ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

MANUSCRIPT-BASED THESIS PRESENTED TO
ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
Ph.D.

BY
Fanny PARZYSZ

A JOINT ANALYSIS OF CODING SCHEMES AND RELAY DEPLOYMENT FOR ENERGY EFFICIENT CELLULAR NETWORKS

MONTREAL, JULY 3, 2015

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

Mr. François Gagnon, Thesis Supervisor
Electrical Engineering Department, École de technologie supérieure

Mrs. Mai Vu, Thesis Co-supervisor
Electrical and Computer Engineering Department, Tufts University

Mr. Christian Desrosiers, President of the Board of Examiners
Software and IT Engineering Department, École de Technologie Supérieure

Mr. Charles Despins, Member of the jury
Electrical Engineering Department, École de technologie supérieure

Mr. Benoît Champagne, External Evaluator
Electrical and Computer Engineering Department, McGill University

THIS THESIS WAS PRESENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND PUBLIC

ON JUNE 4, 2015

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my research supervisors, Professor François Gagnon and Professor Mai Vu, for their patient guidance and useful advice which helped me take this scientific adventure further than I could imagine. This thesis would not have been possible without their valuable support and enthusiastic encouragement over these past five years.

I thank Ultra Electronics TCS and the Natural Science and Engineering Council of Canada for funding this project through the High Performance Emergency and Tactical Wireless Communication Chair at École de Technologie Supérieure. I also thank Chantal Cyr for her constant support through all the administrative procedures.

It would be difficult to name all the people who contributed to this PhD and the uncountable generations of labmates who opened my mind to new research topics and who filled these years with lot of fun. My special acknowledgments go to Safa Saadaoui, for her unreserved support at my arrival in the lab, to my Jedi Master Mehdi Si Moussa, for his precious advice about the IEEE world (and for our vital chocolate dealing), to Normand Gravel, our magical technician, to Mouna Hajir, without whom I would not be packing my luggage for Barcelona, and to Mathieu Boutin, who still doesn't know what he would do if he had children.

But also, I thank my great friends Géraldine and Barbara, who have always been here to cheer me up and with whom I've shared so much, and the little family of the lab: Minh, Zeynab, Parisa, Edena, Mushiar, Pascal, Guy, Mous, Christelle, Adrien, Byron... E não posso esquecer super Nolmar que me suportou mesmo de tão longe, ajudando-me com matematica e Latex, e apoiando-me em momentos difíceis. Enfin, cette page serait pas complète sans remercier toute ma famille, qui a toujours été à mes côtés.

A JOINT ANALYSIS OF CODING SCHEMES AND RELAY DEPLOYMENT FOR ENERGY EFFICIENT CELLULAR NETWORKS

Fanny PARZYSZ

ABSTRACT

To satisfy consumer demand for high rate ubiquitous access, research on next-generation cellular networks mostly focused on capacity and coverage. This focus and the resulting advanced technologies, however, bring along a staggering increase in the energy consumption, creating a pressing shift to energy-efficient designs. Within this perspective, relaying has been considered as a promising and cost-efficient technique, and algorithms for optimal relay deployment have attracted significant attention. To date, analysis for relaying energy consumption is mostly based on suboptimal coding schemes or schemes optimized for rate. In this manuscript-based thesis, we investigate the various advantages provided by energy-optimized coding schemes, notably in terms of efficient relay deployment.

In the first part, we characterize the minimal energy consumption that can be expected from decode-forward relaying. To this end, we design a novel practical half-duplex relaying coding scheme and propose several optimal power allocations to minimize its energy consumption. From a theoretical aspect, an important result is that minimizing the network energy consumption is not equivalent to maximizing the network capacity as often believed. Furthermore, such energy efficiency approach has a certain benefit over the classical maximum-rate approach since it leads to the closed-form solution of the optimal power allocation and allows a comprehensive description of the optimal coding.

In the second part, we explore energy-efficient relay deployment and analyze how the propagation environment, user transmission rate and relay coding scheme affect the choice of relay location and the network performance. We highlight new trade-offs which balance coverage extension, energy consumption and deployment flexibility. We then show that advanced coding schemes can easily overcome harsh radio environments due to suboptimal location, and thus provide a much needed flexibility to the relay deployment.

In the third part, we investigate the maximum energy gain provided by relay stations in shadowing environments, accounting for the relay-generated interference and additional circuitry energy consumption. We propose easily-computable models for the energy-optimized cell area served by relays and define a new performance metric that jointly captures both aspects of energy and interference. We then highlight a trade-off between the relaying energy efficiency and the resulted interference and show that, despite its increased circuitry consumption, the use of partial decode-forward instead of the simple two-hop relaying not only alleviates the interference issue, but also leads to a reduction in the number of relays necessary to reach same performance.

Keywords: Energy efficiency, relay deployment, cellular networks, optimized power allocation, spatial models for performance analysis

ANALYSE DE L'INTERDÉPENDANCE ENTRE CODAGE ET POSITION DES RELAIS POUR DES RÉSEAUX CELLULLAIRES ÉCONOMES EN ÉNERGIE

Fanny PARZYSZ

SUMMARY

Afin de satisfaire la demande des clients à un accès haut débit en tout lieu, les recherches se sont essentiellement concentrées sur la capacité et la couverture des futurs réseaux cellulaires. Mais les avancées technologiques qui ont résulté d'un tel objectif se sont accompagnées d'une très forte hausse de la consommation d'énergie, si bien qu'il est devenu urgent de développer de nouvelles techniques économes en énergie. Dans cette optique, les relais sont considérés comme très prometteurs, pour un coût concurrentiel, et de nombreux algorithmes ont été proposés pour optimiser la position des nouvelles stations. A ce jour pourtant, les analyses de performance des relais se fondent sur des codes sous-optimaux ou optimisés pour le débit. Cette thèse par articles a donc pour but d'étudier les avantages apportés par des codages explicitement optimisés pour l'énergie, en vue de faciliter le déploiement de nouveaux relais.

Dans une première partie, nous déterminons l'énergie minimale consommée par un relai qui décode puis retransmet les données (*decode-forward*), mais qui ne peut transmettre et recevoir simultanément (*half-duplex*). Pour cela, nous développons un nouveau codage, ainsi que plusieurs allocations de puissance pour minimiser la consommation d'énergie. Un résultat théorique important montre que minimiser l'énergie du réseau n'est pas équivalent à maximiser sa capacité, comme souvent admis. De plus, une telle approche énergétique démontre un avantage certain par rapport à l'approche classique par la capacité, puisqu'elle permet d'obtenir une solution explicite pour l'allocation de puissance optimale et de caractériser complètement le schéma de codage correspondant.

Dans une deuxième partie, nous abordons les économies d'énergie qui peuvent être réalisées en optimisant le positionnement des relais. Nous analysons de quelle manière l'environnement de propagation radio, le débit de transmission de l'usager et le codage utilisé affectent le choix de l'emplacement du relai et les performances du réseau. Nous soulignons alors de nouveaux compromis mettant en relation l'extension de couverture, la consommation d'énergie et la facilité de déploiement. L'utilisation de codages optimisés apporte une plus grande robustesse face aux dégradations du canal propagation, dues par exemple à un emplacement peu satisfaisant. De tels codages offrent donc beaucoup de flexibilité dans le positionnement des relais.

Dans une troisième partie, nous étudions une nouvelle fois le gain d'énergie maximale permis par un relai dans un environnement à ombrage log-normal, mais en tenant compte cette fois-ci de l'interférence supplémentaire qu'il génère et de l'énergie additionnelle dissipée par l'électronique. Nous développons pour cela des modèles mathématiques pour calculer aisément la zone de la cellule servie par le relai et proposons une nouvelle métrique de performance qui relie le gain en énergie à la perte en interférence. Puis, sur la base de cette analyse, nous mettons à jour le compromis qui existe entre ces deux mesures de performance des réseaux.

Malgré une dissipation d'énergie dans les circuits électroniques accrue, l'utilisation de codages optimisés, au lieu d'une simple transmission en deux sauts, permet non seulement d'atténuer les dégradations de performance dues aux interférences, mais aussi de réduire le nombre de relais nécessaires pour atteindre les mêmes performances.

Mots-clés: Efficacité énergétique, déploiement de relais, réseaux cellulaires, allocation de puissance optimisée, modèles spatiaux pour l'analyse des performances

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS

2Hop	Two-hop relaying
3G / 4G / 5G	Third / Forth /Fifth Generation
3GPP	3rd Generation Partnership Project
AEP	Asymptotic Equipartition Property
AF	Amplify-Forward
AWGN	Additive White Gaussian Noise
BS	Base Station
CoMP	Coordinated Multi-Point
c.d.f.	Cumulative Distribution Function
CF	Compress-Forward
CSI / CSIT	Channel State Information / Channel State Information at Transmitter
DF / PDF	Decode-Forward / Partial Decode-Forward
DMC	Discrete Memoryless Channel
DTx	Direct Transmission
EEA	Energy Efficiency Area
EO-PDF	Energy-Optimized Partial Decode-Forward
ETSI	European Telecommunications Standards Institute
FD	Full-Duplex
G-EE	Generalized Energy-Efficiency coding scheme
GSM	Global System for Mobile Communications (Groupe Spécial Mobile)

HetNets	Heterogeneous cellular networks
HD	Half-Duplex
ICI	Inter-Cell Interference
ICT	Information and Communication Technology
i.i.d.	Independent and identically distributed
KKT	Karush–Kuhn–Tucker
LOS / NLOS	Line-Of-Sight / Non Line-Of-Sight
LTE / LTE-A	Long-Term Evolution / LTE Advanced
MIMO	Multiple Input Multiple Output
MRC	Maximum Ratio Combining
N-EE	Network Energy-Efficiency coding scheme
OFDM / OFDMA	Orthogonal Frequency-Division Multiplexing / Multiple Access
p.d.f.	Probability Density Function
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
R-EE	Relay Energy-Efficiency coding scheme
REA	Relay Efficiency Area
RF	Radio-Frequency
RRM	Radio Resource Management
RS	Relay Station

RTx	Relay-aided Transmission
S-EE	Source Energy-Efficiency coding scheme
SC	Superposition Coding
SISO	Single Input Single Output
SM	Superposition Modulation
SNR	Signal-to-Noise Ratio
STC	Space-Time Coding
Tx	Transmission
WINNER	Wireless World Initiative New Radio

INTRODUCTION

The wide possibilities offered by wireless networks and real-time applications had the highest impact on everyday life, and consumers have come to rely on wireless for voice communication and always-on access to the Internet wherever they go. While voice application has driven the past growth of cellular systems, data service, e.g. interactive videos, mobile peer-to-peer or streaming, is now taking over with the rapid spread of smartphones and tablets.

