang in [163] established the capacity of the relay channel when the channel from the relay to the destination is a noiseless channel of fixed capacity allowing the relay sending additional information to the destination. Gastpar *et al.* in [164] introduce information theoretic aspects of a point-to-point connection using multiple parallel relays wherein each transceiver has multiple antennas. Borade in [165] applies the cutset theorem [161] to the problem of network information flow involving multiple sources and multiple destinations and derives an information theoretic aspect of the upper bound capacity. S. Zahedi in his Ph.D. thesis [166] studied an energy efficient reliable communication. He showed that the capacity of a class of discrete-memoryless relay channels with orthogonal channels from the sender to the relay receiver and from the sender and relay to the destination is equal to the cutset upper bound.

The AWGN channel with relays has been widely studied in the literature. Schein and Gallager in [167] present upper and lower bounds to capacity of a couple of out of sight nodes through two parallel relays. Gupta and Kumar in [168] present an information theoretic look at the achievable capacity

[1] In relay networks composed of degraded relay channels, the received signals at relay nodes $(Y_1, Y_2, \ldots, Y_R)$ are conditionally distributed as a Markov chain: $Y_2 \to Y_3 \to \cdots \to Y_R \to Y_1$.

of arbitrary size and architecture of wireless networks. Many well-known capacity defining achievable rate regions of cooperation coding, e.g. peer-to-peer network coding in broadcast or multiple access channels, can be derived as special cases of the proposed scheme in [168]. Then, Gastpar and Vetterli in [169] focus on the network presented by Gupta *et al.* [168], but wherein there is only one active transmitter-destination link, and all others nodes act as intermediate relays. They derive lower and upper bounds capacity by allowing arbitrarily complex network coding.

Gupta *et al.* in [170] also discuss the throughput of each randomly placed wireless terminal in an arbitrarily large network. The results should be important for designers to define the number of wireless terminals and the proper range to achieve the highest throughput. Papers [171] and [172] by Sendonaris *et al.* are particularly relevant. They show that cooperation among mobile terminals can help establishing a robust connection irrespective of channel variations, and achieving high channel capacity. The work of Scaglione *et al.* [173] proposes an interaction between the physical layer link and the network layer. A "propagating wave" of relay transmissions is produced from the transmitter to the destination, allowing efficient flooding of a wireless network. Suraweera *et al.* in [174] compute the error performance of amplify and forward technique in a reliable communication. In addition, in recent years, with the advances in MIMO [175], interest in reliable point-to-point reliable MIMO communication has arisen [176].

It is reasonable to assume that relaying is not necessary if the demanded QoS is achievable through the direct-link to destination. In large reliable networks, the questions of "to relay or not to relay" and "the best relay node(s) election policy" is usually based on efficiency of channel parameters. Shan *et al.* in [177] devise a cross-layer protocol to choose the beneficial relays which are able to increase network throughput as much as possible. In [178], the authors propose to choose the best relay depending on its geographic position, are specifically based on the geographic random forwarding (GeRaF) protocol proposed in [179]. Papers [180] in a single relay cooperative network, and [181,182] in multi AF relays communications are some noteworthy works which seek an answer to these questions applying game theory tools.

Notation 1 *Some general notation is needed in the following. Upper case letters X_i, Y_i represent the output and input random vector variables of node i, respectively, and $X_i[b]$ is the bth entry of a random vector. The notations $\mathbb{V}\mathrm{ar}\,(X)$, $\mathbb{E}\,\{X\}$, $\mathrm{H}\,(X)$, and $\mathrm{I}\,(X;Y)$ are used to denote variance, expectation value, entropy, and mutual information, respectively. For the sake of simplicity we will assume that our channels are real-valued. The notation h_{ij} is used for the real-valued channel gain of the link between nodes i and j. The symbol sent by transmitter k in block index b is represented by w_k^b. We use calligraphic letters $\mathcal{X}, \mathcal{Y}, \mathcal{W}$ to indicate (finite) alphabet spaces and lower case letters x_i, y_i for channel distributions. The conventions $\boldsymbol{X}_{\mathcal{A}}$ and $\boldsymbol{x}_{\mathcal{A}}$ denote joint vector variable and channel distribution, respectively of the indices belonging to set $\mathcal{A}$. $\mathcal{A}^C$ represents the complementary set of $\mathcal{A}$. The sign $\widehat{\ }$ is used for estimated parameters, and the notation $d(X, \widehat{X})$ is the average distortion measure between two sequences. The notation $C(\gamma_{ij}) = \frac{1}{2}\log_2(1+\gamma_{ij})$ denotes Shannon channel capacity between nodes i and j with signal-to-noise ratio (SNR) γ_{ij} at the receiver. The symbol $\mathcal{C}$ is used for the outer region channel capacity. Moreover, we use the symbol A_ϵ^m to denote a strongly typical set as defined in [16, p. 59]. To briefly review aspects of network information theory in the context of our problem, see Appendix B.*

Notation 2 *Throughout this thesis, the power gain of the radio link between two nodes i and j at a distance d_{ij}, is scaled by*

$$\Gamma_{ij} = G_{tx} \cdot G_{rx} \cdot \left(\frac{\lambda_0}{4\pi}\right)^2 \cdot d_{ij}^{-\alpha} \tag{4.1}$$

wherein $G_{tx} = G_{rx} = 1$ are the transmit and receive antennas gains assumed omnidirectional, the parameter $\lambda_0 \approxeq 0.12m$ represents carrier wave length, and the path loss exponent is $\alpha = 2$. Consequently, if the node i is placed close to node j and far from node k then: $h_{ij} \gg h_{ik}$.

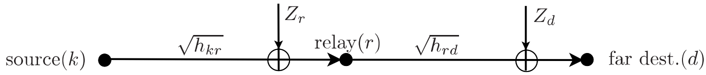

Fig. 4.1: *One source, one relay, one far destination network.*

4.2 Relayed point-to-point communication

The relays are able to outperform a direct-link communication and even actualize some otherwise impossible scenario. For example, Fig. 4.1 depicts the simplest cooperative (relay) communication model to realize a communication between two hidden terminals. The source k wishes to send a message to the far destination d which is out of sight. We assume there is no fading on the wireless channels. One solution is to use an intermediate node. The source k sends a message to the relay r and the received noisy version of the original message is re-transmitted to the far destination node. We assume that the relay is accessible to both the source and destination nodes. In the example, the relay does not do any processing (encoding, decoding, ...) on the receive signal, but its duty is to make possible the information exchange from k to d. This simple two-hop model enlarges significantly the range of the network.

The relaying function between the source and destination nodes can be done in two different signaling ways: *half-duplex* and *full-duplex* mode. In half-duplex mode, the transmission from source to destination is done in two different stages and in each stage the relay acts either as receiver or transmitter. In the first stage, k transmits a stream of information and r operates as a receiver and d is idle. In the second stage, the channel between k and r is kept idle and d is active to receive data from the intermediate node r which acts only as a transmitter. On the other hand, in full-duplex mode, both wireless channels are simultaneously busy and the relay plays the role of a receiver and a transmitter at the same time. The function of the intermediate node is feasible using two distinct antennas, one as receiver and one as transmitter using two orthogonal frequency band. In spite of the the difficulties to apply the full-duplex manner in wireless networks, in this work, we concern with this model for didactic issues. Our goal is to determine how much data we can reliably get from source to destination, placing no importance on delay or computational complexity.

In the scenario of Fig. 4.1, the power expenditure of k and r, p_k and p_r, respectively can be set a priori or can be adjusted given the orthogonal power constraints $\overline{p}_k$ and $\overline{p}_r$ and to the path gain values h_{kr} and h_{rd}. The objective of the power control is to approach the maximum Shannon channel capacity between source and a (far) destination. According to the max-flow min-cut theorem (also referred to as the cutset bound), the maximum end-to-end channel capacity in Fig. 4.1 is achieved when the capacity of the source-relay link is the same as that of the relay-destination link, i.e.

$$C(\gamma_{kr}) = C(\gamma_{rd}) \tag{4.2}$$

We assume that the noises Z_r and Z_d are Gaussian random variables $\mathcal{N}(0, \sigma_w^2)$. The power optimization problem becomes to the equation :

$$h_{kr} p_k = h_{rd} p_r \tag{4.3}$$

We study the final equation in four different cases:

1. if $(h_{kr} \leq h_{rd})$ and $(\overline{p}_k < \overline{p}_r)$ then $p_k = \overline{p}_k \quad , \quad p_r = \frac{h_{kr}}{h_{rd}} \overline{p}_r$;
2. if $(h_{kr} < h_{rd})$ and $(\overline{p}_k \geq \overline{p}_r)$ then $p_k = \overline{p}_r \quad , \quad p_r = \frac{h_{kr}}{h_{rd}} \overline{p}_r$;
3. if $(h_{kr} \geq h_{rd})$ and $(\overline{p}_k > \overline{p}_r)$ then $p_k = \frac{h_{rd}}{h_{kr}} \overline{p}_k \quad , \quad p_r = \overline{p}_r$;
4. if $(h_{kr} > h_{rd})$ and $(\overline{p}_k \leq \overline{p}_r)$ then $p_k = \frac{h_{rd}}{h_{kr}} \overline{p}_k \quad , \quad p_r = \overline{p}_k$.

Case (1) is the situation in which the maximum capacity of the source-relay link is less than that relay-destination link. The source exploits the maximum capacity of the source-relay link, i.e. $p_k = \overline{p}_k$. To adjust p_r it is enough to satisfy Eq. 4.3, because the relay-destination channel capacity is limited to

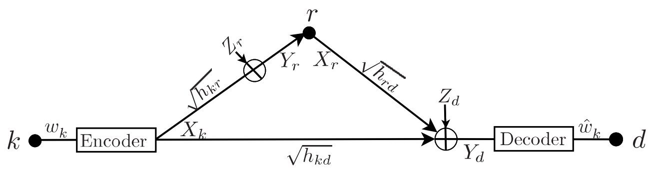

Fig. 4.2: *One source, one relay, one destination network scenario.*

the maximum capacity of source-relay link. The same derivation applies to case (3). In case (2) there is no exact relation between the maximum channel capacity of the two channels. Assigning $p_k = \overline{p}_r$, bounds the source-relay capacity to $h_{kr}\overline{p}_r$ that is less than the maximum capacity of the source-relay link. At this point it is easy to satisfy Eq. 4.3. Case (4) is similar to (2).

Cooperative communication can be efficient also when a direct-link between the source node k and destination d is available. In the network illustrated by Fig. 4.2, when the source node k broadcasts, a noisy version of the data comes to the relay r and another corrupted version of data approaches the destination d. Using an intermediate node is sensible when the received data at d is too weak to be decoded. In this situation, a relay can help communication by transmitting a new version of its own received signal. The direct-link signal is used to help decoding the stronger version of original data sent by relay.

The cooperative network in Fig. 4.2 consists of four finite random spaces: $\mathcal{X}_k$ at the transmitter, $\mathcal{Y}_r$ and $\mathcal{X}_r$ at the relay node, and $\mathcal{Y}_d$ at the destination. The source node wants to transmit a message w_k to the destination through direct and reliable links. The original message w_k is split in a sequence of sub-messages $w_k^1, \ldots, w_k^b, \ldots, w_k^B$ each uniformly and independently drawn from a set with alphabet size m and length R_k, represented by $\mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$. The encoder at the transmitter is a function $\mathcal{W}_k \to \mathcal{X}_k$ which maps w_k^b to $X_k[b]\left(w_k^b\right)$. We assume that the decoder generates a sequence of B Gaussian random codewords given a constraint on average power as $\mathbb{E}\left\{X_k^2\right\} \leq \overline{p}_k$. We also assume $\mathbb{E}\{X_k\} = 0$. In each block with index b, the source k broadcasts the encoded symbols and each $X_k[b]\left(w_k^b\right)$

experiences two different paths to approach the relay and destination. The relay observes $Y_r[b]$ as:

$$Y_r[b] = \sqrt{h_{kr}} X_k[b]\left(w_k^b\right) + Z_r; \qquad Z_r \sim \mathcal{N}\left(0, \sigma_w^2\right) \tag{4.4}$$

The receive signals at r and d are different but statistically correlated. In full-duplex mode, the relay processes the received signal in the previous block index and generates the information $X_r[b]\left(w_k^{b-1}\right)$, to be sent into the relay-destination channel. The information $X_r[b]\left(w_k^{b-1}\right)$ is a re-generated version of $Y_r[b-1]$ and it is the output of the relay r's deterministic function whose input is the sequence of the previous received signals:

$$X_r[b]\left(w_k^{b-1}\right) = f_r^b(Y_r[b-1], Y_r[b-2], \ldots, Y_r[1]); \tag{4.5}$$

The function f depends on the specific cooperative strategy as will be specified later on. Each symbol sequence $\{X_r[b]\left(w_k^{b-1}\right)\}$ is such that $\mathbb{E}\left\{X_r^2\right\} \leq \overline{p}_r$, and $\mathbb{E}\left\{X_r\right\} = 0$. The destination receives a superposition of two different signals:

$$Y_d[b] = \sqrt{h_{kd}} X_k[b]\left(w_k^b\right) + \sqrt{h_{rd}} X_r[b]\left(w_k^{b-1}\right) + Z_d; \quad Z_d \sim \mathcal{N}\left(0, \sigma_w^2\right) \tag{4.6}$$

It is seen that, the destination node receives information both about w_k^b and w_k^{b-1}. It means, the destination receives two different versions of w_k^b in two broadcast and multiple-access stages. The decoder at the destination decodes the message $\hat{w}_k$ canceling the effect of $X_k[b-1]\left(w_k^{b-1}\right)$ from $X_r[b]\left(w_k^{b-1}\right)$. In particular, the decoding function is a mapping function from $\mathcal{Y}_d \rightarrow \mathcal{W}_k$. The probability of error for the destination's decoder is defined as:

$$P_e^m = 2^{-mR_k} \cdot \sum_{w_k} \Pr\left\{\hat{w}_k \neq w_k | w_k \text{ was sent}\right\} \tag{4.7}$$

which is defined based on the assumption that the messages are independent and uniformly distributed over the alphabet space. The rate R_k is achievable if there exists a sequence of codes $\left(m, 2^{mR_k}\right)$ for which P_e^m is arbitrarily close to zero when $m \rightarrow \infty$.

4.3 Cutset upper bound

In this section we derive the upper bound capacity of max-flow min-cut or "cutset". The cutset upper bound is used as a reference to compare the upper bound of the realistic models. M. R. Aref in [161, Th. 3.4] pioneered to establish the cutset bound in a general reliable network with multiple relays. S. Zahedi in [166, Th. 2.2] presents another proof to the cutset upper bound in the one relay case like Fig. 4.2. The following proposition shows that a cooperative system can be decomposed into a broadcast channel from viewpoint of the source node, and a multiple-access channel from the destination point of view.

Prop. 1 *[166, Th. 2.2] For any relay channel* $(\mathcal{X}_k \times \mathcal{X}_r, p(y_d\, y_r \,|\, x_k\, x_r), \mathcal{Y}_r \times \mathcal{Y}_d)$ *the cutset capacity is upper bounded by*

$$\mathcal{C}_{cutset}(k;r;d) = \sup_{p(x_k,x_r)} \min \left\{ \mathrm{I}(X_k \,;\, Y_d\, Y_r|X_r) \,,\, \mathrm{I}(X_k\, X_r \,;\, Y_d) \right\} \tag{4.8}$$

where the supremum is computed over all joint distributions on $\mathcal{X}_k \times \mathcal{X}_r$ *complying with individual power constraints.*

The first term is the mutual information of broadcasting X_k toward r and d with transition probability $p(y_d\, y_r|\, x_k)$. The second term is the mutual information of multiple-access of r and k at the destination node with transition probability $p(y_d \,|\, x_k\, x_r)$. Thus, in general, random variables y_d and y_r are statistically related to both inputs x_k and x_r through $p(y_d\, y_r \,|\, x_k\, x_r)$.

$p(x_k, x_r)$ is a joint Gaussian distribution on $\mathcal{X}_k \times \mathcal{X}_r$ with a cross correlation coefficient of $\rho = \dfrac{\mathbb{E}\{X_k X_r\}}{\sqrt{\mathbb{E}\{X_k^2\}\,\mathbb{E}\{X_r^2\}}}$. In the case of $\rho = 0$ we have: $p(x_k, x_r) = p(x_k) \cdot p(x_r)$. We pause at this point to recall some useful statistics equalities related to Fig. 4.2. To review the algebraic manipulation of the following formulas, see Appendix A.

$$\mathbb{V}\mathrm{ar}\,(Y_d) = h_{kd}\overline{p}_k + h_{rd}\overline{p}_r + 2\rho\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r} + \sigma_w^2 \tag{4.9a}$$

$$\mathbb{V}\mathrm{ar}\,(X_k|X_r) = \overline{p}_k\left(1-\rho^2\right) \tag{4.9b}$$

$$\mathbb{V}\text{ar}\left(Y_r|X_r\right) = h_{kr}\overline{p}_k\left(1-\rho^2\right)+\sigma_w^2 \tag{4.9c}$$

$$\mathbb{V}\text{ar}\left(Y_d|X_r\right) = h_{kd}\overline{p}_k\left(1-\rho^2\right)+\sigma_w^2 \tag{4.9d}$$

$$\mathbf{\Sigma}[Y_dY_r|X_r] = \left(h_{kd}+h_{kr}\right)\overline{p}_k\left(1-\rho^2\right)+\sigma_w^2 \quad \text{(see (A.10))} \tag{4.9e}$$

Reference [183, Proposition 2] shows that the $\mathcal{C}_{cutset}(k;r;d)$ is attained by Gaussian channels. The following theorem presents the capacity of the AWGN relay channel.

Theorem 7 *The AWGN cutset capacity of a point-to-point relayed communication is upper bounded to:*

$$\mathcal{C}_{cutset}(k;r;d) = \sup_{-1\leq\rho\leq 1}\ \min\left\{C\left(\frac{\left(h_{kd}+h_{kr}\right)\overline{p}_k\left(1-\rho^2\right)}{\sigma_w^2}\right),\right.$$
$$\left.C\left(\frac{h_{kd}\overline{p}_k+h_{rd}\overline{p}_r+2\rho\sqrt{h_{kd}\overline{p}_kh_{rd}\overline{p}_r}}{\sigma_w^2}\right)\right\}$$

Proof The mutual information terms are calculated using Eqs. 4.9.

$$\mathrm{I}(X_k\ ;\ Y_d\,Y_r|X_r) = \frac{1}{2}\log_2\frac{\mathbf{\Sigma}[Y_dY_r|X_r]}{\mathbf{\Sigma}[Y_dY_r|X_rX_k]} = \frac{1}{2}\log_2\frac{\mathbf{\Sigma}[Y_dY_r|X_r]}{\sigma_w^2} =$$
$$= C\left(\frac{\left(h_{kd}+h_{kr}\right)\overline{p}_k\left(1-\rho^2\right)}{\sigma_w^2}\right);$$

$$\mathrm{I}(X_k\,X_r\ ;\ Y_d) = \frac{1}{2}\log_2\frac{\mathbb{V}\text{ar}\left(Y_d\right)}{\mathbf{\Sigma}[Y_d|X_k\,X_r]} = \frac{1}{2}\log_2\frac{\mathbb{V}\text{ar}\left(Y_d\right)}{\sigma_w^2} =$$
$$= C\left(\frac{h_{kd}\overline{p}_k+h_{rd}\overline{p}_r+2\rho\sqrt{h_{kd}\overline{p}_kh_{rd}\overline{p}_r}}{\sigma_w^2}\right).$$

■

Performing the maximization over ρ, we can easily obtain the upper bound. In other words, under average power constraints, a jointly Gaussian input distribution simultaneously maximizes both mutual information terms. The important result of Theorem 7 is: If $\gamma_{kr} > \gamma_{rd}$ (i.e. the relay node is in a good position to receive signals from the transmitter rather than to deliver symbols to the destination) then the cutset upper bound capacity is achieved by the maximization of the multiple-access term (the second term), otherwise it is achieved by the broadcast term.

In the AWGN cutset upper bound capacity, when $\rho \geq 0$, the broadcast term, $\mathrm{I}(X_k \,;\, Y_d\, Y_r|X_r)$, is decreasing function of ρ, while the multiple-access term, $\mathrm{I}(X_k\, X_r \,;\, Y_d)$, is increasing. Thus, the maximum of $\mathcal{C}_{cutset}(k;r;d)$ is taken at the point at which two terms are equal:

$$(h_{kr} + h_{kd})\,\overline{p}_k\,\rho^2 + 2\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r}\,\rho + (h_{rd}\overline{p}_r - h_{kr}\overline{p}_k) = 0 \qquad (4.10)$$

We analyze the Eq. 4.10 in different cases. First, we assume the parameter $\rho \geq 0$ is fixed and it is possible for the source and relay nodes to adjust the transmission powers. Then, we assume the source and relay transmit at the maximum power and the network can tune the value of ρ.

1) We assume $\mathcal{X}_k$ and $\mathcal{X}_r$ are statistically independent; i.e. $\rho = 0$. The capacity region of $\mathcal{C}_{cutset}(k;r;d)$ achieves its maximum with adjusting the power expenditure of the source and relay nodes. In Eq. 4.10 with $\rho = 0$, it is enough to satisfy $h_{rd}p_r - h_{kr}p_k = 0$, that is equal to the maximum capacity problem of Fig. 4.1 and Eq. 4.3. With fixed $\rho = 0$ and an appropriate power control, the $\mathcal{C}_{cutset}(k;r;d)$ takes:

$$\mathcal{C}_{cutset}(k;r;d) = C\left((h_{kr} + h_{kd})\,\frac{p_k}{\sigma_w^2}\right) = C\left(\frac{h_{kd}p_k + h_{rd}p_r}{\sigma_w^2}\right) \qquad (4.11)$$

The relation $h_{kr}p_k = h_{rd}p_r$ means that the data rate of the channel between the source and relay nodes is equal to that between the relay and destination. In this case, the $k \to r \to d$ path has the maximum possible efficiency. Choosing two independent random spaces for $\mathcal{X}_k$ and $\mathcal{X}_r$ guarantees minimum processing for relaying's data process. Suppose the destination node consists of two different antennas with orthogonal frequencies which are used simultaneously for receiving data from the source and relay nodes. Thus, there is no interference between the $r \to d$

and $k \to d$ links. So, it results:

$$\mathcal{C}_{cutset}(k;r;d) = C\left(\frac{h_{kd}p_k + h_{rd}p_r}{\sigma_w^2}\right) = C\left(\gamma_{kd} + \gamma_{rd}\right). \tag{4.12}$$

Therefore, when ρ is equal to zero, the requirements to achieve the upper bound cutset capacity is an appropriate power control, and an adder component at destination to sum the received SNRs. Equivalently, the reliable communication form a parallel channel between relay and direct-links.

2) If the sequences of X_k and Y_r are drawn from two correlated code spaces with a strictly positive correlation ($\rho > 0$), the necessary condition for the power control to achieve the capacity $\mathcal{C}_{cutset}(k;r;d)$ is: $\gamma_{kr} > \gamma_{rd}$. This means the data rate of $r \to d$ channel is less that $k \to r$ link. Hence, there must be a delay between sending the broadcast message and the multiple-access message.

3) Finally, we suppose that the source and relay nodes transmit at maximum power and the network is able to tune the correlation parameter. The appropriate value of ρ is found by resolving Eq. 4.10 for ρ.

$$\rho^* = \frac{-\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r} + \sqrt{h_{kr}\overline{p}_k\left(h_{kd}\overline{p}_k + h_{kr}\overline{p}_k - h_{rd}\overline{p}_r\right)}}{\left(h_{kd} + h_{kr}\right)\overline{p}_k} \tag{4.13}$$

on the condition that $\Delta = (h_{kd} + h_{kr})\overline{p}_k - h_{rd}\overline{p}_r > 0$ and $0 \leq \rho^* \leq 1$. If on the contrary $\Delta \leq 0$, then the maximum capacity of $r \to d$ is higher than that of the broadcast channel. This means that the channel between relay and destination must be kept idle for receiving the broadcast message and therefore using the links is not highly efficient. Instead, the condition $\Delta > 0$ means the broadcast capacity is higher than maximum $r \to d$ channel data rate. Using an appropriate memory at the relay node, all channels get busy.

For strictly negative ρ, from the formula of $\mathcal{C}_{cutset}(k;r;d)$ in Theorem 7, it is derived that reducing coefficient ρ toward -1 yields decreasing either broadcast and multiple-access capacities. The solution to compensate the affect of a negative ρ is to raise significantly the upper bound limits of the source and relay power consumption.

4.4 Amplify and forward technique

In the AF technique, the transmit message w_k is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B$ which are independently and uniformly drawn from the message set $\mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$. Each sub-message w_k^b is separately encoded to $X_k[b]\left(w_k^b\right)$ under the constraint that $\mathbb{E}\left\{X_k^2\right\} \leq \overline{p}_k$. In each block index b, the source node broadcasts $X_k[b]\left(w_k^b\right)$, and at the same time the relay just increase the amplitude of the analog observed signal $Y_r[b-1]$ to result a normalized transmit $X_r[b]\left(w_k^{b-1}\right)$ as:

$$\begin{aligned} X_r[b]\left(w_k^{b-1}\right) &= E[b].Y_r[b-1] \\ &= E[b].\left(\sqrt{h_{kr}}\, X_k[b-1]\left(w_k^{b-1}\right) + Z_r\right) \end{aligned} \tag{4.14}$$

wherein E is amplification factor and is chosen to satisfy the relay's power limit. The relay node has its own power constraint as $\mathbb{E}\left\{X_r^2\right\} \leq \overline{p}_r$, so that:

$$|E[b]|^2 \leq \frac{\overline{p}_r}{\sigma_w^2 + h_{kr}\overline{p}_k} \tag{4.15}$$

For simplicity, we assume $E[b] = E$ in every block. As can be observed, if $\sigma_w^2 + h_{kr}\overline{p}_k \gg \overline{p}_r$ the effect of the relay is negligible. Combining Eqs. 4.6 and (4.14) gives:

$$Y_d[b] = \sqrt{h_{kd}}X_k[b]\left(w_k^b\right) + |E|.\sqrt{h_{kr}h_{rd}}\, X_k[b-1]\left(w_k^{b-1}\right) + |E|.\sqrt{h_{rd}}Z_r + Z_d \tag{4.16}$$

By Eq. 4.16, the maximum capacity of AF scheme turns out to be:

$$\mathcal{C}_{AF}(k;r;d) = C\left(\left(\sqrt{h_{kd}} + |E|.\sqrt{h_{kr}h_{rd}}\right)^2 \frac{\overline{p}_k}{\left(1 + |E|^2 h_{rd}\right)\sigma_w^2}\right) \tag{4.17}$$

The relay node does not regenerate any new code, and consequently the complexity of this scheme is low. Since the relay node amplifies whatever it receives, including noise, it is mainly useful in high SNR environments. When the channel between the transmitter and the relay is very noisy, increasing the amplification factor E increases the noise at the destination. The relay should thus not always transmit with maximum power. [184, p. 46] demonstrates that under the condition

$$|E| \leq \gamma_{kr} \tag{4.18}$$

the $\mathcal{C}_{AF}(k;r;d)$ outperforms the capacity of the maximal ratio combining (MRC) technique that is:

$$\mathcal{C}_{MRC}(k;r;d) = C\left(\gamma_{kd} + \frac{\gamma_{kr}\,\gamma_{rd}}{\gamma_{kr} + \gamma_{rd}}\right) \tag{4.19}$$

Comparing (4.17) and (4.19) we find that the MRC technique performs better than AF under the following condition:

$$\frac{\gamma_{kr}\,\gamma_{kd}}{\gamma_{kr} + \gamma_{kd}} < |E|^2.h_{rd} \tag{4.20}$$

4.5 Decode and forward technique

Decode and forward (DF) is a block transmission strategy. There are four Gaussian alphabet spaces: $\mathcal{X}_k$ at source's encoder, $\mathcal{X}_r$ and $\mathcal{Y}_r$ at relay's encoder and decoder respectively and finally $\mathcal{Y}_d$ at destination's decoder. In *regular encoding*, the two codebooks $\mathcal{X}_k$ and $\mathcal{X}_r$ are independent with the same size, but in *irregular coding* there is a correlation between two codebooks e.g. Superposition encoding [159]. Here we study a well known regular encoding called *Block Markov encoding.* The transmit message w_k is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B \in \mathcal{W}_k$ and each sub-message is separately encoded. The encoder at the transmitter a block memory whose size depends upon the decoding strategy at the destination node. The decoder at the destination node applies one of the following two strategies:

1) *Sliding-Window decoding* proposed by Xie *et al.* in [185]. The source codeword in block b is a Gaussian random codebook represented by $X_k[b]\left(w_k^b\right)$.

In block b, the source node broadcasts $X_k[b]\left(w_k^b\right)$ to the relay and destination nodes and signal $Y_r[b]$ approaches the relay node. At the same time, the relay attempts to decode $Y_r[b-1]$ to $\hat{w}_k^{b-1}$, and then re-encode $\hat{w}_k^{b-1}$ and transmit it toward the destination. The decoding process at the relay is correctly done if γ_{kr} is sufficiently high. The relay decodes the previous received signal using $\mathcal{Y}_r$. This is accomplished by removing the effect of the received sub-message in the signaling block number $b-2$. The relay node then encodes $\hat{w}_k^{b-1}$ to $X_r[b]\left(\hat{w}_k^{b-1}\right)$ using $\mathcal{X}_r$, and transmit it to the destination.

$$Y_d[b] = \sqrt{h_{kd}}\, X_k[b]\left(w_k^b\right) + \sqrt{h_{rd}}\, X_r[b]\left(\hat{w}_k^{b-1}\right) + Z_d \tag{4.21a}$$
$$Y_r[b] = \sqrt{h_{kr}}\, X_k[b]\left(w_k^b\right) + Z_r \tag{4.21b}$$

At the destination, decoding of w_k^b is done in block $b+1$ using the received signals $Y_d[b]$ and $Y_d[b+1]$ that depend on w_k^b and $\hat{w}_k^b$, respectively. Of course in most cases $w_k^b = \hat{w}_k^b$. The transmission is completed after $B+1$ th block index. The authors of [185] claim that it is almost impossible to extend this decoding type to a multiple-relays network.

2) *Backward decoding* proposed by Willems in [186, Ch. 7]. In backward decoding the source broadcasts a multiplexed Gaussian random value assigned to $X_k[b]\left(w_k^{b-1}, w_k^b\right)$. The relay's encoder/decoder process is the same as the sliding-window encoding/decoding. So:

$$Y_d[b] = \sqrt{h_{kd}}\, X_k[b]\left(w_k^{b-1}, w_k^b\right) + \sqrt{h_{rd}}\, X_r[b]\left(\hat{w}_k^{b-1}\right) + Z_d \tag{4.22a}$$
$$Y_r[b] = \sqrt{h_{kr}}\, X_k[b]\left(w_k^{b-1}, w_k^b\right) + Z_r \tag{4.22b}$$

The destination's decoder will start to reconstruct message w_k after having collected whole sequence $Y_d[1], \ldots, Y_d[B+1]$. The process starts from the last block and continue backward toward the first block. The delay of sliding-window is therefore much less than that of backward decoding. Backward decoding can be generalized to multiple relays and sources.

The relay's encoder does not care about the type of the combined encoding/decoding strategy at the source/destination. The sub-message w_k^b is fully[2] decoded/re-encoded to one existing index in $\mathcal{Y}_r/\mathcal{X}_r$ with zero probability if [166, Sec. 2.3.1]:

$$R_k \leq \mathrm{I}\left(X_k\,; Y_r|X_r\right) \tag{4.23}$$

In fact, an error occurs when Y_r is so corrupted that there no index exists in the encoder alphabet space. The Markov chains $X_k \leftrightarrow X_r$ and $X_k \leftrightarrow Y_r|X_r$ form a *unique* jointly typical index at relay's encoder according to pmf:

$$p\left(x_k, y_r, x_r, y_d\right) = p\left(x_k\right).p\left(x_r|x_k\right).p\left(y_r|x_k\, x_r\right) \tag{4.24}$$

Then, the relay sends $X_r[b+1](\hat{w}_k^b)$ in block $b+1$. In fact, the whole transmission of B sub-messages is done in $B+1$ blocks and the relay transmits the sequence: $1, ..., X_r[b+1](w_k^b), ..., X_r[B+1](w_k^B)$. Therefore, the overall data rate is $R_k\frac{B}{B+1}$ per use, and obviously it is efficient for $B \to +\infty$.

In each block index b, two types of errors can occur at the destination's decoder: 1) an error occurs because of a very noisy receive information sent by the relay, and 2) an error occurs in the information sent by the source in previous block index. Thus, a bound to the destination data rate is:

$$R_k \leq \mathrm{I}\left(X_r\,; Y_d\right) + \mathrm{I}\left(X_k\,; Y_d|X_r\right) = \mathrm{I}\left(X_k X_r\,; Y_d\right) \tag{4.25}$$

and this guarantees the existence of an index at the destination's decoder. According to the Markov chain $(X_k, X_r) \leftrightarrow Y_d$, the probability of a confusing index error at destination's decoder is prevented if:

$$p\left(x_k, y_r, x_r, y_d\right) = p\left(x_k, x_r\right).p\left(y_d|x_k x_r\right) \tag{4.26}$$

The following theorem presents the upper bound capacity of DF technique:

[2]For generalized block Markov encoding, wherein instead the intermediate relay node only decodes part of the message transmitted by the source node see [166, Sec. 2.3.2].

Theorem 8 *The capacity of point-to-point relayed communication applying the DF technique is upper bounded to:*

$$\mathcal{C}_{DF}(k;r;d) = \sup_{p(x_k,x_r)} \min \left\{ \mathrm{I}(X_k\;;\;Y_r|X_r)\,,\,\mathrm{I}(X_k\,X_r\;;\;Y_d) \right\} \tag{4.27}$$

where supremum is over individual power constraints of source node and relay node and all joint distribution $p\left(x_k,x_r\right).p\left(x_r|x_k\right).p\left(y_d\,y_r|x_k x_r\right)$ *that is obtained by combination of Eqs. 4.24 and 4.26.*

This means that the broadcast data rate of DF technique is limited by the $k \rightarrow r$ channel, and the multiple-access capacity term is the same as that cutset. The DF strategy takes the maximum capacity when the decoding process at the relay is correctly done. Thus, the following relation can be stated:

$$\mathcal{C}_{DF}(k;r;d) = \lim_{\gamma_{kr}\rightarrow+\infty} \mathcal{C}_{cutset}(k;r;d) \tag{4.28}$$

Equivalently, the upper bound capacity of the DF strategy is achieved by that cutset where the relay node is located close to source node. The AWGN upper bound capacity of DF strategy is introduced by following theorem.

Theorem 9 *The AWGN DF capacity is upper bounded by:*

$$\begin{aligned}\mathcal{C}_{DF}(k;r;d) = \sup_{-1\leq\rho\leq1} \min &\left\{ C\left(\frac{h_{kr}\overline{p}_k\left(1-\rho^2\right)}{\sigma_w^2}\right), \right.\\ &\left. C\left(\frac{h_{kd}\overline{p}_k + h_{rd}\overline{p}_r + 2\rho\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r}}{\sigma_w^2}\right)\right\}\end{aligned}$$

Proof The AWGN of mutual information terms are calculated using Eqs. 4.9.

$$\begin{aligned}\mathrm{I}(X_k\;;\;Y_r|X_r) &= \frac{1}{2}\log_2\frac{\mathbb{V}\mathrm{ar}\left(Y_r|X_r\right)}{\mathbf{\Sigma}[Y_r|X_r X_k]} = \frac{1}{2}\log_2\frac{\mathbb{V}\mathrm{ar}\left(Y_r|X_r\right)}{\sigma_w^2} = \\ &= C\left(\frac{h_{kr}\overline{p}_k\left(1-\rho^2\right)}{\sigma_w^2}\right);\end{aligned}$$

$\mathrm{I}(X_k\,X_r\;;\;Y_d)$ was proved in Theorem 7. ■

Theorem 9 implies that, in a reliable network wherein the reception signal level from the source node at destination is stronger than that from the source node at relay, or equivalently $h_{kr} < h_{kd}$, the direct-link channel capacity outperforms the upper bound capacity of DF strategy.