Today, more than 25 million Canadians have a mobile phone or wireless device, a number that is expected to go on increasing significantly. In addition, the amount of traffic generated by each user is growing rapidly. Indeed, a smartphone uses the capacity of as many as 24 "regular" mobile phones and a tablet uses the capacity of as many as 122 mobile phones, thus setting new technical challenges for engineers.

To alleviate traffic hot spots and dead zones, operators have largely invested to improve network availability, both in terms of coverage and capacity. New types of access point have been introduced, thus constituting macro-, micro-, pico- and femto-cells and shaping Heterogeneous Networks (HetNet). In particular, cooperative networks and relaying technologies are likely to play a prominent role in next-generation systems.

However, seeking for capacity and coverage has come along with a staggering increase of the energy consumption, thus becoming one of the major drawbacks of next-generation technologies. From the operators' perspective, energy waste results in significant economic loss. Indeed, mobile operators are already among the largest energy consumers and up to 80% of the energy consumption in a cellular network is consumed by the radio access part, in particular base stations. From the users' perspective, battery lifetime appears to be particularly poor when using 3G services. Without breakthrough in battery technology, this will remain a serious hindrance for the development of new services and markets. Moreover, Information and Communication Technology (ICT) is estimated to contribute to at least 2% of greenhouse gas emissions and to consume 5-10% of global electricity. This consumption is projected to dou-

ble in the next decade. Thus, reducing the energy consumption responds to critical ecological concerns and calls for social responsibility in fighting climate change.

In the urge to shift from capacity-oriented to energy-oriented network designs, research has taken significant steps in modelling the energy efficiency and new adequate metrics have been proposed. Whereas the capacity and coverage improvements achieved through cooperative networks have been widely investigated, the energy impact of relaying is not yet clearly understood and several questions remain open:

- What is the minimal energy consumption that can be expected from relaying ?
- In which conditions such minimal consumption can be achieved?
- How does relay deployment affect the overall energy consumption?
- Does minimizing the energy consumption necessarily lead to interference reduction?

We take in this thesis a new direction towards energy-efficient cellular networks and bring answers to such questions by jointly exploring the impact of the coding scheme and the relay deployment. The Shannon's theory has been extensively used for rate maximization and search for capacity. But information theory has also its role to play in the energy issue and, so far, research has not yet explored such opportunity.

In this research work, we analyze the minimal energy consumption that can be achieved using decode-forward relaying and we show that, in addition to energy gains, more advanced coding schemes bring flexibility to the relay deployment, robustness to channel impairment and can also be used for better interference management.

CHAPTER 1

FROM SPECTRAL EFFICIENCY TO ENERGY EFFICIENCY

The fast escalation of energy demand has urged mobile operators to limit the electric consumption while maintaining high data rates and ubiquitous access. Some early works have shed light on the potential energy saving in wireless networks. However, these efforts were isolated and have mostly focused on some elements or components apart from the rest of network, such that they could not form a global vision of energy-efficient networks [2, Chapter 1]. One of the earliest insights into the potential of existing cellular networks to operate in an energy-aware manner was provided in [3]. And recently, the concept of **Green Networks** [4, 5] has emerged to constitute a comprehensive research area on energy efficiency for the next generation of communication networks. In this first Chapter, we set up the context and the outline of this thesis by introducing green networks, their fundamental trade-offs and specific metrics.

1.1 Overview of energy-efficient techniques

We start with a very general review of energy-efficient techniques that aims at reflecting the importance of such topic in both the academic and industrial worlds. Energy-efficient relaying is then contextualized in this trend.

1.1.1 Technologies for energy savings

A discussion on energy-efficient cellular networks cannot bypass the numerous research actions that have been taken worldwide. Several international projects are being carried out, such as "Energy Aware Radio and NeTwork TecHnologies" (EARTH) [6], OPERA-Net or the consortium GreenTouch [7]. Discussions are also in progress in standardization organizations, e.g. ETSI and 3GPP. Such work groups have proposed system-level approaches that cover all layers in the protocol stack of wireless access networks and their architectures [3, 8, 9]. The main objective is to adjust the transmission parameters to the time-varying radio propagation conditions, the fluctuating traffic load and the QoS requirements. Such techniques are presented in

the following, going from the physical layer to the large-scale network topology, and illustrated in Figure 1.1. In this, the approach of energy-efficient cellular networks which has guided the research work of this thesis is pointed out in red.

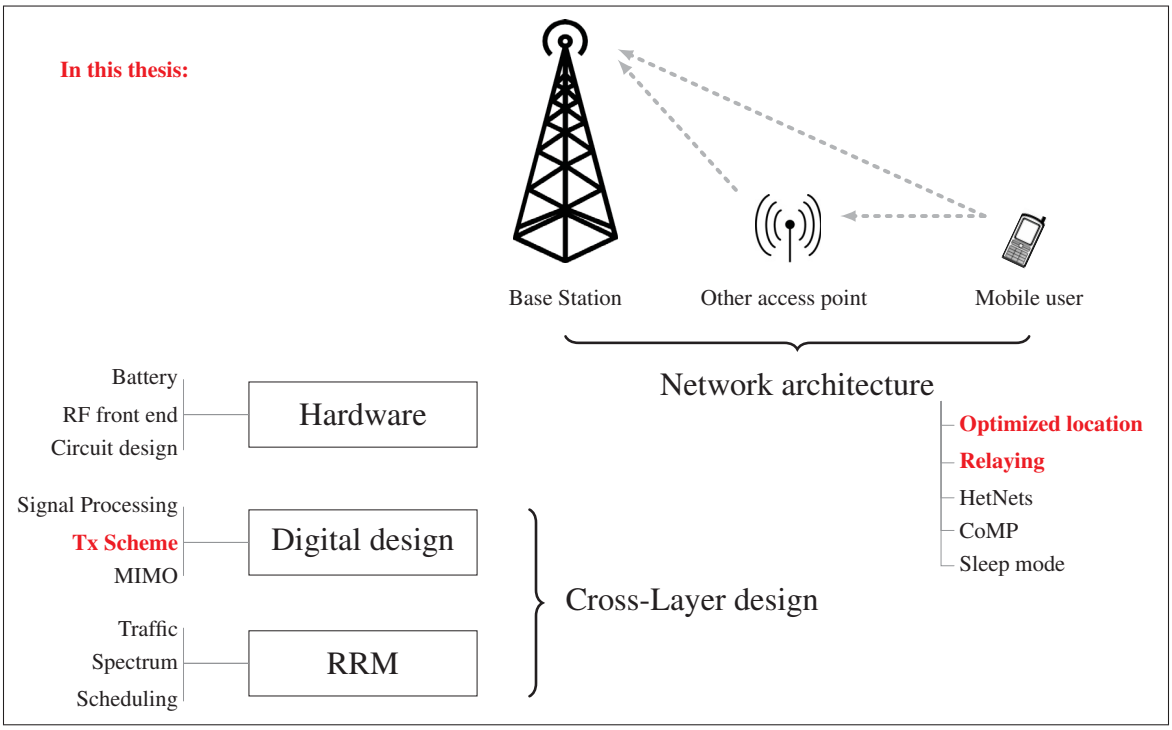


Figure 1.1 Overview of energy-efficient techniques

First, **at the physical layer**, hardware technologies are undoubtedly among the largest energy-consuming blocks of the protocol stack and severely affect the system performance compared to theoretical bounds. For instance, a typical power amplifier usually performs with considerably low efficiency ($< 50\%$). Highlights of energy-efficient hardware technologies include low-power circuit design [10] and high-efficiency digital signal processing [3], but also advanced cooling systems and battery design. Energy-efficient radio transmission schemes are an active research topic as well. Main results regard energy-oriented modulations [8, 11] and multiple-antenna systems (MIMO technologies) [8, 12, 13, 14], which exploit the time and spatial diversity.

Second, **at upper layers**, radio resource management (RRM) plays a prominent role in the race for energy reduction and aim at improving both the traffic and the spectrum managements. Differentiating the QoS requirements of users and performing energy-aware load-balancing allow to adapt the traffic pattern to the radio propagation environment. In addition, energy efficiency can also be reached by a better utilization of the available spectrum. With this perspective, techniques for dynamic spectrum access and cognitive radio are under research. The related literature is quite extensive but has been compiled notably in [15] and [2, Chapter 15].

As nicely expressed in [9], energy efficiency has to do with "assigning the right resource to transmit to the right user at the right time". While techniques for resource management or energy-aware transmission schemes described previously aim at optimizing a single element of the communication system (e.g. the scheduling strategy), **cross-layer design** operates across the different layers of the protocol stack for better adjustment to the user requirements and channel environment [16].

Finally, energy-efficient designs involve as well the **network architecture**. Beyond classical homogeneous cellular networks served by base stations only, the deployment of new types of access points which are coordinated has been spurred by the research community. Multi-tier and cooperative networks, including relaying and Coordinated Multi-Point (CoMP) [17], are the highlights of energy-oriented architectures, along with energy-aware mobility management strategies [18] and base station sleep mode or cell size adjustment [19].

1.1.2 Context of this thesis

In this work, we consider energy-efficient relaying for cellular networks, as depicted in Figure 1.1. Relay stations are usually deployed with the same objectives as small cells, i.e. coverage extension, support for hotspot areas, improved spectrum management and energy reduction. The major difference between both lies in the backbone access. Whereas micro- or pico- base stations are connected to the core network via optical fibre, relay stations rely on wireless transmission, as illustrated in Figure 1.2. Despite the associated performance loss, wireless backhaul can be particularly beneficial when the infrastructure required for optical cables is

not available due to technical, economical or legal reasons. Relay stations are also particularly suited for isolated area and off-grid power systems, which harvest energy from the environment (e.g. solar or wind energy). This also implies that techniques for energy savings are radically different. Such point is further analyzed in the third part of next Chapter.