4.6 Compress and forward technique

In a network where the relay may not be able to correctly decode the message transmitted by the sender, the DF strategy is not efficient. In this subsection we introduce a strategy for a situation wherein the relay observes a more corrupted signal than that observed at the destination. In compress and forward (CF) technique the relay should choose the best X_r to transmit in order to facilitate the source's transmission. CF is a block-transmission and side information encoder strategy. When the relay node has received a (very) noisy signal, instead of decoding it like the DF strategy, it extracts and estimates useful information from its received sequence and re-encodes this estimated signal toward the destination. The destination node will use this additional information as side information to help decode the original message received from the direct channel. In CF strategy the sender does not know about the estimated message, i.e. the source and relay does not cooperate. The CF method achieves its maximum capacity when the signal X_r approaches to the destination with infinite precision. So:

$$\mathcal{C}_{CF}(k;r;d) = \lim_{\gamma_{rd} \to +\infty} \mathcal{C}_{cutset}(k;r;d) \tag{4.29}$$

The CF strategy exploits the fact that the received signals at relay and destination nodes are statistically correlated, since they are both noisy versions of the same signal transmitted by the source node. The relay node, for instance, uses source coding with distortion to compress/quantize the received signal into an index. This index can then be channel coded and sent to the destination node.

The Wyner-Ziv scheme (see Sect. B.2) is the most efficient source coding technique. Wyner-Ziv source coding can be applied to perform rate distortion coding of the relay observation. We suppose the original message w_k is a sequence of B sub-messages $w_k^1, \ldots, w_k^b, \ldots, w_k^B \in \mathcal{W}_k$ and each sub-message is separately and fully encoded. There are four code spaces: the $\mathcal{X}_k$ and $\mathcal{X}_r$ codebooks respectively at the source and relay's encoders, the estimation codebook $\widehat{\mathcal{Y}}_r$ at the relay node, and finally the alphabet space $\mathcal{Y}_d$ at the destination's decoder. The encoder at the source node in block b produces a Gaussian random codeword represented by $X_k[b]\left(w_k^b\right)$. The relay and destination receive $Y_r[b]$ and $Y_d[b]$, respectively. Neither the destination nor the relay attempt to decode this information. The destination stores the received signal, whereas the relay, using $\widehat{\mathcal{Y}}_r$, estimates/quantizes the previous received $Y_r[b-1]$ to $\widehat{Y}_r[b-1]\left(k^b\right)$ where $k^b \in \{0, 1, ..., 2^{m\tilde{R}} - 1\}$ is an estimation/quantization index. The estimation index can be introduced as either the input of an estimation/compression function e.g. Hash function, or the index of a quantization table. Then, the relay encodes the estimated/compressed signal $\widehat{Y}_r[b-1]\left(k^b\right)$ using the Gaussian codebook $\mathcal{X}_r$ and transmit it to the destination. For the last block, the source transmits a termination code word $X_k(1)$.

$$Y_d[b] = \sqrt{h_{kd}}\, X_k[b]\left(w_k^b\right) + \sqrt{h_{rd}}\, X_r[b]\left(\widehat{Y}_r[b-1]\left(k^b\right)\right) + Z_d \tag{4.30a}$$

$$Y_r[b-1] \longrightarrow \widehat{Y}_r[b-1]\left(k^b\right) \tag{4.30b}$$

$$Y_r[b] = \sqrt{h_{kr}}\, X_k[b]\left(w_k^b\right) + Z_r \tag{4.30c}$$

The following factorization of the joint probability distribution function prevents from errors at the relay's encoder [166, Th. 2.5]:

$$p\left(x_k, x_r, y_r, \hat{y}_r, y_d\right) = p\left(x_k\right) \cdot p\left(y_r | x_r\, x_k\right) \cdot p\left(\hat{y}_r | y_r\, x_r\right) \tag{4.31}$$

This is due to the Markov condition $X_k \leftrightarrow Y_r|X_r \leftrightarrow \widehat{Y}_r$ for encoding and quantization processes. The decoding process at destination is perfectly done if the transmit signal X_r is successfully conveyed to the destination and there exists a unique joint decode index for the side information $Y_d|X_r$ and the already received signal X_k in the previous block index, i.e. $X_k \leftrightarrow Y_d|X_r$.

$$p\left(x_k, x_r, y_r, \hat{y}_r, y_d\right) = p\left(x_r\right) \cdot p\left(x_k\right) \cdot p\left(y_d | x_k\, x_r\right) \tag{4.32}$$

Combining Eqs. 4.31 and 4.32 the pmf is given as:

$$p\left(x_k, x_r, y_r, \hat{y}_r, y_d\right) = p\left(x_k\right) \cdot p\left(x_r\right) \cdot p\left(\hat{y}_r | y_r\, x_r\right) \cdot p\left(y_d\, y_r | x_k\, x_r\right) \tag{4.33}$$

In the following theorem we illustrate a relation between compression data rate and later process of encoding and transmitting data rate at the relay. This theorem established in the work of Cover and El. Gamal [159, Th. 6]. Here, we present a different proof based on Wyner-Ziv source coding.

Theorem 10 *Suppose that the full-duplex relay complies with CF strategy. For each joint pmf* $p\left(x_k\right).p\left(x_r\right).p\left(\hat{y}_r|y_r x_r\right).p\left(y_d y_r|x_k x_r\right)$ *there exists a sequence of* $\left(X_k\left(w_k\right), X_r\left(z\right), \widehat{Y}_r\left(k|z\right)\right)$ *whose symbols are randomly derived from the code books* $w_k \in \{0, 1, ..., 2^{mR_k} - 1\}$, $z \in \{0, 1, ..., 2^{m\widehat{R}} - 1\}$, $k \in \{0, 1, ..., 2^{m\tilde{R}} - 1\}$, *such that* $P_e\left(w_k \neq \hat{w}_k\right) \to 0$ *as* $m \to +\infty$ *if:*

$$R_k \leq \mathrm{I}\left(X_k\ ;\ Y_d\,\widehat{Y}_r | X_r\right); \tag{4.34a}$$

$$\textit{Subject to:}\quad \mathrm{I}\left(X_r\ ;\ Y_d\right) \geq \mathrm{I}\left(\widehat{Y}_r\ ;\ Y_r|\, X_r\, Y_d\right). \tag{4.34b}$$

Proof For this theorem, [166, Th. 2.5] and [159, Th. 6] present a demonstration based on block Markov encoding. Here, we look at this theorem as a bit rate encoding problem with side information. The relay node in block b generates the estimate message $\widehat{Y}_r[b-1]$ that is related to sub-message w_k^{b-1}. Then $\widehat{Y}_r[b-1]$ is encoded and transmit to the destination in block b and at the destination's decoder it will be joined with $Y_d[b-1]$. Suppose the destination's decoder is able to correctly decode the received message. The question is the smallest allowable encoding rate at the relay node. We emulate this problem by the following Wyner-Ziv problem: The destination's decoder has side information $Y_d[b-1]$ given $X_r[b]$. The relay is going to encode $Y_r[b-1]$ given $X_r[b]$. There is a tradeoff in terms of correlation between $Y_r[b-1]$ and $Y_d[b-1]$. The network is trying to exploit the side information at the destination's decoder and accordingly, we do not care about the amount of distortion. From the result of Wyner-Ziv we know that the minimum quantization rate $\widehat{R}$ is (Eq. B.32): $\widehat{R} = \inf\left[\mathrm{I}\left(\widehat{Y}_r\,;\, Y_r|X_r\right) - \mathrm{I}\left(\widehat{Y}_r\,;\, Y_d|X_r\right)\right]$

where $Y_d \leftrightarrow Y_r \leftrightarrow \widehat{Y}_r$. The decoding at the destination is correctly done if : $\widehat{R} \leq \mathrm{I}\,(X_r\ ;\ Y_d)$. Therefore:

$$\begin{aligned}
\mathrm{I}\,(X_r\ ;\ Y_d) \geq \widehat{R} \geq \mathrm{I}\left(\widehat{Y}_r\,;\,Y_r|X_r\right)\ -\ \mathrm{I}\left(\widehat{Y}_r\,;\,Y_d|X_r\right) \\
= \mathrm{I}\left(\widehat{Y}_r\,;\,Y_r\,X_r\right) - \mathrm{I}\left(\widehat{Y}_r\,;\,X_r\right) - \mathrm{I}\left(\widehat{Y}_r\,;\,Y_d|X_r\right) \\
= \mathrm{I}\left(\widehat{Y}_r\,;\,Y_r\,X_r\right) - \mathrm{I}\left(\widehat{Y}_r\,;\,Y_d\,X_r\right) \\
= \mathrm{I}\left(\widehat{Y}_r\,;\,Y_r\,Y_d\,X_r\right) - \mathrm{I}\left(\widehat{Y}_r\,;\,Y_d\,X_r\right) \\
= \mathrm{I}\left(\widehat{Y}_r\,;\,Y_r\,|X_r\,Y_d\right)
\end{aligned}$$

The equalities follows from the chain rule. After having received the information $\widehat{Y}_r[b-1]$ about $X_k[b-1]$ given $X_r[b]$, the destination's decoder uses $\left(Y_d[b-1]\,,\,\widehat{Y}_r[b-1]\mid X_r[b]\right)$ to decode the message w_k^b. Decoding will be correct if the achievable rate R_k is upper bounded by:

$$R_k \leq \mathrm{I}\,(X_k\,;\,Y_d) + \mathrm{I}\left(X_k\,;\,\widehat{Y}_r|\,X_r\right) = \mathrm{I}\left(X_k\,;\,Y_d\,\widehat{Y}_r|\,X_r\right).$$

■

The inequality (4.34b), physically means that the compressing rate $Y_r \rightarrow \widehat{Y}_r$, must be smaller than the rate used to transmit data from the relay to the destination. Correspondingly, the h_{rd} channel condition has to be sufficiently good to reliably convey the compressed signal $\widehat{Y}_r$. The CF protocol has at least two interesting properties: 1) It does not require the relay to be in a better reception condition than the destination, and 2) the broadcast capacity reaches that the cutset bound when $Y_r = \widehat{Y}_r$ (inequality 4.34a).

The amount of information extracted and forwarded to the destination depends on the $r \rightarrow d$ channel capacity. The relay transmits $z_b = \widehat{Y}_r[b]\left(k^{b+1}\right)$ to destination, in block $b+1$, via $X_r[b+1]\,(z_b)$. The error probability of decoding X_r at destination vanishes if the data rate between relay and destination is bounded to: $\widehat{R} \leq \mathrm{I}\,(\,X_r\ ;\ Y_d)$. The value z_b is about $X_k[b]\left(w_k^b\right)$ and the relay does not decode the received signal. Therefore: $R_k \leq \widehat{R}$, otherwise the $r \rightarrow d$ channel is not able to transmit all signals. Now, we can derive the upper bound capacity of CF scheme with the following theorem.

Theorem 11 *An upper bound of the capacity of the CF relay channel is given by:*

$$\mathcal{C}_{CF}(k;r;d) = \sup_{p(x_k,x_r)} \min \left\{ \mathrm{I}\left(X_k \; ; \; Y_d \, Y_r | X_r \right) , \mathrm{I}\left(X_r \; ; \; Y_d \right) \right\} \tag{4.35}$$

where the supremum is over individual power constraints of source node and relay node and all joint distributions $p\left(x_k\right).p\left(x_r\right).p\left(\hat{y}_r|y_r\, x_r\right).p\left(y_d\, y_r|x_k\, x_r\right)$.

Theorem 11 results: the multiple-access capacity of the CF strategy is limited to the $r \to d$ channel data rate, whereas the broadcast capacity is equal to that cutset. The AWGN capacity is formulated as:

Theorem 12 *The AWGN CF capacity is upper bounded by:*

$$\mathcal{C}_{CF}(k;r;d) = \sup_{-1\leq\rho\leq1} \min \left\{ C\left(\frac{\left(h_{kd}+h_{kr}\right)\overline{p}_k\left(1-\rho^2\right)}{\sigma_w^2}\right), \right.$$
$$\left. C\left(\frac{h_{kd}\overline{p}_k\rho^2 + h_{rd}\overline{p}_r + 2\rho\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r}}{h_{kd}\overline{p}_k\left(1-\rho^2\right)+\sigma_w^2}\right) \right\}$$

Proof The broadcast term was proved in Theorem 7. The multiple access term is calculated using Eqs. 4.9.

$$\mathrm{I}(X_r\,;\,Y_d) = \frac{1}{2}\log_2\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right)}{\mathbb{V}\mathrm{ar}\left(Y_d|\,X_r\right)} = C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right)-\mathbb{V}\mathrm{ar}\left(Y_d|\,X_r\right)}{\mathbb{V}\mathrm{ar}\left(Y_d|\,X_r\right)}\right) =$$
$$= C\left(\frac{h_{kd}\overline{p}_k\rho^2 + h_{rd}\overline{p}_r + 2\rho\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r}}{h_{kd}\overline{p}_k\left(1-\rho^2\right)+\sigma_w^2}\right).$$

■

If the relay node is located close to the destination, i.e. $\gamma_{kr} < \gamma_{rd}$, the upper bound capacity is achieved by the broadcast term and with $\rho = 0$. On the other hand if the relay node has good reception condition, i.e. $\gamma_{kr} > \gamma_{rd}$ the upper bound capacity is achieved by the second term and with $\rho = 1$.

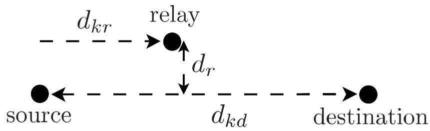

Fig. 4.3: *A single relay communication network scenario.*

4.7 Case study

In this section, we exemplify the various outer region bounds presented so far in this chapter. We consider a point-to-point relayed communication with Gaussian channels wherein the transmitter k, relay r, and sink d are located as sketched in Fig. 4.3. We assume a vertical distance of d_r between the relay and the $k \to d$ direct-link. The path condition values $h_{kr} = h_{rd} = h_{kd} = 1$ are scaled with respect to Notation 2. First, we experiment a high SNR environment and suppose the source and destination are located at a distance of $d_{kd} = 1\,\text{m}$, and the relay is located at a vertical distance of $d_r = 0.1\,\text{m}$ and it is horizontally moving from $d_{kr} = -0.5\,\text{m}$ to $d_{kr} = 1.5\,\text{m}$. Fig. 4.4 plots various data rates for $\overline{p}_k = \overline{p}_r = 100\,\text{mW}$, and $\sigma_w^2 = 1\,\mu\text{W}$. The curve labeled AF shows the outer region of AF strategy with the largest possible scaling factor E in Eq. 4.15. The curve labeled ρ plots a particular value of the correlation coefficient we tried for this example. It is important to note that, changing the correlation function meaningfully results in change of the curves labeled cutset, DF, and CF (strategies).

In our particular case study, the AF and MRC techniques show a very good performance. This is due to the fact of the short distances result in a high SNR. As the relay moves toward the destination ($d_{kr} \to 1$), the achieved signal at the relay becomes weaker and this significantly decreases the AF data rate. Generally, the AF and MRC techniques would be useful when the relay is located so as to be able to perfectly receive and deliver signals, or equivalently, the relay is equidistant from the transmitter and the sink

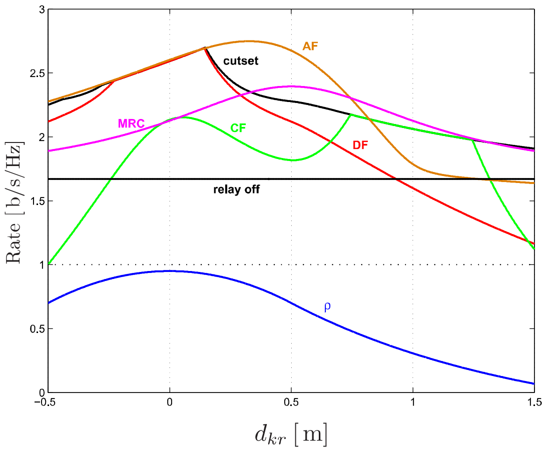

Fig. 4.4: *Rates for one relay with* $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 1\,\mathrm{m}$, *and* $d_r = 0.1\,\mathrm{m}$.

($d_{kr} \to 0.5$).

As the relay moves toward transmitter ($d_{kr} \to 0$), the rates of cutset, DF, and CF become:

$$\mathcal{C}_{cutset}(k;r;d) = \mathcal{C}_{DF}(k;r;d) \to C\left(\frac{h_{kd}\overline{p}_k + h_{rd}\overline{p}_r + 2\rho\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r}}{\sigma_w^2}\right)$$

$$\mathcal{C}_{CF}(k;r;d) \to C\left(\frac{h_{kd}\overline{p}_k\rho^2 + h_{rd}\overline{p}_r + 2\rho\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r}}{h_{kd}\overline{p}_k\left(1-\rho^2\right) + \sigma_w^2}\right)$$

and the DF strategy shows better performance. Correspondingly, as the relay is placed close to the destination ($d_{kr} \to 1$), the various data rates become:

$$\mathcal{C}_{DF}(k;r;d) \to C\left(\frac{h_{kr}\overline{p}_k\left(1-\rho^2\right)}{\sigma_w^2}\right)$$

$$\mathcal{C}_{cutset}(k;r;d) = \mathcal{C}_{CF}(k;r;d) \to C\left(\frac{\left(h_{kd}+h_{kr}\right)\overline{p}_k\left(1-\rho^2\right)}{\sigma_w^2}\right)$$

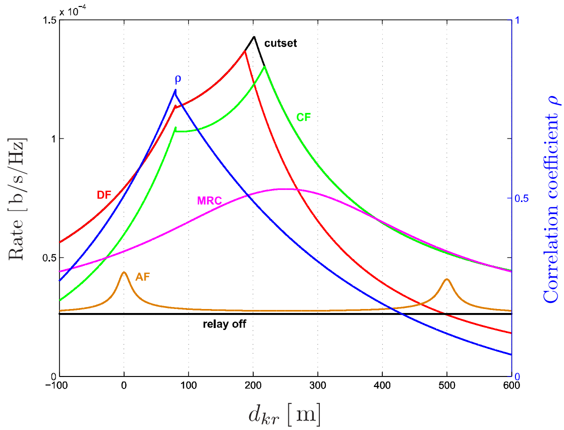

Fig. 4.5: *Rates for one relay with* $\overline{p}_k = \overline{p}_r = 100\,\text{mW}$*,* $\sigma_w^2 = 1\,\mu\text{W}$*,* $d_{kd} = 500\,\text{m}$*, and* $d_r = 10\,\text{m}$.

and the CF technique achieves a higher capacity. In the following chapters we will show that these limiting principles are also generalizable in a multiple sources, multiple parallel relays network.

Now, we consider a point-to-point relayed connection in a low SNR regime. The transmitter and the destination are placed at a distance of $d_{kd} = 500\,\text{m}$, and the relay is moving in a range of $d_{kr} = -100 \div 600\,\text{m}$ with a vertical distance of $d_r = 10\,\text{m}$. We set the same E as the previous simulation. Fig. 4.5 plots various data rates for $\overline{p}_k = \overline{p}_r = 100\,\text{mW}$, and $\sigma_w^2 = 1\,\mu\text{W}$.

We draw a different experimental function for correlation value which is the curve labeled ρ. From Fig. 4.5, it is clearly derived that the AF technique is not quite useful in a low SNR network, whereas the MRC technique performs much better than AF. This is because the received signal at the relay is very noisy, and also the scaling factor is higher than that in the previous scenario. In this situation (almost) no signal perfectly approaches to the destination. The unique point that hold Eq. 4.18 is at $d_{kr} = 0$. Our experiments in a given scenario with different parameters result reducing the amplification factor does not effect the AF data rate.

As can be seen, in low SNR, the DF and CF coding techniques show significantly higher rates than in the direct-link. Like the previous scenario, when the relay is close to the source node, the DF strategy achieves the cutset bound, and instead, when the relay is close to the destination CF technique is much better.

4.8 Summary

The data rate of a relayed communication is a function of both relaying strategy and positions of the nodes. The DF capacity achieves the cutset bound when the channel between the source and relay is sufficiently good, i.e. the relay is close to the source, and accordingly the received information is perfectly decoded at the relay. The CF scheme can outperform DF when decoding is less reliable at the relay node and the channel between the relay and destination is sufficiently good, i.e. the relay is closer to the destination and more information can be conveyed to the destination through the relayed channel. Moreover, CF also provides a more general form of compression compared to the simple analog scaling done in AF. In a low SNR environment, DF and CF technique achieve a significantly high data rate, whereas AF shows a good performance in high SNR regime.

Chapter 5

Multiple parallel relayed point-to-point communication

In this chapter, we investigate the upper bound capacity of point-to-point cooperative communication assisted by more than one relay in particular by a bank of R "parallel" relays. As the number of relays increases, more radio resources and more degrees of freedom can be jointly utilized to assist the source's transmission. However, to exploit these advantages, one must overcome the challenges posed by the individual power constraints. We will show that increasing the number of relays does not necessarily raise the upper bound of the capacity region, and may even deteriorate data rate and harmfully increase the network cost in terms of power expenditure.

The main focus of this chapter is to derive the highest achievable data rate of the transmitter, applying different strategies at the relay nodes. First, in Sect. 5.1 we describe the channel model of a multiple parallel relayed point-to-point communication. The cutset theorem is used in Sect. 5.2. The following three sections study the three relay strategies. Sect. 5.3 is devoted to AF strategy, Sect. 5.4 to DF, and Sect. 5.5 to CF. In Sect. 5.6 we suppose some relay adopt the DF strategy and the others relays apply CF technique. We exemplify the results in Sect. 5.7. Finally, we conclude in Sect. 5.8.

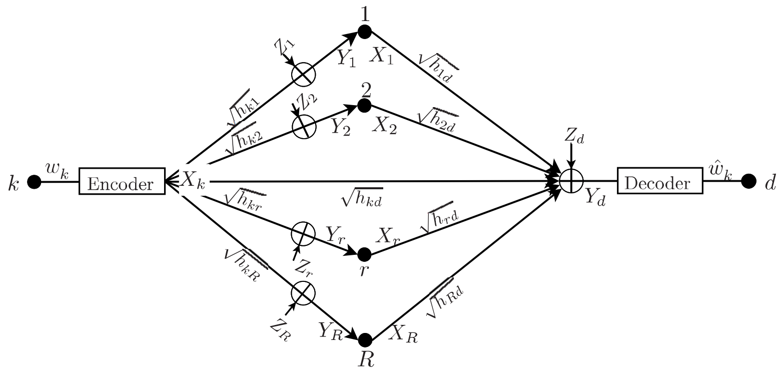

Fig. 5.1: *One source, multiple parallel relays, one destination cooperative communication network.*

5.1 System model

We study a cooperative network that consists of one source k, one destination d, and an arbitrary number of parallel relays belonging to set $\mathcal{R} = [1, \ldots, r, \ldots, R]$. We assume there is no link between relays and the communication is full-duplex mode, so that relay assisted transmissions must be conducted over one phase and all channels are always busy. The cooperative network of Fig. 5.1 consists of $2R+2$ alphabet spaces: $\mathcal{X}_k$ at the transmitter's encoder, $\mathcal{X}_r$ and $\mathcal{Y}_r$ at each relay r, and finally $\mathcal{Y}_d$ at destination's decoder.

The sent message w_k is a sequence of sub-messages $w_k^1, \ldots, w_k^b, \ldots, w_k^B$ uniformly drawn from a message set with size m and rate R_k represented by $\mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$. First, the transmitter encodes the message w_k in a symbol sequence $X_k[1]\left(w_k^1\right), \ldots, X_k[b]\left(w_k^b\right), \ldots, X_k[B]\left(w_k^B\right)$ under the constraint that $\mathbb{E}\left\{X_k^2\right\} \leq \overline{p}_k$, with $\mathbb{E}\left\{X_k\right\} = 0$. Each $X_k[b]\left(w_k^b\right)$ represents the bth random Gaussian codeword in the alphabet space $\mathcal{X}_k$. In each block with index b, the $X_k[b]\left(w_k^b\right)$ is broadcast to the destination and to all relays. A noisy version of the transmitted signal approaches the relay r as:

$$Y_r[b] = \sqrt{h_{kr}}\, X_k[b]\left(w_k^b\right) + Z_r \tag{5.1}$$

where $Z_r \sim \mathcal{N}\left(0, \sigma_w^2\right)$ is additive white Gaussian noise at the proper relay. At the same time, each relay r executes a function of $\mathcal{Y}_r \to \mathcal{X}_r$. That is, every relay r processes the received signal in the previous block index, generates a new signal and then forward the new symbol $X_r[b]\left(w_k^{b-1}\right)$ to destination. The information $X_r[b]\left(w_k^{b-1}\right)$ is the output of the relay r's deterministic function that depends on the specific cooperative strategy whose inputs are the previous received signals:

$$X_r[b]\left(w_k^{b-1}\right) = f_r^b(Y_r[b-1], Y_r[b-2], ..., Y_r[1]); \tag{5.2}$$

Each relay r has its own individual power constraint as $\mathbb{E}\left\{X_r^2\right\} \le \overline{p}_r$. The symbol $X_r[b]\left(w_k^{b-1}\right)$ is forwarded to the destination simultaneously with broadcasting a new symbol by the source. The receive signal at the destination is given by:

$$Y_d[b] = \sqrt{h_{kd}} X_k[b]\left(w_k^b\right) + \sum_{r\in\mathcal{R}} \sqrt{h_{rd}}\, X_r[b]\left(w_k^{b-1}\right) + Z_d \tag{5.3}$$

where $Z_d \sim \mathcal{N}\left(0, \sigma_w^2\right)$. The decoding function at the destination is a Gaussian de-mapping function of $\mathcal{Y}_d \to \mathcal{W}_k$ with error probability:

$$P_e^m = 2^{-mR_k} \cdot \sum_{w_k} \Pr\left\{\hat{w}_k \neq w_k | w_k \text{ was sent}\right\} \tag{5.4}$$

based on the assumption that the messages are independent, and uniformly distributed over the alphabet space $\mathcal{W}_k$. The minimum rate R_k is achievable if there exists a sequence of code $\left(m, 2^{mR_k}\right)$ for which P_e^m is arbitrarily close to zero when $m \to \infty$.

To go on with our analysis, we define the parameter ρ_{kr} as the correlation between the output sequences of X_k and X_r where $r \in \mathcal{R}$, i.e. $\rho_{kr} = \frac{\mathbb{E}\left\{X_k X_r\right\}}{\sqrt{\mathbb{E}\left\{X_k^2\right\}\mathbb{E}\left\{X_r^2\right\}}}$, and using a similar formulation, we represent the parameter τ_{rj} as the correlation between the outputs of two relay nodes $r, j \in \mathcal{R}$, i.e. the statistical correlation between X_r and X_j.

5.2 Cutset upper bound

To compute the achievable cutset capacity of one source, multiple parallel relays, one destination cooperative network, we resort on the general representation of M. R. Aref's tool [161, Th. 3.4] .

Prop. 2 *[161, Th. 3.4] A network consisting of one transmitter, multiple relays, and one destination satisfies:*

$$\mathcal{C}_{cutset}(k;\mathcal{R};d) = \sup_{p(x_k,\boldsymbol{x}_{\mathcal{R}})} \min_{\mathcal{T}\subseteq\mathcal{R}} \{\mathrm{I}\,(X_k\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C})\} \tag{5.5}$$

where maximization is subject to the power constraints defined by the network, and the channel condition $p(y_d\,\boldsymbol{y}_{\mathcal{R}}\,|\,x_k\,\boldsymbol{x}_{\mathcal{R}})$.

The aim of Prop. 2 is to find the minimum data flow between the source and receiver via one of the $\sum_{r=0}^{R}\binom{R}{r}$ possible relay selections. There are 2^R possibilities to select subset $\mathcal{T}\subset\mathcal{R}$, known as *network cuts.* In Eq. 5.5, let us suppose the mutual information term achieves its minimum with an $\mathcal{A}\subseteq\mathcal{R}$, i.e. $\mathcal{C}_{cutset}(k;\mathcal{R};d) = \sup_{p(x_k,\boldsymbol{x}_{\mathcal{R}})} \mathrm{I}\,(X_k\boldsymbol{X}_{\mathcal{A}}\,;\,Y_d\boldsymbol{Y}_{\mathcal{A}^C}|\boldsymbol{X}_{\mathcal{A}^C})$. This means that the members of $\mathcal{A}$ are in a sufficiently good conditions to (perfectly) receive signals from the transmitter rather than to deliver signals to the destination. On the other hand, the relays in $\mathcal{A}^C$ are in a good positions to (error freely) convey signals to the destination.

The relation stated in Prop. 2 was proved in a general case where the relays are physically connected to each other. For the cooperative network scenario of Fig. 5.1, where there is no cooperation among the relays, references [183, 187, 188] represent remarkable contributions which concern the maximum achievable capacity. In [188, Sec. 5.4] and [187, Sec. 3.2.2] only two broadcast ($\mathcal{T}=\emptyset$) and multiple-access ($\mathcal{T}=\mathcal{R}$) cuts have been considered and the power constraint was not identified for each node. Here, for the scenario of Fig. 5.1, first we demonstrate that $\mathcal{C}_{cutset}(k;\mathcal{R};d)$ takes the maximum when the Gaussian variables $\boldsymbol{X}_{\mathcal{R}}$ are fully correlated, i.e. $\tau_{rj}=1\ \ \forall r,j\in\mathcal{R}$. Then, we approach the upper bound capacity considering all 2^R network cuts.

Theorem 13 *In the network of Fig. 5.1, the* $\mathcal{C}_{cutset}(k;\mathcal{R};d)$ *takes the maximum value when the Gaussian variables* $\boldsymbol{X}_{\mathcal{R}}$ *are fully correlated.*

Proof We expand Eq. 5.5 as:

$$\mathrm{I}\,(X_k\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}) = \mathrm{H}\,(\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}) - \mathrm{H}\,(\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}X_k)\,.$$

In the second term, $\mathrm{H}\,(\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}X_k)$ is not a function of τ_{rj}. To prove the theorem, it is enough to show that the maximum of $\mathrm{H}\,(\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C})$ is achieved when $\tau_{rj} = 1\ \forall\,r, j \in \mathcal{R}$, irrespective of the value ρ_{kr}. This is somehow obvious for Gaussian random variables because conditioning reduces entropy except the case of fully correlated given variables condition. That is $\mathrm{H}\,(A|B) \geq \mathrm{H}\,(A|B, C)$, with equality whenever B and C are fully correlated, or A is conditionally independent of C given B, i.e. $A \leftrightarrow B \leftrightarrow C$. ■

The following results can be derived from Theorem 13. First, it is reasonable if we assume the same correlation coefficient between X_k and every X_rs. Thus, in the rest of this chapter, we suppose $\rho = \rho_{kr}\ \forall\,r \in \mathcal{R}$. The second result is that $\boldsymbol{\Sigma}[Y|\boldsymbol{X}_{\mathcal{R}}] = \mathbb{V}\mathrm{ar}(Y|X_r)$.

In the following, we find a general formula, considering all 2^R cuts, for a relayed communication consisting of one source, multiple parallel relays and one destination where each node has its own individual power constraint.

Theorem 14 *For one source, multiple parallel relays, one destination network, the cutset upper bound of Prop. 2 is shortened to:*

$$\mathcal{C}_{cutset}(k;\mathcal{R};d) = \sup_{p(x_k,\boldsymbol{x}_{\mathcal{R}})} \min\left\{\min_{r\in\mathcal{R}}\{\mathrm{I}\,(X_k\,;\,Y_d\,Y_r|X_r)\}\,;\ \mathrm{I}\,(X_k\,\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d)\right\} \tag{5.6}$$

Proof The second term is the multiple-access capacity of all relays and the source node at destination, i.e. $\mathcal{T} = \mathcal{R}$. To prove the theorem, it is enough to show that the following equality holds:

$$\min_{\mathcal{T}\subsetneq\mathcal{R}}\ \{\mathrm{I}\,(X_k\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C})\} = \min_{r\in\mathcal{R}}\{\mathrm{I}\,(X_k\,;\,Y_d\,Y_r|X_r)\}.$$

The expansion of the left-hand side gives:

$$\mathrm{I}\,(X_k\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}) = \underbrace{\mathrm{I}\,(\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C})}_{(1)\,=\,0} + \underbrace{\mathrm{I}\,(X_k\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}})}_{(2)}$$

$$(1):\ \mathrm{I}\,(\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}) = \frac{1}{2}\log_2\frac{\boldsymbol{\Sigma}[\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}]}{\boldsymbol{\Sigma}[\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}]} = $$

$$= \frac{1}{2}\log_2\frac{\boldsymbol{\Sigma}[\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|X_{r\in\mathcal{T}^C}]}{\boldsymbol{\Sigma}[\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|X_{r\in\mathcal{T}^C}]} = 0.$$

$$(2):\ \mathrm{I}\,(X_k\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}) = \mathrm{I}\,(X_k\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|X_r)$$

Until now, we demonstrated that:

$$\min_{\mathcal{T}\subsetneq\mathcal{R}}\ \{\mathrm{I}\,(X_k\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C})\} = \min_{\mathcal{T}\subsetneq\mathcal{R}}\ \{\mathrm{I}\,(X_k\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|X_r)\}$$

Since $\mathcal{T}$ is every strictly subset of $\mathcal{R}$, it is clear to conclude that:

$$\min_{\mathcal{T}\subsetneq\mathcal{R}}\ \{\mathrm{I}\,(X_k\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|X_r)\} = \min_{r\in\mathcal{R}}\{\mathrm{I}\,(X_k\,;\,Y_d\,Y_r|X_r)\}.$$

■

The result of Eq. 5.6 is the minimum value between $R+1$ mutual information terms. The first term of the $\mathcal{C}_{cutset}(k;\mathcal{R};d)$ is the broadcast capacity from the source node to the destination and the relay with which has the minimum information flow rate. The second term is the multiple-access capacity of all relays and the transmitter at destination. The novel result of Theorem 14 is: *the upper bound cutset capacity of a point-to-point multiple parallel relayed network is achieved either using one relay only or using all relays together.* If all relays are in good conditions to transmit signals to the destination rather than to receive signals from the source, $\mathcal{C}_{cutset}(k;\mathcal{R};d) = \mathcal{C}_{cutset}(k;r;d)$ wherein r is the relay with which the source node achieves the smallest broadcast capacity of $\mathrm{I}\,(X_k\,;\,Y_d\,Y_r|X_r)$. *In such a network the existence of all others relays is useless.* On the other hand, if all relays are in good conditions to receive data from the transmitter rather than to convey signals to the destination, the $\mathcal{C}_{cutset}(k;\mathcal{R};d)$ is achieved with maximization of the multiple-access capacity of all relays and the source node at the destination, i.e. $\mathrm{I}\left(X_k\, \boldsymbol{X}_{\mathcal{R}}\,;\, Y_d\right)$. As can be seen, the capacity of the broadcast term of $\mathcal{C}_{cutset}(k;\mathcal{R};d)$ is much less than that presented by [187, Sec. 3.2.2] and [188, Sec. 5.4]. This is because, [187] and [188] consider only two broadcast and multiple-access cuts rather than all 2^R possible cuts.