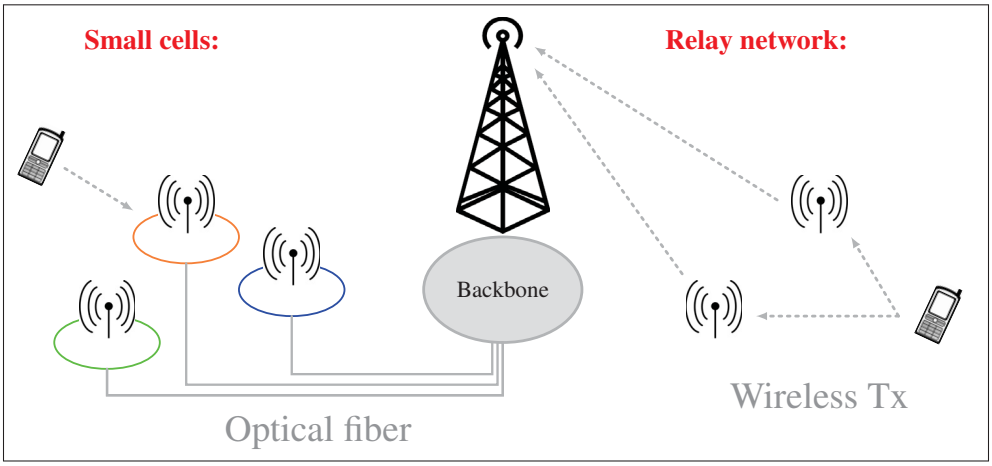


Figure 1.2 Comparison of relay-aided network and small cells

Relay-assisted cellular networks have generally been featured with dedicated relay stations, which do not act as sources nor destinations, but whose sole purpose is to facilitate the communication of other users. This is the only case considered in this thesis but it must be noted that peer nodes can act as relays as well. As envisaged in 5G cellular networks, the possibility to perform device-to-device communications will lead to a hybrid of conventional cellular and ad-hoc designs. Besides the technical challenges and security issues of such architecture, this would imply convincing consumers to use their own resources, notably the battery, for the benefit of the traffic of others.

As guiding line, we bring together two different aspects of energy-efficient relaying: the **geographical deployment** of relay stations and the **choice of the relaying coding scheme**, i.e. how to encode and decode data, with regards to Information Theory. While such aspects may seem disconnected at first, we show that, in fact, the coding scheme has major impact on the

optimal relay deployment and can be considered as a fundamental key to unlock energy and cost efficiency in relay-aided cellular networks.

1.2 Challenges of Green Networks

New challenges have been raised from energy-efficient communications and have necessitated the definition of specific network performance and cost indicators. Analyzing those challenges and indicators clarifies how we have addressed the energy issue in this thesis.

1.2.1 Fundamental trade-offs

The comprehensive work of [9] gives a substantial description of the questions and compromises which have to be considered to reach energy efficiency. The four fundamental trade-offs proposed in this paper pertain to spectrum allocation, energy consumption, QoS requirements and deployment cost as follows:

- **Spectrum versus transmit energy:** The literature has largely investigated this trade-off with regards to capacity maximization. The attainable spectrum efficiency is defined as the system throughput reached per unit of bandwidth (b/s/Hz) and is generally known given the available bandwidth and energy consumption [20, 21, 22, 23]. It is a widely accepted criterion for wireless network optimization but investigation is still required with specific focus on the minimal energy necessary to transmit data at a given rate.
- **Overall power versus required bandwidth:** This trade-off further characterizes the relationship between spectrum and energy. Given a target transmission rate, bandwidth and power are the most important but limited resources in wireless communications. However, due to the energy dissipated in the circuitry and the additional bandwidth required for channel state information at transmitter (CSIT) and signaling, the full utilization of both power and bandwidth resources may not be the optimal energy- and spectrum-efficient approach. As highlighted in [9], the relationship between the signal bandwidth (MHz), the power density (W/MHz) and the energy efficiency (bits/s/W) for a given user rate requires better understanding.

- **Energy versus quality of service (QoS):** This uncommon trade-off associates the system energy efficiency and the user experience, through the service latency. Such metric becomes essential since the heterogeneity of latency requirements and the time-varying traffic load have considerably modify the management of buffers and schedulers. For example, delay-tolerant traffic can be stored when fading occurs, and transmitted only when the channel conditions are strong, i.e. through the best nodes within the cell and at the best instant within the message delivery deadline.
- **CapEx versus OpEx:** Last but not least, one cannot ignore the infrastructure and operating costs, also called capital expenditure (CapEx) and operational expenditure (OpEx). Both are critical performance indicators from the point of view of a mobile operator and are generally in conflict with each other. Indeed, reducing the transmit energy, and thus the OpEx, can be efficiently achieved by shortening the distance between the base station and mobile users, which comes at the price of increasing the infrastructure cost. And conversely, spacing out base stations reduces the infrastructure cost but more energy has to be consumed to reach cell-edge and provide ubiquitous coverage.

In the next subsection, we analyze how these fundamental trade-offs on energy-efficient communications have modified the standard performance metrics and how they have guided our research work.

1.2.2 Definition of new energy-oriented metrics

As highlighted by Auer *et al.* in [2, Chapter 16], "what metric we use to capture efficiency determines how we think and act". To achieve energy efficiency, measuring the energy consumption, in Joules, at the component level is quite straightforward but cannot quantify properly the potential gains of new approaches and configurations at the system level. Thus, new energy metrics are required.

Energy per transmitted bit: One of the most popular and earliest metric is expressed in Joule-per-bit [J/b], or equivalently in Watt-per-bit-per-second [W/bps]. It refers to the network energy consumption divided by the number of bits delivered during a given observation period and quantifies the first trade-off defined previously, between the transmit energy and spectrum.

The reversed metric, in bit-per-joule, has been widely used for maximum-rate analysis. Yet, research work on the joule-per-bit metric has mostly regarded the fundamental information-theoretic concept that quantifies the minimum energy required to reliably transmit one bit over a noisy communication channel given infinite bandwidth, as in [24, 25]. More practical analyses have been proposed for particular simulation settings and relaying coding schemes. However, it is not yet clear what is the minimal theoretical energy consumption that can be expected from relaying, given a user rate and available bandwidth. Such result is relevant since, when designing a new practical system, it can tell how far the proposed technique is from the theoretical lower bound on energy.

Energy per covered unit area: The joule-per-bit metric is generally valid for long-range transmissions or high traffic load. While cellular operators should provide consumers with wireless support for high data rate services, they should as well offer the possibility to access the network everywhere, at any time. To this end, a large part of the network is primarily providing coverage and does not operate at full load even at peak traffic hours. Due to the energy cost necessary to maintain operational such network, the system energy efficiency appears to be particularly poor under low-traffic loads [3]. Including the aspect of the cell coverage is essential to evaluate the network performance, even when no data is transmitted.

More recently, new metrics have emerged to capture this second aspect of the network energy efficiency and complement analysis in joule-per-bit. One of these metrics refers to the average power consumption per unit area, expressed in Watt-per-square-meter [W/m^2], given a minimum performance requirement, e.g. outage probability or user rate requirement. Also note that such metric is rather formulated in terms of Watt-per-user [W/user], when only the average user density, in [user/m^2], is known.

Analyzing the energy gain jointly with the coverage has direct link with the last trade-off, between CapEx and OpEx. Concerning relaying, it brings out the issue of the number of relays that should be deployed and of the coverage extension, which is not proved to be energy efficient for all situations.

1.3 Measuring energy efficiency

1.3.1 Need for unanimous metrics definition

Although the metrics previously presented are now widely accepted in the research community, the various analyses proposed in the literature are still difficult to compare. Indeed, no single definition for performance measurement won unanimous support so far and the considered simulation settings depend on the network and service that is investigated.

For example, in the Joule-per-bit metric, it is not yet clear whether only the data bits should be considered, or if the overhead bits should be accounted for as well. In the Watt-per-square-meter metric, the coverage can be defined with regards to outage requirement, maximal power consumption or QoS, implying that the performance measurement of a same system can be significantly different depending on the considered definition. Similarly, the energy consumption can be measured using different approaches, as explained in next section.

Note that research is being carried out to develop a standardized methodology for global evaluation of cellular networks. As example, a holistic simulation-based framework, named the Energy-Efficiency Evaluation Framework (E^3F), is being developed as part of the European EARTH project (Energy-Aware Radio and neTwork tecHnologies [6]). Such methodology aims at bringing together the various metrics existing at each layer of the protocol stack and proposing for each a single definition. A second objective is to propose a framework that allows to compare radically different networks, such as dense urban and sparse rural areas, and that considers the system as a whole, accounting for both short-term and long-term (over a day) traffic model, small-scale and large-scale (over a whole city or even larger area) deployment scenarios, and so on. In this thesis however, such system-level performance evaluation is not required, since our problem formulation is focused on specific scenarios.

1.3.2 Accounting for the side energy consumption

Analyzing the **transmit power**, i.e. the power radiated at the output of the RF circuitry and antenna, allows a fair characterization of the network upper-bound performance which can be

achieved with regards to classical indicators such as spectrum efficiency. However, accounting for the useful transmit energy only is generally not sufficient for an accurate energy-efficient analyzis and setting up a system-wide energy efficient metric is imperative.

A comprehensive energy analyzis should include the **overall energy consumption**, i.e. :

- the transmission power,
- the circuitry power of each node, which includes both a transmission-independent and a transmission-dependent part,
- the energy of the signaling overhead necessary to perform the transmission,

as depicted in Figure 1.3. Trade-offs on the energy consumption must be made such that the energy savings in one part would not be counteracted by increased energy consumption in another part. This is particularly true when analyzing how the number of relays in a cell affects the network energy efficiency.