At this point, we recall some useful equalities relevant to the equations of Fig. 5.1, with $\tau_{rj} = 1\ \forall\, r, j \in \mathcal{R}$. To review the algebraic manipulation of the following formulas see Appendix A.

$$\mathbb{V}\mathrm{ar}\left(Y_d\right) = h_{kd}\overline{p}_k + \sum_{r\in\mathcal{R}} 2\rho\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r} + \sum_{r\in\mathcal{R}}\sum_{j\in\mathcal{R}} \sqrt{h_{rd}\overline{p}_r h_{jd}\overline{p}_j} + \sigma_w^2 \tag{5.7a}$$

$$\boldsymbol{\Sigma}[X_k|\boldsymbol{X}_{\mathcal{R}}] = \mathbb{V}\mathrm{ar}\left(X_k|X_r\right) = \overline{p}_k\left(1-\rho^2\right) \qquad \forall\, r\in\mathcal{R} \tag{5.7b}$$

$$\boldsymbol{\Sigma}[Y_r|\boldsymbol{X}_{\mathcal{R}}] = \mathbb{V}\mathrm{ar}\left(Y_r|X_r\right) = h_{kr}\overline{p}_k\left(1-\rho^2\right) + \sigma_w^2 \tag{5.7c}$$

$$\boldsymbol{\Sigma}[Y_d|\boldsymbol{X}_{\mathcal{R}}] = \mathbb{V}\mathrm{ar}\left(Y_d|X_r\right) = h_{kd}\overline{p}_k\left(1-\rho^2\right) + \sigma_w^2 \tag{5.7d}$$

$$\boldsymbol{\Sigma}[Y_dY_r|\boldsymbol{X}_{\mathcal{R}}] = \boldsymbol{\Sigma}[Y_dY_r|X_r] = \left(h_{kd}+h_{kr}\right)\overline{p}_k\left(1-\rho^2\right) + \sigma_w^2 \tag{5.7e}$$

In the following we introduce the capacity of the cooperative framework depicted by Fig. 5.1 with Gaussian channels.

Theorem 15 *The AWGN cutset upper bound of the capacity one source, multiple parallel relays is:*

$$\mathcal{C}_{cutset}(k;\mathcal{R};d) = \sup_{-1\le\rho\le1}\ \min\left\{\min_{r\in\mathcal{R}}\left\{C\left(\frac{\boldsymbol{\Sigma}[Y_dY_r|X_r]-\sigma_w^2}{\sigma_w^2}\right)\right\},\ C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right)-\sigma_w^2}{\sigma_w^2}\right)\right\}$$

which is calculated using Eqs. 5.7.

Proof

$$\mathrm{I}(X_k\,;\,Y_d\,Y_r|X_r) = \frac{1}{2}\log_2\frac{\boldsymbol{\Sigma}[Y_dY_r|X_r]}{\boldsymbol{\Sigma}[Y_dY_r|X_rX_k]} \overset{(1)}{=} \frac{1}{2}\log_2\frac{\boldsymbol{\Sigma}[Y_dY_r|X_r]}{\sigma_w^2} = $$
$$= C\left(\frac{\boldsymbol{\Sigma}[Y_dY_r|X_r]-\sigma_w^2}{\sigma_w^2}\right);$$

(1) is derived from the fact that $\mathbf{\Sigma}[Y_dY_r|X_rX_k] = \mathbf{\Sigma}[Y_dY_r|\boldsymbol{X}_{\mathcal{R}}X_k]$ and $Y_dY_r|\boldsymbol{X}_{\mathcal{R}}X_k$ is a function of independent variables Z_d and Z_r.

$$\mathrm{I}(X_k\, \boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d) = \frac{1}{2}\log_2 \frac{\mathbb{V}\mathrm{ar}\,(Y_d)}{\mathbf{\Sigma}[Y_d|X_k\, \boldsymbol{X}_{\mathcal{R}}]} = \frac{1}{2}\log_2 \frac{\mathbb{V}\mathrm{ar}\,(Y_d)}{\sigma_w^2} = \\ = C\left(\frac{\mathbb{V}\mathrm{ar}\,(Y_d) - \sigma_w^2}{\sigma_w^2}\right).$$

■

If all relays are much closer to the destination than source node, or equivalently $\gamma_{kr} < \gamma_{rd}\ \forall\ r \in \mathcal{R}$, then $\mathcal{C}_{cutset}(k;\mathcal{R};d) = \mathcal{C}_{cutset}(k;r;d)$ wherein r is the relay with the weakest γ_{kr} channel. This means that the others relays make only "crowd" and increase the overall power consumption. In such a relay assisted network the supremum is achieved with $\rho = 0$. On the other hand, if all relays are located in the contrary positions, i.e. they are much closer to the transmitter than destination, $\mathcal{C}_{cutset}(k;\mathcal{R};d)$ is achieved with the multiple-access term (the second term) and the supremum is achieved when X_k and $\boldsymbol{X}_{\mathcal{R}}$ are fully correlated, i.e. $\rho = 1$.

In a multiple parallel relay network wherein all relays are well located to correctly receive signals from the transmitter, i.e. $\gamma_{kr} > \gamma_{rd}$, the $\mathcal{C}_{cutset}(k;\mathcal{R};d)$ dominates the parallel channels capacity that is $C(\gamma_{kd} + \sum_{r\in\mathcal{R}}\gamma_{rd})$. This is because Gaussian variables $\boldsymbol{X}_{\mathcal{R}}$ are fully correlated. In such a reliable network, adding a new relay, always close to the source node, expands the upper bound of the cutset capacity region, up to the number of relays until which the multiple-access capacity does not exceed that broadcast (the first term).

Another interesting result of the Theorem 15 is: *adding a new relay does not always increase the cutset upper bound capacity.* In a one source-multiple parallel relays-one destination system, locating a new relay very close to the destination may decrease the upper bound capacity, and adding a new relay very close to the transmitter may do not change upper bound capacity.

Let us start to study different relaying strategies in a point-to-point multiple parallel relayed network. First, we try to calculate the upper bound capacity of the simplest relaying strategy, i.e. amplify and forward (AF) technique.

5.3 Amplify and forward technique

In a reliable network applying AF technique, the relays do not have any code space. There exist two Gaussian codebooks: $\mathcal{X}_k$ at the source's encoder, and $\mathcal{Y}_d$ at the destination's decoder. Suppose a transmit message w_k which is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B$ and each sub-message w_k^b is uniformly drawn from $\mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$. Each sub-message w_k^b is separately encoded to $X_k[b]\left(w_k^b\right)$ under the constraint that $\mathbb{E}\left\{X_k^2\right\} \leq \overline{p}_k$. In each block index b, each relay r scales the amplitude of the analog observed signal $Y_r[b-1]$ as:

$$\begin{aligned} X_r[b]\left(w_k^{b-1}\right) &= E_r[b].Y_r[b-1] \\ &= E_r[b].\left(\sqrt{h_{kr}}\,X_k[b-1]\left(w_k^{b-1}\right) + Z_r\right) \end{aligned} \tag{5.8}$$

wherein the amplification factor E_r is chosen so as to satisfy the proper relay's power constraint. Each relay node has its own power constraint as $\mathbb{E}\left\{X_r^2\right\} \leq \overline{p}_r$. We assume all channels are slow time-varying. So that we can assume $E_r[b] = E_r$ and

$$|E_r|^2 \leq \frac{\overline{p}_r}{\sigma_w^2 + h_{kr}\overline{p}_k} \tag{5.9}$$

As can be seen, if $\sigma_w^2 + h_{kr}\overline{p}_k \gg \overline{p}_r$ the (amplification of the) relay r is useless. Replacing Eq. 5.8 into (5.3), the received signal at the destination is:

$$\begin{aligned} Y_d[b] = \sqrt{h_{kd}}\,X_k[b]\left(w_k^b\right) + \sum_{r\in\mathcal{R}} |E_r|\sqrt{h_{kr}\,h_{rd}}\,X_k[b-1]\left(w_k^{b-1}\right) + \\ + \sum_{r\in\mathcal{R}} |E_r|\sqrt{h_{rd}}\,Z_r + Z_d \end{aligned} \tag{5.10}$$

The relay nodes do not regenerate any new code, and consequently the complexity of this scheme is low. Since every relay node amplifies whatever it receives, including noise, it is mainly useful in high SNR environments. Increasing the amplification factor E increases the noise of ISI at the destination, and also significantly increase the noise of ICI where there is only one antenna at the destination. The relay should thus transmit with an appropriate power where the network is able to adjust the power of the relay. By Eq. 5.10, the maximum data rate of the AF scheme is formulated as:

$$\mathcal{C}_{AF}(k;\mathcal{R};d) = C\left(\left(\sqrt{h_{kd}} + \sum_{r\in\mathcal{R}} |E_r|\sqrt{h_{kr}h_{rd}}\right)^2 \frac{\overline{p}_k}{\left(1+\sum_{r\in\mathcal{R}}|E_r|^2 h_{rd}\right)\sigma_w^2}\right) \quad (5.11)$$

[184, p. 46] demonstrates that under the condition

$$|E_r| \leq \gamma_{kr} \quad (5.12)$$

the $\mathcal{C}_{AF}(k;\mathcal{R};d)$ outperforms the capacity of the maximal ratio combining (MRC) technique that is:

$$\mathcal{C}_{MRC}(k;\mathcal{R};d) = C\left(\gamma_{kd} + \sum_{r\in\mathcal{R}} \frac{\gamma_{kr}\,\gamma_{rd}}{\gamma_{kr}+\gamma_{rd}}\right) \quad (5.13)$$

5.4 Decode and forward technique

Here we study a well known regular encoding called Block Markov encoding, that is concerned with $2R+2$ Gaussian code spaces: $\mathcal{X}_k$ at the transmitter's encoder, $\mathcal{Y}_r$ and $\mathcal{X}_r$ $\forall\, r \in \mathcal{R}$ at the relays' decoders and encoders respectively, and finally $\mathcal{Y}_d$ at the destination's decoder. The transmit message w_k is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B$. Each sub-message w_k^b is uniformly drawn from alphabet space $\mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$, and then it is separately and fully encoded to a Gaussian code.

The encoder at the transmitter has one block memory that, applying multiplex coding, assigns a Gaussian random number $X_k[b](w_k^{b-1}, w_k^b)$ to each w_k^b. In each block b, the source node broadcasts $X_k[b](w_k^{b-1}, w_k^b)$ to all relays and to the destination. Each relay r receives the signal:

$$Y_r[b] = \sqrt{h_{kr}}\, X_k[b](w_k^{b-1}, w_k^b) + Z_r \tag{5.14}$$

As the communication is performed in full-duplex mode, every relay decodes and then re-encodes the previous received message and sends it to the destination. Each relay $r \in \mathcal{R}$ tries to correctly and fully decode the received message using the code space $\mathcal{Y}_r$, and then re-encodes it using $\mathcal{X}_r$. A DF reliable channel can achieve the highest possible capacity when the decoding processes at all relays are done with vanishing small BER. To this end, for every $k \to r$ channel it must guaranteed that:

$$R_k \leq \mathrm{I}\,(X_k\,;\,Y_r|X_r) \qquad \forall\, r \in \mathcal{R}. \tag{5.15}$$

At each relay r, an error occurs when Y_r is corrupted such that there exist no index in encoder alphabet space $\mathcal{X}_r$. The Markov chains $X_k \leftrightarrow X_r \;\forall\, r \in \mathcal{R}$ and $X_k \leftrightarrow Y_r|X_r \;\forall\, r \in \mathcal{R}$ form a unique jointly typical index at every relays' encoders with the probability function:

$$p\,(x_k, \boldsymbol{y}_{\mathcal{R}}, \boldsymbol{x}_{\mathcal{R}}, y_d) = p\,(x_k)\,.\prod_{r\in\mathcal{R}} p\,(x_r|x_k).\prod_{r\in\mathcal{R}} p\,(y_r|x_r\, x_k) \tag{5.16}$$

The destination tries to re-construct the sent message w_k using backward decoding. The source transmits the sequence

$$X_k[1](0, w_k^1), \ldots, X_k[b](w_k^{b-1}, w_k^b), \ldots, X_k[B](w_k^{B-1}, w_k^B), X_k[B+1](w_k^B, 0)$$

in $B+1$ blocks. The destination collects the sequence $Y_d[1], \ldots, Y_d[B+1]$ as:

$$Y_d[b] = \sqrt{h_{kd}}\, X_k[b](w_k^{b-1}, w_k^b) + \sum_{r\in\mathcal{R}} \sqrt{h_{rd}}\, X_r[b](\hat{w}_k^{b-1}) + Z_d \tag{5.17}$$

The destination's decoder starts from the last block index and proceeds backward to the first block. Suppose it has properly decoded w_k^{b+1}. It can then decode w_k^b from $Y_d[b-1]$ if the multiple-access data rate guarantees:

$$R_k \leq \mathrm{I}\,(X_k\, \boldsymbol{X}_{\mathcal{R}}\,;\,Y_d) \tag{5.18}$$

The error probability at the destination's decoder is prevented if all inputs, i.e. the outputs of all relays and the source node, are error-free received. According to the Markov chain $(X_k, \boldsymbol{X}_{\mathcal{R}}) \leftrightarrow Y_d$, and taking into account that $p\,(y_d|\boldsymbol{x}_{\mathcal{R}}) = p\,(y_d|x_r)$, the probability of an index error at destination's decoder vanishes if:

$$p\,(x_k, \boldsymbol{y}_{\mathcal{R}}, \boldsymbol{x}_{\mathcal{R}}, y_d) = p\,(x_k, \boldsymbol{x}_{\mathcal{R}}) \cdot p\,(y_d|x_r\, x_k) \tag{5.19}$$

All above discussion summarizes in the following theorem:

Theorem 16 *The capacity of DF technique for a network consisted by one source, multiple parallel relays and one destination is upper bounded to:*

$$\mathcal{C}_{DF}(k;\mathcal{R};d) = \sup_{p(x_k,\boldsymbol{x}_{\mathcal{R}})} \min \left\{ \min_{r\in\mathcal{R}} \left\{ \mathrm{I}(X_k\,;\,Y_r|X_r) \right\}\,,\, \mathrm{I}(X_k\,\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d) \right\}$$

where the supremum is over individual power constraints of the source node and all relays, with joint pmf:

$$p\,(x_k, \boldsymbol{x}_{\mathcal{R}}) \cdot \prod_{r\in\mathcal{R}} p\,(x_r|x_k) \cdot \prod_{r\in\mathcal{R}} p\,(y_r|x_r\, x_k) \cdot p\,(y_d|x_r\, x_k)\,.$$

The result is that in a multiple parallel relay network the transmission rate of DF strategy is limited by the worst channel between the source node and the relays. In the following, we introduce capacity of Gaussian channels in a cooperative framework based on DF technique.

Theorem 17 *The AWGN DF upper bound of one source, multiple parallel relays, one destination achieves the rate:*

$$\mathcal{C}_{DF}(k;\mathcal{R};d) = \sup_{-1\le\rho\le1} \min \left\{ \min_{r\in\mathcal{R}} \left\{ C\left(\frac{\mathbb{V}\mathrm{ar}\,(Y_r|X_r) - \sigma_w^2}{\sigma_w^2} \right) \right\}\,, \right.$$
$$\left. C\left(\frac{\mathbb{V}\mathrm{ar}\,(Y_d) - \sigma_w^2}{\sigma_w^2} \right) \right\}$$

which is calculated using Eqs. 5.7.

Proof

$$\mathrm{I}(X_k\ ;\ Y_r|X_r) = \frac{1}{2}\log_2 \frac{\mathbb{V}\mathrm{ar}\,(Y_r|X_r)}{\mathbf{\Sigma}[Y_r|X_r X_k]} = \frac{1}{2}\log_2 \frac{\mathbb{V}\mathrm{ar}\,(Y_r|X_r)}{\sigma_w^2} = $$
$$= C\left(\frac{\mathbb{V}\mathrm{ar}\,(Y_r|X_r) - \sigma_w^2}{\sigma_w^2}\right);$$

$\mathrm{I}(X_k\,\boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d)$ was proved in Theorem 15. ■

Substituting Eq. 5.7c in the first term of Theorem 17 implies that: *in a $(k;\mathcal{R};d)$ reliable network wherein there exists a relay r such that $\gamma_{kr} < \gamma_{kd}$, the direct-link channel capacity outperforms the upper bound capacity of DF strategy.* In such a network, the relays make only "crowd" and reduce the capacity of the direct-link between the transmitter and destination.

5.5 Compress and forward technique

In decode and forward technique, the broadcast capacity is limited. In compress and forward (CF) technique, each relay uses a source coding framework to estimate (or compress or quantize) the received signal with a certain distortion, and then uses channel coding to forward it to the destination. What may happen instead is that the broadcast limitation of DF disappears and multiple-channel access capacity will be restricted to the source coding rate.

In the CF relaying, there are $2R+2$ Gaussian code books: $\mathcal{X}_r$, $\widehat{\mathcal{Y}}_r$ at every relay, $\mathcal{X}_k$ at the source's encoder, and the $\mathcal{Y}_d$ at the destination's decoder. The transmit message w_k is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B$, and each sub-message $w_k^b \in \mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$ is separately encoded. In block index b, the transmitter broadcasts the encoded message $X_k[b](w_k^b)$ toward relays and the destination. Each relay r and destination receive $Y_r[b]$ and $Y_d[b]$ respectively. Neither the destination nor the relays attempt to decode the received information. At the same time, each relay independently estimates its own previous observed signal and then encodes it and transmits toward the destination. In block index b, each relay r using $\widehat{\mathcal{Y}}_r = \{0, 1, ..., 2^{m\widehat{R}_r} - 1\}$ estimates the received signal $Y_r[b-1]$ to $\widehat{Y}_r[b-1]\left(k_r^b\right)$

where $k_r^b \in \{0, 1, ..., 2^{m\tilde{R}_r} - 1\}$ is the estimation index of the relay in the block index b. Then, every relay using the proper codebook $\mathcal{X}_r$ encode the $\widehat{Y}_r[b-1]\left(k_r^b\right)$ and re-converge it to the destination. For the last block, the source transmits a default $X_k(1)$. The following formulas apply:

$$Y_r[b] = \sqrt{h_{kr}}\, X_k[b](w_k^b) + Z_r \tag{5.20a}$$

$$Y_r[b-1] \longrightarrow \widehat{Y}_r[b-1]\left(k_r^b\right) \tag{5.20b}$$

$$Y_d[b] = \sqrt{h_{kd}}\, X_k[b]\left(w_k^b\right) + \sum_{r\in\mathcal{R}} \sqrt{h_{rd}}\, X_r[b]\left(\widehat{Y}_r[b-1]\left(k_r^b\right)\right) + Z_d \tag{5.20c}$$

The network applies a distributed Wyner-Ziv coding with multiple sources and a common decoder at the destination. The input of each individual Wyner-Ziv source coder is its observed signal Y_r. An auxiliary random $\widehat{Y}_r \in \widehat{\mathcal{Y}}_r$ is drawn such that $\widehat{Y}_r$ and Y_r be jointly typical, conventionally $(\widehat{Y}_r, Y_r) \in A_\epsilon^m$, according to $p\left(\hat{y}_r|x_k\right)$. The side information of the Wyner-Ziv network is the already collected signal (in the previous block index) at the destination. By using the fact that the side information values Y_d and Y_r are statistically correlated, we can state the Markov condition $Y_d \leftrightarrow Y_r \leftrightarrow \widehat{Y}_r$. This condition applies other restriction to $\widehat{Y}_r$ at each relay. The estimation variable $\widehat{Y}_r$ must be drawn to satisfy $(\widehat{Y}_r, Y_r) \in A_\epsilon^m$ for a given $(Y_d, Y_r) \in A_\epsilon^m$. Accordingly, the Markov condition guarantees $(Y_d, \widehat{Y}_r, Y_r) \in A_\epsilon^m$. Using induction it is straightforward to show that $(Y_d, \widehat{Y}_1, \ldots, \widehat{Y}_R, Y_1, \ldots, Y_R) \in A_\epsilon^m$. Moreover, for every relay there exists a function $g_r(\cdot)$ such that:

$$\mathbb{E}\left\{d(X_k,\ g_r(Y_d, \widehat{Y}_1, \ldots, \widehat{Y}_R))\right\} \leq D_r \tag{5.21}$$

We do not concern about the amount of distortion D_r, justifying that the decoder is trying to exploit the side information as much as possible.

An error event may be that the encoders do not find any pair of jointly typical codewords. This is prevented if [16, Lemma 10.6.2]:

$$\tilde{R}_r > \mathrm{I}(\widehat{Y}_r\,;\, Y_r|X_r) \qquad \forall\, r \in \mathcal{R} \tag{5.22}$$

This results show that there is a Markov condition of $X_k \leftrightarrow Y_r|X_r \leftrightarrow \widehat{Y}_r$ at each relay r, for encoders and quantization processes. Thus, an error event at all relays' encoders is prevented if:

$$p\left(x_k, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}\right) = p\left(x_k\right) . \prod_{r\in\mathcal{R}} p\left(y_r | x_r x_k\right) . \prod_{r\in\mathcal{R}} p\left(\hat{y}_r | y_r x_r\right) \qquad (5.23)$$

Now, we find a channel probability distribution that prevents from errors at destination's decoder. In each block index b, the decoder uses the receive symbols of $\boldsymbol{X}_{\mathcal{R}}$ to decode the already received symbol X_k in block $b-1$. Decoding is successfully accomplished if all $\boldsymbol{X}_{\mathcal{R}}$ are error free at the destination. Next, $Y_d[b-1]$ given $\boldsymbol{X}_{\mathcal{R}}[b]$ is exploited as the side information to decode $X_k[b-1]$. The error of assigning a unique index at the decoder is prevented if $(X_k\ ,\ Y_d|\boldsymbol{X}_{\mathcal{R}}) = (X_k\ ,\ Y_d|X_r) \in A_\epsilon^m$. These conditions are satisfied with a probability of:

$$p\left(\boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}, y_d\right) = p\left(x_k\right) \cdot p\left(\boldsymbol{x}_{\mathcal{R}}\right) \cdot p\left(y_d | x_k x_r\right) \qquad (5.24)$$

The combination of error probabilities of Eqs. 5.23 and 5.24 gives the pmf of CF technique in a $(k;\mathcal{R};d)$ network as:

$$\begin{aligned} &p\left(\boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}, y_d\right) = \\ &\quad p\left(x_k\right) \cdot p\left(\boldsymbol{x}_{\mathcal{R}}\right) \cdot p\left(y_d | x_k x_r\right) \cdot \prod_{r\in\mathcal{R}} p\left(y_r | x_r x_k\right) \cdot \prod_{r\in\mathcal{R}} p\left(\hat{y}_r | y_r x_r\right) \qquad (5.25) \end{aligned}$$

Another type of error that occurs whenever the decoding is mistaking with another existing index in the alphabet space $\mathcal{Y}_d$. There is a condition under which this error event occurs with arbitrary small probability, as stated by the following theorem.

Theorem 18 *The destination's decoder will not mistake the correct index with any other admissible one if:*

$$\tilde{R}_r - \widehat{R}_r \ \leq\ \mathrm{I}\left(\widehat{Y}_r\ ;\ Y_d | X_r\right) \qquad \forall\, r \in \mathcal{R} \qquad (5.26)$$

Proof The variable Y_d is a function of $\boldsymbol{X}_{\mathcal{R}}$ and X_k. A very corrupted received X_r harms the decoding process. We assume an error at the decoder events due to a very noisy received symbols from the relays belonging to $\mathcal{T} \subseteq \mathcal{R}$, while the sent signals from the relays in $\mathcal{T}^C$, and the side information $Y_d|\boldsymbol{X}_\mathcal{R}$ are correctly approached. To find the error probability at the decoder, we resort to [16, Lemma 10.6.2], and we extend the proof of the joint AEP theorem [16, Th. 7.6.1], especially equations (7.51)-(7.53).

$$\Pr\left\{\left(\widehat{\boldsymbol{Y}}_\mathcal{R}\,,\,Y_d|\boldsymbol{X}_\mathcal{R}\right)\in A_\epsilon^m\right\}\leq$$
$$\prod_{r\in\mathcal{T}}2^{-m\left(\tilde{R}_r-\widehat{R}_r\right)}\cdot\prod_{r\in\mathcal{T}}2^{-m\left(\mathrm{H}\left(\widehat{Y}_r\right)-\epsilon\right)}\cdot 2^{-m\left(\mathrm{H}\left(\,(Y_d|\boldsymbol{X}_\mathcal{R})\,,\,\widehat{\boldsymbol{Y}}_{\mathcal{T}^C}\right)-\epsilon\right)}\cdot$$
$$\cdot\,2^{m\left(\mathrm{H}\left(\,(Y_d|\boldsymbol{X}_\mathcal{R})\,,\,\widehat{\boldsymbol{Y}}_{\mathcal{T}^C}\,,\,\widehat{\boldsymbol{Y}}_\mathcal{T}\,\right)+\epsilon\right)}$$

The above inequality is an extension of [16, Eq. 7.52] (see also Eq. B.18, and Sect. B.1.2). So, as m tends toward infinity, the decoding error probability caused by relays $\mathcal{T}\subseteq\mathcal{R}$ is arbitrarily small if:

$$\sum_{r\in\mathcal{T}}\left(\tilde{R}_r-\widehat{R}_r\right)\leq$$
$$\leq\sum_{r\in\mathcal{T}}\mathrm{H}\left(\widehat{Y}_r\right)+\mathrm{H}\left((Y_d|\boldsymbol{X}_\mathcal{R})\,,\,\widehat{\boldsymbol{Y}}_{\mathcal{T}^C}\right)-\mathrm{H}\left((Y_d|\boldsymbol{X}_\mathcal{R})\,,\,\widehat{\boldsymbol{Y}}_{\mathcal{T}^C}\,,\,\widehat{\boldsymbol{Y}}_\mathcal{T}\right)$$
$$=\sum_{r\in\mathcal{T}}\mathrm{H}\left(\widehat{Y}_r\right)-\mathrm{H}\left(\widehat{\boldsymbol{Y}}_\mathcal{T}\mid(Y_d|\boldsymbol{X}_\mathcal{R}),\widehat{\boldsymbol{Y}}_{\mathcal{T}^C}\right)$$
$$=\sum_{r\in\mathcal{T}}\mathrm{H}\left(\widehat{Y}_r\right)-\sum_{r\in\mathcal{T}}\mathrm{H}\left(\widehat{Y}_r\mid\widehat{Y}_{r-1},\widehat{Y}_{r-2},\ldots,(Y_d|\boldsymbol{X}_\mathcal{R}),\widehat{\boldsymbol{Y}}_{\mathcal{T}^C}\right)$$
$$=\sum_{r\in\mathcal{T}}\mathrm{I}\left(\widehat{Y}_r\,;\,\widehat{Y}_{r-1},\widehat{Y}_{r-2},\ldots,(Y_d|\boldsymbol{X}_\mathcal{R}),\widehat{\boldsymbol{Y}}_{\mathcal{T}^C}\right)$$
$$\stackrel{(2)}{=}\mathrm{I}\left(\widehat{Y}_\mathcal{T}\,;\,Y_d|\boldsymbol{X}_\mathcal{R}\right)$$
$$\stackrel{(3)}{=}\mathrm{I}\left(\widehat{Y}_\mathcal{T}\,;\,Y_d|X_r\right)$$

Equality (2) comes from the Markov chain condition: $Y_d\leftrightarrow Y_r\leftrightarrow\widehat{Y}_r$ at every relays, and equality (3) comes from the fact that $\boldsymbol{X}_\mathcal{R}$ are fully correlated. The proof concludes noting that Eq. 5.26 is satisfied $\forall\,\mathcal{T}\subseteq\mathcal{R}$. ■

The combination of Eqs. 5.22 and 5.26 implies:

$$\widehat{R}_r\geq\mathrm{I}\left(\widehat{Y}_r\,;\,Y_r|X_r\right)-\mathrm{I}\left(\widehat{Y}_r\,;\,Y_d|X_r\right)\qquad\forall\,r\in\mathcal{R}\tag{5.27}$$

This means a $(k, \mathcal{R}, d)$ network applying CF technique forms a distributed Wyner-Ziv source coding network wherein each relay acts as an individual source coder whose input and side information are $Y_r|X_r$ and $Y_d|X_r$ respectively. Following the steps in Theorem 10 we get:

$$\widehat{R}_r \geq \mathrm{I}\left(\widehat{Y}_r\,;\, Y_r|X_r Y_d\right) \qquad \forall\, r \in \mathcal{R} \tag{5.28}$$

Consequently, it is possible to establish an expanded form of Theorem 10 for multiple parallel relays. The following theorem states that the data rate in every $r \to d$ link must be higher than the source coding rate at the proper relay.

Theorem 19 *Suppose that the full-duplex relay complies with CF strategy. For each joint pmf* $p\,(x_k) \cdot p\,(\boldsymbol{x}_{\mathcal{R}}) \cdot p\,(y_d|x_k x_r) \cdot \prod_{r\in\mathcal{R}} p\,(y_r|x_r x_k) \cdot \prod_{r\in\mathcal{R}} p\,(\hat{y}_r|y_r x_r)$ *there exist a sequences of code books* $\left(X_k\,(w_k)\,, X_r\,(z_r)\,, \widehat{Y}_r\,(k_r|z_r)\right)$ *whose symbols are chosen from sets* $w_k \in \{0, 1, ..., 2^{mR_k}-1\}$, $z_r \in \{0, 1, ..., 2^{m\widehat{R}_r}-1\}$, *and* $k_r \in \{0, 1, ..., 2^{m\tilde{R}_r}-1\}$, $\forall\, r \in \mathcal{R}$ *such that* $P_e\,(w_k \neq \hat{w}_k) \to 0$ *as* $m \to +\infty$ *if:*

$$R_k \leq \min_{r\in\mathcal{R}} \mathrm{I}\left(X_k\,;\, Y_d\,\widehat{Y}_r|X_r\right); \tag{5.29a}$$

$$\textit{Subject to:}\quad \mathrm{I}\,(X_r\,;\, Y_d) \geq \mathrm{I}\left(\widehat{Y}_r\,;\, Y_r|\,X_r\,Y_d\right) \qquad \forall\, r \in \mathcal{R} \tag{5.29b}$$

The upper capacity bound of a multiple parallel CF relays is given by the following theorem:

Theorem 20 *For one source, multiple parallel relays, one destination network, the CF capacity is upper bounded by:*

$$\mathcal{C}_{CF}(k;\mathcal{R};d) = \sup_{p(x_k,\,\boldsymbol{x}_{\mathcal{R}})} \min\left\{\min_{r\in\mathcal{R}}\{\mathrm{I}\,(X_k\,;\,Y_d\,Y_r|X_r)\}\,;\, \mathrm{I}\,(\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d)\right\}$$

where the supremum is over individual power constraints of every nodes, and pmf represented in Theorem 19.

Proof After having broadcast a message by the source node, the relays compress their observations in a distributed and individual manner. The destination's input is a combination of the received message on the direct-link, and the relayed versions of the previous sent message. The received signal at the destination can be perfectly decoded if:

$$R_k \leq \mathrm{I}\left(X_k\,;\,Y_d\right) + \mathrm{I}\left(X_k\,;\,Y_r|X_r\right) = \mathrm{I}\left(X_k\,;\,Y_d\,Y_r|X_r\right) \qquad \forall\, r \in \mathcal{R}$$

Thus, the broadcast data rate is not limited to the relays' encoders data rate. Each relay sends the channel coded of its own received signal towards the destination. To do so, the relays converge them onto multiple-access channel after having done estimation and encoding. In the meantime, the destination has already received $X_k[b-1]\left(w_k^{b-1}\right)$, and the capacity is limited by the relays' multiple-access channel, i.e. $R_k \leq \mathrm{I}\left(\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d\right)$. ■

The result is that in a point-to-point multiple parallel relayed network the transmission rate of CF strategy is limited by the multiple-access capacity of the relays at the destination. If all relays are closer to the destination than the source node, the upper bound capacity is achieved by only one relay and others relays can be turned off. The following theorem evaluates the AWGN upper bound capacity.

Theorem 21 *The AWGN CF upper bound of one source, multiple parallel relays, one destination achieves the rate:*

$$\mathcal{C}_{CF}(k;\mathcal{R};d) = \sup_{-1\leq\rho\leq 1} \min\left\{ \min_{r\in\mathcal{R}}\left\{ C\left(\frac{\boldsymbol{\Sigma}[Y_d Y_r|X_r] - \sigma_w^2}{\sigma_w^2}\right)\right\}\,,\right.$$
$$\left. C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right) - \mathbb{V}\mathrm{ar}\left(Y_d\,|\,X_r\right)}{\mathbb{V}\mathrm{ar}\left(Y_d\,|\,X_r\right)}\right)\right\}$$

that is calculated using Eqs. 5.7.

Proof The broadcast term was proved in Theorem 15. For the second term we have:

$$\mathrm{I}\left(\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d\right) = \frac{1}{2}\log_2\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right)}{\mathbb{V}\mathrm{ar}\left(Y_d\,|\,X_r\right)} = C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right) - \mathbb{V}\mathrm{ar}\left(Y_d\,|\,X_r\right)}{\mathbb{V}\mathrm{ar}\left(Y_d\,|\,X_r\right)}\right).$$

■

If all relays are closer to the source node than the destination, or equivalently $\gamma_{kr} > \gamma_{rd}\ \forall r \in \mathcal{R}$, the minimum term is the multiple-access term (the second term), and the supremum is achieved by $\rho = 1$. In such a network, a negative value for correlation ρ significantly reduces the upper bound capacity. In such a network adding a new relay, near to the source node, improves the upper bound capacity as long as the multiple-access capacity has not outperformed the broadcast capacity (the first term). On the other hand, if $\gamma_{kr} < \gamma_{rd}$ $\forall r \in \mathcal{R}$, the minimum term is the broadcast term and the supremum is achieved by $\rho = 0$. In such a network, placing a new relay near to the destination can even deteriorate the upper bound capacity.

5.6 Mixed DF & CF techniques

In this section, we study a multiple parallel relays network wherein the relays in $\mathcal{T} \subset \mathcal{R}$ act as Block Markov DF relays and others relays belonging to $\mathcal{T}^C = \mathcal{R} \setminus \mathcal{T}$ perform CF technique. The network comprises $2R + 2$ code spaces. The code space $\mathcal{X}_k$ at source node's encoder, $\mathcal{X}_r\ \forall r \in \mathcal{R}$ at every relay's encoder, $\mathcal{Y}_r\ \forall r \in \mathcal{T}$ at DF relays' decoders, $\widehat{\mathcal{Y}}_r\ \forall r \in \mathcal{T}^C$ at CF relays' estimators, and $\mathcal{Y}_d$ code space at the destination's decoder. The transmit message w_k is a sequence of B sub-messages of $w_k^1, ..., w_k^b, ..., w_k^B$. Each w_k^b is independently and randomly drawn from the alphabet space $\mathcal{W}_k = \{0, 1, \ldots, 2^{mR_k} - 1\}$, and each sub-message is separately encoded. The encoder at the transmitter has one block memory that applying multiplex coding assigns a Gaussian random number $X_k[b](w_k^{b-1}, w_k^b)$ to each w_k^b. The transmission is done in $B + 1$ blocks and the source transmits the sequence:

$$X_k[1](0, w_k^1), \ldots, X_k[b](w_k^{b-1}, w_k^b), \ldots, X_k[B](w_k^{B-1}, w_k^B), X_k[B+1](w_k^B, 0)$$

In each block b, the source node broadcasts $X_k[b](w_k^{b-1}, w_k^b)$ to all relays and destination. All relays $r \in \mathcal{R}$ receive the signal:

$$Y_r[b] = \sqrt{h_{kr}}\, X_k[b](w_k^{b-1}, w_k^b) + Z_r \qquad (5.30)$$

At the same time, every relay independently performs its processing function (DF or CF) on the previous received signal. The CF relays estimate their own previous received signal and then encodes it to $X_r[b]\left(\widehat{Y}_r[b-1]\left(k_r^b\right)\right)$. The DF relays separately decode and fully re-encode the previous observed message to $X_r[b]\left(w_k^{b-1}\right)$. Then, all relays simultaneously converge the result signal in multiple-access mode at the destination. The destination collects the sequence $Y_d[1], \ldots, Y_d[B+1]$ as:

$$Y_d[b] = \sqrt{h_{kd}}\, X_k[b](w_k^{b-1}, w_k^b) + \sum_{r\in\mathcal{T}} \sqrt{h_{rd}}\, X_r[b](\hat{w}_k^{b-1}) + \\ + \sum_{r\in\mathcal{T}^C} \sqrt{h_{rd}}\, X_r[b]\left(\widehat{Y}_r[b-1]\left(k_r^b\right)\right) + Z_d \tag{5.31}$$

The decoding processes at the DF relays are done with arbitrary small error probability if:

$$R_k \leq \mathrm{I}\left(X_k\,;\,Y_r|X_r\right) \qquad r\in\mathcal{T} \tag{5.32}$$

For the CF relays, R_k is not limited to the relays' encoders process data rate; i.e.

$$R_k \leq \mathrm{I}\left(X_k\,;\,Y_d\,Y_r|X_r\right) \qquad r\in\mathcal{T}^C \tag{5.33}$$

Therefore, the broadcast message data rate in the decode-compress and forward technique is limited to:

$$R_k \leq \min\left\{\min_{r\in\mathcal{T}}\left\{\mathrm{I}\left(X_k\,;\,Y_r|X_r\right)\right\},\ \min_{r\in\mathcal{T}^C}\left\{\mathrm{I}\left(X_k\,;\,Y_d\,Y_r|X_r\right)\right\}\right\} \tag{5.34}$$

If there exists a DF relay $r\in\mathcal{T}$ such that $\gamma_{kr} < \gamma_{kd}$, the result of Eq. 5.34 is less than the direct-link channel capacity. In such a network the existence of all relays is harmful.