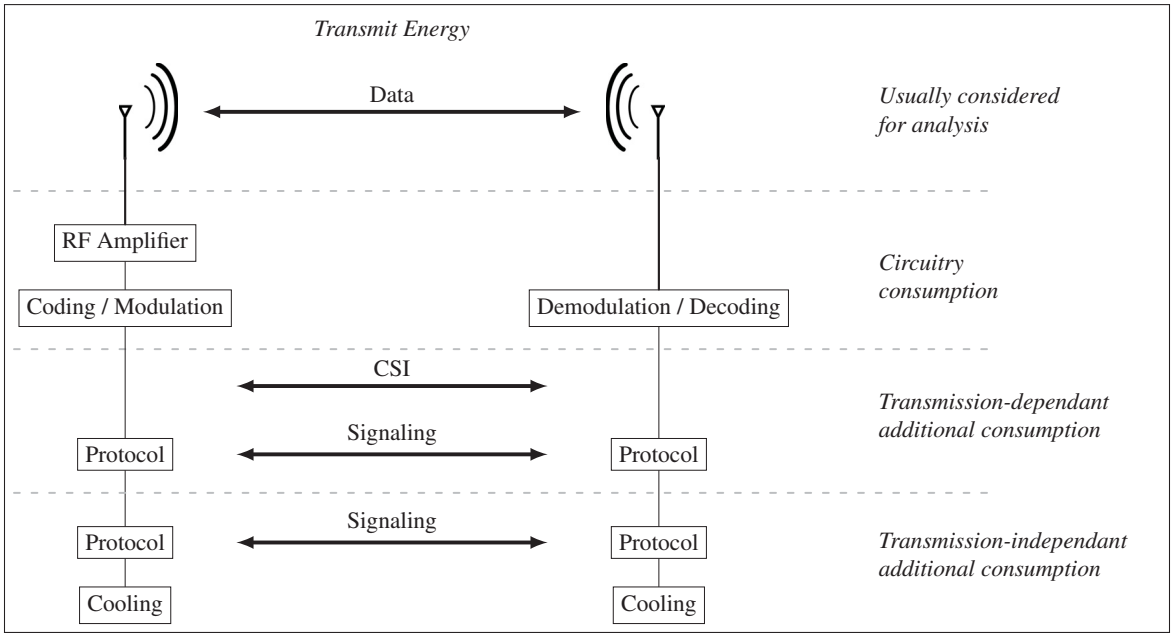


Figure 1.3 Description of the overall energy consumption

First, the **transmission-independent circuitry consumption** accounts for the energy dissipated whether or not data is transmitted for site cooling or network maintenance. It is usually

modelled by an energy offset. The consumption of macro base station is generally dominated by this offset, but it largely depends on the type of the considered station, as stated in [8]. Models for the circuitry offset consumption are available in [2, Chapter 6] and related references.

Second, the **transmission-dependent circuitry consumption** refers to the energy dissipated for signal processing during data transmission. It accounts for the feeder loss, the extra loss in transmission-related cooling, but above all, the energy loss of the physical RF amplifier which usually performs with considerably low efficiency ($< 50\%$). The energy consumption of small cells (e.g. pico / femto-cells) is generally dominated by this energy part. To accurately model the power cost attributed to the node power amplifier and other components, many factors should be included, such as the quality of electronics, the complexity of data processing, but also the resource management and buffering strategy.

Third, the **energy consumption of signaling exchange, CSIT acquisition and protocol headers** can significantly affect the system performance, as highlighted in [26], and there may exist some trade-off between the amount of signaling that is necessary and the overall system energy efficiency. This is particularly true regarding CSIT-hungry techniques such as multi-user MIMO or coordinated multipoint (CoMP) for interference mitigation. However, it is still difficult to quantify the overhead energy consumption, especially at the physical layer. This depends on the considered protocol and is yet to be defined in standards.

1.4 Thesis framework and outline

So far, we have introduced the context of energy-efficient networks and highlighted the principal trade-offs and metrics of such issue. By jointly exploring the impact of the coding scheme and the relay deployment, we propose a new analytical approach on energy efficiency.

In spite of steadily increasing interest on energy efficiency for relay networks, the available research is largely based on coding schemes and power allocation optimized for rate, as given by classical results from information theory, but not optimized for energy. Thus, as a first step, we focus on transmitting data at a desired user rate and answer the following question:

What is the minimal transmit energy consumption
that can be achieved with decode-forward relaying ?

To this end, we explore coding techniques and power allocations explicitly optimized for energy, using information-theoretical tools. The literature review on this topic is covered by "PART I: Coding for energy efficient relaying" in the next Chapter.

Whereas theoretically-optimal coding schemes may achieve significant energy gain, they may rapidly turn out to be computationally complex. Thus, it is essential to make wise use of advanced coding schemes, such that the provided gains overcompensate their increased complexity. This can be done by carefully choosing the relay location within the cell. Algorithms for optimal placement have been extensively studied to improve the system performance, but we address this question from a different point of view:

What do energy-optimized coding schemes bring
to energy-efficient relay deployment ?

In particular, we aim at analyzing the flexibility of the relay deployment and the impact of the radio propagation environment. The literature review on this topic is covered by "PART II: A new perspective on energy efficient relay deployment" in the next Chapter.

Finally, we explore relaying in the context of larger networks. To this end, we refine our analysis on energy-efficient relay deployment by explicitly considering the circuitry consumption and by investigating the impact of relaying on the interference received in neighbourhood cells. We question the following point:

Does energy-efficient relay deployment
necessarily lead to interference reduction ?

This question highlights a new trade-off on energy efficiency. While using energy-optimized coding schemes necessitate additional signal processing, and thus extra energy consumption,

the reduction of the transmit energy benefits in return the rest of the network. Related literature review is covered by "PART III: Reducing the energy consumption and interference: an equivalence ?" in the next Chapter.

This research work is presented as a manuscript-based thesis and is composed of our IEEE Information Theory Workshop (ITW) paper [27], our IEEE Transaction on Communications [28], our IEEE Transaction on Wireless Communications [29] and our manuscript recently submitted to the IEEE Journal on Selected Areas in Communications [30]. In addition, three other peer-reviewed papers [31, 32, 33] have been respectively published in the 34th IEEE Sarnoff Symposium, the IEEE International Conference on Communication (ICC) and the IEEE Vehicular Technology Conference (VTC-Fall).

CHAPTER 2

LITERATURE REVIEW

This second chapter raises out the particular challenges which have been addressed in this thesis and surveys the related literature. It is divided into three parts, in line with the three manuscripts that compose this thesis. The first part overviews relaying strategies and explains how information theory can be used for the purpose of energy saving. Then, the second part clarifies our choice to analyze relay deployment and radio propagation environments. Finally, in the last part, we take a broader perspective and investigate the challenges of relaying in multi-cell networks.

2.1 Part I: Coding for energy efficient relaying

Among all the energy-aware techniques proposed in the literature, from hardware to network architecture, the impact of the relaying coding scheme has been little explored and the gap between information-theoretical results and more practical analysis remains wide. This section introduces such topic and specifically focuses on three-node relay networks, which consist of a source, a relay and a destination, or equivalently a base station, a relay and a user. Information-theoretical calculations will be further detailed in Chapter 3.

2.1.1 Techniques for relaying

The basic principle of relaying, and more generally cooperative networks, is that some nodes overhear part of the transmitted data and forward it to the destination, instead of treating it as interference. Since the destination receives multiple independently faded copies of the transmitted data, decoding is more reliable and less energy is required for same performance. However, the performance of relaying is significantly affected by the collaborative strategy, i.e. the decision on when to relay, how to relay and with which relay station.

2.1.1.1 Duplexing mode and carrier frequency

First, a relay operates either in a half-duplex or a full-duplex mode:

- **Half-duplex:** Here, the relay cannot receive and transmit simultaneously and the transmissions are generally multiplexed in time. In a cellular system, a specific scheduling policy is required between the users and the base station. For example, time slots can be statically pre-assigned or dynamically allocated for better flexibility and optimization.
- **Full-duplex:** In this case, the relay can transmit and receive at the same time, at the same frequency: it receives the signal, processes it and then forwards it, allowing only a small delay within the frame duration. This is usually possible by physically separating the transmit and receive antenna radio patterns.

Whereas full-duplex is attracting more and more attention for potential implementation in 5G networks, it is still facing severe issues. Techniques for isolating the received signal from the transmitted one remains under research, particularly for compact antenna dispositions as analyzed in [34]. Moreover, the trade-off between the performance enhancement and the increased cost should be considered. In this thesis, half-duplex relaying is the only considered mode. Another important feature of a relay station is the carrier frequency it operates on:

- **In-band relaying:** Here, the user-to-relay link and the relay-to-base station link are carried over the same frequency,
- **Out-band relaying:** the transmissions between the relay and the base station are carried over a dedicated frequency, as for a conventional backhaul link.

Active cooperation for energy reduction, such as beamforming gain between the relay and the source transmissions, can be achieved more naturally with in-band relaying, while out-band relaying is more suitable for coverage extension. In this research work, we only consider in-band relaying.

2.1.1.2 Relaying strategies

Relaying can be implemented using many different strategies, as depicted in Figure 2.1:

- **Amplify-forward (AF):** The relay simply amplifies and retransmits the received signal, such that the destination receives two independently faded versions of the source signal for decoding. AF relays may be simple repeaters, without any further processing, or may perform more advanced processing, including beamforming and MIMO.
- **Decode-forward (DF):** This scheme is the closest to the idea of a traditional relay: the relay decodes the received signal, re-encodes it and forwards it to the destination. In a Full-DF scheme, the whole source message is decoded, whereas, in partial DF, only part of it is decoded and re-encoded at the relay station.
- **Demodulate-forward:** Here, the relay solely makes hard decision on received symbols, but without decoding data. This way of relaying makes more frugal use of resources and decreases the processing delay at the relay, compared to decode-forward [35, 36].
- **Compress-forward (CF):** This strategy is particularly beneficial when the link from source to relay is weak compared to the link from source to destination [37]. Even if the relay may not be able to fully decode the source message, the signal received at the relay station remains an independent observation of the source signal. Such observation is hence compressed and forwarded to help decoding at the destination.
- **Distributed space-time coded schemes (STC):** With a perspective different from processing and forwarding, a pool of relays cooperate to simultaneously transmit the different columns of a space-time coding matrix.

There exists numerous variations of these relaying strategies. In particular, decode-forward has been derived into the following sub-schemes, which are used when CSIT is not fully available:

- **Selective decode-and-forward (SDF):** This is an adaptive DF scheme where the relay forwards the re-encoded signal only if some performance criterion is satisfied, e.g. a SNR threshold [38].
- **Incremental decode-and-forward (IDF):** Here, the relay only cooperates when it receives a Non-Acknowledgement from the destination. Thus, a feedback channel between the destination and the relay is required [39, 40].
- **Two-hop relaying:** This is a degraded version of Full decode-forward. The relay station acts as a router and the broadcast nature of the wireless medium is not exploited, i.e.

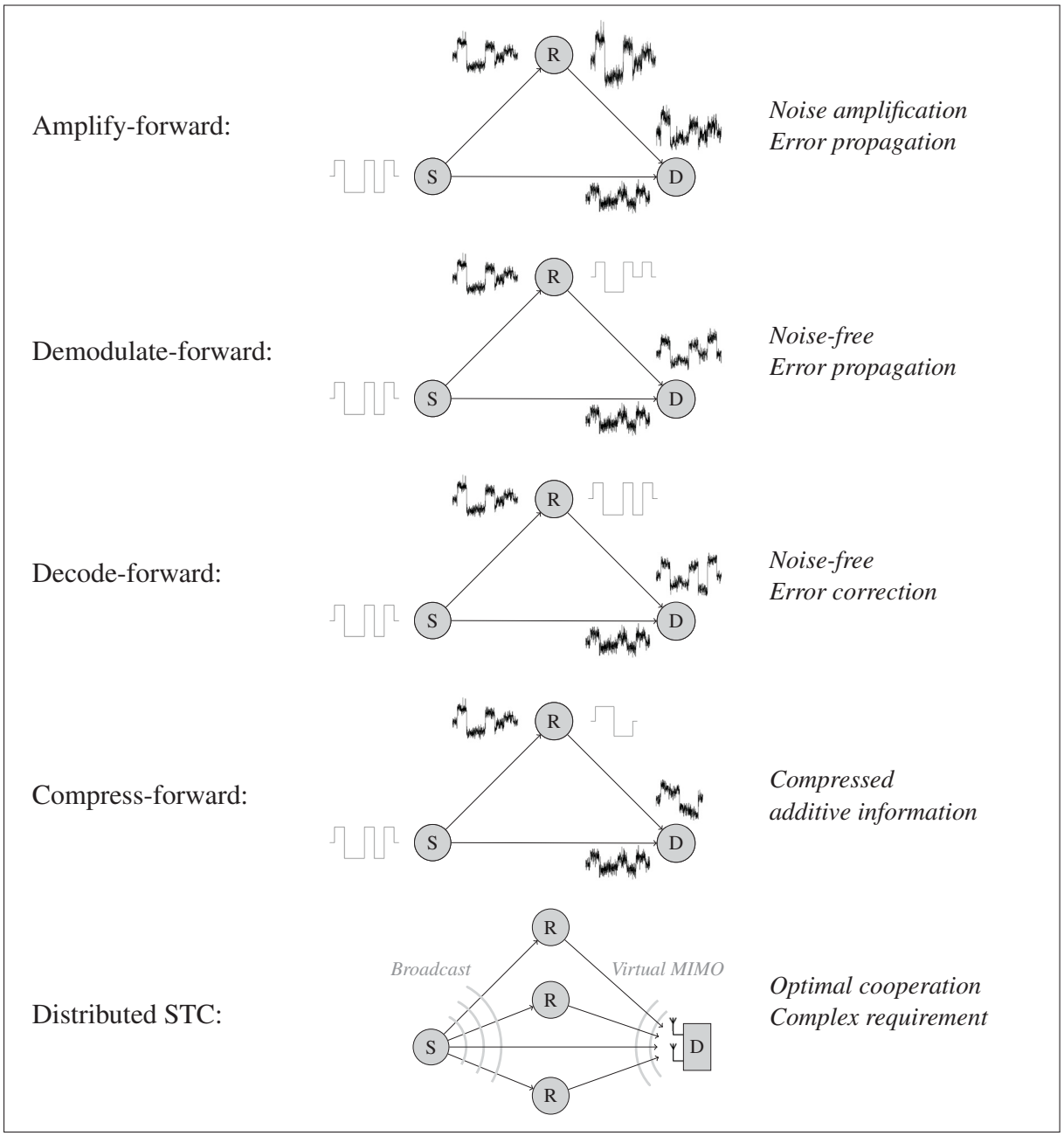


Figure 2.1 Simplified illustration of the main relaying strategies,
Inspired from Figure 4 of [1]

the signal received from the direct link is not considered for decoding at the destination. Note that there is no single definition of two-hop relaying in the literature, and such denomination is sometimes used in place of Full-DF.

We now compare the described relaying strategies. First, distributed STC schemes show interesting performance but require fine synchronization and co-phased transmission from each relay involved in the transmission. Moreover, due to mobility or channel variations, the pool of relays able to cooperate may change from one slot to another. This unfixed number of participating antennas make such scheme challenging for implementation [2, 41]. Due to its complexity, Compress-Forward remains hard to implement on practical systems as well.

On the contrary, AF relays are transparent for both the source and the destination and offers the lowest complexity, with the shortest processing delay. Yet, its main drawback is that it amplifies and retransmits a noisy version of the received signal. By mitigating the channel fading at the first hop and restoring the original signal at the second hop, decode-forward generally outperforms both amplify-forward and demodulate-forward under the channel conditions of cellular networks and using practical modulations, as shown in [42]. For these reasons, we consider in this thesis **decode-forward relaying**. Whereas multi-hop relaying provides even wider coverage and are also concerned by the energy issue [43, 44], we focus in this work on single-hop transmissions.

2.1.2 Relaying in LTE

Regarding practical cellular networks, relaying techniques are considered in the standardization process of LTE-advanced and IEEE 802.16j networks. The main reason for implementation is two-fold. First, MIMO-OFDM technologies and advanced error correction techniques allow data transmission at very high rates under many conditions, but they are often not sufficient to fully mitigate the critical problems experienced at the cell edge, where signal strength is poor and interference is high. Second, relay stations are easier and cheaper to deploy than macro or micro base stations. They require no separate wired backhaul and their small size allows installation in many convenient areas, e.g. on street lamps or on walls. Relays can increase the network density, coverage extension and performance enhancement, to alleviate hotspots and traffic dead zones. There are two types of LTE relays:

- **Type 1 (or non-transparency):** Here, relays can be considered as "light base stations", since they control their own cell with separate identity number and transmit their own

synchronisation channels and reference symbols. Usually, half-duplex in-band relaying is considered. The principal objective of type-1 relays is to provide network access to remote users, located outside of the macro base station coverage.

- **Type 2 (or transparency):** In this case, the relay shares the same identification and control channels that the base station it serves, such that users cannot distinguish them. In particular, signaling can be broadcast by the main base station, where the relay only conveys users data. Type-2 relays allows to improve the service quality of users located within the base station coverage by achieving multipath diversity.

Most analysis on energy-efficient LTE networks are based on two-hop relaying. The advantages brought by more advanced coding schemes have been little explored with respect to the energy issue. In the following, we introduce the topic of coding design, i.e. information theory, and explain how it can be used for performance enhancement.

2.1.3 Shannon's theory and energy efficiency

The pioneer work of Claude Elwood Shannon in 1948 has set up the basis principles of communication theory and has shaped our digital era. By applying thermodynamic laws to communication [45], Shannon has established a whole new discipline, Information Theory, which has in return redefined thermodynamics and finds application in telecommunication, data compression, cryptography or even gambling [37, 46]. Information theory is still guiding research on next-generation networks.

Information and channel capacity: For an uncoded system with a given transmit power, increasing the transmission rate implies the use of more complex modulation schemes having multiple bits per symbol, thus leading to an inevitable increase of the error probability. Shannon's theory has shown that, by adding redundancy in the transmitted data (i.e. encoding) and by using error correction, the transmission rate can be increased without increasing the final error probability, up to a certain rate. Beyond, data cannot be recovered without errors. This critical rate is called the channel capacity, measured in bits/symbols, or in bits/sec given that the number of symbols per second is determined by the channel's Nyquist bandwidth. Infor-

mation theory has allowed a better understanding on how the transmit power, the bandwidth and the noise may affect the channel capacity.

Nowadays, the research community continues to expand the Shannon's theory. New guidance has been proposed for future telecommunication systems, particularly regarding multi-user channels, and new mathematical methods have been developed to better understand their achievable performance. Multi-user channels are particularly challenging and the capacity of some of them, e.g. the broadcast channel and the relay channel, remains unknown, despite extensive research over the years.

Coding to reach fundamental limits: Coding has to do with adding redundancy into the transmitted signal to get low error probability at the destination. When designing a code, information theory can tell how much redundancy is required and how far it is from the theoretical limit, i.e. the channel capacity.

Even if Information Theory does not always provide the theoretically optimal code (and only shows its existence), it suggests means of achieving these ultimate limits of communication and tells how a practical scheme should be structured. As shown in [47], the highest performance is achieved when the signal structure given by Information Theory is emulated the most closely. Such observation has lead to significant progress in integrated circuits, signaling methods and code design. New techniques have been proposed, such as superposition coding, rate splitting, multiple-antennas systems or multiple-access designs.

What about energy consumption ? Since C.E. Shannon stated Information Theory in 1948, researchers and engineers have mostly targeted the channel capacity, by maximizing the communication rate, given a power budget. References within this direction are numerous and we just evoke here the landmark works of [48] for the relay channel or [49, 50] for the relay broadcast channel. Yet, using the Shannon's theory with regards to energy consumption has attracted little attention and often, it is implicitly admitted that minimizing energy and maximizing data rate are equivalent problems. Such equivalence is immediate for single-source

single-destination transmissions, where the channel capacity for the Gaussian case is given by

$$C = B \log_2 \left(1 + \frac{P}{N_0 B}\right) \tag{2.1}$$

with P the transmit power, B the bandwidth and N_0 the noise power spectral density. However, it is far less obvious that this equivalence still holds for multi-user transmissions. To the bast of our knowledge, no coding scheme with optimal power allocation has been derived specifically for energy-efficient decode-forward relaying and most of existing energy analyses are based on rate-optimal schemes [51, 52, 53, 54, 55].

2.1.4 Thesis methodology and contribution

As first step of this thesis, we question the use of maximum-rate coding schemes for the purpose of energy-efficiency analysis and define the minimal energy consumed by decode-forward relaying. To this end, we design the new half-duplex relaying coding scheme depicted in Figure 2.2 and propose for it three power allocations, all dedicated to energy minimization. This coding scheme and power allocations will be considered throughout this whole thesis, as reference to define the best achievable performance.

This first step is covered by Chapters 3 and 4. Chapter 3 describes the new relaying coding scheme that has been published in our conference paper [27], with further details on the main tools of Information Theory. Chapter 4, which is an unedited version of our journal paper [28], proposes optimal resource allocations for energy efficiency. In addition to these two chapters, the conference paper [32] extends the proposed resource allocation to distributed networks, where the channel state information is known only partially at transmitters.