At the destination, a multiplex decoding process starts after having collected $B+1$ samples of $Y_d[b]$. The decoding starts from the last block and proceeds backward to the first block. For each $Y_d[b]$, the destination uses the alphabet space $\mathcal{Y}_d$ for multiplex them and calculate $\hat{w}_k^{b-1}$. If the destination's decoder has properly decoded w_k^{b-1}, then the decoding of $Y_d[b]$ is correctly done.

According to Theorem 16, the pmf of the channels of DF relays is:

$$p\left(x_k, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}, y_d\right) = \\ p\left(x_k, \boldsymbol{x}_{\mathcal{R}}\right) \cdot \prod_{r\in\mathcal{T}} p\left(x_r|x_k\right) \cdot \prod_{r\in\mathcal{T}} p\left(y_r|x_r\, x_k\right) \cdot p\left(y_d|x_r\, x_k\right) \quad (5.35)$$

Equality (5.25) allows to evaluate the joint pmf of the CF channels as:

$$p\left(x_k, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}, y_d\right) = \\ p\left(x_k\right) \cdot p\left(\boldsymbol{x}_{\mathcal{R}}\right) \cdot p\left(y_d|x_k x_r\right) \cdot \prod_{r\in\mathcal{T}^C} p\left(y_r|x_r x_k\right) \cdot \prod_{r\in\mathcal{T}^C} p\left(\hat{y}_r|y_r x_r\right) \quad (5.36)$$

The combination of Eqs. 5.35 and 5.36 yields the pmf of the decode-compress and forward technique in a network consisted by one source, multiple parallel relays and one destination as:

$$p\left(x_k, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}, y_d\right) = \\ p\left(x_k, \boldsymbol{x}_{\mathcal{R}}\right) \cdot p\left(y_d|x_k x_r\right) \cdot \prod_{r\in\mathcal{T}} p\left(x_r|x_k\right) \cdot \prod_{r\in\mathcal{T}^C} p\left(\hat{y}_r|y_r x_r\right) \cdot \prod_{r\in\mathcal{R}} p\left(y_r|x_r\, x_k\right) \quad (5.37)$$

According to Theorem 16 the MAC channel capacity of the relays belong to $\mathcal{T}$ is upper bounded by:

$$R_k \leq \mathrm{I}(X_k\, \boldsymbol{X}_{\mathcal{T}}\ ;\ Y_d) = \mathrm{I}(X_k\, \boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d) \quad (5.38)$$

Theorem 20 shows that the MAC channel capacity of the CF relays is upper bounded to:

$$R_k \leq \mathrm{I}(\boldsymbol{X}_{\mathcal{T}^C}\ ;\ Y_d) = \mathrm{I}(\boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d) \quad (5.39)$$

Therefore, the MAC channel message data rate in the decode-compress and forward technique is limited by:

$$R_k \leq \min\{\mathrm{I}(X_k\, \boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d)\ ,\ \mathrm{I}(\boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d)\} \quad (5.40)$$

If $\rho = 1$ then the two terms are equal, otherwise the minimum term is the second term. In a $(k;\mathcal{R};d)$ reliable network wherein all relays are close to the transmitter with $\rho \neq 1$, the decode-compress and forward (DCF) upper bound capacity is equal to that the CF strategy, i.e. that is limited to the multiple access capacity of (only) all relays at the destination.

With the above discussion, it is straightforward to derive the upper bound capacity of a DCF reliable communication as in the following theorem.

Theorem 22 *For one source, multiple parallel relays, one destination network, wherein the relays in subset $\mathcal{T} \subset \mathcal{R}$ perform DF technique and the others relays perform CF technique, the upper bound capacity is:*

$$\mathcal{C}_{DCF}(k;\mathcal{R};d) = \min\left\{\mathcal{C}_{DF}(k;\mathcal{T};d)\ ,\ \mathcal{C}_{CF}(k;\mathcal{T}^C;d)\right\}$$

where supremum of each term is over individual power constraints of source node and each relay node and the joint pmf presented by Eq. 5.37.

This means, the $\mathcal{C}_{DCF}(k;\mathcal{R};d)$ fails both $\mathcal{C}_{DF}(k;\mathcal{T};d)$ and $\mathcal{C}_{CF}(k;\mathcal{T}^C;d)$. The capacity achieves its maximum if all DF relays are (very) close to the transmitter and all CF relays are (very) close to the destination. In such a network wherein all relays are close to the transmitter, the upper bound capacity is limited to the multiple access capacity of all relays at the destination. If all relays are close to the destination the outer region capacity is bounded by the most weak channel between the transmitter and DF relays.

5.7 Case study

Here we illustrate the various outer region bounds of a point-to-point multiple relayed communication with Gaussian channels wherein the transmitter k, two relays r, and the destination d are located as depicted in Fig. 5.2. We assume a vertical equidistance of d_r between two relays and the $k \to d$ direct-link. The path condition values $h_{kr} = h_{rd} = h_{kd} = 1$ are scaled with respect to Notation 2. First, we experiment a high SNR environment and suppose that the source and destination are located at a distance of $d_{kd} = 1\,\text{m}$, and the relays are located in vertical distances of $d_r = 0.1\,\text{m}$ and they are

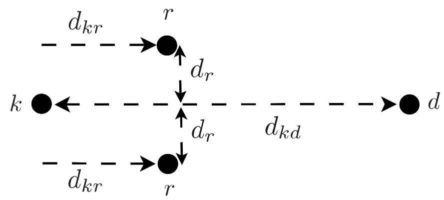

Fig. 5.2: *Point-to-point two relayed communication network scenario.*

simultaneously and horizontally moving from $d_{kr} = -0.5\,\text{m}$ to $d_{kr} = 1.5\,\text{m}$. Fig. 5.3 plots various data rates for $\overline{p}_k = \overline{p}_r = 100\,\text{mW}$, and $\sigma_w^2 = 1\,\mu\text{W}$. The curve labeled AF shows the outer region of AF strategy with the largest possible scaling factor E_r in Eq. 5.9. The curve labeled ρ plots a particular value of the correlation coefficient which is the same as Fig. 4.4. Like the one-relay case, as the relays moves toward the transmitter the DF strategy shows better performance, and the CF technique instead performs better where the relays are closer to the destination. As the relays moves toward the destination the received signal at the relay becomes weaker and this significantly reduces AF data rate. The comparison of Fig. 5.3 to Fig. 4.4 reveals that when the relays are close to the transmitter, the data rates of two-relays network exhibits almost 25% higher performance. On the contrary, when the relays are close to the destination, data rates of cutset, DF, and CF techniques are equal to that of the one-relay network. This is because, in such a situation the maximum data rate of coding techniques is achieved by one relay only.

Now, we consider a point-to-point relayed connection in a low SNR regime. The transmitter and the destination are placed at a distance of $d_{kd} = 500\,\text{m}$, and the relays are simultaneously and horizontally moving in a range of $d_{kr} = -100 \div 600\,\text{m}$ with vertical distance of $d_r = 10\,\text{m}$. Fig. 5.4 plots various data rates for $\overline{p}_k = \overline{p}_r = 100\,\text{mW}$, $\sigma_w^2 = 1\,\mu\text{W}$, and the same E_r as the previous simulation.

We draw a different experimental function for correlation value which is the

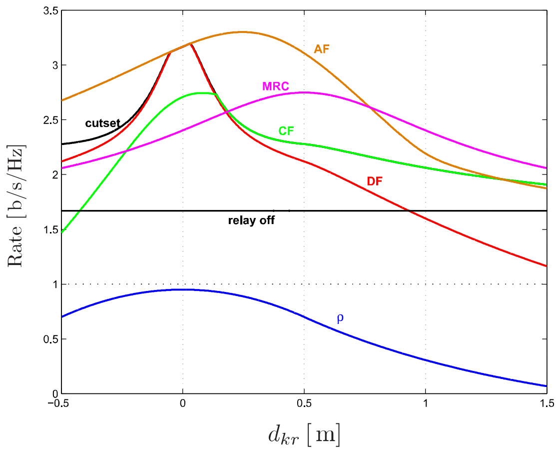

Fig. 5.3: *Rates for two relays with* $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 1\,\mathrm{m}$, *and* $d_r = 0.1\,\mathrm{m}$.

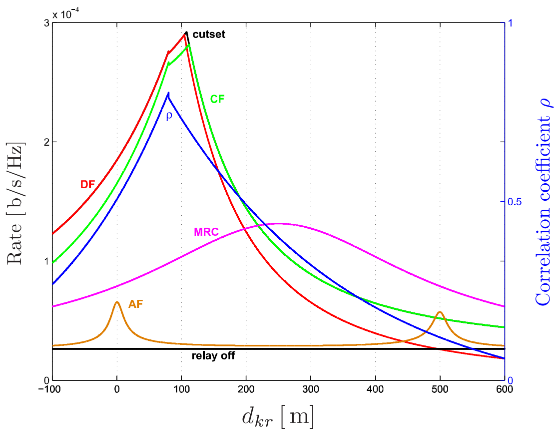

Fig. 5.4: *Rates for one relay with* $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$, $\sigma_w^2 = 1\,\mu\mathrm{W}$, $d_{kd} = 500\,\mathrm{m}$, *and* $d_r = 10\,\mathrm{m}$.

curve labeled ρ, and that is the same as the correlation function in Fig. 4.5. From Fig. 5.4, it is clearly derived that, like in the one-relay network, the AF technique is not quite useful in a low SNR network, whereas MRC technique performs much better than AF. Our experiments in a given scenario with different parameters reveals that reducing the amplification factor does not increase the AF data rate. The comparison between Fig. 5.4 and (4.5) shows that the data rates of coding techniques in two-relays network is almost 50% higher than those of one-relay network. This means that adding a new relay in a low SNR network is much more useful than that in a high SNR network. Like one-relay case in the previous chapter, in a low SNR scene, the DF and CF coding techniques show significant higher rates than the relay off mode.

5.8 Summary

We studied point-to-point communication aided by multiple parallel relays with full-duplex signaling and AWGN channels. We focused on finding the maximum achievable capacity applying three well-known relay strategies: AF, DF, and CF. First, we showed that the maximum data rate is achieved when the output signals of the relays are fully correlated. The first interesting result is that the maximum source to destination flow rate is approached by either only one relay or all relays together.

Like in the single-relay networks, the performance of multiple parallel relays channels basically depends upon both the strategy of the relays and their positions. If all relays are located so as to perfectly receive signals from the source the DF strategy is useful, and the maximum capacity is limited by the weakest source-to-relay channel data rate. In this situation, the others relays can be turned off. Placing a DF relay very far from the source node can even decrease the source to destination direct-link capacity. On the other hand, the CF strategy is useful when all relays are located so as to perfectly deliver signals to the destination, and the maximum capacity is equal to the multiple-access capacity of all relays at the destination. In such a network, adding a new relay may expand the outer region of capacity. We further showed that, in a multiple parallel relays network, applying the same strategy at every relays achieves larger data rate, rather than applying different strategies.

Chapter 6

Multiple sources, parallel relays, one destination

Up to now, we have focused on cooperative systems where a pair of nodes is allowed to communicate with each other, while all of the other nodes act as parallel relays of the source. However, in wireless multiuser systems, multiple sources may be accessing the cooperative channel simultaneously. In this chapter, we study a cooperative network scenario wherein multiple sources, without cooperating together, transmit to one destination via multiple parallel relays. The main focus of this chapter is to derive the upper bound capacity of each transmitter and of the entire network, applying different relay strategies. We will examine whether multiple parallel relays can significantly increase the maximum capacity region of a multiuser network. We will address such questions as: Is increasing the number of relays always helpful? Does adding a new source node always increase the overall flow data rate?

First, in Sect. 6.1 we describe the channel model of a multiple parallel relayed multiple-access communication. The cutset theorem is used in Sect. 6.2 to introduce the largest possible data rate regardless of the relaying strategy. Sect. 6.3 is devoted to AF strategy, Sect. 6.4 to DF, and Sect. 6.5 to CF. In Sect. 6.6 we suppose some relays acting DF strategy and the others relays apply CF technique. Then, we illustrate the results in Sect. 6.7, and finally we conclude in Sect. 6.8.

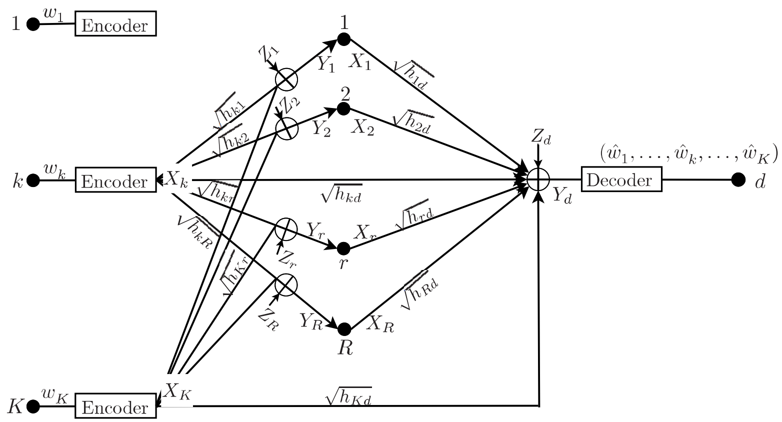

Fig. 6.1: *Multiple sources, multiple parallel relays, one destination communication network scenario.*

6.1 System model

In this chapter, we will study a reliable network that consists of multiple sources belonging to set $\mathcal{K} = [1, \ldots, k, \ldots, K]$ which transmit to the unique destination d via a direct-link, and are assisted by an arbitrary number of relays represented by the set $\mathcal{R} = [1, \ldots, r, \ldots, R]$. We assume there is no connection between relays and no cooperation among source nodes. The source node transmissions are simultaneous. All transmitters are connected to all relays, and the destination node is in sight of all source nodes and relays. The communication is performed in full-duplex mode, so that cooperative communication must be carried on over one phase, and all channels are always busy. All wireless channels are time invariant and frequency flat. The cooperative network of Fig. 6.1 consists of $K+2R+1$ alphabet spaces denoted by $\mathcal{X}_k$ at each transmitter's encoder, $\mathcal{X}_r$ and $\mathcal{Y}_r$ at each relay r, and finally $\mathcal{Y}_d$ at destination's decoder. Every relay is shared by all source nodes and thus all transmitters' outputs interferer at every relay nodes. This scenario is suggestive of the uplink of a cell in a cellular network, within a number of wireless mobile terminals separating the sources, and another (idle) set playing the role of relays.

Each source node k possesses its own individual message set $\mathcal{W}_k$ with size m and rate R_k. Each source node k chooses its message w_k which is a sequence of sub-messages $w_k^1, \ldots, w_k^b, \ldots, w_k^B$ which are going to be reliably conveyed to the unique destination d. All the w_ks are independent and uniformly distributed over their respective alphabet sets $\mathcal{W}_k = \{0, 1, 2, \ldots, 2^{mR_k} - 1\}$. Each source node's encoder is a function $\mathcal{W}_k \rightarrow \mathcal{X}_k$ which maps a Gaussian random codeword $X_k[b]\left(w_k^b\right)$ to each sub-message w_k^b. Each symbol sequence $X_k[1]\left(w_k^1\right), \ldots, X_k[b]\left(w_k^b\right), \ldots, X_k[B]\left(w_k^B\right)$ is bounded to an individual average power of $\mathbb{E}\left\{X_k^2\right\} \leq \overline{p}_k$, with $\mathbb{E}\left\{X_k\right\} = 0$. In each block index b, all transmitters in $\mathcal{K}$ simultaneously broadcast encoded sub-messages $X_k[b]\left(w_k^b\right)$ to all relays and the destination. Since all source nodes interfere at all relays, the receive message at the relay r is:

$$Y_r[b] = \sum_{k \in \mathcal{K}} \sqrt{h_{kr}}\, X_k[b]\left(w_k^b\right) + Z_r \tag{6.1}$$

where $Z_r \sim \mathcal{N}\left(0, \sigma_w^2\right)$ is additive white Gaussian noise at the respective relay. The receive signal $Y_r[b]$ at the relay r is composed of the sent symbols by all transmitters. The received signals at the relays and destination are statistically correlated. Since the data transfer is done in full-duplex mode, at the same time, each relay r performs a relaying function $\mathcal{Y}_r \rightarrow \mathcal{X}_r$. That is, every relay $r \in \mathcal{R}$ processes the proper received signal in the previous block index and generates the information $X_r[b]\left(\boldsymbol{w}_{\mathcal{K}}^{b-1}\right)$ where $\boldsymbol{w}_{\mathcal{K}}^{b-1} = (w_1^{b-1}, \ldots, w_K^{b-1})$, and then converge it into the $r \rightarrow d$ link. The symbol $X_r[b]\left(\boldsymbol{w}_{\mathcal{K}}^{b-1}\right)$ is a regenerated version of $Y_r[b-1]$ and it is the output of the relay r's deterministic function whose inputs is the previous received signals:

$$X_r[b]\left(\boldsymbol{w}_{\mathcal{K}}^{b-1}\right) = f_r^b(Y_r[b-1], Y_r[b-2], \ldots, Y_r[1]); \tag{6.2}$$

The function f depends on the specific cooperative strategy. Each symbol sequence $X_r[1]\left(1\right), \ldots, X_r[b]\left(\boldsymbol{w}_{\mathcal{K}}^{b-1}\right), \ldots, X_r[B+1]\left(\boldsymbol{w}_{\mathcal{K}}^{B}\right)$ is generated under a limitation on average power $\mathbb{E}\left\{X_r^2\right\} \leq \overline{p}_r$, supposing that $\mathbb{E}\left\{X_r\right\} = 0$.

In each block index, the destination receives a combination of the signals $X_k[b]\left(w_k^b\right)\ \forall\, k \in \mathcal{K}$ and $X_r[b]\left(\boldsymbol{w}_{\mathcal{K}}^{b-1}\right)\ \forall\, r \in \mathcal{R}$ over multiple-access channel:

$$Y_d[b] = \sum_{k\in\mathcal{K}} \sqrt{h_{kd}}\, X_k[b]\left(w_k^b\right) + \sum_{r\in\mathcal{R}} \sqrt{h_{rd}}\, X_r[b]\left(\boldsymbol{w}_{\mathcal{K}}^{b-1}\right) + Z_d \qquad (6.3)$$

where $Z_d \sim \mathcal{N}\left(0, \sigma_w^2\right)$. In fact, each received $Y_d[b]$ is a linear combination of the sent signals about $w_1^b, \ldots, w_K^b$ over the direct channels, and R different relayed signals about $\boldsymbol{w}_{\mathcal{K}}^{b-1}$.

The destination's decoder tries to precisely reconstruct the whole B sequences of $\boldsymbol{w}_{\mathcal{K}} = (w_1, \ldots, w_K)$. The decoding process is a (Gaussian) joint de-mapping function $\mathcal{Y}_d \to \mathcal{W}_1 \times \cdots \times \mathcal{W}_K$. The decoder at the destination jointly estimates the messages $\hat{\boldsymbol{w}}_{\mathcal{K}} = (\hat{w}_1, \ldots, \hat{w}_K)$. The decoding function can be executed either in each block index b to estimate $\hat{w}^{b-1}$, or can be done after having collected the whole sequence of $Y_d[1], \ldots, Y_d[B+1]$. The probability of error at destination's decoder is formulated as:

$$P_e^m = \prod_{k\in\mathcal{K}} 2^{-mR_k} \cdot \sum_{\boldsymbol{w}_{\mathcal{K}}=(w_1,\ldots,w_K)} \Pr\left\{\hat{\boldsymbol{w}}_{\mathcal{K}} \neq \boldsymbol{w}_{\mathcal{K}} | \boldsymbol{w}_{\mathcal{K}} \text{ was sent}\right\} \qquad (6.4)$$

which is based on the assumption that the messages are independent, and uniformly distributed over their respective alphabet ranges. The K-tuple $(R_1, \ldots, R_K)$ is achievable if there exists a sequence of code spaces of $\big(m, 2^{mR_1}, \ldots, 2^{mR_K}\big)$ for which P_e^m is arbitrarily close to zero when $m \to \infty$. The destination's decoder reveals all sent messages through a joint decoding function. Thus, each received signal affects decoding process and a very noisy channel can make to perfectly detecting of all messages even impossible.

We have to consider the joint statistics between outputs of different nodes. We define the parameter $\rho_{kr} = \frac{\mathbb{E}\{X_k X_r\}}{\sqrt{\mathbb{E}\{{X_k}^2\}\,\mathbb{E}\{{X_r}^2\}}}$ as the correlation coefficient between the outputs X_k and X_r for $k \in \mathcal{K}$ and $r \in \mathcal{R}$. With a similar formulation, we define τ_{rj} as the correlation between the outputs of two relays $r, j \in \mathcal{R}$, and finally let σ_{km} be the correlation between the outputs of two source nodes $k, m \in \mathcal{K}$. For instance, $\rho_{kr} = 0$ statistically means $p(x_k\,,\,x_r) = p(x_k) \cdot p(x_r)$.

6.2 Cutset upper bound

To find the cutset outer region of a network with multiple sources, multiple parallel relays and one destination, first and foremost, we establish an extended version of general M. R. Aref's seminal formula [161, Th. 3.4].

Theorem 23 *For the general discrete memoryless reliable network with multiple sources, multiple relays and one sink, represented by* $(\mathcal{K}, \mathcal{R}, d)$*, the maximum possible flow rate is upper bounded by:*

$$R_k \leq \sup_{p(\boldsymbol{x}_\mathcal{K}, \boldsymbol{x}_\mathcal{R})} \min_{\mathcal{T} \subseteq \mathcal{R}} \left\{ \mathrm{I}\left(X_k \, \boldsymbol{X}_\mathcal{T} \, ; \, Y_d \, \boldsymbol{Y}_{\mathcal{T}^C} | \boldsymbol{X}_{\mathcal{T}^C} \, \boldsymbol{X}_{\{k\}^C} \right) \right\} \tag{6.5}$$

for each transmitter k*, and for the whole network:*

$$\mathcal{C}_{cutset}(\mathcal{K}; \mathcal{R}; d) = \sup_{p(\boldsymbol{x}_\mathcal{K}, \boldsymbol{x}_\mathcal{R})} \min_{\mathcal{T} \subseteq \mathcal{R}} \left\{ \mathrm{I}\left(\boldsymbol{X}_\mathcal{K} \, \boldsymbol{X}_\mathcal{T} \, ; \, Y_d \, \boldsymbol{Y}_{\mathcal{T}^C} | \boldsymbol{X}_{\mathcal{T}^C} \right) \right\} \tag{6.6}$$

where maximization is subject to the power constraints defined by the network and channel condition $p(y_d \, \boldsymbol{y}_\mathcal{R} \,|\, \boldsymbol{x}_\mathcal{K} \, \boldsymbol{x}_\mathcal{R})$*.*

Proof For $K = 1$, we know [161, Th. 3.4] [183, Prop. 1]:

$$\mathcal{C}_{cutset}(\mathcal{K} = \{k\}; \mathcal{R}; d) = \sup_{p(x_k, \boldsymbol{x}_\mathcal{R})} \min_{\mathcal{T} \subseteq \mathcal{R}} \left\{ \mathrm{I}\left(X_k \, \boldsymbol{X}_\mathcal{T} \, ; \, Y_d \, \boldsymbol{Y}_{\mathcal{T}^C} | \boldsymbol{X}_{\mathcal{T}^C} \right) \right\}$$

For $K > 1$, the combination of the "General multi terminal networks" [16, Sec. 15.10] and "Multiple-access channel" [16, Th. 15.3.6] theorems tells us that the maximum capacity of each transmitter in a multiple sources, multiple relays and one destination is the closure of the achievable rate that is characterized by:

$$R_k \leq \sup_{p(\boldsymbol{x}_\mathcal{K}, \boldsymbol{x}_\mathcal{R})} \min_{\mathcal{T} \subseteq \mathcal{R}} \left\{ \mathrm{I}\left(X_k \, \boldsymbol{X}_\mathcal{T} \, ; \, Y_d \, \boldsymbol{Y}_{\mathcal{T}^C} | \boldsymbol{X}_{\mathcal{T}^C} \, \boldsymbol{X}_{\{k\}^C} \right) \right\} \quad \forall \, k \in \mathcal{K}$$

Consequently, the outer region of the achievable rate by all sources turns out to be (see Sect. B.1.3):

$$\mathcal{C}_{cutset}(\mathcal{K}; \mathcal{R}; d) = \sum_{k \in \mathcal{K}} R_k \leq \sup_{p(\boldsymbol{x}_\mathcal{K}, \boldsymbol{x}_\mathcal{R})} \min_{\mathcal{T} \subseteq \mathcal{R}} \left\{ \mathrm{I}\left(\boldsymbol{X}_\mathcal{K} \, \boldsymbol{X}_\mathcal{T} \, ; \, Y_d \, \boldsymbol{Y}_{\mathcal{T}^C} | \boldsymbol{X}_{\mathcal{T}^C} \right) \right\}.$$

■

The goal of Eq. 6.6 is finding a subset $\mathcal{T} \subseteq \mathcal{R}$ such that the average mutual information between the transmitters $\{\mathcal{K}\,,\,\mathcal{T}\}$ and the receivers $\left\{d\,,\,\mathcal{T}^C\right\}$ is the minimum. Then, mutual information is maximized under a given pmf. The general multi terminal networks theorem (Sect. B.1.3) is established for a general network with arbitrary connections between intermediate nodes. Our framework is a network with parallel relays. In the next theorem we show that the supremum of $\mathcal{C}_{cutset}(\mathcal{K};\mathcal{R};d)$ is achieved when there is no cooperation among transmitters themselves, and outputs of the relays are fully correlated.

Theorem 24 *In the reliable network of Fig. 6.1, $\mathcal{C}_{cutset}(\mathcal{K};\mathcal{R};d)$ achieves the maximum capacity when the Gaussian output variables $\boldsymbol{X}_{\mathcal{K}}$ are statistically independent, and $\boldsymbol{X}_{\mathcal{R}}$ are fully correlated.*

Proof The expansion of Eq. 6.6 yields:

$$\mathrm{I}\left(\boldsymbol{X}_{\mathcal{K}}\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}\right) = \mathrm{H}\left(\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}\right) - \mathrm{H}\left(\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}\boldsymbol{X}_{\mathcal{K}}\right)$$

The second term $\mathrm{H}\left(\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}\boldsymbol{X}_{\mathcal{K}}\right)$ is a function neither τ_{rj} nor σ_{km}. To conclude the proof, it is enough to show that the entropy $\mathrm{H}\left(\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}\right)$ is maximized when $\tau_{rj} = 1\,\forall\,r,j \in \mathcal{R}$ and $\sigma_{km} = 0\,\forall\,k,m \in \mathcal{K}$, regardless of the value of ρ_{kr}. This is true because $\mathrm{H}\left(A|B\right) \geq \mathrm{H}\left(A,C|B,D\right)$ with equality iff A and C are statistically independent and also B and D are fully correlated. ∎

Since $\tau_{rj} = 1\;\;\forall\,r,j \in \mathcal{R}$ and $\sigma_{km} = 0\;\;\forall\,k,m \in \mathcal{K}$, it is reasonable if we assume the same correlation coefficient between X_k and all $\boldsymbol{X}_{\mathcal{R}}$. So, in the rest of this work, for each $k \in \mathcal{K}$, we suppose $\rho_k = \rho_{kr}\;\;\forall\,r \in \mathcal{R}$. A full uncorrelation between variables $\boldsymbol{X}_{\mathcal{K}}$ statistically results: $p(\boldsymbol{x}_{\mathcal{K}}) = \prod_{k\in\mathcal{K}} p(x_k)$. A full correlation between variables $\boldsymbol{X}_{\mathcal{R}}$ results: $\boldsymbol{\Sigma}[Y|\boldsymbol{X}_{\mathcal{R}}] = \mathbb{V}\mathrm{ar}(Y|X_r)\,\forall\,r \in \mathcal{R}$.

The results of Theorem 24 allow to simplify the cutset upper bound represented by Theorem 23:

Theorem 25 *In a $(\mathcal{K};\mathcal{R};d)$ network, the cutset upper bound of each source node k is:*

$$R_k \leq \sup_{p(\boldsymbol{x}_{\mathcal{K}},\boldsymbol{x}_{\mathcal{R}})} \min\left\{\min_{r\in\mathcal{R}}\left\{\mathrm{I}\left(X_k\,;\,Y_d\,Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}\right)\right\}\,,\,\mathrm{I}\left(X_k\boldsymbol{X}_{\mathcal{R}};\,Y_d|\boldsymbol{X}_{\{k\}^C}\right)\right\}$$

Proof As a result of Theorem 24, the second term is equal to the cutset bound with $\mathcal{T}=\mathcal{R}$. Taking into account the correlation coefficients $\tau_{rj}=1\,\forall\, r,j\in\mathcal{R}$ and $\sigma_{km}=0\,\forall\, k,m\in\mathcal{K}$, Eq. 6.5 with every $\mathcal{T}\subsetneq\mathcal{R}$ simplified to:

$$\begin{aligned}\mathrm{I}\left(X_k\,\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}\,\boldsymbol{X}_{\{k\}^C}\right) &=\\ &=\underbrace{\mathrm{I}\left(\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}\,\boldsymbol{X}_{\{k\}^C}\right)}_{(1)\,=\,0}+\underbrace{\mathrm{I}\left(X_k\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}\,\boldsymbol{X}_{\{k\}^C}\right)}_{(2)}\end{aligned}$$

$$\begin{aligned}(1):\mathrm{I}\left(\boldsymbol{X}_{\mathcal{T}}\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}\,\boldsymbol{X}_{\{k\}^C}\right) &= \frac{1}{2}\log_2\frac{\boldsymbol{\Sigma}\left[Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}\,\boldsymbol{X}_{\{k\}^C}\right]}{\boldsymbol{\Sigma}\left[Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}\,\boldsymbol{X}_{\{k\}^C}\right]}=\\ &=\frac{1}{2}\log_2\frac{\boldsymbol{\Sigma}\left[Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|X_{r\in\mathcal{T}^C}\,\boldsymbol{X}_{\{k\}^C}\right]}{\boldsymbol{\Sigma}\left[Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|X_{r\in\mathcal{T}^C}\,\boldsymbol{X}_{\{k\}^C}\right]}=0.\\ (2):\mathrm{I}\left(X_k\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{R}}\,\boldsymbol{X}_{\{k\}^C}\right) &= \mathrm{I}\left(X_k\,;\,Y_d\,\boldsymbol{Y}_{\mathcal{T}^C}|X_r\,\boldsymbol{X}_{\{k\}^C}\right).\end{aligned}$$

Until now we demonstrated that:

$$R_k \leq \sup_{p(\boldsymbol{x}_{\mathcal{K}},\boldsymbol{x}_{\mathcal{R}})} \min\left\{\min_{\mathcal{T}\subsetneq\mathcal{R}}\left\{\mathrm{I}\left(X_k;Y_d\boldsymbol{Y}_{\mathcal{T}^C}|X_r\boldsymbol{X}_{\{k\}^C}\right)\right\},\mathrm{I}\left(X_k\boldsymbol{X}_{\mathcal{R}};Y_d|\boldsymbol{X}_{\{k\}^C}\right)\right\}$$

Since $\mathcal{T}$ is a strict subset of $\mathcal{R}$, it is obvious that the minimum of R_k is achieved with:

$$R_k \leq \sup_{p(\boldsymbol{x}_{\mathcal{K}},\boldsymbol{x}_{\mathcal{R}})} \min\left\{\min_{r\in\mathcal{R}}\left\{\mathrm{I}\left(X_k\,;\,Y_dY_r|X_r\boldsymbol{X}_{\{k\}^C}\right)\right\}\,,\,\mathrm{I}\left(X_k\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d|\boldsymbol{X}_{\{k\}^C}\right)\right\}.$$

■

Each data rate R_k is decomposed into $R+1$ mutual information terms. The novel result of Theorem 25 is that the term R_k achieves its maximum either by virtue of only one relay or by virtue of all relays together. In the general network of Fig. 6.1 wherein the nodes are located in a random position, the minimum mutual information term of R_k is found by calculating of $R+1$ mutual information terms: R broadcast terms, and one multiple access term. If all relays are able to deliver symbols without errors to the destination, i.e. they are (very) close to destination, the minimum of the mutual information terms is one of the first terms which is the broadcast capacity of the source node k to the destination d and the relay r with which the broadcast capacity is the minimum, while the other senders $\{k\}^C$ may not be sending any information. Instead, in the very special case when all relays are in such positions so as to perfectly receive symbols, i.e. the relays are (very) close to the source node k, the minimum of the terms is the multiple-access channel capacity of all relays, and the source k at the destination, while others source nodes may be silent. The following theorem introduce a simplified version of formula (6.6) as the cutset upper bound of a whole network $(\mathcal{K};\mathcal{R};d)$:

Theorem 26 *The cutset upper bound of a* $(\mathcal{K};\mathcal{R};d)$ *network, Eq. 6.6, is shortened to:*

$$\mathcal{C}_{cutset}(\mathcal{K};\mathcal{R};d) = \sup_{p(\boldsymbol{x}_{\mathcal{K}},\boldsymbol{x}_{\mathcal{R}})} \min \left\{ \min_{r\in\mathcal{R}} \{\mathrm{I}\,(\boldsymbol{X}_{\mathcal{K}}\,;\,Y_d\,Y_r|X_r)\}\,,\,\mathrm{I}\,(\boldsymbol{X}_{\mathcal{K}}\,\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d) \right\}$$

Proof The proof mirrors that of Theorem 25. ■

The cutset upper bound $\mathcal{C}_{cutset}(\mathcal{K};\mathcal{R};d)$ is achieved either by virtue of one relay only, or by all relays together. Like R_k, the minimum mutual information term of $\mathcal{C}_{cutset}(\mathcal{K};\mathcal{R};d)$ is found by calculating $R+1$ mutual information terms. If all relays are in good positions to perfectly deliver symbols to the destination, i.e. they are (very) close to the destination, the minimum of mutual information terms is equal to the broadcast capacity of all transmitters $\mathcal{K}$ to d and the relay r with which the broadcast capacity is the minimum (the first term). Instead, in a very special case when all relays are in such positions so as to receive sent symbols from all source nodes, i.e. the relays are (very) close to all source nodes, the minimum of $R+1$ terms is equal to the second term which is the multiple-access channel capacity of all relays $\mathcal{R}$ and all source nodes $\mathcal{K}$ from the destination viewpoint.