2.2 Part II: A new perspective on energy efficient relay deployment

While advanced coding schemes and resource allocations, as introduced in Part I, can overcome the weaknesses of simpler schemes as two-hop relaying, they should be used wisely given their increased hardware complexity and particular requirements, e.g. CSIT and fine synchronization between the source and the relay. It is thus important to investigate in which conditions

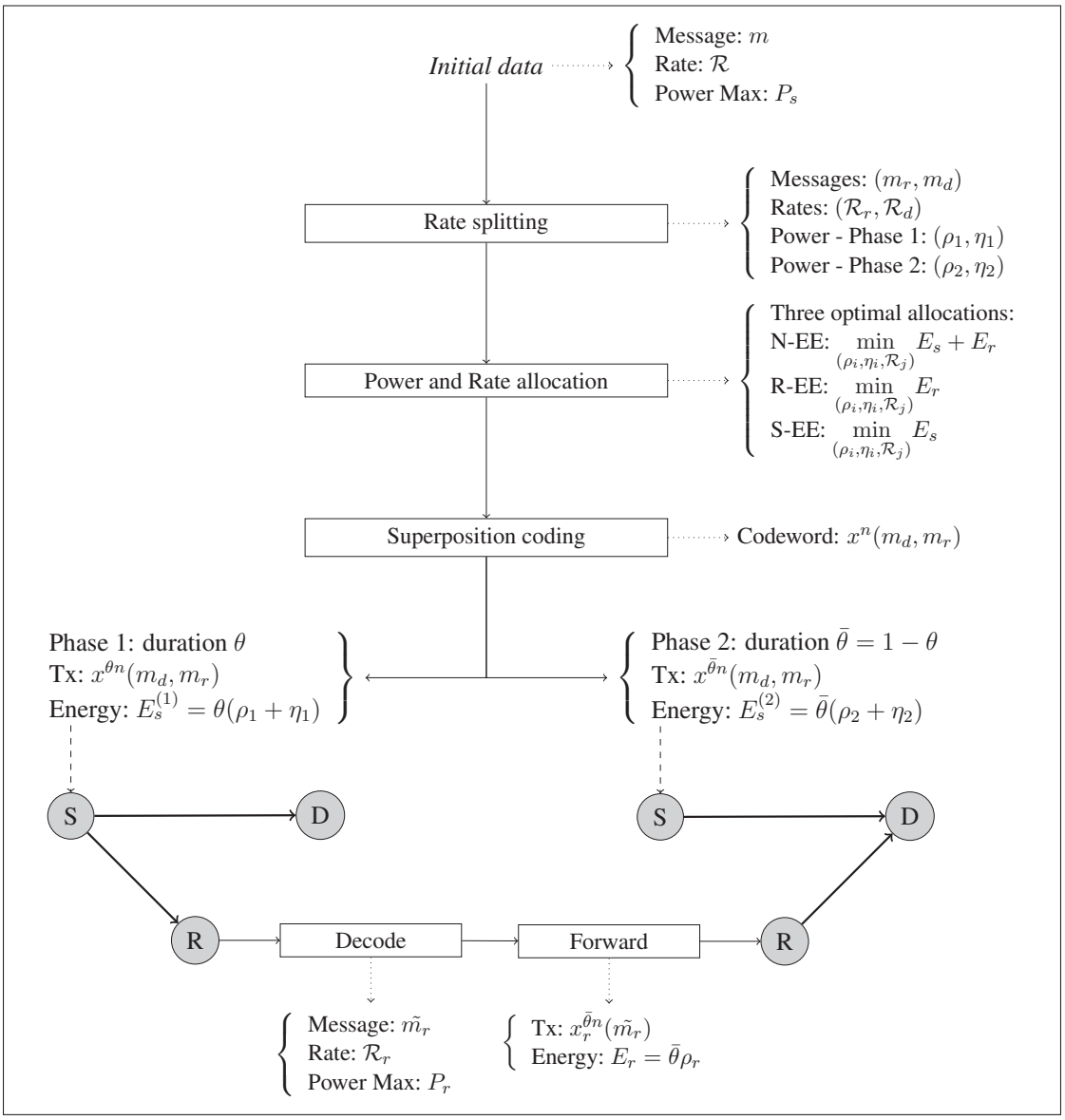


Figure 2.2 Proposed half-duplex partial decode-forward scheme and power allocations

relaying can provide worthwhile energy gain, i.e. how the radio propagation environment impacts the relay performance.

2.2.1 Just a matter of technology ?

With this perspective, efficiently positioning the relay station and avoiding poor locations with negligible gain is highly relevant. The literature related to optimal relay placement algorithms

is quite extensive and proposes a wide range of analyses regarding coverage [56, 57, 58], capacity [59, 60], outage [61] or energy [53, 54, 62] requirements. The recommended relay locations are numerous, depending on the considered network management and transmission technologies (e.g. the coding scheme, the antenna specification, the scheduling technique...).

However, in practice, deploying new wireless stations has a cost which can lead to considering sub-optimal solutions. In addition, a new site for relay station is as well selected based on external factors, such as the legal procedures, the preferences of affected populations, the urban topography and, of course, the radio frequency characteristics of the potential location. In the following, we analyze: 1) the signal propagation and how it can affect relay deployment, 2) the process for site selection and how it should modify our search for optimal relay location.

2.2.2 Statistical description of the wireless propagation channel

A relay location is first selected based on the radio frequency characteristics. The wireless radio channel is exposed to absorption, reflection, scattering, diffraction, thermal noise and interference, which randomly change over time, space and frequency as a result of user movement and environment dynamics. This constitutes a severe challenge for reliable communication. This subsection describes the main channel characteristics affecting the signal propagation. Then, the different channel models used to analyze and simulate the network performance are presented. The objective is not to be exhaustive about such a vast topic. Rather, we emphasize how the urban topology and environment affect the channel statistical distribution.

2.2.2.1 Signal propagation mechanisms

The amplitude of a channel response results from three main attenuation factors, which apply at different space and time scales: path-loss, shadowing and multipath fading. They have been extensively presented in numerous reference books, such as [63, 64, 65, 66].

First, **path-loss** refers to the dissipation of the power radiated by the transmitter. It is commonly assumed that it only depends on the transmitter-receiver distance and general environment characteristics. As opposed to free space, in a typical urban or indoor scenario, a signal

transmitted from a fixed source encounters multiple objects and reflectors in the environment which generate many replicas of the original signal. This can bring diversity gain, but may also degrade the transmitted signal. A dense environment and the loss of line-of-sight (LOS), i.e. the possibility for a signal to propagate along a straight line between the transmitter and the receiver, generally intensify the channel attenuation.

If the transmitter and the receiver are fixed and the environment is non-mobile, the radio-frequency characteristics of the received signal are deterministic and time-invariant. However, as a result of transceivers mobility and environment dynamics, the channel amplitude suffers from other attenuation factors, which may vary with time, such as **shadowing**. Shadowing results from obstacles between the transmitter and receiver that attenuate signal power through absorption. Generally, shadowing is slowly-varying with time and occurs over distances that are proportional to the length of the electromagnetically important obstructions (10–100 m in outdoor environments and less in indoor environments). It is often considered as the most challenging issue in terms of resource allocation and interference management.

Both path-loss and shadowing occur over relatively large distances and are referred as large-scale propagation attenuation. On the contrary, **multipath fading** may be referred as small-scale propagation effects since it occurs over very short distances, of the order of the signal wavelength. More precisely, when a single pulse is transmitted over a multipath channel, the received signal appears as a pulse train, where each pulse of the train has been carried over a distinct path between the transmitter and the receiver, via different reflectors, scatterers or cluster of scatterers. Multipath fading results thereby from the constructive and destructive addition of scattered signal components and can cause significant distortion to the transmitted signal. Moreover, the multipath channel is highly time-varying, due to the environment dynamics and scatterer mobility.

We do not further describe the multipath channels. Such small-scale channel effects are more relevant for signal processing analysis. We can reasonably assume that the considered systems are robust enough against multipath signal distortions, as it is the case for OFDM systems which transform a wideband frequency-selective channel into several flat narrow band

sub-channels. Therefore, the analysis and models proposed throughout this thesis can be understood as the mean performance result over a sufficient period of time such that fading is averaged out.

2.2.2.2 Models for wireless channels

Whereas the channel amplitude could be theoretically determined, the number, location, and dielectric properties of reflectors and scattering objects are generally unknown. In particular, shadowing and multipath fading become unpredictable. Thus, statistical models have been proposed, based on empirical measurements over a given distance, in a given frequency range and various transmission scenarios. Such models tend to be sufficiently general in that they are not restricted to environments in which the measurements have been made.

At time t, the gain of a channel h can be expressed as:

$$|h(t)|^2 = \frac{s(t)\ r(t)}{\gamma} = \frac{s(t)\ r(t)}{K d^{\alpha}},$$

where $s(t)$ stands for shadowing, $r(t)$ for multipath fading and γ for path-loss. The shadowing coefficient is most commonly modelled by a log-normal random variable, whose standard deviation varies between 3 and 8dB, depending on the considered environment. For multipath fading, several models have been proposed, using the Rayleigh or Rice distributions. Finally, path-loss is given by the transmitter-receiver distance d, the path-loss exponent α and a factor K which accounts for the impact of the environment, the transmission frequency and the transceivers height. Numerous models have been developed over the years to compute K, α and the standard deviations of s and r in typical wireless environments, e.g. urban macro-cells and micro-cells, suburban cells and also indoors scenarios. However, the main drawback of such approach is that the proposed models are not always consistent with each other.

Existing channel models: The Okumura model, and then the Hata model, are some of the earliest works on propagation models for large urban macro-cells based on extensive measurement campaigns and giving closed-form formulations for path-loss [64]. Although they have

provided good models for first-generation cellular systems, they are not adapted to current networks, that operate at higher frequencies and over smaller cell sizes (eventually indoors).

Subsequently, standardized channel models have been developed for the purpose of simulation and calibration for 3G and 4G systems. The four main models are the European Cooperative for Scientific and Technical model (COST-231) of [67], which extends the Hata model, the 3GPP spatial channel model (SCM) of [68], the WINNER II model of [69] and the IEEE 802.16j channel model of [70], which partly reuses the WINNER II model.

Thanks to its accuracy and flexibility, the WINNER II model is often referred as the most comprehensive state-of-the-art in wireless channel models [71]. This model includes 13 scenarios for indoor, urban as well as rural environments with both LOS ans NLOS cases for each of them. Moreover, the WINNER II model includes scenario specific to fixed relay stations (B5 stationary feeder scenario), and further divides into sub-scenarios, where the relay is located from street-level to rooftop.

Considered models for relay placement: As emphasized in [71], a large part of theoretical results on relay placement are based on oversimplified channel models, implying that some algorithms proposed in the literature can be inappropriate for practical relay placement. For example, works in [53, 56, 57, 58, 59, 60, 61, 72] consider AWGN channels with only path-loss exponent or assume the same model for all links in the network. Some research work, e.g. [59, 61], even consider linear models, where the relay is located on a straight line between the source and the destination. So far, little analysis has considered three-dimensional geometrical models, accounting for the impact of the relay height. Indeed, locating the relay station closer to the user or, on the contrary, at high height offers significantly different propagation environment, as illustrated in Figure 2.3.

Understanding how the wireless radio propagation characteristics affect the relay deployment opens new perspectives and provides an approach complementary to designing algorithms for optimal placement as existing in the literature.

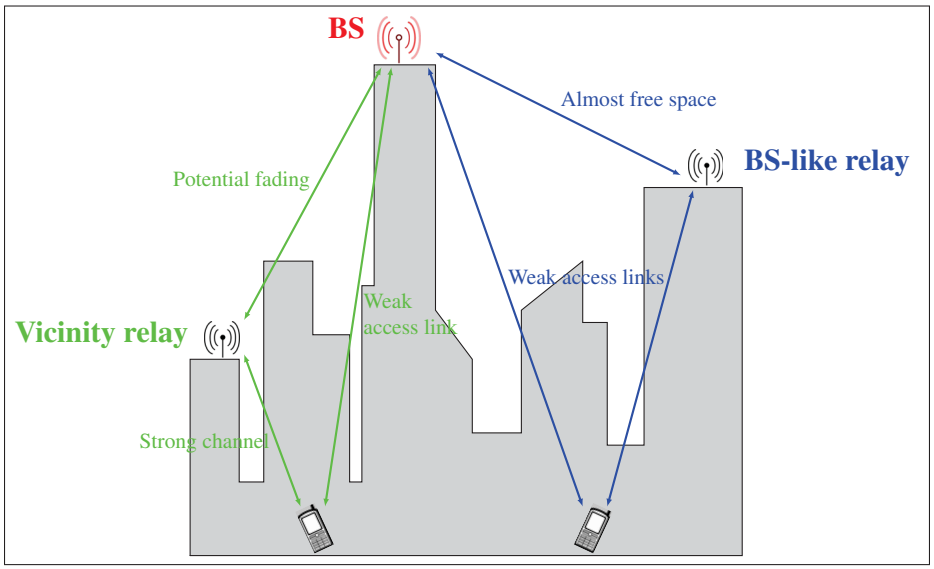


Figure 2.3 Impact of the relay height on the propagation environment

2.2.3 Deploying relay stations in urban environments

This subsection introduces the process for wireless antenna tower site selection and deployment, as described in [73]. Such discussion may first look as a digression to our topic. However, being aware of such procedure highlights the complexity of balancing technological requirements, capital expenditure, authorities and affected populations. Deploying a new antenna structure raises concerns in a community with regards to aesthetics, property, but also health care, due to the exposure to electromagnetic fields. This is particularly true for relay stations, which are connected to the main network via a wireless backhaul link.

The operator's point of view: The operator's strategy to locate a new antenna site, for example a relay station, accounts for both technical and performance requirements and capital expenditure, which includes the infrastructure cost, the property values, the access facilities for maintenance and so on. Sharing or colocating a new antenna on an existing tower is usually preferred by operators, since it is considerably more cost effective, with reduced time to service and smaller impact on neighbourhood. However, finding an existing structure with the suitable location or height may not always be feasible, thus demanding for a free-standing antenna system (built for example on a dedicated tower or mast).

The municipality's point of view: Consultation of local communities potentially affected is one of the most important part of the antenna siting process, and results in an information package for the operator including preferred and discouraged locations, as well as design and landscaping requirements or environmental impact statements. For example, industrial, agricultural and commercial areas are usually preferred for antenna siting, rather than residential areas. Proximity to transportation and utility corridors, or institutions necessitating improved telecommunication access, such as emergency services or university, is also considered. On the contrary, the municipality may declare inappropriate ecologically sensitive areas, historical sites or locations affecting aesthetics (e.g. in front of doors, windows, balconies or residential frontages). Although colocating antennas on existing towers may be practical and cost efficient for an operator, this comes in return with more massive structures, with additional equipment shelters at the base of the tower, such that some municipality may discourage such colocation.

The Government's point of view: In Canada, the final decision to approve and license the location of a new antenna system is made by the Federal Government, which guarantees compliance with a standard protocol. In this, various rules on the station deployment are prescribed about, for example, safety distances, minimal and maximal heights, but also the physical dimensions of the antenna and associate equipment. Moreover, Industry Canada ensures that the new antenna system respects all the radio-frequency emission safety levels (Health Canada's Safety Code 6 [74]). Its role includes as well ensuring that local authorities and nearby populations are consulted throughout the negotiation of a new antenna site. In case of litigation between the different actors involved in the process, the Federal Government has exclusive and comprehensive jurisdiction over the wireless antenna tower siting and may potentially render inoperative municipal site-plan regulation and local restrictive by-laws [73, 75].

Searching for optimal relay location ? On one side, and even if practical relay deployment is done on a case-by-case basis, theoretical analysis and simulations are meaningful. They can give insights on relay configurations which improve the network performance and help setting basis for future site negotiation. On the other side, numerous algorithms for relay deployment propose theoretically optimal locations which may not be feasible due to the urban topography

and administrative by-laws. It is thus critical to investigate the performance of a relay station when not optimally located and to look for transmission techniques which can offer more deployment flexibility while ensuring satisfying performance.

2.2.4 Thesis methodology and contribution

As second step in this thesis, we investigate relay deployment with regards to the channel propagation characteristics and the relaying coding scheme. To propose meaningful analysis without requiring time-consuming extensive simulations, we develop a geometrical model describing the cell area where relaying is more energy-efficient than direct transmission. Such an area is illustrated in Figure 2.4, in comparison with other patterns for the relay serving area.

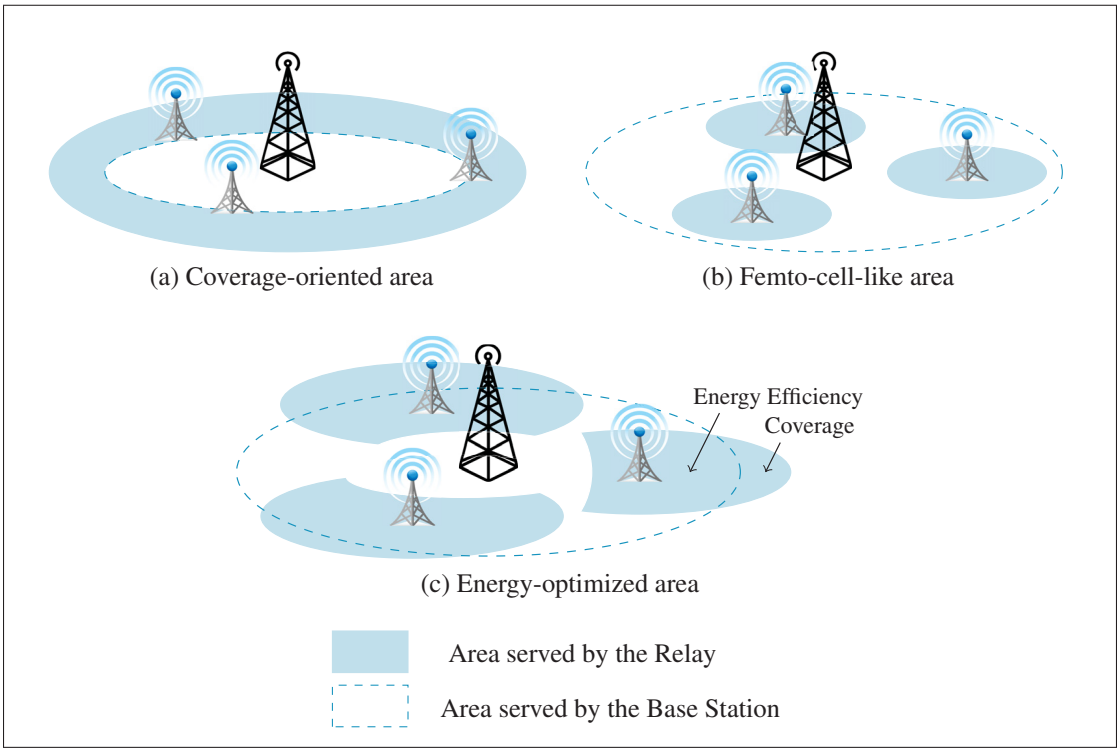


Figure 2.4 Relay serving areas for difference performance criteria

The proposed model allows to predict the average energy consumed to transmit data at a given rate, for a given relay location. Yet, it only needs a few characteristic distances and is valid for various propagation environments. Then, using the proposed model, we highlight new

trade-offs which balance the coverage extension, the energy consumption and the deployment flexibility. We show that advanced coding schemes can easily overcome harsh radio environments due to suboptimal location, and thus provide much flexibility to the relay deployment.

This second step is described in Chapter 5, which is an edition of [29]. The proposed trade-offs are further analyzed in our conference paper [33]. In this, the issue of the number of relay stations is addressed, considering their circuitry energy consumption and their economic cost.

2.3 Part III: Reducing the energy consumption and interference: an equivalence ?

So far, the literature review has focused on the different relaying techniques in networks composed of only three nodes: the user, the base station and the relay. We now introduce the possibilities and challenges of relaying in multi-cell networks, where many base stations, along with relays and other access points, co-exist in the same area and share the available radio resources. In particular, we identify the issue of interference in multi-cell relay-aided networks optimized for energy saving.

2.3.1 Challenges of multi-cell networks

Multi-cell networks can be classified in three distinct domains as follows [15]:

- **Homogeneous networks**, which consist only of large scale macro-base stations operating over licensed spectrum and managed by operators;
- **Heterogeneous networks (HetNets)**, which are composed of multiple tiers of small-scale micro, pico and femto-cells. Each of them differs with regards to transmit power level, hardware attributes and signal processing capability and is overlaid by a primary tier of large-scale macro-cells. HetNets help enhancing the network capacity or closing the coverage holes in areas with high traffic demand, e.g. hotspots.
- **Cooperative networks**, where several nodes of the same tier and/or different tiers (base stations, relays...) pool their resources -in time, frequency, power but also antennas and computing facilities- to jointly schedule transmissions and provide capacity, coverage and/or energy gains.

Due to their practicality and backward compatibility with previous generations, multi-cell networks are considered for implementation in next cellular standards, such as LTE-A [76]. Yet, a number of important new challenges have emerged from such technologies: How to allocate resource and schedule users ? When to proceed hand-off ? How to mitigate interference ? But also which metric should be considered to guarantee satisfying QoS for each user and good performance of the network in general ?