Here, we give some useful equalities which are derived from the formulas pertinent to Fig. 6.1, and the results of Theorem 24. These equalities will be used to compute the capacity of Gaussian channels. To review the algebraic manipulation of the following equalities, see Appendix A.

$$\mathbb{V}\mathrm{ar}\left(Y_d\right) = \sum_{k\in\mathcal{K}} h_{kd}\overline{p}_k + \sum_{k\in\mathcal{K}}\sum_{r\in\mathcal{R}} 2\rho_k\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r} + \\ + \sum_{r\in\mathcal{R}}\sum_{j\in\mathcal{R}}\sqrt{h_{rd}\overline{p}_r h_{jd}\overline{p}_j} + \sigma_w^2 \tag{6.7a}$$

$$\mathbf{\Sigma}[Y_d|\boldsymbol{X}_\mathcal{R}] = \mathbb{V}\mathrm{ar}\left(Y_d|X_r\right) = \sum_{k\in\mathcal{K}} h_{kd}\overline{p}_k\left(1-\rho_k^2\right) + \sigma_w^2 \tag{6.7b}$$

$$\mathbf{\Sigma}\left[Y_d|\boldsymbol{X}_{\{k\}^C}\right] = h_{kd}\overline{p}_k + \sum_{r\in\mathcal{R}} 2\rho_k\sqrt{h_{kd}\overline{p}_k h_{rd}\overline{p}_r} + \\ + \sum_{r\in\mathcal{R}}\sum_{j\neq r}\sqrt{h_{rd}\overline{p}_r h_{jd}\overline{p}_j} + \left(1 - \sum_{m\in\{k\}^C}\rho_m^2\right)\cdot\sum_{r\in\mathcal{R}} h_{rd}\overline{p}_r + \sigma_w^2 \tag{6.7c}$$

$$\mathbf{\Sigma}[Y_r|X_r\,\boldsymbol{X}_{\mathcal{A}\subseteq\mathcal{K}}] = \sum_{k\in\mathcal{A}^C} h_{kr}\overline{p}_k\left(1-\rho_k^2\right) + \sigma_w^2 \tag{6.7d}$$

$$\mathbf{\Sigma}[Y_d\,Y_r|X_r\,\boldsymbol{X}_{\mathcal{A}\subseteq\mathcal{K}}] = \sum_{k\in\mathcal{A}^C}\left(h_{kr}+h_{kd}\right)\overline{p}_k\left(1-\rho_k^2\right) + \sigma_w^2 \tag{6.7e}$$

From Eq. 6.7d the following equalities are derived:

$$\mathbf{\Sigma}\left[Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}\right] = h_{kr}\overline{p}_k\left(1-\rho_k^2\right) + \sigma_w^2 \tag{6.7f}$$

$$\mathbb{V}\mathrm{ar}\left(Y_r|X_r\right) = \sum_{k\in\mathcal{K}} h_{kr}\overline{p}_k\left(1-\rho_k^2\right) + \sigma_w^2 \tag{6.7g}$$

From Eq. 6.7e the following equalities are derived:

$$\mathbf{\Sigma}\left[Y_d\,Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}\right] = \left(h_{kr}+h_{kd}\right)\overline{p}_k\left(1-\rho_k^2\right) + \sigma_w^2 \tag{6.7h}$$

$$\mathbf{\Sigma}[Y_d\,Y_r|X_r] = \sum_{k\in\mathcal{K}}\left(h_{kr}+h_{kd}\right)\overline{p}_k\left(1-\rho_k^2\right) + \sigma_w^2 \tag{6.7i}$$

The following two theorems introduce the AWGN cutset upper bound of each transmitter, and all transmitters together in a multiple access channel by way of multiple parallel relays.

Theorem 27 *The AWGN cutset upper bound of each transmitter in a multiple sources, multiple parallel relays, one destination network is:*

$$R_k \leq \sup_{-1 \leq \rho_1, \ldots, \rho_K \leq 1} \min \left\{ \min_{r \in \mathcal{R}} \left\{ C\left(\frac{\boldsymbol{\Sigma}\left[Y_d\, Y_r | X_r\, \boldsymbol{X}_{\{k\}^C}\right] - \sigma_w^2}{\sigma_w^2} \right) \right\}, \right.$$
$$\left. C\left(\frac{\boldsymbol{\Sigma}\left[Y_d | \boldsymbol{X}_{\{k\}^C}\right] - \sigma_w^2}{\sigma_w^2} \right) \right\}$$

that is computed using equalities (6.7).

Proof

$$\begin{aligned} \mathrm{I}(X_k\ ;\ Y_d\, Y_r | X_r\, \boldsymbol{X}_{\{k\}^C}) &= \frac{1}{2} \log_2 \frac{\boldsymbol{\Sigma}\left[Y_d\, Y_r | X_r\, \boldsymbol{X}_{\{k\}^C}\right]}{\boldsymbol{\Sigma}[Y_d\, Y_r | X_r\, \boldsymbol{X}_{\mathcal{K}}]} \\ &\overset{(1)}{=} \frac{1}{2} \log_2 \frac{\boldsymbol{\Sigma}\left[Y_d\, Y_r | X_r\, \boldsymbol{X}_{\{k\}^C}\right]}{\sigma_w^2} \\ &= C\left(\frac{\boldsymbol{\Sigma}\left[Y_d\, Y_r | X_r\, \boldsymbol{X}_{\{k\}^C}\right] - \sigma_w^2}{\sigma_w^2} \right) \end{aligned}$$

(1) follows from the fact that $\boldsymbol{\Sigma}[Y_d\, Y_r | X_r\, \boldsymbol{X}_{\mathcal{K}}] = \boldsymbol{\Sigma}[Y_d\, Y_r | \boldsymbol{X}_{\mathcal{R}}\, \boldsymbol{X}_{\mathcal{K}}]$ and multivariate $Y_d\, Y_r | \boldsymbol{X}_{\mathcal{R}}\, \boldsymbol{X}_{\mathcal{K}}$ is a function independent variables Z_r and Z_d.

$$\begin{aligned} \mathrm{I}(X_k\, \boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d | \boldsymbol{X}_{\{k\}^C}) = \frac{1}{2} \log_2 \frac{\boldsymbol{\Sigma}\left[Y_d | \boldsymbol{X}_{\{k\}^C}\right]}{\boldsymbol{\Sigma}[Y_d | \boldsymbol{X}_{\mathcal{K}}\, \boldsymbol{X}_{\mathcal{R}}]} &= \frac{1}{2} \log_2 \frac{\boldsymbol{\Sigma}\left[Y_d | \boldsymbol{X}_{\{k\}^C}\right]}{\sigma_w^2} = \\ &= C\left(\frac{\boldsymbol{\Sigma}\left[Y_d | \boldsymbol{X}_{\{k\}^C}\right] - \sigma_w^2}{\sigma_w^2} \right). \end{aligned}$$

■

The interesting result is: in a $(\mathcal{K}; \mathcal{R}; d)$ network wherein $\gamma_{kr} < \gamma_{rd}\ \forall\, r \in \mathcal{R}$, i.e. the data rate of every $k \to r$ channel is smaller than that at the $r \to d$ channel for all relays, the minimum term of R_k is equivalent to the broadcast term (one of the first terms) which is achieved by the relay r which has the smallest $h_{kr} \overline{p}_k \left(1 - \rho_k^2\right)$. As can be seen, the power constraint of $\overline{p}_r$ does not play any role, and the best correlation value is $\rho_k = 0$. Instead, in a network wherein $\gamma_{kr} > \gamma_{rd}\ \forall\, r \in \mathcal{R}$, the minimum term of R_k is equal to the multiple-access term (the second term), and it is maximized with $\rho_k = 1$, and $\rho_{m\neq k} = 0$. In such a network it should be satisfied: $\sum_{m\neq k} \rho_m^2 < 1$.

Theorem 28 *The overall AWGN cutset upper bound capacity of a network consisting of multiple sources, multiple parallel relays, and one destination is bounded by:*

$$\mathcal{C}_{cutset}(\mathcal{K};\mathcal{R};d) = \sup_{-1\leq\rho_1,\ldots,\rho_K\leq 1} \min\left\{\min_{r\in\mathcal{R}}\left\{C\left(\frac{\mathbf{\Sigma}[Y_d\, Y_r|X_r] - \sigma_w^2}{\sigma_w^2}\right)\right\}, \right. \left. C\left(\frac{\mathbb{V}\mathrm{ar}\,(Y_d) - \sigma_w^2}{\sigma_w^2}\right)\right\}$$

that is computed applying the equalities (6.7)*.*

Proof

$$\mathrm{I}(\boldsymbol{X}_{\mathcal{K}}\ ;\ Y_d\, Y_r|X_r) = \frac{1}{2}\log_2 \frac{\mathbf{\Sigma}[Y_d\, Y_r|X_r]}{\mathbf{\Sigma}[Y_d\, Y_r|X_r \boldsymbol{X}_{\mathcal{K}}]} \overset{(1)}{=} \frac{1}{2}\log_2 \frac{\mathbf{\Sigma}[Y_d\, Y_r|X_r]}{\sigma_w^2} = \\ = C\left(\frac{\mathbf{\Sigma}[Y_d\, Y_r|X_r] - \sigma_w^2}{\sigma_w^2}\right)$$

(1) follows from the fact that $\mathbf{\Sigma}[Y_d\, Y_r|X_r \boldsymbol{X}_{\mathcal{K}}] = \mathbf{\Sigma}[Y_d\, Y_r|\boldsymbol{X}_{\mathcal{R}} \boldsymbol{X}_{\mathcal{K}}]$ and multivariate $Y_d\, Y_r|\boldsymbol{X}_{\mathcal{R}} \boldsymbol{X}_{\mathcal{K}}$ is a function independent variables Z_r and Z_d.

$$\mathrm{I}(\boldsymbol{X}_{\mathcal{K}}\, \boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d) = \frac{1}{2}\log_2 \frac{\mathbb{V}\mathrm{ar}\,(Y_d)}{\mathbf{\Sigma}[Y_d|\boldsymbol{X}_{\mathcal{K}}\, \boldsymbol{X}_{\mathcal{R}}]} = \frac{1}{2}\log_2 \frac{\mathbb{V}\mathrm{ar}\,(Y_d)}{\sigma_w^2} = \\ = C\left(\frac{\mathbb{V}\mathrm{ar}\,(Y_d) - \sigma_w^2}{\sigma_w^2}\right).$$

■

In a $(\mathcal{K};\mathcal{R};d)$ network wherein $\gamma_{kr} < \gamma_{rd}\ \forall\, k \in \mathcal{K}, \forall\, r \in \mathcal{R}$, the upper bound capacity $\mathcal{C}_{cutset}(\mathcal{K};\mathcal{R};d)$ is achieved by the relay $r \in \mathcal{R}$ with the minimum $\sum_{k\in\mathcal{K}} h_{kr}\overline{p}_k\left(1-\rho_k^2\right)$. The $\overline{p}_r$ does not play any role. In such a reliable network, the activity of all others relays is useless and the AWGN upper bound of the whole network is supremum with $\rho_k = 0$ for all $k \in \mathcal{K}$. Instead, in a $(\mathcal{K};\mathcal{R};d)$ network, wherein $\gamma_{kr} > \gamma_{rd}\ \forall\, k \in \mathcal{K},\ \forall\, r \in \mathcal{R}$, the AWGN upper bound capacity is the multiple-access channel capacity of all relays and source nodes at the destination, and the best correlation value is $\rho_k = 1\ \forall\, k \in \mathcal{K}$.

6.3 Amplify and forward technique

In this approach, there is no regeneration code at relay nodes. There are $K+1$ Gaussian code spacfes: $\mathcal{X}_k$ at each source node and $\mathcal{Y}_d$ at the destination's decoder. Each source k wishes to transmit the message w_k which is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B$. Each sub-message $w_k^b \in \mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$ is separately and fully encoded by the respective transmitter. In each block index b, each relay r scales the amplitude of the analog observed signal $Y_r[b-1]$ as:

$$\begin{aligned} X_r[b]\left(\boldsymbol{w}_{\mathcal{K}}^{b-1}\right) &= E_r[b] \cdot Y_r[b-1] \\ &= E_r[b] \cdot \left(\sum_{k \in \mathcal{K}} \sqrt{h_{kr}}\, X_k[b-1]\left(w_k^{b-1}\right) + Z_r \right) \end{aligned} \tag{6.8}$$

wherein E_r is chosen to satisfy the proper relay's power constraint. Each relay node has its own power constraint as $\mathbb{E}\left\{X_r^2\right\} \leq \overline{p}_r$. Hence, in AWGN mode, it must be fulfilled:

$$|E_r[b]|^2 \leq \frac{\overline{p}_r}{\sigma_w^2 + \sum_{k \in \mathcal{K}} h_{kr} \overline{p}_k} \tag{6.9}$$

So, we can suppose $E_r[b] = E_r$. As can be seen, if $\sigma_w^2 + \sum_{k \in \mathcal{K}} h_{kr}\overline{p}_k \gg \overline{p}_r$, the AF technique performance is equal to the relay r off mode. Combining Eqs. 6.3 and (6.8) yields the observed signal at the destination as:

$$\begin{aligned} Y_d[b] = \sum_{k \in \mathcal{K}} \sqrt{h_{kd}}\, X_k[b]\left(w_k^b\right) + \sum_{k \in \mathcal{K}} \sum_{r \in \mathcal{R}} |E_r| \sqrt{h_{kr}\, h_{rd}}\, X_k[b-1]\left(w_k^{b-1}\right) + \\ + \sum_{r \in \mathcal{R}} |E_r| \sqrt{h_{rd}}\, Z_r + Z_d \end{aligned} \tag{6.10}$$

The relay nodes do not regenerate any new code, and consequently the complexity of this scheme is low. Since every relay node amplifies whatever it receives, including noise and interference, and noting that all source nodes interferer at every relays, AF technique is mainly useful in sufficiently high SNR environments. Increasing the amplification factor E_r increases ISI noise at the destination, and also significantly increase the noise of ICI where there is only one antenna at the destination. If the network is able to adjust the power of the relay, the relay should thus transmit with a properly fine-tuned power. The challenge of adapting the best value of E_r makes AF technique almost useless when there are a large number of source nodes and relays.

Each wireless terminal k can look at $Y_d[b]$ as:

$$\begin{aligned} Y_d[b] = & \underbrace{\sqrt{h_{kd}}\, X_k + \sum_{r \in \mathcal{R}} |E_r| \sqrt{h_{kr}\, h_{rd}}\, X_k +}_{(1)} \\ & + \underbrace{\sum_{k \neq m \in \mathcal{K}} \sqrt{h_{md}}\, X_m + \sum_{r \in \mathcal{R}} \sum_{k \neq m \in \mathcal{K}} |E_r| \sqrt{h_{mr}\, h_{rd}}\, X_m +}_{(2)} \\ & + \underbrace{\sum_{r \in \mathcal{R}} |E_r| \sqrt{h_{rd}}\, Z_r + Z_d}_{(3)} \end{aligned} \tag{6.11}$$

where (1) can be described as signal term, (2) as that multiple access interference (MAI), and (3) as Gaussian thermal noise. Consequently, in AWGN mode, the flow rate of every transmitter k is calculated as:

$$R_k = C\left(\frac{\left(\sqrt{h_{kd}} + \sum_{r \in \mathcal{R}} |E_r| \sqrt{h_{kr} h_{rd}} \right)^2 \overline{p}_k}{\sum_{m \neq k} \left(\sqrt{h_{md}} + \sum_{r \in \mathcal{R}} |E_r| \sqrt{h_{mr} h_{rd}} \right)^2 \overline{p}_m + \left(1 + \sum_{r \in \mathcal{R}} |E_r|^2 h_{rd} \right) \cdot \sigma_w^2} \right) \tag{6.12}$$

By Eq. 6.10, the overall capacity of AF scheme is formulated as:

$$\mathcal{C}_{AF}(\mathcal{K};\mathcal{R};d) =$$
$$C\left(\left(\sum_{k\in\mathcal{K}}\sqrt{h_{kd}\overline{p}_k}+\sum_{k\in\mathcal{K}}\sum_{r\in\mathcal{R}}|E_r|\sqrt{h_{kr}h_{rd}\overline{p}_k}\right)^2\frac{1}{\left(1+\sum_{r\in\mathcal{R}}|E_r|^2h_{rd}\right)\cdot\sigma_w^2}\right) \tag{6.13}$$

According to [184, p. 46], if the condition

$$|E_r| \leq \sum_{k\in\mathcal{K}}\gamma_{kr} \tag{6.14}$$

holds, then the $\mathcal{C}_{AF}(\mathcal{K};\mathcal{R};d)$ outperforms the capacity of the maximal ratio combining (MRC) technique that is:

$$\mathcal{C}_{MRC}(\mathcal{K};\mathcal{R};d) = C\left(\sum_{k\in\mathcal{K}}\gamma_{kd}+\sum_{r\in\mathcal{R}}\frac{\sum_{k\in\mathcal{K}}\gamma_{kr}\cdot\gamma_{rd}}{\sum_{k\in\mathcal{K}}\gamma_{kr}+\gamma_{rd}}\right) \tag{6.15}$$

6.4 Decode and forward technique

Here, like Sect. 5.4, we study a well known regular encoding called block Markov encoding, that is concerned with $K+2R+1$ Gaussian code spaces: $\mathcal{X}_k$ at every transmitter's encoder, respectively $\mathcal{Y}_r$ and $\mathcal{X}_r$ at the proper relay's decoder and encoder, and finally $\mathcal{Y}_d$ at the destination's decoder. Every source node k wishes to transmit a message w_k which is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B$. Each sub-message $w_k^b \in \mathcal{W}_k = \{0, 1, ..., 2^{mR_k}-1\}$ is separately and fully encoded by the proper source node.

The encoder at each transmitter has one block memory that using multiplex coding maps a Gaussian random number $X_k[b](w_k^{b-1}, w_k^b)$ to each w_k^b. In each block b, every source node k emits $X_k[b](w_k^{b-1}, w_k^b)$ to all relays and the destination. Each relay r receives the signal:

$$Y_r[b] = \sum_{k\in\mathcal{K}}\sqrt{h_{kr}}\,X_k[b](w_k^{b-1}, w_k^b) + Z_r \tag{6.16}$$

Since the communication is performed in full-duplex mode, every relay decodes and then re-encode the previous received message, and then transmits it toward the destination. Each relay $r \in \mathcal{R}$ tries to correctly and fully decodes the received signal using $\mathcal{Y}_r$ and then re-encode it using $\mathcal{X}_r$. A DF reliable channel can achieve to the highest possible capacity when decoding processes at all relays are correctly done. To this end, and taking into account that the transmissions from the relays are simultaneously done, every $k \to r$ link data rate is limited to:

$$R_k \leq \mathrm{I}\left(X_k\,;\,Y_r|X_r\boldsymbol{X}_{\{k\}^C}\right) \qquad \forall\, r \in \mathcal{R} \tag{6.17}$$

Each R_k is achieved maximum while others transmitters $\{k\}^C$ are not transmitting. This inequality allows all relays to precisely decode the received signals. At each encoder of the relay r, error occurs when Y_r is such noisy that the decoding process can not be correctly performed and then therefore there exists no index in encoder alphabet space $\mathcal{X}_r$. The Markov chains $(X_1, \ldots, X_K) \leftrightarrow X_r$ and $(X_1, \ldots, X_K) \leftrightarrow Y_r|X_r$ at every relays form a unique jointly typical index at each relay's encoder according to the following pmf:

$$p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{y}_{\mathcal{R}}, \boldsymbol{x}_{\mathcal{R}}, y_d\right) = p\left(\boldsymbol{x}_{\mathcal{K}}\right) \cdot \prod_{r\in\mathcal{R}} p\left(x_r|\boldsymbol{x}_{\mathcal{K}}\right) \cdot \prod_{r\in\mathcal{R}} p\left(y_r|x_r\, \boldsymbol{x}_{\mathcal{K}}\right) \tag{6.18}$$

Each source k transmits the sequence:

$$X_k[1](0, w_k^1), \ldots, X_k[b](w_k^{b-1}, w_k^b), \ldots, X_k[B](w_k^{B-1}, w_k^B), X_k[B+1](w_k^B, 0)$$

in $B+1$ blocks. The DF relaying strategy is useful when $B \to \infty$. The destination collects the sequence of $Y_d[1], \ldots, Y_d[B+1]$ as:

$$Y_d[b] = \sum_{k\in\mathcal{K}} \sqrt{h_{kd}}\, X_k[b](w_k^{b-1}, w_k^b) + \sum_{r\in\mathcal{R}} \sqrt{h_{rd}}\, X_r[b](\hat{\boldsymbol{w}}_{\mathcal{K}}^{b-1}) + Z_d \tag{6.19}$$

At the destination, the decoding process starts from the last block and proceeds backward to the first block. If the destination has properly decoded w_k^{b+1}. It can then decode w_k^b from $Y_d[b-1]$ if the multiple-access data rate achieves to:

$$R_k \leq \sum_{r\in\mathcal{R}} \mathrm{I}\left(X_k\, X_r\,;\,Y_d|\boldsymbol{X}_{\{k\}^C}\right) = \mathrm{I}\left(X_k\, \boldsymbol{X}_{\mathcal{R}}\,;\,Y_d|\boldsymbol{X}_{\{k\}^C}\right) \tag{6.20}$$

The decoding function at the destination considers all received signals from all relays and source nodes. Accordingly, the error probability at the decoder vanishes if all inputs received error free. According to the Markov chain $(\boldsymbol{X}_{\mathcal{K}}, \boldsymbol{X}_{\mathcal{R}}) \leftrightarrow Y_d$, the probability of occurring a confusing index error at the destination's decoder is prevented if:

$$p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{y}_{\mathcal{R}}, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_d\right) = p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}}\right) \cdot p\left(y_d | x_r\, \boldsymbol{x}_{\mathcal{K}}\right) \tag{6.21}$$

Note that $p\left(y_d | \boldsymbol{x}_{\mathcal{R}}\, \boldsymbol{x}_{\mathcal{K}}\right) = p\left(y_d | x_r\, \boldsymbol{x}_{\mathcal{K}}\right)$. All above discussion results:

Theorem 29 *The upper bound capacity of each source k in a $(\mathcal{K}; \mathcal{R}; d)$ reliable network applying DF technique is limited by:*

$$R_k \leq \sup_{p(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}})} \min \left\{ \min_{r \in \mathcal{R}} \left\{ \mathrm{I}(X_k\, ;\, Y_r | X_r\, \boldsymbol{X}_{\{k\}^C}) \right\}, \mathrm{I}(X_k\, \boldsymbol{X}_{\mathcal{R}}\, ;\, Y_d | \boldsymbol{X}_{\{k\}^C}) \right\}$$

where the supremum is over individual power constraints of the nodes, with joint pmf: $p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}}\right) \cdot \prod_{r \in \mathcal{R}} p\left(x_r | \boldsymbol{x}_{\mathcal{K}}\right) \cdot \prod_{r \in \mathcal{R}} p\left(y_r | x_r\, \boldsymbol{x}_{\mathcal{K}}\right) \cdot p\left(y_d | x_r\, \boldsymbol{x}_{\mathcal{K}}\right)$.

The number of relays and their positions significantly influence R_k. The overall DF capacity of a $(\mathcal{K}; \mathcal{R}; d)$ network is upper bounded to:

Theorem 30 *The capacity of DF technique for a whole $(\mathcal{K}; \mathcal{R}; d)$ network is upper bounded by:*

$$\mathcal{C}_{DF}(\mathcal{K}; \mathcal{R}; d) = \sup_{p(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}})} \min \left\{ \min_{r \in \mathcal{R}} \left\{ \mathrm{I}(\boldsymbol{X}_{\mathcal{K}}\, ;\, Y_r | X_r) \right\}, \mathrm{I}(\boldsymbol{X}_{\mathcal{K}}\, \boldsymbol{X}_{\mathcal{R}}\, ;\, Y_d) \right\}$$

where the supremum is over individual power constraints of the nodes, with joint pmf: $p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}}\right) \cdot \prod_{r \in \mathcal{R}} p\left(x_r | \boldsymbol{x}_{\mathcal{K}}\right) \cdot \prod_{r \in \mathcal{R}} p\left(y_r | x_r\, \boldsymbol{x}_{\mathcal{K}}\right) \cdot p\left(y_d | x_r\, \boldsymbol{x}_{\mathcal{K}}\right)$.

Proof Since $\mathcal{C}_{DF}(\mathcal{K}; \mathcal{R}; d) = \sum_{k \in \mathcal{K}} R_k$, and applying R_k from Theorem 29, the proof is somehow obvious. ■

The first term shows that the upper bound capacity is limited by the smallest multiple-access capacity from all source nodes at the relays. Equivalently, if all relays are well located to deliver signals without errors to the destination, the conditions of direct-links between the transmitters and destination do not matter at all. In such a reliable network, the upper bound capacity of $\mathcal{C}_{DF}(\mathcal{K};\mathcal{R};d)$ is equal to the one of the broadcast terms of Theorem 30 which depends only on one relay. The second term (the multiple-access term) of $\mathcal{C}_{DF}(\mathcal{K};\mathcal{R};d)$ is equal to that $\mathcal{C}_{cutset}(\mathcal{K};\mathcal{R};d)$. This means that if all relays are able to perfectly receive and precisely decode the transmitted symbols by all source nodes, the DF capacity approaches the cutset upper bound capacity. In the following, we introduce the capacity of cooperative framework based on DF technique for a network with Gaussian channels.

Theorem 31 *The AWGN DF upper bound of a source node k, in a multiple sources, multiple parallel relays, one destination achieves the rate:*

$$R_k \leq \sup_{-1\leq\rho_1,\ldots,\rho_K\leq 1} \min \left\{ \min_{r\in\mathcal{R}} \left\{ C\left(\frac{\boldsymbol{\Sigma}\big[Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}\big]-\sigma_w^2}{\sigma_w^2}\right)\right\}, \right.$$
$$\left. C\left(\frac{\boldsymbol{\Sigma}\big[Y_d|\boldsymbol{X}_{\{k\}^C}\big]-\sigma_w^2}{\sigma_w^2}\right)\right\}$$

that is computed applying the equalities (6.7).

Proof

$$\begin{aligned} \mathrm{I}(X_k\,;\,Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}) &= \frac{1}{2}\log_2\frac{\boldsymbol{\Sigma}\big[Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}\big]}{\boldsymbol{\Sigma}[Y_r|X_r\,\boldsymbol{X}_{\mathcal{K}}]} \\ &= \frac{1}{2}\log_2\frac{\boldsymbol{\Sigma}\,\big[Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}\big]}{\sigma_w^2} \\ &= C\left(\frac{\boldsymbol{\Sigma}\big[Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}\big]-\sigma_w^2}{\sigma_w^2}\right) \end{aligned}$$

For AWGN of the second term , $\mathrm{I}(X_k\,\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d|\boldsymbol{X}_{\{k\}^C})$, see Theorem 27. ■

The first term implies that, if there exists a relay r such that $h_{kr} < h_{kd}$, then the direct $k \to d$ channel capacity dominates R_k. This means, in such a network it is reasonable to inactivate all relays.

Theorem 32 *The AWGN upper bound of a $(\mathcal{K};\mathcal{R};d)$ reliable network applying DF technique achieves the rate:*

$$\mathcal{C}_{DF}(\mathcal{K};\mathcal{R};d) = \sup_{-1\leq\rho_1,\dots,\rho_K\leq 1} \min\left\{\min_{r\in\mathcal{R}}\left\{C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_r|X_r\right)-\sigma_w^2}{\sigma_w^2}\right)\right\},\right. \left. C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right)-\sigma_w^2}{\sigma_w^2}\right)\right\}$$

that is computed using the equalities (6.7).

Proof

$$\mathrm{I}(\boldsymbol{X}_{\mathcal{K}}\ ;\ Y_r|X_r) = \frac{1}{2}\log_2\frac{\mathbb{V}\mathrm{ar}\left(Y_r|X_r\right)}{\boldsymbol{\Sigma}[Y_r|X_r\,\boldsymbol{X}_{\mathcal{K}}]} = \frac{1}{2}\log_2\frac{\mathbb{V}\mathrm{ar}\left(Y_r|X_r\right)}{\sigma_w^2} = C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_r|X_r\right)-\sigma_w^2}{\sigma_w^2}\right)$$

The AWGN of the multiple-access capacity term, $\mathrm{I}(\boldsymbol{X}_{\mathcal{K}}\,\boldsymbol{X}_{\mathcal{R}}\ ;\ Y_d)$, was proved in Theorem 28. ■

Consequently, in a reliable network wherein all relays are in good positions to error freely deliver signals to the destination, the $\mathcal{C}_{DF}(\mathcal{K};\mathcal{R};d)$ is achieved by the relay $r\in\mathcal{R}$ with which $\sum_{k\in\mathcal{K}} h_{kr}\overline{p}_k\left(1-\rho_k^2\right)$ is the minimum. In such a reliable network, the supremum of $\mathcal{C}_{DF}(\mathcal{K};\mathcal{R};d)$ is achieved by $\rho_k = 0$ $\forall\, k\in\mathcal{K}$. Correlation between X_ks and X_rs decreases the capacity. As can be derived, adding a new transmitter, always near others transmitters and with correlation value $\rho_k \neq \pm 1$, increases the upper bound capacity of the whole network. Adding a new relay (very) far from all transmitters can even degrade the total capacity. On the other hand, if the minimum AWGN capacities is equal to the second term, adding a new transmitter, always well located to perfectly deliver symbols to all relays with $\rho_k > 0$, also expands the outer region of $\mathcal{C}_{DF}(\mathcal{K};\mathcal{R};d)$. In such a reliable network, adding a new relay, always close to all transmitters, also increase the upper bound capacity of the whole network. In such a network, a negative correlation value for ρ_k degrades the overall data rate.

6.5 Compress and forward technique

Applying compress and forward (CF) technique, each relay uses a source coding scheme to compress (or estimate) the received signals with a certain distortion to a joint index, and then use channel coding to transmit it to the destination. Therefore, the multiple-channel access capacity must guarantee required source coding rate. This technique is useful where the decoding process at relays can not be error freely done.

There are $K+2R+1$ finite Gaussian code books: $\mathcal{X}_r$, $\widehat{\mathcal{Y}}_r$ at the proper relay's encoder and estimation space, respectively, $\mathcal{X}_k$ at the respective source node's encoder, and $\mathcal{Y}_d$ at the destination's decoder. The transmit message w_k of the transmitter k is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B$. Each w_k^b is uniformly drawn from $\mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$, and it is separately encoded. In each block b, every source nodes simultaneously emit the proper message $X_k[b](w_k^b)$ toward relays and the destination. Each relay r and the destination receive $Y_r[b]$ and $Y_d[b]$, respectively. Neither the destination nor the relays try to decode the received signals. At the same time, each relay independently estimates its own previous observed signal to an index, and then encode it and transmit toward the destination. In the block index b, each relay r using the alphabet space $\widehat{\mathcal{Y}}_r = \{0, 1, ..., 2^{m\widehat{R}_r} - 1\}$, estimates the received signal $Y_r[b-1]$ to a symbol $\widehat{Y}_r[b-1]\left(k_r^b\right)$ where $k_r^b \in \{0, 1, ..., 2^{m\tilde{R}_r} - 1\}$ is the individual estimation/quantization parameter of the proper relay in the block index b. Then, every relays using the proper codebook $\mathcal{X}_r$ encode the $\widehat{Y}_r[b-1]\left(k_r^b\right)$ to an index and simultaneously transmit them to the destination. For the last block, the sources broadcast a termination codeword $X_k(1)$. The following formulas apply:

$$Y_r[b] = \sum_{k \in \mathcal{K}} \sqrt{h_{kr}}\, X_k[b](w_k^b) + Z_r \tag{6.22a}$$

$$Y_r[b-1] \longrightarrow \widehat{Y}_r[b-1]\left(k_r^b\right) \tag{6.22b}$$

$$Y_d[b] = \sum_{k \in \mathcal{K}} \sqrt{h_{kd}} X_k[b]\left(w_k^b\right) + \sum_{r \in \mathcal{R}} \sqrt{h_{rd}} X_r[b]\left(\widehat{Y}_r[b-1]\left(k_r^b\right)\right) + Z_d \tag{6.22c}$$

A $(\mathcal{K};\mathcal{R};d)$ network applying CF technique, applies distributed Wyner-Ziv coding with multiple sources and a common decoder at the destination. The input of each individual Wyner-Ziv source coder is its observed signal Y_r. An auxiliary random $\widehat{Y}_r \in \widehat{\mathcal{Y}}_r$ is drawn such that $\widehat{Y}_r$ and Y_r to be jointly typical with respect to $p\left(\hat{y}_r|\boldsymbol{x}_{\mathcal{K}}\right)$. The side information of the Wyner-Ziv network is the already received signal (in the previous block index) at the destination. By using the fact that in every block index the side information Y_d and Y_r are statistically correlated, the Markov conditions $Y_d \leftrightarrow Y_r \leftrightarrow \widehat{Y}_r$ are formed. This condition applies other restriction to $\widehat{Y}_r$ at each relay. The estimation variable $\widehat{Y}_r \in \widehat{\mathcal{Y}}_r$ must be drawn to satisfy $(\widehat{Y}_r, Y_r) \in A_\epsilon^m$ for a given $(Y_d, Y_r) \in A_\epsilon^m$. Accordingly, the Markov condition guarantees $(Y_d, \widehat{Y}_r, Y_r) \in A_\epsilon^m$. Using induction it is straightforward to show that $(Y_d, \widehat{Y}_1, \ldots, \widehat{Y}_R, Y_1, \ldots, Y_R) \in A_\epsilon^m$. Moreover, for every relay there exists a function $g_r(\cdot)$ such that:

$$\mathbb{E}\left\{d(X_1,\ldots,X_K,\ \ g_r(Y_d,\widehat{Y}_1,\ldots,\widehat{Y}_R))\right\} \leq D_r \tag{6.23}$$

Whereas the network tries to exploit the side information as much as possible, we do not care the amount of distortion D_r.

The destination's decoding process is a function $\mathcal{Y}_d \to \mathcal{X}_1 \times \cdots \times \mathcal{X}_K$. If it finds a unique K-tuple index in $\mathcal{Y}_d$ space, it calculates $(\hat{w}_1, \ldots, \hat{w}_k, \ldots, \hat{w}_K)$. The condition:

$$\tilde{R}_r > \mathrm{I}\left(\widehat{Y}_r\,;\,Y_r|X_r\right) \qquad \forall\, r \in \mathcal{R} \tag{6.24}$$

guarantees to find a jointly typical codewords at relay r's encoder [16, Lemma 10.6.2]. Thus, an error event at all relays' encoders is prevented if:

$$p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}\right) = \prod_{k\in\mathcal{K}} p\left(x_k\right).\prod_{r\in\mathcal{R}} p\left(y_r|x_r\boldsymbol{x}_{\mathcal{K}}\right).\prod_{r\in\mathcal{R}} p\left(\hat{y}_r|y_r x_r\right) \tag{6.25}$$

That is derived from uncorrelated output signals X_ks, and the probability of the Markov condition $\boldsymbol{X}_{\mathcal{K}} \leftrightarrow Y_r|X_r \leftrightarrow \widehat{Y}_r$ at each relay r, for encoders and quantization processes.

Now, we find a channel probability joint distribution for impeding error at the destination's decoder. In each block index b, the decoder uses the receive symbols $\boldsymbol{X}_{\mathcal{R}}$ to decode the already received symbols $\boldsymbol{X}_{\mathcal{K}}$ in block $b-1$. The decoding process is successfully fulfilled if all messages $\boldsymbol{X}_{\mathcal{R}}$ are successfully approached the destination. Next, $Y_d[b-1]$ given $\boldsymbol{X}_{\mathcal{R}}[b]$ is exploited as the side information to decode $\boldsymbol{X}_{\mathcal{K}}[b-1]$. The error of assigning a unique index at the decoder is prevented if $(\boldsymbol{X}_{\mathcal{K}}\,,\,Y_d|\boldsymbol{X}_{\mathcal{R}}) \in A_\epsilon^m$. Taking into account that X_ks are uncorrelated, these conditions are satisfied with probability:

$$p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}, y_d\right) = \prod_{k \in \mathcal{K}} p\left(x_k\right) \cdot p\left(\boldsymbol{x}_{\mathcal{R}}\right) \cdot p\left(y_d | \boldsymbol{x}_{\mathcal{K}} x_r\right) \tag{6.26}$$

The combination of error probabilities of Eqs. 6.25 and (6.26) achieves the probability mass function of CF technique in a one source, multiple parallel relays and one destination network as:

$$\begin{aligned} &p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}, y_d\right) = \\ &\quad \prod_{k \in \mathcal{K}} p\left(x_k\right) \cdot p\left(\boldsymbol{x}_{\mathcal{R}}\right) \cdot \prod_{r \in \mathcal{R}} p\left(y_r | x_r \boldsymbol{x}_{\mathcal{K}}\right) \cdot \prod_{r \in \mathcal{R}} p\left(\hat{y}_r | y_r x_r\right) \cdot p\left(y_d | \boldsymbol{x}_{\mathcal{K}} x_r\right) \end{aligned} \tag{6.27}$$

Other type of error at the destination's decoder may occur is confusing with other existing index in alphabet space $\mathcal{Y}_d$. There is a condition under which this error event occurs with arbitrary small probability, as stated by the following theorem.

Theorem 33 *The destination's decoder will not mistake the correct index with any other admissible one if:*

$$\tilde{R}_r - \widehat{R}_r \;\leq\; \mathrm{I}\left(\widehat{Y}_r\,;\, Y_d|\boldsymbol{X}_{\mathcal{R}}\right) = \mathrm{I}\left(\widehat{Y}_r\,;\, Y_d|X_r\right) \qquad \forall\, r \in \mathcal{R}$$

Proof The same proof of Theorem 18. ■

Consequently, the same results of Theorem 19 are derived. In a $(\mathcal{K}, \mathcal{R}, d)$ reliable network applying CF technique, each relay acts as an individual Wyner-Ziv source coder whose input and side information respectively are $Y_r|X_r$ and $Y_d|\boldsymbol{X}_{\mathcal{R}}$. Following the steps in Theorem 10 approaches:

$$\tilde{R}_r \;\geq\; \mathrm{I}\left(\widehat{Y}_r\,;\, Y_r|X_r Y_d\right) \qquad \forall\, r \in \mathcal{R} \tag{6.28}$$

So, for a $(\mathcal{K}; \mathcal{R}; d)$ reliable network scenario, an extension of Theorem 10 can be established as follows:

Theorem 34 *Suppose that the full-duplex relay complies with CF strategy. For the pmf* $\prod_{k\in\mathcal{K}} p(x_k)\cdot p(\boldsymbol{x}_{\mathcal{R}})\cdot \prod_{r\in\mathcal{R}} p(y_r|x_r\boldsymbol{x}_{\mathcal{K}})\cdot \prod_{r\in\mathcal{R}} p(\hat{y}_r|y_r x_r)\cdot p(y_d|x_k x_r)$ *there exist sequences of code books* $\left(X_k(w_k), X_r(z_r), \widehat{Y}_r(k_r|z_r)\right)$ $\forall\, r\in\mathcal{R}$ $\forall\, k\in\mathcal{K}$*, whose symbols are independently and uniformly derived from the* $w_k\in\{0,1,...,2^{mR_k}-1\}$*,* $z_r\in\{0,1,...,2^{m\widehat{R}_r}-1\}$*, and* $k_r\in\{0,1,...,2^{m\tilde{R}_r}-1\}$ *such that* $P_e((w_1,\ldots,w_k,\ldots,w_K)\neq(\hat{w}_1,\ldots,\hat{w}_k,\ldots,\hat{w}_K))\to 0$ *as* $m\to+\infty$ *if:*

$$R_k \leq \min_{r\in\mathcal{R}} \mathrm{I}\left(X_k\,;\,Y_d\,\widehat{Y}_r|X_r\right); \qquad \forall\, k\in\mathcal{K} \tag{6.29a}$$

$$\textit{Subject to:}\quad \mathrm{I}(X_r\,;\,Y_d) \geq \mathrm{I}\left(\widehat{Y}_r\,;\,Y_r|\,X_r\,Y_d\right) \qquad \forall\, r\in\mathcal{R} \tag{6.29b}$$

The following two theorems introduce the upper bound capacities.

Theorem 35 *The CF capacity of each source* k *in a reliable network consisted by multiple sources, multiple parallel relays, one destination is limited to:*

$$R_k \leq \sup_{p(\boldsymbol{x}_{\mathcal{K}},\,\boldsymbol{x}_{\mathcal{R}})} \min\left\{\min_{r\in\mathcal{R}}\left\{\mathrm{I}\left(X_k\,;\,Y_d\,Y_r|X_r\boldsymbol{X}_{\{k\}^C}\right)\right\}\,;\,\mathrm{I}\left(\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d\right)\right\}$$

where supremum is given over individual input constraints, and the channel probability distribution (6.27)*.*

Proof After having broadcast a message from the source node k, while the others transmitters in $\{k\}^C$ are silent, the relays compress their proper observations in a distributed and individual manner. The destination receives a message from k and the relayed messages of the previous sent symbol. Thus, the broadcast data rate is limited to:

$$\begin{aligned} R_k &\leq \mathrm{I}\left(X_k\,;\,Y_d|\boldsymbol{X}_{\{k\}^C}\right) + \min_{r\in\mathcal{R}}\left\{\mathrm{I}\left(X_k\,;\,Y_r|X_r\boldsymbol{X}_{\{k\}^C}\right)\right\} \\ &= \min_{r\in\mathcal{R}}\left\{\mathrm{I}\left(X_k\,;\,Y_d\,Y_r|X_r\boldsymbol{X}_{\{k\}^C}\right)\right\}. \end{aligned}$$

That means, the broadcast data rate is not limited to the relays' encoders data rate. Next, each relay sends the channel coded of its own received signal towards the destination. To do so, the relays map them onto $r\to d$ channels after they have accomplished the estimation and encoding. Whereas, the destination has already received $X_k[b-1]\left(w_k^{b-1}\right)$ and the data rate is limited to the relays' multiple-access channel, i.e. $R_k \leq \mathrm{I}\left(\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d\right)$. ■

Theorem 36 *The overall capacity of CF technique for a* $(\mathcal{K};\mathcal{R};d)$ *network is upper bounded by:*

$$\mathcal{C}_{CF}(\mathcal{K};\mathcal{R};d) = \sup_{p(\boldsymbol{x}_{\mathcal{K}},\boldsymbol{x}_{\mathcal{R}})} \min\left\{\min_{r\in\mathcal{R}}\left\{\mathrm{I}(\boldsymbol{X}_{\mathcal{K}}\,;\,Y_dY_r|X_r)\right\},\, \mathrm{I}(\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d)\right\}$$

where the supremum is given over individual input constraints, and the probability channel distribution (6.27).

Proof The proof is somehow obvious because: $\mathcal{C}_{CF}(\mathcal{K};\mathcal{R};d) = \sum_{k\in\mathcal{K}} R_k$, and R_k is derived from Theorem 35. ■

Theorem 36 implies that the maximum achievable capacity of a $(\mathcal{K};\mathcal{R};d)$ network applying CF strategy is limited by the multiple-access capacity of all relays at the destination. In such a reliable network, wherein all relays are in such conditions so as to perfectly receive signals from all source nodes, i.e. all relays are very close to all source nodes, increasing the number of relays, always close to the transmitters, yields making upper bound capacity better off, up to the number of relays with which the multiple-access capacity (the second term) dominates the first term. From this moment on, the overall capacity is equal to the broadcast term which is achieved by only one relay.

The AWGN capacities are calculated using the following theorems:

Theorem 37 *The AWGN CF upper bound of each source node* k*, in a* $(\mathcal{K};\mathcal{R};d)$ *network can be achieved the data rate:*

$$R_k \leq \sup_{-1\leq\rho_1,\ldots,\rho_K\leq 1} \min\left\{\min_{r\in\mathcal{R}}\left\{C\left(\frac{\boldsymbol{\Sigma}\left[Y_d\,Y_r|X_r\,\boldsymbol{X}_{\{k\}^C}\right]-\sigma_w^2}{\sigma_w^2}\right)\right\},\right.$$
$$\left. C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right)\;-\;\mathbb{V}\mathrm{ar}\left(Y_d\,|\,X_r\right)}{\mathbb{V}\mathrm{ar}\left(Y_d\,|\,X_r\right)}\right)\right\}$$

that is calculated using Eqs. 6.7.

Proof The broadcast term was proved in Theorem 27.

$$\mathrm{I}\left(\boldsymbol{X}_{\mathcal{R}}\,;\,Y_d\right) = \frac{1}{2}\log_2 \frac{\mathbb{V}\mathrm{ar}\left(Y_d\right)}{\mathbb{V}\mathrm{ar}\left(Y_d \,|\, X_r\right)} = C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right) \;-\; \mathbb{V}\mathrm{ar}\left(Y_d \,|\, X_r\right)}{\mathbb{V}\mathrm{ar}\left(Y_d \,|\, X_r\right)}\right).$$

■

Theorem 38 *The AWGN upper bound of the entire $(\mathcal{K};\mathcal{R};d)$ CF network achieves the data rate:*

$$\mathcal{C}_{CF}(\mathcal{K};\mathcal{R};d) = \sup_{-1\leq\rho_1,\ldots,\rho_K\leq 1} \min\left\{\min_{r\in\mathcal{R}}\left\{C\left(\frac{\boldsymbol{\Sigma}[Y_dY_r|X_r]-\sigma_w^2}{\sigma_w^2}\right)\right\},\; C\left(\frac{\mathbb{V}\mathrm{ar}\left(Y_d\right) \;-\; \mathbb{V}\mathrm{ar}\left(Y_d \,|\, X_r\right)}{\mathbb{V}\mathrm{ar}\left(Y_d \,|\, X_r\right)}\right)\right\}$$

that is calculated using Eqs. 6.7.

Proof The broadcast term was proved in Theorem 28 and the multiple-access term in Theorem 37. ■

In a reliable network wherein $\gamma_{kr} < \gamma_{rd}\ \forall\, k \in \mathcal{K},\ \forall\, r \in \mathcal{R}$, the $\mathcal{C}_{CF}(\mathcal{K};\mathcal{R};d)$ is achieved by the relay r with which $\sum_{k\in\mathcal{K}} h_{kr}\overline{p}_k(1-\rho_k^2)$ is the minimum. Setting up $\rho_k = 0\ \forall\ k \in \mathcal{K}$, significantly increases the upper bound capacity. In such a network, adding a new transmitter k, always $\gamma_{kr} < \gamma_{rd}\ \forall\, r \in \mathcal{R}$, expands the upper bound capacity of the whole network. On the other hand, in a network wherein $\gamma_{kr} > \gamma_{rd}\ \forall\, k \in \mathcal{K},\ \forall\, r \in \mathcal{R}$, adding a new transmitter may expand the total data rate, on the condition that $\rho_k > 0\ \forall\, k \in \mathcal{K}$. *In such a network with $\rho_k = 0$, increasing the number of transmitters significantly degrades the overall data rate of the network.*

6.6 Mixed DF & CF techniques

In this section, we study a multiple parallel relays network wherein the relays in $\mathcal{T} \subset \mathcal{R}$ execute block Markov DF strategy, and others relays belonging to $\mathcal{T}^C = \mathcal{R} \setminus \mathcal{T}$ perform CF technique. The network consists of $K + 2.R + 1$ code spaces. The alphabet space $\mathcal{X}_k\ \forall\, k \in \mathcal{K}$ at every transmitter's encoder,

$\mathcal{X}_r$ $\forall r \in \mathcal{R}$ at each relay's encoder, $\mathcal{Y}_r$ $\forall r \in \mathcal{T}$ at DF relays' decoders, $\widehat{\mathcal{Y}}_r$ $\forall r \in \mathcal{T}^C$ at CF relays' estimators, and $\mathcal{Y}_d$ code space at the destination's decoder. The transmit message of each source node w_k is a sequence of B sub-messages $w_k^1, ..., w_k^b, ..., w_k^B$ and each w_k^b is independently and uniformly drawn from the respective alphabet space $\mathcal{W}_k = \{0, 1, \ldots, 2^{mR_k} - 1\}$, and then it is fully encoded separately. The encoder at each transmitter k has one block memory that applying multiplex coding assigns a Gaussian random number $X_k[b](w_k^{b-1}, w_k^b)$ to each w_k^b. The transmission is done in $B+1$ blocks and each source node k transmits the sequence:

$$X_k[1](0, w_k^1), \ldots, X_k[b](w_k^{b-1}, w_k^b), \ldots, X_k[B](w_k^{B-1}, w_k^B), X_k[B+1](w_k^B, 0)$$

In each block b, the source node broadcasts $X_k[b](w_k^{b-1}, w_k^b)$ to all relays and destination. All relays $r \in \mathcal{R}$ receive the signal:

$$Y_r[b] = \sum_{k \in \mathcal{K}} \sqrt{h_{kr}}\, X_k[b](w_k^{b-1}, w_k^b) + Z_r \tag{6.30}$$

At the same time, every relay independently performs its processing function (DF or CF) on the previous received signal. The CF relays independently estimate their own previous received signal and then encode it to $X_r[b]\left(\widehat{Y}_r[b-1]\left(k_r^b\right)\right)$, as it was explained in the previous section. Every DF relay separately decodes and fully re-encode the previous observed message to $X_r[b]\left(\boldsymbol{w}_{\mathcal{K}}^{b-1}\right)$. Then, all relays simultaneously converge the result signal in multiple-access mode at the destination. The destination collects the sequence $Y_d[1], \ldots, Y_d[B+1]$ as:

$$\begin{aligned} Y_d[b] = \sum_{k \in \mathcal{K}} \sqrt{h_{kd}}\, X_k[b](w_k^{b-1}, w_k^b) + \sum_{r \in \mathcal{T}} \sqrt{h_{rd}}\, X_r[b](\boldsymbol{w}_{\mathcal{K}}^{b-1}) + \\ + \sum_{r \in \mathcal{T}^C} \sqrt{h_{rd}}\, X_r[b]\left(\widehat{Y}_r[b-1]\left(k_r^b\right)\right) + Z_d \end{aligned} \tag{6.31}$$

The decoding processes at the DF relays are successfully accomplished if:

$$R_k \leq \mathrm{I}\left(X_k\, ; Y_r | X_r \boldsymbol{X}_{\{k\}^C}\right) \qquad \forall\, r \in \mathcal{T} \tag{6.32}$$

For the CF relays, R_k is not limited to the relays' encoders process data rate, and it is limited to:

$$R_k \leq \mathrm{I}\left(X_k\,;\,Y_d\,Y_r|X_r\boldsymbol{X}_{\{k\}^C}\right) \qquad \forall\, r \in \mathcal{T}^C \tag{6.33}$$

Therefore, each transmitter's broadcast message data rate in the decode-compress and forward technique is limited to:

$$R_k \leq \min\{\min_{r\in\mathcal{T}}\{\mathrm{I}\left(X_k\,;\,Y_r|X_r\boldsymbol{X}_{\{k\}^C}\right)\}\,,\,\min_{r\in\mathcal{T}^C}\{\mathrm{I}\left(X_k\,;\,Y_d\,Y_r|X_r\boldsymbol{X}_{\{k\}^C}\right)\}\} \tag{6.34}$$

Consequently,

$$\mathcal{C}_{DCF}(\mathcal{K};\mathcal{R};d) \leq \min\{\min_{r\in\mathcal{T}}\{\mathrm{I}\left(\boldsymbol{X}_{\mathcal{K}}\,;\,Y_r|X_r\right)\}\,,\,\min_{r\in\mathcal{T}^C}\{\mathrm{I}\left(\boldsymbol{X}_{\mathcal{K}}\,;\,Y_d\,Y_r|X_r\right)\}\} \tag{6.35}$$

If there exists a DF relay $r \in \mathcal{T}$ such that $\sum_{k\in\mathcal{K}}\gamma_{kr} < \sum_{k\in\mathcal{K}}\gamma_{kd}$, then Eq. 6.35 results that it is better to turn off all relays and using direct-link connections.

At the destination, a multiplex decoding process starts after having collected $B+1$ samples of $Y_d[1],\ldots,Y_d[B+1]$. The decoding starts from the last block and proceeds backward to the first block. For each $Y_d[b]$, the destination uses the alphabet space $\mathcal{Y}_d$ for multiplex them and calculate $\hat{\boldsymbol{w}}_{\mathcal{K}}^{b-1} = (\hat{w}_1^{b-1},\ldots,\hat{w}_K^{b-1})$. If the destination's decoder has properly decoded $\boldsymbol{w}_{\mathcal{K}}^{b-1} = (w_1^{b-1},\ldots,w_K^{b-1})$, then the decoding of $Y_d[b]$ is successfully accomplished. According to Theorem 16 the pmf of the channels of the DF relays is:

$$\begin{aligned} &p\left(\boldsymbol{x}_{\mathcal{K}},\boldsymbol{x}_{\mathcal{R}},\boldsymbol{y}_{\mathcal{R}},\hat{\boldsymbol{y}}_{\mathcal{R}},y_d\right) = \\ &\qquad p\left(\boldsymbol{x}_{\mathcal{K}},\boldsymbol{x}_{\mathcal{T}}\right)\cdot\prod_{r\in\mathcal{T}}p\left(x_r|\boldsymbol{x}_{\mathcal{K}}\right).\prod_{r\in\mathcal{T}}p\left(y_r|x_r\,\boldsymbol{x}_{\mathcal{K}}\right).\,p\left(y_d|x_r\,\boldsymbol{x}_{\mathcal{K}}\right) \end{aligned} \tag{6.36}$$

Eq. 5.25 allows to evaluate the joint pmf of the CF channels as:

$$\begin{aligned} p\left(\boldsymbol{x}_{\mathcal{K}},\boldsymbol{x}_{\mathcal{R}},\boldsymbol{y}_{\mathcal{R}},\hat{\boldsymbol{y}}_{\mathcal{R}},y_d\right) = &\prod_{k\in\mathcal{K}}p\left(x_k\right)\cdot p\left(\boldsymbol{x}_{\mathcal{T}^C}\right)\cdot p\left(y_d|\boldsymbol{x}_{\mathcal{K}}x_r\right)\cdot \\ &\cdot\prod_{r\in\mathcal{T}^C}p\left(y_r|x_r\boldsymbol{x}_{\mathcal{K}}\right)\cdot\prod_{r\in\mathcal{T}^C}p\left(\hat{y}_r|y_rx_r\right) \end{aligned} \tag{6.37}$$

The combination of Eqs. 6.36 and 6.37 yields the pmf of the decode-compress and forward technique in a network consisted by one source, multiple parallel relays and one destination as:

$$p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}}, \boldsymbol{y}_{\mathcal{R}}, \hat{\boldsymbol{y}}_{\mathcal{R}}, y_d\right) = p\left(\boldsymbol{x}_{\mathcal{K}}, \boldsymbol{x}_{\mathcal{R}}\right) \cdot p\left(y_d | \boldsymbol{x}_{\mathcal{K}} x_r\right) \cdot \\ \cdot \prod_{r \in \mathcal{T}^C} p\left(\hat{y}_r | y_r x_r\right) \cdot \prod_{r \in \mathcal{R}} p\left(y_r | x_r \, \boldsymbol{x}_{\mathcal{K}}\right) \quad (6.38)$$

According to Theorem 16 the MAC channel capacity of the relays belong to the $\mathcal{T}$ is upper bounded to:

$$R_k \leq \mathrm{I}(X_k \, \boldsymbol{X}_{\mathcal{T}} \; ; \; Y_d | \boldsymbol{X}_{\{k\}^C}) = \mathrm{I}(X_k \, \boldsymbol{X}_{\mathcal{R}} \; ; \; Y_d | \boldsymbol{X}_{\{k\}^C}) \quad (6.39)$$

Theorem 20 shows that the MAC channel capacity of the CF relays is upper bounded to:

$$R_k \leq \mathrm{I}(\boldsymbol{X}_{\mathcal{T}^C} \; ; \; Y_d) = \mathrm{I}(\boldsymbol{X}_{\mathcal{R}} \; ; \; Y_d) \quad (6.40)$$

Therefore, the MAC channel message data rate in the decode-compress and forward technique is limited to:

$$R_k \leq \min_{r \in \mathcal{R}} \left\{ \mathrm{I}(X_k \, \boldsymbol{X}_{\mathcal{R}} \; ; \; Y_d | \boldsymbol{X}_{\{k\}^C}) \; , \; \mathrm{I}(\boldsymbol{X}_{\mathcal{R}} \; ; \; Y_d) \right\} \quad (6.41)$$

Consequently,

$$\mathcal{C}_{DCF}(\mathcal{K}; \mathcal{R}; d) \leq \min_{r \in \mathcal{R}} \left\{ \mathrm{I}(\boldsymbol{X}_{\mathcal{K}} \, \boldsymbol{X}_{\mathcal{R}} \; ; \; Y_d) \; , \; \mathrm{I}(\boldsymbol{X}_{\mathcal{R}} \; ; \; Y_d) \right\} \quad (6.42)$$

If $\rho_k = 1 \; \forall \, k \in \mathcal{K}$, then the two terms are equal, otherwise the minimum term is the second term. In a $(\mathcal{K}; \mathcal{R}; d)$ reliable network wherein all relays are in good reception conditions from all transmitters with $\rho_k \neq 1 \; \forall \, k \in \mathcal{K}$, the mixed DF and CF (DCF) upper bound capacity is equal to that the CF strategy, i.e. that is limited to the multiple access capacity of (only) all relays at the destination.

With the above discussion, it is straightforward to derive the upper bound capacity of a DCF reliable communication as in the following theorem.

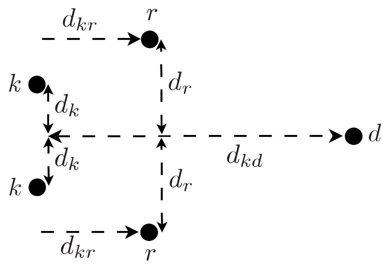

Fig. 6.2: *Multiple access multiple relayed communication network scenario.*

Theorem 39 *For one source, multiple parallel relays, one destination network, wherein the relays in subset $\mathcal{T} \subset \mathcal{R}$ perform DF technique and the others relays perform CF technique, the upper bound capacity is:*

$$\mathcal{C}_{DCF}(\mathcal{K};\mathcal{R};d) = \min\left\{\mathcal{C}_{DF}(\mathcal{K};\mathcal{T};d)\ ,\ \mathcal{C}_{CF}(\mathcal{K};\mathcal{T}^{C};d)\right\}$$

where supremum of each term is over individual power constraints of source nodes and each relay node and the joint pmf presented by Eq. 6.38.

This means, the $\mathcal{C}_{DCF}(\mathcal{K};\mathcal{R};d)$ reduces both $\mathcal{C}_{DF}(\mathcal{K};\mathcal{T};d)$ and $\mathcal{C}_{CF}(\mathcal{K};\mathcal{T}^{C};d)$. The $\mathcal{C}_{DCF}(\mathcal{K};\mathcal{R};d)$ achieves its maximum if all DF relays are in good condition to perfectly decode the proper observed signals and all CF relays error freely deliver signals to the destination. In a network wherein all relays are close to all source nodes, the upper bound capacity is limited to the multiple access capacity of all relays at the destination. If all relays are close to the destination, the outer region capacity is bounded by a DF relay at which the multiple-access capacity of all transmitters is the minimum.

6.7 Case study

Now, we illustrate the various outer region bounds of a multiple access multiple relayed communication with Gaussian channels wherein two transmitters

k, two relays r, and one destination d are located as depicted in Fig. 6.2. We assume a vertical distance of d_k between transmitters and the destination, and a vertical distance of d_r between the relays and the destination. The path condition values $h_{kr} = h_{rd} = h_{kd} = 1$ are scaled with respect to Notation 2.

First, we experiment a high SNR environment and suppose a horizontal distance of $d_{kd} = 1\,\text{m}$, and a vertical distance of $d_k = 0.05\,\text{m}$ between the source nodes and the destination. The relays are located in vertical distances of $d_r = 0.1\,\text{m}$ from the destination, and they are simultaneously and horizontally moving from $d_{kr} = -0.5\,\text{m}$ to $d_{kr} = 1.5\,\text{m}$. Fig. 6.3 plots various data rates for $\overline{p}_k = \overline{p}_r = 100\,\text{mW}$, and $\sigma_w^2 = 1\,\mu\text{W}$. The curve labeled AF shows the outer region of AF strategy with the largest possible scaling factor E_r in Eq. 6.9. The correlation coefficient $\rho_k = \rho$ is the same as that in Fig. 5.3. Like the one-relay case, as the relays moves toward the transmitter the DF strategy exhibits better performance, and the CF technique instead performs better when the relays are closer to the destination. The overall capacity of the network in the relays off mode is calculated using Eq. B.26. The comparison of Fig. 6.3 to Fig. 5.3 reveals that the overall data rates of two transmitters shows almost 20% better performance than that one transmitter.

Now, we consider a multiple-access multiple relayed communication in a low SNR regime. The transmitters and the destination are placed at horizontal distances of $d_{kd} = 500\,\text{m}$, and vertical distances of $d_k = 5\,\text{m}$. The relays are simultaneously and horizontally moving in a range of $d_{kr} = -100 \div 600\,\text{m}$ with vertical distances of $d_r = 10\,\text{m}$ from the destination. Fig. 6.4 plots various data rates for $\overline{p}_k = \overline{p}_r = 100\,\text{mW}$, $\sigma_w^2 = 1\,\mu\text{W}$, and the same E_r as the previous simulation.

The correlation coefficient $\rho_k = \rho$ is the same as that in Fig. 5.4. The comparison between Fig. 6.4 and (5.4) shows the overall data rate of AF technique in a low SNR network is almost 4 times higher than that in a point-to-point two relayed network. Adding new relays increases MAI at the destination and this reduces the performance of AF technique. For cutset, DF, CF, and MRC schemes, the comparison between Fig. 6.4 and (5.4) shows that the overall data rates of two transmitters aided by two relays are at most

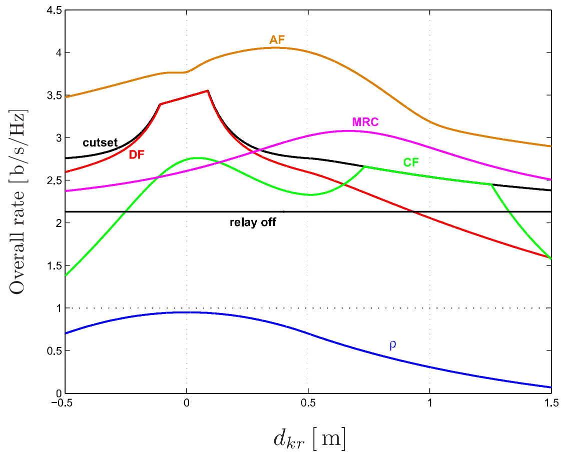

Fig. 6.3: *Overall rates of two transmitters relayed by two relays with* $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$*,* $\sigma_w^2 = 1\,\mu\mathrm{W}$*,* $d_{kd} = 1\,\mathrm{m}$*,* $d_k = 0.05\,\mathrm{m}$*, and* $d_r = 0.1\,\mathrm{m}$*.*

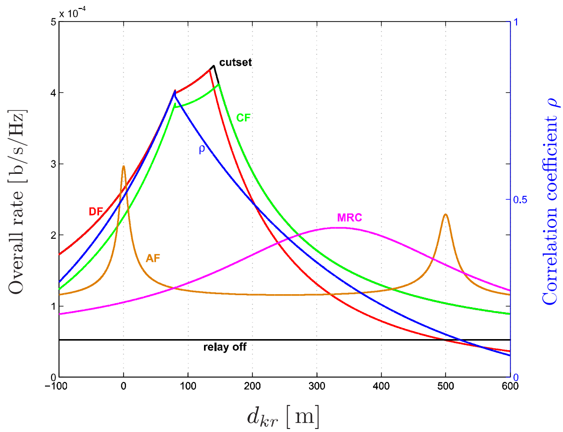

Fig. 6.4: *Overall rates of two transmitters aided by two relays with* $\overline{p}_k = \overline{p}_r = 100\,\mathrm{mW}$*,* $\sigma_w^2 = 1\,\mu\mathrm{W}$*,* $d_{kd} = 500\,\mathrm{m}$*,* $d_k = 5\,\mathrm{m}$*, and* $d_r = 10\,\mathrm{m}$*.*

50% higher than those of point-to-point two relayed network. Like in low SNR scenes in the previous chapters, the DF and CF coding techniques show significant higher rates than the relay off mode.

6.8 Summary

We studied multiple-access communication through an arbitrary number of multiple parallel relays, with full-duplex signaling and AWGN channels. We focused on finding the maximum achievable flow rate applying three well-known relay strategies of AF, DF and CF. First, we showed that the information rate achieve the maximum when the output signals of the relays are fully correlated, whereas the output of the transmitters are fully uncorrelated. The first interesting result is that the maximum channel capacity region is approached by either only one relay or all relays together.

Like in the one-relay networks, the performance of a multiple sources, multiple parallel relays, and one destination network basically depends upon both relays strategy and the their positions. If all relays are located so as to perfectly receive signals from all source nodes the DF strategy is useful, and the maximum capacity is limited by a relay at which the multiple-access capacity of all source nodes is the minimum. In this situation, the others relays can be turned off. Placing a DF relay at a distance of very far from all source nodes can even achieve a total data rate less than that the sources-to-destination direct-links. On the other hand, if all relays are appropriately placed to perfectly deliver signals to the destination the CF strategy is useful, and the maximum capacity is limited to the multiple-access capacity of all relays at the destination. In such a network, adding a new source node can even degrade the outer region capacity of the entire network, whereas adding a new relay may expand the outer region of capacity.

We further showed that, in a multiple sources, multiple parallel relays, and one destination network, applying the same strategy at every relays achieves larger data rate, rather than applying different strategies.

Chapter 7

Summary and perspective

In the first part of this thesis, we used cooperative game theoretic solutions to different wireless network engineering problems, focusing in particular on the issue of fairness. In particular, we studied resource allocation techniques in the uplink direction of orthogonal frequency division multiple access (OFDMA) systems.

The main concern in the identification of the game has been the best utilization of the network resources applying a low complexity algorithm. This has led us to introduce a utility function with which each active wireless terminal achieves its request data rate exactly. This fairness criterion satisfies the expectation of both the wireless service provider and each user's terminal. To cope with the non-convexity and non-concavity of the utility function we proposed a dynamic learning algorithm. The algorithm enforces every terminal to adapt its transmit power to approach the maximum value of the utility function. The proposed algorithm approaches a distribution of transmit powers over different assigned subcarriers for each user from which no terminal wishes to unilaterally deviate (the core set). We demonstrated that the core set coincides with a Nash equilibrium point. The convergence and stability of the algorithm was proved based on Markov modeling, and the complexity of the algorithm favorably compared to the existing literature.

The second part of this dissertation has investigated the role of relaying in communication networks. We started from the simplest relaying communication in which a couple of terminals communicate with each other through one relay, and then we extended it to deal with large networks consisting of multiple sources and multiple parallel relays. We assume the signaling is performed in full-duplex mode.

Firstly, we studied the simplest example of a communication network with three nodes, i.e. a relay channel where one intermediate relay node facilitates the communication from the transmitter to the receiver. We studied different relaying strategies, amplify and forward (AF), block Markov decoding and forward (DF), and compress and forward (CF). We established upper bounds on the capacity of these channels subject to average power constraints. With two different case studies we also showed that the AF scheme is only useful in practice in high SNR regime. When the relay is close to the transmitter the DF technique is to be preferred, and instead when the relay is close to the destination the CF technique achieves higher performance. When the relay is very far from the transmitter, the capacity of DF relaying strategy can even be less than that of the direct-link. These rules are also valid in a large reliable network with multiple relays and multiple transmitters.

We have extended in fact the previous scenario to point-to-point communication aided by multiple parallel relays. We proved that the largest capacity is achieved when the relays output symbols are fully correlated. Then, we showed that the capacity of such a network is upper bounded by either only one relay or all relays together. When all relays are in a good condition to deliver signals to the destination with low BER, the cutset upper bound capacity is equal to the broadcast capacity of the source node to the destination and to the relay with the weakest channel condition. In such a network, the other relays can be turned off. Instead, when all relays are located so as to receive signals from the transmitter with good quality, the cutset upper bound capacity is equal to the multiple access capacity of all relays and the transmitter at the destination. The DF capacity is limited by the link between the transmitter and the relay with the minimum data rate. The CF capacity is bounded by the multiple access capacity of all relays at the destination. A point-to-point communication assisted by multiple parallel relays which perform CF technique form a distributed Wyner-Ziv coding network with different but correlated inputs, and a common decoder at the destination. We also showed that adding a new relay does not necessarily result expanding the upper bound capacity.

Finally, we extended the previous scenario to the multiple access channel with multiple parallel relays, i.e. multiple transmitters communicate to a single sink aided by multiple parallel relays. We proved that the largest capacity is achieved when the relays output symbols are fully correlated, and the output of the source nodes are fully independent. Similarly, we showed that the cutset is upper bounded by either only one relay or all relays together. As before, adding a new relay does not necessarily result in expanding the upper bound capacity, and adding a new transmitter can even degrade the overall data rate. The overall upper bound capacity of DF technique is limited by the relay at which the multiple access capacity of all transmitters is the minimum. The overall upper bound capacity of CF technique is limited to the multiple access capacity of all relays at the destination.

For the future work, we plan a cooperative resource allocation game in a multiple-access OFDMA wireless network consisting of multiple terminals and multiple relays. The algorithm consists of subcarrier assignment and power allocation at each node for maximizing the frequency spectral efficiency at the minimum cost of power consumption. The communication is full-duplex mode. At the relays and the base-station each subcarrier is allowed be shared by more terminals. There are different subcarrier assignments at each transmitter, and the relays are not constrained to transmit the same subcarriers over which receive the symbols. There are also separate power constraints on each subcarrier at every nodes. Taking into account the relaying strategy, we introduce a condition under which source nodes discern the needed subcarriers to be assisted. The transmitted symbols by the source nodes achieve the base-station and all relays. Each relay decodes/estimates the received signals on poor channels, encodes them, and forwards toward the base-station.

Appendix A

Multivariate Gaussian distribution

The multivariate Gaussian (normal) distribution of a M-dimensional real-valued random vector $\boldsymbol{X} = [X_1, \ldots, X_M]$ can be written in the following notation:

$$\boldsymbol{X} \sim \mathcal{N}(\mu, \Sigma), \tag{A.1}$$

with M-dimensional mean vector:

$$\mu = \begin{bmatrix} \mathbb{E}\{X_1\} \\ \vdots \\ \mathbb{E}\{X_M\} \end{bmatrix} \in \mathbb{R}^M, \tag{A.2}$$

and $M*M$-dimensional covariance matrix:

$$\Sigma = [\mathbb{C}\mathrm{ov}(X_i, X_j)] \in \mathbb{R}^{M*M} \quad i = 1, \ldots, M; \;\; j = 1, \ldots, M. \tag{A.3}$$

The entropy of $\boldsymbol{X}$ is formulated as:

$$\mathrm{H}(\boldsymbol{X}) = \frac{1}{2}\log_2\left(2\pi e.\det\{\Sigma\}\right) \tag{A.4}$$

If we consider another multivariate Gaussian random $\boldsymbol{Y} = [Y_1, \ldots, Y_N]$ jointly Gaussian with $\boldsymbol{X}$, then mean vector and covariance matrix of $\boldsymbol{XY}$ is identified by mean vector:

$$\mu = \begin{bmatrix} \mu_{\boldsymbol{X}} \\ \mu_{\boldsymbol{Y}} \end{bmatrix} \text{ with sizes } \begin{bmatrix} M*1 \\ N*1 \end{bmatrix} \tag{A.5}$$

and covariance matrix:

$$\Sigma = \begin{bmatrix} \Sigma_{\boldsymbol{XX}} & \Sigma_{\boldsymbol{XY}} \\ \Sigma_{\boldsymbol{YX}} & \Sigma_{\boldsymbol{YY}} \end{bmatrix} \text{with sizes} \begin{bmatrix} M*M & M*N \\ N*M & N*N \end{bmatrix} \tag{A.6}$$

where

$$\Sigma_{\boldsymbol{XY}} = [\mathbb{C}\text{ov}(X_i, Y_j)] \in \mathbb{R}^{M*N} \quad i = 1, \ldots, M; \;\; j = 1, \ldots, N. \tag{A.7}$$

and similarly for $\Sigma_{\boldsymbol{XX}}$, $\Sigma_{\boldsymbol{YX}}$ and $\Sigma_{\boldsymbol{YY}}$.

Then, the conditional distribution of $\boldsymbol{X}|\boldsymbol{Y}=\boldsymbol{y}$ is a multivariate Gaussian distribution $\mathcal{N}\left(\overline{\mu}, \overline{\Sigma}\right)$, where:

$$\overline{\mu} \;=\; \mu_{\boldsymbol{X}} + \Sigma_{\boldsymbol{XY}} \cdot \left(\Sigma_{\boldsymbol{YY}}\right)^{-1} \cdot (\boldsymbol{y} - \mu_{\boldsymbol{Y}}) \tag{A.8}$$

$$\overline{\Sigma} \;=\; \Sigma_{\boldsymbol{XX}} - \Sigma_{\boldsymbol{XY}} \cdot \left(\Sigma_{\boldsymbol{YY}}\right)^{-1} \cdot \Sigma_{\boldsymbol{YX}} \tag{A.9}$$

To compute (conditional) mutual information, we need the statistical parameter:

$$\boldsymbol{\Sigma}[\boldsymbol{X}|\boldsymbol{Y}] \triangleq \det\{\overline{\Sigma}\} \tag{A.10}$$

Appendix B

Some known results in information theory

B.1 Memoryless channels

In this section we give some basic definitions and theorems mainly on capacity for memoryless channels. A discrete memoryless channel consists of six entities, grouped into: transmitter, channel, and destination. The transmitted symbol w_k is selected in a randomly and equiprobable manner from an input alphabet set $\mathcal{W}_k = \{0, 1, ..., 2^{mR_k} - 1\}$. As sketched in Fig. B.1, we add an encoder before the channel and a decoder after it. The encoder encodes the source message w_k into a transmitted message X_k, adding redundancy to the original message in some way. When the signal is transmitted through the channel, it is distorted in a random way which depends on the channel characteristics. As such, the signal received may be different from the signal transmitted. A simple model for wireless transmission is the channel that adds noise Z to the transmitted message, yielding a received message Y_d.

$$Y_d = \sqrt{h_{kd}}\, X_k + Z \tag{B.1}$$

where h_{kd} represents the value of channel condition between source and destination. The decoder uses the known redundancy introduced by the encoding system to infer both the original signal w_k and the added noise.

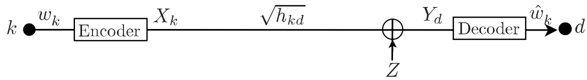

Fig. B.1: *The basic communication system.*

Even though the source and the channel are memoryless, the channel input sequence X_k is not a sequence of independent and identically distributed (i.i.d.) random variables in general. The encoder is a function $\mathcal{W}_k \rightarrow \mathcal{X}$ and its outputs X_ks are usually a sequence of zero mean (Gaussian) random values with a bounded variance $\mathbb{E}\left\{X_k^2\right\} \leq \overline{p}_k$. The sequence of $\{X_k\}$ is produced with a probability (mass) function of $p(x)$. The parameter $\overline{p}_k$ is defined as the physical power constraint on the channel, that the transmitter can spend to convey the data symbol to receiver. The decoder is a function of $\mathcal{Y} \rightarrow \mathcal{W}_k$ that maps the received signal Y_d to an estimated symbol $\hat{w}_k$. The average probability of error is defined as:

$$P_e = \Pr\{w_k \neq \hat{w}_k \,|\, w_k \text{ was sent}\} \tag{B.2}$$

A transmit rate R_k is said to be achievable if:

$$\lim_{m\rightarrow\infty} 2^{-mR_k}. \Pr\{w_k \neq \hat{w}_k \,|\, w_k \text{ was sent}\} = 0 \tag{B.3}$$

A discrete memoryless channel is defined as:

Definition 22 *Let $\mathcal{X}$ and $\mathcal{Y}$ be discrete alphabets, and $p(y|x)$ be a transition matrix from $\mathcal{X}$ to $\mathcal{Y}$. A discrete channel $p(y|x)$ is a single-input single-output system with input random variable X_k taking values in $\mathcal{X}$ and output random variable Y_d taking values in $\mathcal{Y}$ such that:*

$$\Pr\{X_k = x\,,\, Y_d = y\} = \Pr\{X_k = x\}\, p(y|x)$$

for all $(x\,,\,y)$ in $\mathcal{X} \times \mathcal{Y}$.

The fundamental starting point of digital communication is entropy that is defined as the measure of the uncertainty of a multidimensional set of discrete random variables.

$$\begin{aligned} \mathrm{H}(X_1,\ldots,X_K) &= -\sum_{\mathcal{X}_1\times\cdots\times\mathcal{X}_K} p(x_1,\ldots,x_K)\log_2 p(x_1,\ldots,x_K) \\ &= -\,\mathbb{E}\left\{\log_2 p(x_1,\ldots,x_K)\right\} \end{aligned} \tag{B.4}$$

where $(X_1,\ldots,X_K)$ follow the probability function $p(x_1,\ldots,x_K)$, and $\mathcal{X}_k$ is the limited support set of the random variables X_k. The *chain rule* for joint entropy states that:

$$\mathrm{H}(X_1,\ldots,X_K) = \sum_{i=1}^{K}\mathrm{H}\left(X_i|X_{i-1},X_{i-2},\ldots,X_1\right) \tag{B.5}$$

For continuous-valued random variables, definition (B.4) is replaced by that of "differential entropy" that we do not report here. In the particular case of Gaussian variables, as usual it is assumed in digital communication, the (differential) entropy of a Gaussian random variable X is equal to:

$$\mathrm{H}(X) = \frac{1}{2}\log_2\left(2\pi e.\,\mathbb{V}\mathrm{ar}(X)\right) \tag{B.6}$$

Shannon also introduced the notion of the (average) mutual information between the two random variables as follows:

$$\begin{aligned} \mathrm{I}\left(X_k\,;\,Y_d\right) &= \sum_{\mathcal{X}\times\mathcal{Y}} p(x,y)\log_2\frac{p(x,y)}{p(x)p(y)} = \mathbb{E}\left\{\log_2\frac{p(x,y)}{p(x)p(y)}\right\} \\ &= \mathrm{H}(X_k)+\mathrm{H}(Y_d)-\mathrm{H}(X_k,Y_d) \\ &= \mathrm{H}(X_k)-\mathrm{H}(X_k|Y_d) = \mathrm{H}(Y_d)-\mathrm{H}(Y_d|X_k) \end{aligned} \tag{B.7}$$

where the conditional entropy of X_k given Y_d is defined as:

$$\mathrm{H}(X_k\,|\,Y_d) = -\sum_{\mathcal{X}\times\mathcal{Y}} p(x,y)\log_2 p(x|y) = -\,\mathbb{E}\left\{\log_2 p(x|y)\right\} \tag{B.8}$$

Mutual information also satisfies a chain rule:

$$\mathrm{I}\left(X_1,\ldots,X_K\,;\,Y_d\right) = \sum_{i=1}^{K}\mathrm{I}\left(X_i\,;\,Y_d\mid X_{i-1},X_{i-2},\ldots,X_1\right) \tag{B.9}$$

Shannon in 1948 [189] has shown that, theoretically, it is possible to transmit information over a given channel with an arbitrary small error probability if the data rate is not greater than the *channel capacity* (*information rate*).

Definition 23 *(Shannon channel capacity function) The capacity function of a single-input, single-output channel is calculated as:*

$$\mathcal{C} = \sup_{X_k:\,\mathbb{E}\left\{X_k^2\right\}\leq\overline{p}_k} \mathrm{I}\left(X_k\,;\,Y_d\right)$$

Therefore, the channel capacity is the ultimate limit of the data rate for reliable communication. Channel capacity is the maximum rate of information for which an arbitrary small error probability can be achieved. For the Gaussian random input signal X_k, the Shannon channel capacity is equal to:

$$\begin{aligned}\mathcal{C} = \mathrm{I}\left(X_k\,;\,Y_d\right) &= \frac{1}{2}\log_2\left(2\pi e.\,\mathbb{V}\mathrm{ar}(Y_d)\right) - \frac{1}{2}\log_2\left(2\pi e.\,\mathbb{V}\mathrm{ar}(Y_d|X_k)\right) \\ &= \frac{1}{2}\log_2\frac{\mathbb{V}\mathrm{ar}(Y_d)}{\mathbb{V}\mathrm{ar}(Y_d|X_k)} \quad \text{bit/use}\end{aligned} \tag{B.10}$$

We assume Z is an additive white Gaussian noise (AWGN) of $\mathcal{N}(0,\sigma_w^2)$, and also $\mathbb{E}\left\{X_k\right\} = 0$. Hence, the Shannon channel capacity is equal to:

$$\frac{1}{2}\log_2\left(1+\frac{h_{kd}\overline{p}_k}{\sigma_w^2}\right) \quad \text{bit/use} \tag{B.11}$$

We will indicate the capacity of a Gaussian memoryless channel as a function of the signal-to-noise ratio (SNR) by the C-function:

$$C(\gamma) = \frac{1}{2}\log_2\left(1+\gamma\right) \quad \text{bit/use} \tag{B.12}$$

In other words, there is a maximal rate, called the capacity of the channel C, for which this can be done: if one attempts to communicate at rates above the channel capacity, then it is impossible to drive the error probability to zero.

In communication engineering, we are interested in conveying messages reliably through a noisy channel at the maximum possible rate. However, if we allow P_e to be any small quantity, Shannon showed that there exists a block code whose coding rate is arbitrarily close to $\mathrm{H}(X_k)$ when m is sufficiently large. This is the direct part of Shannon's source coding theorem, and in this sense the source sequence X_k is said to be reconstructed almost perfectly.

Theorem 40 *(AEP) [16, Theorem 3.1.2] If the sequence $\{X_k\}$ is generated with probability mass function $p(x)$, the following holds for any small $\epsilon > 0$, and m sufficiently large:*

$$2^{-m(\mathrm{H}(X_k)+\epsilon)} \leq p(x) \leq 2^{-m(\mathrm{H}(X_k)-\epsilon)}$$

The joint AEP will enable us to calculate the probability of error for jointly typical decoding for the various coding schemes:

Theorem 41 *(Joint AEP) [16, Theorem 7.6.1] If the sequence $\{(X_1, X_2)\}$ is generated with probability function $p(x_1, x_2)$, the following holds for any small $\epsilon > 0$, and $m \to \infty$:*

$$2^{-m(\mathrm{H}(X_1,X_2)+\epsilon)} \leq p(x_1, x_2) \leq 2^{-m(\mathrm{H}(X_1,X_2)-\epsilon)}$$

Let $\{(X_1, X_2)\}$ be sequences of length m drawn i.i.d. according to the joint probability distribution $p(x_1, x_2) = p(x_1).p(x_2)$, i.e. X_1 and X_2 are independent with the same marginal as $p(x_1, x_2)$. Then, the probability of jointly typical $m \to +\infty$ sequences of $\{(X_1, X_2)\}$ is bounded to [16, Eq. 7.53]:

$$\begin{aligned} Pr\{(X_1,\, X_2) \in A_\epsilon^m\} &\leq 2^{-m(\mathrm{H}(X_1)-\epsilon)}.\,2^{-m(\mathrm{H}(X_2)-\epsilon)}.\,2^{m(\mathrm{H}(X_1\,,\,X_2)+\epsilon)} \\ &= 2^{-m(\mathrm{I}(X_1;\,X_2)-3\epsilon)} \end{aligned} \tag{B.13}$$

where the typical set (i.e., the set of jointly typical sequence) A_ϵ^m is defined by:

Definition 24 *The set A_ϵ^m of jointly ϵ-typical m sequences of $(X_1, \ldots, X_K)$ with respect to the distribution $p\,(x_1, \ldots, x_K)$ is defined as follows:*

$$A_\epsilon = \left\{(X_1, \ldots, X_K) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_K : \left| -\frac{1}{m} \log_2 p(\boldsymbol{x}) - \mathrm{H}(\boldsymbol{X}) \right| \leq \epsilon \right\} \tag{B.14}$$

whatever $\boldsymbol{X} \subseteq \{X_1, \ldots, X_K\}$*, and where* $\boldsymbol{x}$ *represents the ordered set of sequence* $\{x_1, \ldots, x_K\}$ *corresponding to* $\boldsymbol{X}$*.*

In the memoryless channel of Fig. B.1, the joint AEP theorem tells us that the probability that X_k and Y_d are jointly typical at destination's decoder is $2^{-m(\mathrm{I}(X_k;\, Y_d)-3\epsilon)}$. Hence, by sending 2^{mR_k} codes over the channel, the probability that every independent sent codeword X_k is jointly typical with received Y_d is bounded to [16, Eq. 7.78]:

$$\Pr\{(X_k,\, Y_d) \in A_\epsilon^m\} \leq 2^{mR_k}.2^{-m(\mathrm{I}(X_k;\, Y_d)-3\epsilon)} \tag{B.15}$$

which goes to zero as $m \to +\infty$ if:

$$R_k < \mathrm{I}(X_k\,;\, Y_d). \tag{B.16}$$

Suppose that a finite collection of discrete random variables $\boldsymbol{X}_{\mathcal{K}} = (X_1, \ldots, X_K)$ is generated on the space set of $\mathcal{X}_1 \times \cdots \times \mathcal{X}_K$ with some fixed joint distribution $p\,(\boldsymbol{x}_{\mathcal{K}})$. The following holds for any arbitrary small $\epsilon > 0$, and $m \to \infty$ [16, Th. 15.2.1]:

$$2^{-m(\mathrm{H}(\boldsymbol{X}_{\mathcal{K}})+\epsilon)} \leq p\,(\boldsymbol{x}_{\mathcal{K}}) \leq 2^{-m(\mathrm{H}(\boldsymbol{X}_{\mathcal{K}})-\epsilon)} \tag{B.17}$$

If the variables of $\boldsymbol{X}_{\mathcal{K}}$ are drawn independently with the same marginal, i.e. $(\boldsymbol{X}_{\mathcal{K}}) \sim p\,(\boldsymbol{x}_{\mathcal{K}}) = \prod_{k\in\mathcal{K}} p\,(x_k)$, then the joint AEP theorem is extended to:

$$\Pr\{(\boldsymbol{X}_{\mathcal{K}}) \in A_\epsilon^m\} \leq \prod_{k\in\mathcal{K}} 2^{mR_k}. \prod_{k\in\mathcal{K}} 2^{-m(\mathrm{H}(X_k)-\epsilon)}.\, 2^{m(\mathrm{H}(\boldsymbol{X}_{\mathcal{K}})+\epsilon)} \tag{B.18}$$

B.1.1 Broadcast communication

A broadcast channel consists of a single broadcaster and more receivers, for example, communication from a satellite to several ground stations. The properties of the channel are defined by a conditional distribution $p\,(y_1 \ldots y_D \,|\, x)$. Again, in the Gaussian channel for every signal X_k transmitted by the broadcaster, receiver d receives:

$$Y_d = \sqrt{h_{kd}}\, X_k + Z_d \qquad Z_d \sim \mathcal{N}(0,\, \sigma_w^2) \tag{B.19}$$

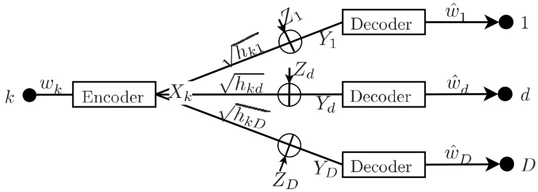

Fig. B.2: *Broadcast communication network.*

A simple benchmark for such a channel is given by time-sharing (time division multiplexing, TDM). That is, devoting a fraction of the transmission time to each channel. Reference [16, Ch. 15.1.3] introduces a degraded channel which can do better than TDM. In a degraded broadcast channel the conditional probabilities are such that the random variables have the structure of a Markov chain, $X_k \leftrightarrow Y_1 \leftrightarrow Y_2 \leftrightarrow \cdots$, i.e. Y_d is a further degraded version of Y_{d-1}. In this special case, it turns out that whatever information is getting through to receiver d can also be recovered by receiver $d-1$.

The capacity region of a degraded broadcast channel is computed in [16, Ch. 15.1.3] for $D = 2$. Consider hereafter, without loss of generality, that $h_{k1} > h_{k2}$. The source randomly selects two codewords $w_k \in \mathcal{W}_k$ $k = 1, 2$ where $\mathcal{W}_k = \left\{0, 1, ..., 2^{mR_k} - 1\right\}$ for transmission to the respective destinations. The idea is to divide the total power constraint $\overline{p}_k$ into two fractions $p_1 = \alpha \overline{p}_k$ and $p_2 = (1-\alpha)\overline{p}_k$, $\alpha \in (0, 1)$ and to construct two independent Gaussian codebooks for the two destinations with powers p_1 and p_2, respectively. To send two independent messages, one codeword is chosen from each codebook, and their sum is transmitted. Because Y_2 is a degraded version of Y_1, the codeword intended for Y_2 can also be decoded by Y_1. Each individual decoding function $\mathcal{Y}_{d=k} \rightarrow \mathcal{W}_k$ $k = 1, 2$ de-maps the received signal Y_d into an estimated message $\hat{w}_d$. A transmission rate pair (R_1, R_2) [bits/symbol] is said to be achievable if there exists a sequence of codes $(m, 2^{mR_1}, 2^{mR_2})$ for which no errors occur:

$$P_e^m =$$
$$\lim_{m\to+\infty}\prod_{k=1}^{2} 2^{-mR_k}\cdot \sum_{(w_1,w_2)} \Pr\ \{(\hat{w}_1,\hat{w}_2)\neq(w_1,w_2)|\,(w_1,w_2)\text{was sent}\} = 0 \tag{B.20}$$

The capacity region of the Gaussian broadcast channel with source power constraint $\overline{p}_k$ and noise power σ_w^2 at both receivers is:

$$\begin{aligned} R_1 &\leq C\left(\frac{h_{k1}p_1}{\sigma_w^2}\right) \\ R_2 &\leq C\left(\frac{h_{k2}p_2}{\sigma_w^2 + h_{k2}p_1}\right) \end{aligned} \tag{B.21}$$

B.1.2 Multiple-access communication

Consider a noisy channel with several inputs and one output, for example, a shared telephone line. Users cannot communicate with each other, and they cannot hear the output of the channel. A simple system model has $K \geq 2$ inputs $X_k \in \mathcal{X}_k$ and an output Y_d equal to the sum of the inputs. Every X_k is the encoded of codeword w_k which belong to the proper alphabet space $\mathcal{W}_k = \{0, 1, \ldots, 2^{mR_k} - 1\}$, under constraint that $\mathbb{E}\{X_k^2\} \leq \overline{p}_k$. Each encoder consists of an individual function of $\mathcal{W}_k \rightarrow \mathcal{X}_k$. The destination receives:

$$Y_d = \sum_{k=1}^{K} \sqrt{h_{kd}}X_k + Z_d \tag{B.22}$$

A joint function of $\mathcal{Y}_d \rightarrow \mathcal{W}_1 \times \cdots \times \mathcal{W}_K$ at the receiver's decoder reveal K-tuple of symbols $\hat{\boldsymbol{w}}_k = (\hat{w}_1, \ldots, \hat{w}_K)$ for the sent $\boldsymbol{w}_k = (w_1, \ldots, w_K)$. The code space $\mathcal{Y}_d$ consists of 2^{mR_d} code words. The probability transition matrix is $p\,(y_d|x_1 \ldots x_K)$. The error probability for each received signal is defined as:

$$P_e = \Pr\{\hat{\boldsymbol{w}}_k \neq \boldsymbol{w}_k|\boldsymbol{w}_k \text{ was sent}\} \tag{B.23}$$

A transmission K-tuple $(R_1, \ldots, R_K)$ is said to be achievable if there exists

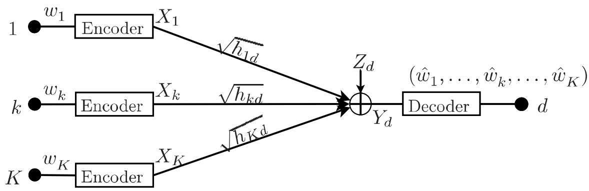

Fig. B.3: *Multiple access communication network.*

a sequence of codes $\left(m, 2^{mR_1}, \ldots, 2^{mR_K}\right)$ for which:

$$P_e^m = \lim_{m\to\infty} \prod_{k=1}^{K} 2^{-mR_k} \cdot \sum_{\boldsymbol{w}_k = (w_1,\ldots,w_K)} \Pr\{\hat{\boldsymbol{w}}_k \neq \boldsymbol{w}_k | \boldsymbol{w}_k \text{ was sent}\} = 0 \tag{B.24}$$

All source nodes interferer at the destination and obviously each transmitter achieves its maximum data rate while others transmitters are not sending signals. The capacity of the Gaussian multiple-access channel is:

Prop. 3 *[16, Sec. 15.3.6] The capacity of a multiple-access channel described by* $\left(\times_{k\in\mathcal{K}} \mathcal{X}_k\,,\, p\left(y_d|\boldsymbol{x}_\mathcal{K}\right),\, \mathcal{Y}_d\right)$ *given individual source power constraints* $\overline{p}_k$*, and noise power* σ_w^2*, is the closure of the convex hull of all* R_k *satisfying:*

$$\sum_{k\in\mathcal{A}} R_k \leq \mathrm{I}\left(\boldsymbol{X}_\mathcal{A}\,;\, Y_d|\boldsymbol{X}_{\mathcal{A}^C}\right) \quad \forall\, \mathcal{A} \subseteq \mathcal{K} \tag{B.25}$$

for some product distribution $\prod_{k\in\mathcal{K}} p(x_k)$ *on* $\times_{k\in\mathcal{K}} \mathcal{X}_k$*. The sum-rate of the channel obeys:*

$$\sum_{k\in\mathcal{A}} R_k \leq C\left(\sum_{k\in\mathcal{A}} \frac{h_{kd}\overline{p}_k}{\sigma_w^2}\right) \tag{B.26}$$

As can be derived from Eq. B.22, a very noisy channel affects decoding of the others codewords sent. Assume that the receiver's decoder can not de-map the received signal Y_d to a jointly typical of $(X_1, \ldots, X_K)$. Let us suppose that a subset of signals $\mathcal{A} \subseteq \{\boldsymbol{X}_\mathcal{K}\}$ caused this error because their SNR condition is very weak. Whatever transmit symbols $(w_1, \ldots, w_K)$, at the decoder, the jointly typical error probability caused by the sources $\mathcal{A}$ is bounded to [16, Eq. 7.52]:

$$\Pr\{(\boldsymbol{X}_{\mathcal{K}}, Y_d) \in A_\epsilon\} \leq \prod_{k \in \mathcal{A}} 2^{-m(\mathrm{H}(X_k)-\epsilon)} \cdot 2^{-m\left(\mathrm{H}\left(\boldsymbol{X}_{\mathcal{K}\setminus\mathcal{A}}\, Y_d\right)-\epsilon\right)} \cdot 2^{m(\mathrm{H}(\boldsymbol{X}_{\mathcal{K}}\, Y_d)+\epsilon)} \tag{B.27}$$

For a sequence of $m \to \infty$ number of $\{(X_1, \ldots, X_K)\}$, the probability of finding a jointly typical index at destination's decoder is defined by:

$$\Pr\{(\boldsymbol{X}_{\mathcal{K}}, Y_d) \in A_\epsilon^m\} \leq \prod_{k \in \mathcal{A}} 2^{-m(R_d-R_k)} \cdot \prod_{k \in \mathcal{A}} 2^{-m(\mathrm{H}(X_k)-\epsilon)} \cdot 2^{-m\left(\mathrm{H}\left(\boldsymbol{X}_{\mathcal{K}\setminus\mathcal{A}}\, Y_d\right)-\epsilon\right)} \cdot 2^{m(\mathrm{H}(\boldsymbol{X}_{\mathcal{K}}\, Y_d)+\epsilon)} \tag{B.28}$$

where R_d indicates the data rate process at destination's decoder. Consequently, the decoding process can guarantee of finding a jointly typical index if the following formula holds for every $\mathcal{A} \subseteq \mathcal{K}$:

$$\sum_{k \in \mathcal{A}} (R_d - R_k) \leq \sum_{k \in \mathcal{A}} \mathrm{H}(X_k) + \mathrm{H}\left(\boldsymbol{X}_{\mathcal{K}\setminus\mathcal{A}}\, Y_d\right) - \mathrm{H}(\boldsymbol{X}_{\mathcal{K}}\, Y_d) \tag{B.29}$$

B.1.3 General multi node network

Now, we consider a general multi terminal network consisting of one source k, one receiver d, and a finite number of intermediate nodes represented by $\mathcal{R} = [1, \ldots, r \ldots, R]$. The network can be considered as a graph where each intermediate node represents a potential transmitter or receiver. A *cut* is a partition of the intermediate node set $\mathcal{R}$ into two subsets $\mathcal{T}$ and $\mathcal{T}^C = \mathcal{R} \setminus \mathcal{T}$ that separates the network into two disjoint parts $\{k\,, \mathcal{T}\}$ and $\{d\,, \mathcal{T}^C\}$. The flow across a cut is just the sum of the capacities of the links that the cut cuts. It is shown in [190], that the maximum value of the flow from the source node k to the sink node d (the "max-flow") equals the minimum capacity among

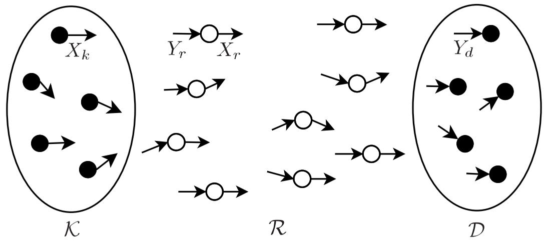

Fig. B.4: *General communication network.*

all $\mathcal{T} \subseteq \mathcal{R}$ cuts (the "min-cut"), often referred to as the max-flow min-cut theorem.

A general upper bound on the capacity was given by M.R. Aref [161, Th. 3.4]. We assume that all intended messages of the source node are independently and uniformly drawn from the alphabet space $\mathcal{W}_k = \{0, 1, \ldots, 2^{mR_k} - 1\}$. Suppose each intermediate node $r \in \mathcal{R}$ observes Y_r and transmits X_r with an arbitrary coding scheme. The theorem established by M.R. Aref can be stated as follows:

Theorem 42 *For the general discrete memoryless network described by* $\left(\mathcal{X}_k \underset{r \in \mathcal{R}}{\times} \mathcal{X}_r \, , \, p\left(y_d \, \boldsymbol{y}_{\mathcal{R}} | x_k \, \boldsymbol{x}_{\mathcal{R}}\right) \, , \, \mathcal{Y}_d \underset{r \in \mathcal{R}}{\times} \mathcal{Y}_r \right)$ *and satisfying the network input (power) constraints by all nodes, the maximum capacity between source k and destination d is bounded by:*

$$R_k \leq \sup_{p(y_d \boldsymbol{y}_{\mathcal{R}} | x_k \boldsymbol{x}_{\mathcal{R}})} \min_{\mathcal{T} \subseteq \mathcal{R}} \left\{ \mathrm{I}\left(X_k \boldsymbol{X}_{\mathcal{T}} \, ; \, Y_d \boldsymbol{Y}_{\mathcal{T}^C} | \boldsymbol{X}_{\mathcal{T}^C} \right) \right\} \tag{B.30}$$

The goal to find a subset $\mathcal{T}$ among 2^R different subsets of $\mathcal{R}$, such that the rate of information flow from the transmitters $\{k \, , \, \mathcal{T}\}$ to the receivers $\{d \, , \, \mathcal{T}^C\}$ is the minimum. The supremum is achieved under a given probability (mass) function (pmf).

Theorem 42 can be easily extended to a general case where there are several transmitters represented by $\mathcal{K} = [1, \ldots, k, \ldots, K]$, and several destinations denoted by $\mathcal{D} = [1, \ldots, d, \ldots, D]$. Each node has its own input constraint.

In the general network scenario of Fig. B.4, we assume that all intended messages w_{kd} from source k to the receiver d, which are going to be transmitted over $\mathcal{R}$ network uses, are uniformly derived from the respective set $\mathcal{W}_{kd} = \left\{0, 1, \ldots, 2^{mR_{(kd)}} - 1\right\}$, where $R_{(kd)}$ represents the rate at which transmitter k transmits symbols to destination d. A cut separates the network in a way that one half contains all sources $\mathcal{K}$ and the other half all destinations $\mathcal{D}$. Again, there exist 2^R different cuts. Applying Eq. B.30 it is straightforward to introduce the *general cutset upper bound theorem* as [16, Sec. 15.10]:

Theorem 43 *Suppose* $\left\{R_{(kd)}\right\}$ *be achievable data rates in a general discrete memoryless network* $\left(\underset{k\in\mathcal{K}}{\times}\mathcal{X}_k \underset{r\in\mathcal{R}}{\times}\mathcal{X}_r\, , \, p\left(\boldsymbol{y}_{\mathcal{D}}\,\boldsymbol{y}_{\mathcal{R}}|\boldsymbol{x}_{\mathcal{K}}\,\boldsymbol{x}_{\mathcal{R}}\right), \underset{d\in\mathcal{D}}{\times}\mathcal{Y}_d \underset{r\in\mathcal{R}}{\times}\mathcal{Y}_r\right)$ *and satisfying the network input constraints by all nodes. The sum of the data rates between sources and destinations is a bounded above by:*

$$\sum_{k\in\mathcal{K}}\sum_{d\in\mathcal{D}} R_{(kd)} \leq \sup_{p(\boldsymbol{y}_{\mathcal{D}}\boldsymbol{y}_{\mathcal{R}}|\boldsymbol{x}_{\mathcal{K}}\boldsymbol{x}_{\mathcal{R}})} \min_{\mathcal{T}\subseteq\mathcal{R}} \left\{\mathrm{I}\left(\boldsymbol{X}_{\mathcal{K}}\boldsymbol{X}_{\mathcal{T}}\,;\,\boldsymbol{Y}_{\mathcal{D}}\boldsymbol{Y}_{\mathcal{T}^C}|\boldsymbol{X}_{\mathcal{T}^C}\right)\right\} \tag{B.31}$$

where maximization is subject to the power constraints defined by the network.

B.2 Wyner-Ziv source coding

Here we briefly review the general set-up and some results about Wyner-Ziv (WZ) [191] coding. Wyner-Ziv coding is a special case of Slepian-Wolf (SW) coding theorem [16, Ch. 15.4], which is sometimes called "binning". In general, a Wyner-Ziv coding scheme is obtained by adding a quantizer and a de-quantizer to the Slepian-Wolf coding scheme.

A rate $R_1 > \mathrm{H}\left(w_1\right)$ is sufficient to encode $w_1 \in \mathcal{W}_1 = \{0, 1, \ldots, 2^{mR_1} - 1\}$ with marginal distribution $p\left(x_1\right)$. What if the $\mathcal{W}_1 \times \mathcal{W}_2 \ni \left(w_1, w_2\right) \sim p\left(x_1, x_2\right)$ must be separately distinguished for a user who intends to reconstruct w_1 and w_2? Slepian-Wolf theorem is concerned with lossless source coding with side information at the decoder. The surprising result of Slepian-Wolf theorem is that the total rate $R_{SW} = \mathrm{H}\left(w_1, w_2\right)$ suffices for separate encoding and joint decoding of X_1 and X_2 like depicted in Fig. B.5. In fact, when the

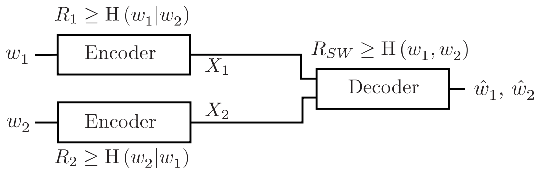

Fig. B.5: *Slepian-Wolf source coding scheme.*

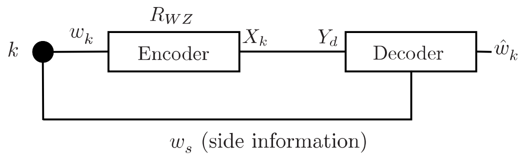

Fig. B.6: *Wyner-Ziv source coding scheme.*

side information w_2 is correctly available, the compress rate of w_1 is upper bounded to $\mathrm{H}(w_1|w_2)$.

The Wyner-Ziv theorem determines instead how many bits are needed to encode $w_k \in \mathcal{W}_k = \{0, 1, \ldots, 2^{mR_k} - 1\}$ under an average distortion constraint $\mathbb{E}\{d(w_k, \hat{w}_k)\} \leq D$ on the set $\mathcal{W}_k \times \widehat{\mathcal{W}}$, and under assumption that the side information $w_s \in \mathcal{W}_s$ is available at the decoder but not at the encoder and the decoder reconstructs only the sequence $\hat{w}_k$ as an estimation of w_k. As a mnemonic of remembering the above, we can think of w_k, w_s as being generated by the configuration in Fig. B.6. When the decoder has the knowledge of a signal correlated with the source, the latter can be encoded at a lower rate for a given distortion. The intuitive difference between Wyner-Ziv and Slepian-Wolf theorem is that encoding of w_k to X_k is lossy with a distortion measure rather than lossless. Wyner and Ziv in [191, Formula 15] give the rate distortion function for this problem as:

$$R_{WZ}(D) = \min_{p(z|x)} \min_{g} \left[\mathrm{I}(w_k\,;\,w_a) \;-\; \mathrm{I}(w_s\,;\,w_a)\right] \tag{B.32}$$

where $w_a \in \mathcal{W}_a$ is an auxiliary random variable sequence satisfying:

$$\sum_{z \in \mathcal{W}_a} p\left(x, y, z\right) = \Pr\{w_k = x, w_s = y\}\,, \qquad \text{where } p\left(x, y, z\right) = \Pr\{w_k = x, w_s = y\} \cdot p\left(z|x\right) \tag{B.33}$$

and w_s and w_a are conditionally independent given w_k and $p\left(z|x\right)$ is transition probability of a channel whose input is w_k and whose output is w_a. The minimization in (B.32) is performed over all $g : \mathcal{W}_s \times \mathcal{W}_a \longrightarrow \widehat{\mathcal{W}}$ satisfying:

$$\mathbb{E}\left\{d(w_k,\ g(w_s, w_a))\right\} \leq D \tag{B.34}$$

It means there is a general rate loss in comparison to the Slepian-Wolf problem. When D is close to zero, Wyner-Ziv degenerates to Slepian-Wolf. If w_a is an auxiliary random variable which creates the Markov chain $w_s \leftrightarrow w_k \leftrightarrow w_a$, i.e. w_s, w_a are conditionally independent given w_k, formula (B.32) can be expressed as [191, Formula 16]:

$$R_{WZ}(d) = \min_{p(z|x)}\ \min_{g}\ \mathrm{I}\left(w_k\,;\, w_a|w_s\right) \tag{B.35}$$

The distortion constraint $R_{WZ}(d) \leq D$ is achievable if and only if [188, Th. 1.10]:

$$R_{WZ}(D) \leq\ R_k\ \leq \sup_{\mathbb{E}\left\{X_k^2\right\} \leq \overline{p}_k} \mathrm{I}(X_k\,;\,Y_d) \tag{B.36}$$

So that we in back to lossless encoding, the probability of an error at the decoder is bounded to [16, Lemma 10.6.2]:

$$\Pr\left\{(w_k, \hat{w}_k) \in A_\epsilon^m\right\} \leq e^{-\left(2^{mR_k} 2^{-m\left(\mathrm{I}\left(w_k;\hat{w}_k\right)+\epsilon\right)}\right)} \tag{B.37}$$

which goes to zero as $m \to \infty$ if:

$$R_k > \mathrm{I}\left(w_k\,;\,\hat{w}_k\right) + \epsilon \tag{B.38}$$

Bibliography

[1] A. J. Goldsmith, *Wireless Communications.* Cambridge, UK: Cambridge University Press, 2005.

[2] A. R. Bahai, B. R. Saltzberg, and M. Ergen, *Multi-Carrier Digital Communications Theory and Applications of OFDM.* Boston, MA: Springer, 2nd ed., 2004.

[3] M. Doelz, E. Heald, and D. Martin, "Binary data transmission techniques for linear systems," *Proceedings of the IRE*, vol. 45, pp. 656–661, May 1957.

[4] B. Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Trans. on Commun. Technology*, vol. 15, pp. 805–811, Dec. 1967.

[5] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Systems Technical Journal*, vol. 45, pp. 177–1796, Dec. 1966.

[6] M. Zimmerman and A. Kirsch, "The AN/GSC-10 (KATHRYN) variable rate data modem for HF radio," *IEEE Trans. on Commun. Technology*, vol. 15, pp. 197–204, Apr. 1967.

[7] G. Porter, "Error distribution and diversity performance of a frequency-differential PSK HF modem," *IEEE Trans. on Commun. Technology*, vol. 16, pp. 567–575, Aug. 1968.

[8] "Orthogonal Frequency Division Multiplexing." U.S. Patent No. 3,488,4555, field Nov. 14, 1996, issued Jan. 6, 1970.

[9] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. on Commun. Technology*, vol. 19, pp. 628–634, Oct. 1971.

[10] A. Peled and A. Ruiz, "Frequency domain data transmission using reduced computational complexity algorithms," in *IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, pp. 964–967, Apr. 1980.

[11] ANSI T1.413-1998, "Network and customer installation interfaces - asymmetric digital subscriber line (ADSL) metallic interface." (American National Standards Institute 1998).

[12] C. Arasu, P. Hyanki, H. Sharma, G. Lakshminarayanan, M. H. Lee, and S.-B. Ko, "PAPR reduction for improving performance of OFDM system," in *IEEE International Conf. on Communication Control and Computing Technologies (ICCCCT)*, (Tamilnadu, India), pp. 77–82, Oct. 2010.

[13] A.-B. Salberg and A. Swami, "Doppler and frequency-offset synchronization in wideband OFDM," *IEEE Transactions on Wireless Commun.*, vol. 4, pp. 2870–2881, Nov. 2005.

[14] A. Bo, Y. Zhi-xing, P. Chang-yong, G. Jian-hua, W. Yong, and L. Zhen, "On the synchronization techniques for wireless OFDM systems," *IEEE Trans. on Broadcasting*, vol. 52, pp. 236–244, June 2006.

[15] H. Sari and G. Karam, "Orthogonal frequency-division multiple access and its application to CATV networks," *European Trans. on Telecommun.*, vol. 9, no. 6, pp. 507–516, 1998.

[16] T. M. Cover and J. A. Thomas, *Elements of information theory.* New York, NY, USA: Wiley-Interscience, second ed., 2006.

[17] P. Chow, J. Cioffi, and J. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. on Commun.*, vol. 43, pp. 773–775, Feb./Mar./Apr. 1995.

[18] J. Forney, G.D. and M. Eyuboğlu, "Combined equalization and coding using precoding," *IEEE Commun. Magazine*, vol. 29, pp. 25–34, Dec. 1991.

[19] D. Hughes-Hartogs, "Ensemble modem structure for imperfect transmission media." U.S Patents Nos. 4,679,227 (July 1988) 4,731,816 (March 1988) and 4,833,796 (May 1989).

[20] J. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Commun. Magazine*, vol. 28, pp. 5–14, May 1990.

[21] I. Kalet, "The multitone channel," *IEEE Trans. on Commun.*, vol. 37, pp. 119–124, Feb. 1989.

[22] A. Czylwik, "Adaptive OFDM for wideband radio channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 1, (San Francisco, CA), pp. 713–718, Nov. 1996.

[23] R. Fischer and J. Huber, "A new loading algorithm for discrete multitone transmission," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 1, (London, England), pp. 724–728, Nov. 1996.

[24] L. Van-der Perre, S. Thoen, P. Vandenameele, B. Gyselinckx, and M. Engels, "Adaptive loading strategy for a high speed OFDM-based WLAN," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 4, (Sydney, Australia), pp. 1936–1940, Nov. 1998.

[25] V. Ramamurthi and C. Wei-Peng, "Mobility based MIMO link adaptation in LTE-advanced cellular networks," in *Intern. Conf. on Broadband,*

Wireless Computing, Communication and Applications (BWCCA), (Fukuoka, Japan), pp. 235–241, Nov. 2010.

[26] Z. Han and K. J. R. Liu, "Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions," *IEEE Trans. on Commun.*, vol. 53, pp. 1366–1376, Aug. 2005.

[27] Z. Han, Z. Ji, and K. J. R. Liu, "Non-cooperative resource competition game by virtual referee in multi-cell OFDMA networks," *IEEE Journal on Selected Areas in Commun.*, vol. 25, Aug. 2007.

[28] D. Yu, D. Wu, Y. Cai, and W. Zhong, "Power allocation based on power efficiency in uplink OFDMA systems: A game theoretic approach," in *Proc. IEEE Intern. Conf. Commun. Systems (ICCS)*, (Singapore, Singapore), pp. 92–97, Nov. 2008.

[29] M. Moretti, N. Dimitriou, G. Danielli, V. Corvino, R. Verdone, F. Brah, L. Vanderdorpe, and V. Tralli, "Channel and interference based cross layer OFDMA scheduling," in *Newcom++/ACORN workshop*, (Barcelona, Spain), pp. 1–5, Mar. 2009.

[30] Y. J. Zhang and K. B. Letaief, "An efficient resource-allocation scheme for spatial multiuser access in MIMO/OFDM systems," *IEEE Trans. on Communications*, vol. 53, pp. 107–116, Jan. 2005.

[31] L. Yang, Z. He, L. Wang, and W. Wu, "A new utility based subcarrier and power joint allocation scheme in uplink OFDMA systems," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, (Dublin, Ireland), pp. 2756–2760, Apr. 2007.

[32] L. J. Cimini, B. Daneshrad, and N. Sollenberger, "Clustered OFDM with transmitter diversity and coding," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 1, (San Francisco, CA), pp. 703–707, Nov. 1996.

[33] M. Johansson, "Benefits of multiuser diversity with limited feedback," in *IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, (Rome, Italy), pp. 155–159, June 2003.

[34] D. Gesbert and M. S. Alouini, "Selective multi-user diversity," in *Proc. IEEE Intern. Symp. on Signal Processing and Information Technology (ISSPIT)*, (Darmstadt, Germany), pp. 162–165, Dec. 2003.

[35] P. Svedman, S. Wilson, J. Cimini, L.J., and B. Ottersten, "A simplified opportunistic feedback and scheduling scheme for OFDM," in *IEEE Vehicular Technology Conference (VTC)*, vol. 4, (Milan, Italy), pp. 1878–1882, May 2004.

[36] M. Pischella and J. Belfiore, "Distributed margin adaptive resource allocation in MIMO OFDMA networks," *IEEE Trans. on Commun.*, vol. 58, pp. 2371–2380, Aug. 2010.

[37] C. Tan, T. Chuah, and S. Tan, "Adaptive multicast scheme for OFDMA based multicast wireless systems," *Electronics Letters, 28th*, vol. 47, pp. 570–572, Apr. 2011.

[38] H. Tabassum, Z. Dawy, and M. Alouini, "Sum rate maximization in the uplink of multi-cell OFDMA networks," in *Intern. Wireless Commun. and Mobile Computing Conf. (IWCMC)*, (Istanbul, Turkey), pp. 1152–1157, July 2011.

[39] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Trans. on Wireless Commun.*, vol. 2, pp. 1150–1158, Nov. 2003.

[40] X. Wang, M. Tao, J. Mo, and Y. Xu, "Power and subcarrier allocation for physical-layer security in OFDMA-based broadband wireless networks," *IEEE Trans. on Information Forensics and Security*, vol. 6, pp. 693–702, Sept. 2011.

[41] R. Cheng and S. Verdù, "Gaussian multiaccess channels with ISI: capacity region and multiuser water-filling," *IEEE Trans. on Information Theory*, vol. 39, pp. 773–785, May 1993.

[42] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. on Information Theory*, vol. 43, pp. 1986–1992, Nov. 1997.

[43] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proc. IEEE Veh. Technol. Conf. (VTC)-Spring*, (Tokyo, Japan), pp. 1085–1089, May 2000.

[44] G. Debreu, *Theory of Value: An Axiomatic Analysis of Economic Equilibrium.* John Wiley and Sons Inc., New York, 1959.

[45] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Commun.*, vol. 17, pp. 1747–1758, Oct. 1999.

[46] W. Yu and J. Cioffi, "FDMA capacity of the Gaussian multiple access channel with ISI," in *IEEE Intern. Conf. on Commun. (ICC)*, vol. 3, (New Orleans, Louisiana), pp. 1365–1369, 2000.

[47] D. Bertsekas and R. Gallager, *Data Networks.* Prentice Hall, second ed., 1992. ISBN: 0-13-200916-1.

[48] F. Kelly, "Charging and rate control for elastic traffic," *European Trans. on Telecommun.*, vol. 8, pp. 33–37, 1997.

[49] P. Tsiaflakis, *Resource management and optimization in multi-user DSL systems.* PhD thesis, Department of Electrical Eng., Katholieke Universiteit Leuven, Leuven, Belgium, 2009.

[50] I. Toufik and R. Knopp, "Channel allocation algorithms for multi-carrier systems," in *IEEE Vehicular Technology Conf. (VTC)*, vol. 2, (Los Angeles, CA), pp. 1129–1133, Sept. 2004.

[51] Kelly F.P. and Maulloo A.K. and Tan D.K.H., "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, 1998.

[52] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge, UK: Cambridge University Press, first ed., 2005.

[53] Z. Shen, J. Andrews, and B. Evans, "Optimal power allocation in multiuser OFDM systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 1, (San Francisco, CA), pp. 337–341, Dec. 2003.

[54] I. C. Wong, Z. Shen, B. L. Evans, and J. G. Andrews, "A low complexity algorithm for proportional resource allocation in OFDMA systems," in *IEEE Workshop on Signal Processing Systems (SIPS)*, (Austin, Texas), pp. 1–6, Oct. 2004.

[55] Z. Shen, J. Andrews, and B. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. on Wireless Commun.*, vol. 4, pp. 2726–2737, Nov. 2005.

[56] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. on Networking*, vol. 8, pp. 556–567, Oct. 2000.

[57] J.-W. Lee, R. Mazumdar, and N. Shroff, "Nonconvexity issues for Internet rate control with multiclass services: stability and optimality," in *IEEE Computer and Commun. Societies Conf. (INFOCOM)*, vol. 1, (Hong Kong), p. 4, Mar. 2004.

[58] S. Shenker, "Fundamental design issues for the future Internet," *IEEE Journal on Selected Areas in Commun.*, vol. 13, pp. 1176–1188, Sept. 1995.

[59] L. Badia, S. Merlin, and M. Zorzi, "Resource management in IEEE 802.11 multiple access networks with price-based service provisioning," *IEEE Trans. on Wireless Commun.*, vol. 7, pp. 4331–4340, Nov. 2008.

[60] Z. Cao and E. Zegura, “Utility max-min: an application-oriented bandwidth allocation scheme,” in *Proc. IEEE Computer and Commun. Societies Conf. (INFOCOM)*, vol. 2, (New York, USA), pp. 793–801, Mar. 1999.

[61] J.-W. Cho and S. Chong, “Utility max-min flow control using slope-restricted utility functions,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 2, (St. Louis, MO), pp. 6 pp. –824, Dec. 2005.

[62] A. Dewey and E. Jury, “A stability inequality for a class of nonlinear feedback systems,” *IEEE Trans. on Automatic Control*, vol. 11, pp. 54–62, Jan. 1966.

[63] C. Zhou, G. Wunder, and T. Michel, “Utility maximization for OFDMA systems over discrete sets,” in *IEEE Intern. Conf. on Commun. (ICC)*, (Glasgow, Scotland), pp. 5726–5731, June 2007.

[64] S. Kim and J.-W. Lee, “Joint resource allocation for uplink and downlink in wireless networks: A case study with user-level utility functions,” in *IEEE Vehicular Technology Conf. (VTC)*, (Barcelona, Spain), Apr. 2009.

[65] M. Chiang, *Geometric Programming for Communication Systems.* Hanover, MA: now Publishers Inc., first ed., 2005. ISBN: 1-933019-09-3.

[66] A. W. Marshall and I. Olkin, *Inequalities: theory of majorization and its applications.* New York: Academic Press, 1979.

[67] D. Y. Zhangyu Guan and H. Zhang, “Optimal and fair resource allocation for multiuser wireless multimedia transmissions,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 17, pp. 1–10, 2009.

[68] A. Tölli, *Resource management in cooperative MIMO-OFDM cellular systems*. PhD thesis, Department of Electrical and Information Eng., University of Oulu, Oulu, Finland, 2008.

[69] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends Networking*, vol. 1, pp. 1–144, Apr. 2006.

[70] I. Koutsopoulos and L. Tassiulas, "Cross-layer adaptive techniques for throughput enhancement in wireless OFDM-based networks," *IEEE/ACM Trans. Networking*, vol. 14, pp. 1056–1066, Oct. 2006.

[71] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks-part I: theoretical framework," *IEEE Trans. on Wireless Commun.*, vol. 4, pp. 614–624, Mar. 2005.

[72] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks-part II: algorithm development," *IEEE Trans. on Wireless Commun.*, vol. 4, pp. 625–634, Mar. 2005.

[73] M. Morelli, M. Moretti, G. Imbarlina, and N. Dimitriou, "Low complexity SNR estimation for transmissions over time-varying flat-fading channels," in *IEEE Wireless Communications and Networking Conf. (WCNC)*, (Budapest, Hungary), pp. 1–4, Apr. 2009.

[74] C. M. Yen, C. J. Chang, and L. C. Wang, "A utility-based TMCR scheduling scheme for downlink multiuser MIMO-OFDMA systems," *IEEE Trans. on Vehicular Technology,*, vol. 59, pp. 4105–4115, Oct. 2010.

[75] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.

[76] B. Peleg and P. Sudhölter, *Introduction to the Theory of Cooperative Games*. Berlin, Germany: Springer-Verlag, 2nd ed., 2007.

[77] D. Wu, D. Yu, and Y. Cai, "Subcarrier and power allocation in uplink OFDMA systems based on game theory," in *Intern. Conf. on Neural Networks and Signal Processing*, (Nanjing , China), pp. 522–526, June 2008.

[78] V. Shah, N. Mandayam, and D. Goodman, "Power control for wireless data based on utility and pricing," in *Proc. IEEE Intern. Symp. Personal, Indoor, and Mobile Radio Commun.*, vol. 3, (Boston, MA), pp. 1427–1432, Sept. 1998.

[79] H. Kwon and B. G. Lee, "Distributed resource allocation through noncooperative game approach in multi-cell OFDMA systems," in *IEEE Intern. Conf. on Commun. (ICC)*, vol. 9, (Istanbul, Turkey), pp. 4345–4350, June 2006.

[80] A. Zappone, G. Alfano, S. Buzzi, and M. Meo, "Energy-efficient noncooperative resource allocation in multi-cell OFDMA systems with multiple base station antennas," in *IEEE Online Conf. on Green Commun. (GreenCom)*, pp. 82–87, Sept. 2011.

[81] G. Bacci, A. Bulzomato, and M. Luise, "Uplink power control and subcarrier assignment for an OFDMA multicellular network based on game theory," in *Proc. Int. Conf. on Performance Evaluation Methodologies and Tools (ValueTools)*, (Paris, France), May 2011.

[82] D. Monderer and L. S. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[83] R. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems.* Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2009.

[84] K.-D. Lee and V. Leung, "Fair allocation of subcarrier and power in an OFDMA wireless Mesh network," *IEEE Journal on Selected Areas in Commun.*, vol. 24, pp. 2051–2060, Nov. 2006.

[85] H. Raiffa, *Contributions to the Theory of Game II.* Princeton, NJ: Princeton University Press, 1953.

[86] E. Kalai and M. Smorodinsky, "Other solutions to Nash's bargaining problem," *Econometrica*, vol. 4, pp. 513–518, 1975.

[87] A. E. Roth, "An impossibility result concerning n-person bargaining games," *International Journal of Game Theory*, vol. 8, pp. 129–132, 1979.

[88] T. K. Chee, C.-C. Lim, and J. Choii, "A cooperative game theoretic framework for resource allocation in OFDMA systems," in *Proc. IEEE Intern. Conf. Commun. Systems (ICCS)*, (Singapore, Singapore), Oct. 2006.

[89] W. Noh, "A distributed resource control for fairness in OFDMA systems: English-auction game with imperfect information," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (New Orleans, LA), pp. 1–6, Dec. 2008.

[90] K. Yang, N. Prasad, and X. Wang, "An auction approach to resource allocation in uplink OFDMA systems," *IEEE Trans. on Signal Processing*, vol. 57, pp. 4482–4496, Nov. 2009.

[91] M. A. McHenry and D. McCloskeyk, "Spectrum occupancy measurements Chicago, Illinois," Nov. 2005.

[92] C. Hyun-Ho, L. Jong Bu, H. Hyosun, and J. Kyunghun, "Optimal handover decision algorithm for throughput enhancement in cooperative cellular networks," in *IEEE Vehicular Technology Conf. (VTC)*, (Ottawa, Canada), pp. 1–5, Sep. 2010.

[93] S. Hew and L. White, "Cooperative resource allocation games in shared networks: symmetric and asymmetric fair bargaining models," *IEEE Trans. on Wireless Commun.*, vol. 7, pp. 4166–4175, Nov. 2008.

[94] F. Shams, G. Bacci, and M. Luise, "An OFDMA resource allocation algorithm based on coalitional games," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 1, p. 46, 2011.

[95] H. Kwon and G. B. Lee, "Cooperative power allocation for broadcast/multicast services in cellular OFDM systems," *IEEE Trans. on Commun.*, vol. 57, pp. 3092–3102, Oct. 2009.

[96] E. Zeydan, D. Kivanc, U. Tureli, and C. Comaniciu, "Joint iterative beamforming and power adaptation for MIMO ad hoc networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 1, p. 79, 2011.

[97] D. Li, Y. Xu, X. Wang, and M. Guizani, "Coalitional game theoretic approach for secondary spectrum access in cooperative cognitive radio networks," *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 844–856, Mar. 2011.

[98] Z. Khan, S. Glisic, L. DaSilva, and J. Lehtomändki, "Modeling the dynamics of coalition formation games for cooperative spectrum sharing in an interference channel," *IEEE Trans. on Computational Intelligence and AI in Games*, vol. 3, pp. 17–30, Mar. 2011.

[99] I. Stanojev, O. Simeone, U. Spagnolini, Y. Bar-Ness, and R. Pickholtz, "Cooperative ARQ via auction-based spectrum leasing," *IEEE Trans. on Commun.*, vol. 58, pp. 1843–1856, June 2010.

[100] F. Javadi, M. Kibria, and A. Jamalipour, "Bilateral shapley value based cooperative gateway selection in congested wireless mesh networks," in *IEEE Global Telecommun. Conf. (GLOBECOM)*, (New Orleans, LA), pp. 1–5, Dec. 2008.

[101] J. Huang, Z. Han, M. Chiang, and H. Poor, "Auction-based resource allocation for multi-relay asynchronous cooperative networks," in *IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, (Las Vegas, Nev.), pp. 5356–5359, Apr. 2008.

[102] H. Deng, Y. Wang, and J. Lu, "Auction based resource allocation for balancing efficiency and fairness in OFDMA relay networks with service differentiation," in *IEEE Vehicular Technology Conf. (VTC)*, (Ottawa, Canada), pp. 1–5, Sept. 2010.

[103] E. Hossain, D. I. Kim, and V. K. Bhargava, *Cooperative Cellular Wireless Networks.* New York, NY, USA: Cambridge University Press, 2011.

[104] L. S. Shapley, "Cores of convex games," *International Journal of Game Theory*, vol. 1, pp. 11–26, 1971.

[105] A. ParandehGheibi, A. Eryilmaz, A. Ozdaglar, and M. Medard, "Resource allocation in multiple access channels," in *Conf. on Signals, Systems and Computers (ACSSC)*, (Pacific Grove, CA), pp. 878–882, Nov. 2007.

[106] D. B. Gillies, *Some theorems on n-person games.* PhD thesis: Department of Mathematics, Princeton University, N.J., 1953.

[107] M. M. Madiman, "Cores of cooperative games in information theory," *EURASIP Journal on Wireless Commun. and Networking*, vol. abs/0901.0062, 2009.

[108] D. Li, Y. Xu, J. Liu, and X. Wang, "Relay assignment and cooperation maintenance in wireless networks," in *IEEE Wireless Commun. and Networking Conf. (WCNC)*, (Sydney, Australia), pp. 1–5, Apr. 2010.

[109] D. Niyato and E. Hossain, "A cooperative game framework for bandwidth allocation in 4G heterogeneous wireless networks," in *IEEE Intern. Conf. on Commun. (ICC)*, vol. 9, (Istanbul, Turkey), pp. 4357–4362, June 2006.

[110] T. W. Sandholm and V. R. Lesser, "Coalitions among computationally bounded agents," *Artificial Intelligence*, vol. 94, pp. 99–137, 1997.

[111] O. Bondareva, "Some applications of the methods of linear programming to the theory of cooperative games," *Problemy Kibernetiki*, vol. 10, pp. 119–139, 1963, (Russian).

[112] L. S. Shapley, "On balanced sets and cores," *Naval Research Logistics Quarterly*, vol. 14, no. 4, pp. 453–560, 1967.

[113] A. E. Roth, *The Shapley Value. Essays in Honor of Lloyd S. Shapley.* Cambridge University Press, 1988.

[114] L. S. Shapley, "A value for n-person games. Contribution to the theory of games," *Annals of Mathematics Studies*, vol. 2, p. 28, 1953.

[115] S. Hart, "A comparison of non-transferable utility values," discussion paper series, Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem, 2003.

[116] G.-J. Otten and H. J. Peters, "The Shapley transfer procedure for NTU-games," open access publications from maastricht university, Maastricht University, 2002.

[117] H. P. Young, "Monotonic solutions of cooperative games," *International Journal of Game Theory*, vol. 14, pp. 65–72, 1985.

[118] S. Kim, "Cooperative game theoretic online routing scheme for wireless network managements," *Commun., IET*, vol. 4, pp. 2074–2083, June 2010.

[119] M. Khouzani and S. Sarkar, "Economy of spectrum access in time varying multichannel networks," *IEEE Trans. on Mobile Computing*, vol. 9, pp. 1361–1376, Oct. 2010.

[120] H. Park and M. Van-Der Schaar, "Coalition-based resource negotiation for multimedia applications in informationally decentralized networks," *IEEE Trans. on Multimedia*, vol. 11, pp. 765–779, June 2009.

[121] S. Pechersky, “On proportional excess for NTU games,” Tech. Rep. Ec-02/01, European University at St. Petersburg, Department of Economics, 2001.

[122] T. S. H. Driessen, “A note on the inclusion of the kernel in the core of the bilateral assignment game,” *International Journal of Game Theory*, vol. 27, no. 2, pp. 301–303, 1998.

[123] Z. Han and H. V. Poor, “Coalition games with cooperative transmission: a cure for the curse of boundary nodes in selfish packet-forwarding wireless networks,” *IEEE Trans. on Commun.*, vol. 57, pp. 203–213, Jan. 2009.

[124] M. Maschler, “The bargaining set, kernel, and nucleolus,” in *Handbook of Game Theory with Economic Applications* (R. Aumann and S. Hart, eds.), vol. 1, ch. 18, pp. 591–667, Elsevier, 1 ed., 1992.

[125] G. Demange and M. Wooders, *Group Formation in Economics: Networks, Clubs, and Coalitions.* UK: Cambridge University Press, 2005.

[126] R. J. Aumann, “Acceptable points in general cooperative n-person games,” in *Annals of Mathematics Studies, 40, in Contributions to the Theory of Games*, (Princeton, NJ), pp. 287–324, Princeton University Press, 1959.

[127] B. D. Bernheim, B. Peleg, and M. D. Whinston, “Coalition-proof Nash equilibria I. concepts,” *Journal of Economic Theory*, vol. 42, no. 1, pp. 1–12, 1987.

[128] M. Félegyházi and J. P. Hubaux, “Game theory in wireless networks: A tutorial,” *ACM computing surveys*, 2006. Technical Report: LCA-REPORT-2006-002, EPFL.

[129] M. Jackson and V. D. Nouweland, “Strongly stable networks,” *Games and Economic Behavior*, vol. 51, no. 2, pp. 420–444, 2005. Special Issue in Honor of Richard D. McKelvey.

[130] S. Zhong and F. Wu, "A collusion-resistant routing scheme for noncooperative wireless ad hoc networks," *IEEE/ACM Trans. on Networking*, vol. 18, pp. 582–595, Apr. 2010.

[131] E. Altman, T. Basar, I. Menache, and H. Tembine, "A dynamic random access game with energy constraints," in *Intern. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, (Seoul, Korea), pp. 1–7, June 2009.

[132] B. Moldovanu, "Coalition-proof Nash equilibria and the core in three-player games," *Games and Economic Behavior*, vol. 4, pp. 565–581, Oct. 1992.

[133] H. Konishi, M. Le Breton, and S. Weber, "Equivalence of strong and coalition-proof Nash equilibria in games without spillovers," *Economic Theory*, vol. 9, pp. 97–113, 1997.

[134] M. Félegyházi, M. Cagalj, S. Bidokhti, and J.-P. Hubaux, "Non-cooperative multi-radio channel allocation in wireless networks," in *Proc. IEEE Computer and Commun. Societies Conf. (INFOCOM)*, pp. 1442–1450, May 2007.

[135] L. Gao, X. Wang, and Y. Xu, "Multiradio channel allocation in multihop wireless networks," *IEEE Trans. on Mobile Computing*, vol. 8, pp. 1454–1468, Nov. 2009.

[136] IEEE 802.16 Broadband Wireless Access Working Group, "IEEE 802.16m Evaluation Methodology Document (EMD)," *Tech. Rep. IEEE 802.16m-08/004r5*, Jan. 2009.

[137] R. J. Aumann, "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, vol. 1, no. 1, pp. 67–96, 1974.

[138] Y. Heller, *Correlated Equilibrium and Seemingly-Iterational Behavior.* PhD thesis, Faculty of exact sciences, Tel-Aviv University, Israel, 2011.

[139] G. Bacci, M. Luise, and H. V. Poor, "Game theory and power control in ultrawideband networks," *Physical Commun.*, vol. 1, no. 1, pp. 21–39, 2008.

[140] N. Bonneau, E. Altman, and M. Debbah, "Correlated equilibrium in access control for wireless communications," in *Intern. Conf. on Networking (IFIP)*, (Coimbra, Portugal), pp. 15–19, May 2006.

[141] N. Bonneau, M. Debbah, E. Altman, and A. Hjorungnes, "Non-Atomic games for multi-user systems," *IEEE Journal on Selected Areas in Commun.*, vol. 26, pp. 1047–1058, Sept. 2008.

[142] M. Charafeddine, Z. Han, A. Paulraj, and J. Cioffi, "Crystallized rates region of the interference channel via correlated equilibrium with interference as noise," in *IEEE Intern. Conf. on Commun. (ICC)*, (Dresden, Germany), pp. 1–6, June 2009.

[143] R. Dilts and T. A. Epstein, *Dynamic Learning.* Capitola, CA: MeTa Publications, 1995.

[144] M. Van-Der Schaar and F. Fu, "Spectrum access games and strategic learning in cognitive radio networks for delay-critical applications," *Proceedings of the IEEE*, vol. 97, pp. 720–740, Apr. 2009.

[145] H.-P. Shiang and M. Van-Der Schaar, "Online learning in autonomic multi-hop wireless networks for transmitting mission-critical applications," *IEEE Journal on Selected Areas in Commun.*, vol. 28, pp. 728–741, June 2010.

[146] Z. Lin and M. van der Schaar, "Autonomic and distributed joint routing and power control for delay-sensitive applications in multi-hop wireless networks," *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 102–113, Jan. 2011.

[147] C. J. C. H. Watkins and P. Dayan, "Q-Learning," *Machine Learning*, vol. 8, no. 3, pp. 279–292, 1992.

[148] C. Claus and C. Boutilier, "The dynamics of reinforcement learning in cooperative multiagent systems," in *National Conf. on Artificial Intelligence (AAAI-98)*, (Madison, Wisconsin), pp. 746–752, July 1998.

[149] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, May 1996.

[150] E. Biglieri, J. Proakis, and S. S. (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. on Information Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[151] M. J. Osborne and A. Rubinstein, *Bargaining and Markets*. San Diego, CA: Academic Press, 1990.

[152] M. Agastya, "Perturbed adaptive dynamics in coalition form games," *J. Economic Theory*, vol. 89, pp. 207–233, Dec. 1999.

[153] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Berlin, Germany: Springer-Verlag, 2nd ed., 1976.

[154] G. Dainelli, M. Moretti, A. Zalonis, and N. Dimitriou, "Distributed subcarrier allocation schemes for OFDMA multiple access systems," in *Proc. Newcom++/Cost2100 Workshop*, (Vienna, Austria), Sept. 2009.

[155] T. Wang and L. Vandendorpe, "Resource allocation for maximizing weighted sum min-rate in downlink cellular OFDMA systems," in *Proc. IEEE Int. Conf. Commun.*, (Cape Town, South Africa), May 2010.

[156] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, vol. 42, pp. 74–80, Oct. 2004.

[157] E. C. Van-der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.

[158] H. Sato, "Information transmission through a channel with relay. the ALOHA system," in *Technical Rep.*, (University of Hawaii, Honolulu), 1976.

[159] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. on Information Theory*, vol. 25, pp. 572–584, Sep. 1979.

[160] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. on Information Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[161] M. R. Aref, *Information Flow in Relay Networks.* PhD thesis, Department of Electrical Eng., Stanford University, 1980.

[162] A. A. El Gamal and M. R. Aref, "The capacity of the semi-deterministic relay channel," *IEEE Trans. on Information Theory*, vol. 28, p. 536, May 1982.

[163] Z. Zhang, "Partial converse for a relay channel," *IEEE Trans. on Information Theory*, vol. 34, pp. 1106–1110, Sept. 1988.

[164] M. Gastpar, G. Kramer, and P. Gupta, "The multiple-relay channel: coding and antenna-clustering capacity," in *Proc. IEEE Intern. Symposium on Information Theory*, (Lausanne, Switzerland), p. 136, June/July 2002.

[165] S. P. Borade, "Network information flow: Limits and achievability," in *Proc. IEEE Intern. Symposium on Information Theory*, (Lausanne, Switzerland), p. 139, June-July 2002.

[166] S. Zahedi, *On Reliable Communication Over Relay Channel.* PhD thesis, Department of Electrical Eng., Stanford University, 2005.

[167] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Proc. IEEE Intern. Symposium on Information Theory*, (Sorrento, Italy), p. 22, July 2000.

[168] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: An achievable rate region," *IEEE Trans. on Information Theory*, vol. 49, pp. 1877–1894, Aug. 2003.

[169] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. IEEE Intern. Conf. on Computer Commun. (INFOCOM)*, pp. 1577–1586, Nov. 2002.

[170] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. on Information Theory*, vol. 46, pp. 388–404, Mar. 2000.

[171] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part i. system description," *IEEE Trans. on Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[172] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity - part ii: Implementation aspects and performance analysis," *IEEE Trans. on Commun.*, vol. 51, pp. 1939–1948, Nov. 2003.

[173] A. Scaglione and Y. W. Hong, "Opportunistic large arrays: Cooperative transmission in wireless multihop ad hoc networks to reach far distances," *IEEE Trans. on Signal Processing*, vol. 51, pp. 2082–2092, Aug. 2003.

[174] H. Suraweera, D. Michalopoulos, and G. Karagiannidis, "Performance of distributed diversity systems with a single amplify-and-forward relay," *IEEE Trans. on Vehicular Technology*, vol. 58, pp. 2603–2608, June 2009.

[175] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and V. H. Poor, *MIMO Wireless Communications.* Cambridge University Press, Jan. 2007.

[176] S. Simoens, *Cooperative MIMO Communications - Information Theoretical Limits and Practical Coding Strategies.* PhD thesis, Dept. of Signal Theory and Communications, UPC, Barcelona, Spain, 2009 July.

[177] H. Shan, W. Zhuang, and H. T. Cheng, "Cross-layer protocol design for distributed wireless networks with novel relay selection," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, (Miami, FL), pp. 1–6, Dec. 2010.

[178] B. Zhao and M. Valenti, "Practical relay networks: A generalization of hybrid-ARQ," *IEEE Journal on Selected Areas in Commun.*, vol. 23, pp. 7–18, Jan. 2005.

[179] M. Zorzi and R. Rao, "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance," *IEEE Trans. on Mobile Computing*, vol. 2, pp. 337–348, Oct.-Dec. 2003.

[180] J. Huang, Z. Han, M. Chiang, and H. Poor, "Auction-based distributed resource allocation for cooperation transmission in wireless networks," in *IEEE Global Telecommunications Conf. (GLOBECOM)*, (Washoington DC), pp. 4807–4812, Nov. 2007.

[181] J. Huang, Z. Han, M. Chiang, and H. Poor, "Auction-based resource allocation for multi-relay asynchronous cooperative networks," in *IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, (Las Vegas, Nev), pp. 5356–5359, Apr. 2008.

[182] B. Wang, Z. Han, and K. Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using stackelberg game," *IEEE Trans. on Mobile Computing*, vol. 8, pp. 975–990, July 2009.

[183] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. on Information Theory*, vol. 51, pp. 3037–3063, Sept. 2005.

[184] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking.* New York, NY: Cambridge University Press, 2009.

[185] L. L. Xie and P. R. Kumar, “A network information theory for wireless communication: Scaling laws and optimal operation,” *IEEE Trans. on Information Theory*, vol. 50, pp. 748–767, 2002.

[186] F. M. J. Willems, *Information theoretical Results for the Discrete Memoryless Multiple Access Channel.* PhD thesis, Department of Electrical Eng., Katholieke Universiteit Leuven, Leuven, Belgium, 1982.

[187] A. del Coso Sánchez, *Achievable Rates for Gaussian Channles with Multiple Relays.* PhD thesis, Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC), Barcelona, 2008.

[188] M. Gastpar, *To Code or Not To Code.* PhD thesis, Faculty of Information and Communications Eng., École polytechnique fédérale de Lausanne, Switzerland, 2002.

[189] C. E. Shannon, “A mathematical theory of communication,” *Bell system technical journal*, vol. 27, 1948.

[190] L. R. Ford and D. R. Fulkerson, *Flows in Networks.* Princeton, NJ: Princeton University Press, 1962.

[191] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. on Information Theory*, vol. 22, pp. 1–10, Jan. 1976.