Techniques for energy savings in HetNets and in cooperative networks are radically different. On one side, small cells are mostly user-deployed, with possibly restricted access, and are designed for point-to-point access. Energy-oriented techniques thus focus on efficient hand-offs, scheduling and resource management, in particular frequency allocation for interference avoidance. On the other side, cooperative communications are performed using more advanced schemes which exploit the spatial diversity offered by the joint transmission of neighbouring nodes (base stations, relays or others cooperative access points). By coordinating transmissions, the inter-cell interference (ICI) can be reduced and the network energy consumption minimized. Relaying is not the only cooperative technique. In addition, cooperation principally includes:

- **Multi-user MIMO (MU-MIMO):** mobile user devices and small-scale access points are usually equipped with a single antenna only, implying that traditional MIMO cannot be performed. However, by pooling antennas and adapting the precoding weights of each transmission, cooperative nodes can emulate a distributed MIMO system;
- **Coordinated Multi-Point (CoMP):** CoMP has been designed as an integral part of the LTE-A standard. As for MU-MIMO, geographically separated base stations dynamically cooperate, via optical-fiber backhaul links, to provide joint scheduling and processing of the received signal. It is particularly beneficial for cell-edge users.

Despite the significant performance enhancement they can provide, MU-MIMO and CoMP are severely affected by synchronization errors and outdated channel state information. For example, it is shown in [77, 78, 79] that a deviation of only 5ms in the feedback delay may result in a 20% capacity decrease for a downlink CoMP system. Moreover, CoMP and MU-MIMO both require substantial extra traffic load and energy consumption due to the additional

signaling, channel state information exchange and signal processing required for cooperation. For example, up to 30% of the downlink transmissions is dedicated to channel state information exchange in a CoMP system with four antennas, as shown in [80]. The performance of MIMO has been analyzed in [81, 82]. Because more antenna devices are used for transmission, the energy dissipated in circuitry may actually negate the benefits in the transmit energy and interference level, such that MIMO is not always more energy-efficient than single-antenna transmission.

Relaying thus appears as a particularly interesting alternative to implement cooperation in cellular networks. It is more robust against imperfect synchronization and delayed CSI, requires low computational and deployment costs and furthermore, does not necessitate major changes in the existing infrastructure [15]. When a user has the possibility to communicate with several relay stations, a first solution is to maximize the transmission gain by involving each of them. However, relaying would then suffer from the same drawbacks as MU-MIMO and CoMP, especially in terms of CSIT. A second option is to exploit spatial diversity by choosing the best relay. Relay selection is still an active research topic [83, 84] but it is not considered in this thesis. Rather, we propose a new perspective on single-relay transmissions by analyzing in which way an energy-oriented relay deployment may impact the inter-cell interference.

2.3.2 Analyzing interference in energy-efficient systems

Reducing interference has been largely considered as a major key to unlock energy efficiency. This subsection describes the major techniques for interference mitigation in multi-cell networks and the assumptions which are commonly considered for investigation. Based on such literature review, we highlight the necessity of designing a new framework for performance analysis in energy-optimized relay-aided networks.

2.3.2.1 Techniques for interference mitigation

The most obvious approach to reduce interference is simply to avoid it by orthogonalizing transmissions. Such approach has been spurred for next-generation cellular OFDMA systems and encompasses frequency reuse techniques, for which adjacent cells use different frequen-

cies. In particular, Fractional Frequency Reuse (FFR) is under consideration for implementation in the LTE-A standard [85, 86]. In this framework, a cell is divided in several regions, each using a different frequency reuse factor. The inner cell region has a frequency reuse factor of 1, whereas the outer region has a factor of N. For example, in the Soft Frequency Reuse pattern (as illustrated in Figure 2.5 with a reuse factor of 3), a cell-edge user receives interference only from the interior of the adjacent cell, where users have low transmit power.

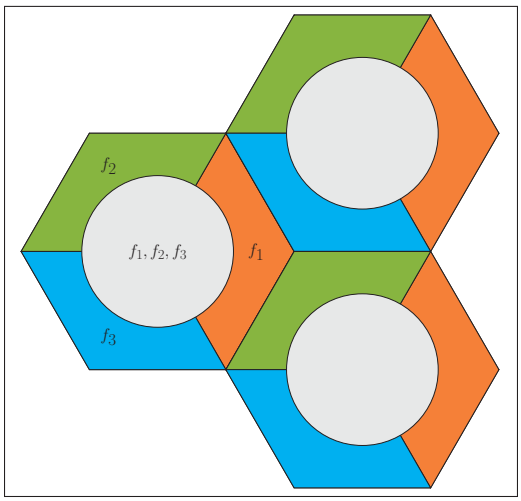


Figure 2.5 Soft Frequency Reuse, with N=3

However, such frequency reuse pattern may not be the spectrum-efficient. The well-known Shannon's capacity formula (2.1) shows that the capacity increases linearly with the available bandwidth, but only logarithmically with power. This implies that a better utilization of the available spectrum potentially provides higher performance gains than simply increasing the signal-to-noise ratio. Techniques of cognitive radio may bring partial answer to such question.

Other approaches have been proposed for interference mitigation like MU-MIMO and CoMP, as described earlier, but also user scheduling with inter-base-station coordination, dynamic antenna down-tilt and base-station switch-off when the traffic load is low. Such interference-aware techniques call for full coordination between nodes and intensive CSI exchange to accurately track the fluctuations of the time-varying wireless channel conditions, the traffic load

and QoS requirements. Moreover, it is necessary to accurately understand how the interference is distributed across the network. In the next subsection, we thus review existing models for interference analysis.

2.3.2.2 Models for interference analysis

To design practical interference-aware techniques, it is necessary to differentiate between the time scale at which they are performed:

- **Long-term techniques:** they consider the load fluctuations on a base of tens of minutes to several hours and account for the average channel behavior, i.e. the average path loss. Switching-off base stations [87, 88] is an example of such long-term techniques.
- **Short-term techniques:** they are based on smaller time granularity, on the order of few time slots (i.e. in the range of seconds or minutes). Base station switching-off is not adapted to short-time channel variations due to the transitory wake-up period, but switching-off OFDM sub-carrier only can be considered.

On one side, long-term techniques are designed to fit the large-scale traffic movements over the day, for instance the consumers going from home to work place. Modelling long-term interference is thus rather a matter of user behaviour than of technology. On the other side, short-term techniques require faster adaptation and lighter computation at user terminals and base stations. Optimal performance can only be reached by considering realistic network topology and energy distribution models, together with small-scale channel fading, i.e. path-loss, shadowing and sometimes even fast fading.

Multi-cells models for interference analysis are already well-developed concerning femtocells. They consider both regular hexagonal cellular network patterns and stochastic-geometry-based patterns, where base stations are randomly located according to a Poisson point distribution such as in [89] and references therein. Such models are however non applicable to relay-aided networks, where a cell has to manage both direct and two-hop transmissions and where the cost of the wireless backhaul has to be considered in the interference analysis. Moreover, as opposed to base station, relays are generally equipped with omni-directional antennas.

Most of existing analysis for relay-aided networks assume very pessimistic models for interference. For example in [56, 57], the signal-to-noise ratio of a reference user is too severely degraded since each neighbouring node (i.e. base stations and all relays of each cell) is transmitting at each time slot and with full transmit power. The obtained results can thus be understood for a worst case but do not reflect the performance standard of the network.

Furthermore, assuming that neighbouring stations are transmitting at a given power cannot be used for the energy minimization issue considered in this thesis. When optimized for energy, the actual consumption is generally much less than the considered power constraint, implying that the interference received in neighbourhood cells is not as high as expected. To the best of our knowledge, no analysis has been provided to describe how the transmit energy consumption is distributed across the cell and where the highest interference is generated, depending on the user location and rate, the relaying coding scheme, the cell coverage, for the specific context of energy minimization in relay-assisted cellular networks.

2.3.3 Thesis methodology and contribution

As the last step of this thesis, we characterize the impact of omni-directional relays in terms of both the achieved energy gain and the resulted performance loss in neighbourhood cells, due to the additional interference. Whereas it seems quite intuitive that reducing the transmit energy allows to mitigate the generated interference, such statement does not account for the relay potential proximity with cell-edge, as illustrated in Figure 2.6. Thereby, it is not clear whether or not an energy-efficient relay deployment is also interference-aware.

To answer this issue, we refine our previous geometrical models to include the effect of shadowing and the extra energy dissipated in the circuitry for signal processing and network maintenance. We then propose new definitions for the relaying efficiency which allow the prediction of the probability of energy-efficient relaying, the overall energy consumption and the generated interference. The proposed framework has wide application and can be used for resource management (in particular frequency reuse techniques), load-balancing or base-station switch-off. In this thesis, we focus on the relay deployment. We show that energy-efficient

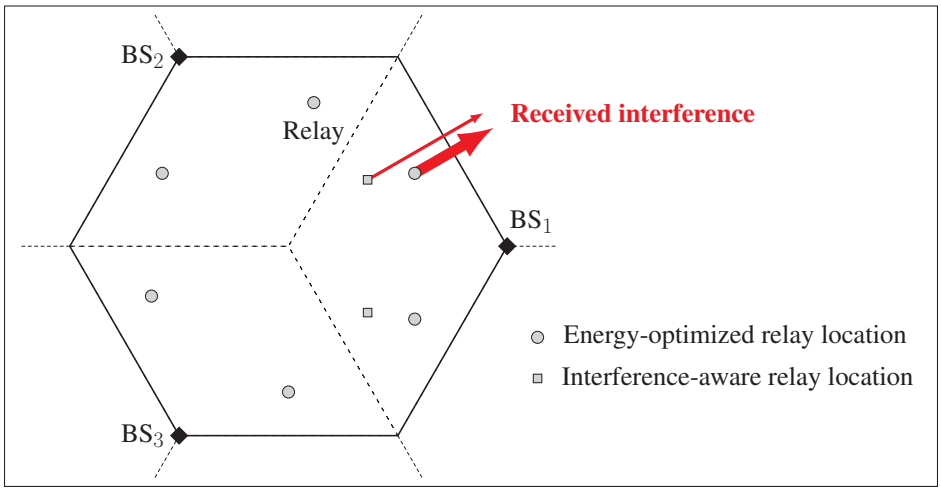


Figure 2.6 Energy-efficient vs. interference-aware relay deployment

relay deployment does not necessarily lead to interference reduction and reversely, that an interference-aware deployment is suboptimal for energy. In addition, a new performance metric is proposed to account for both aspects. This last step is covered by Chapter 6, which is an edition of [30].

CHAPTER 3

DESIGN OF A NEW HALF-DUPLEX DECODE-FORWARD RELAYING SCHEME

Summary of this Chapter: Our research work has been initiated by the design of a new half-duplex decode-forward coding scheme for energy-efficient performance analysis. In this, the source splits its data message into two parts and transmits in two phases. Both message parts are sent during both phases. At the end of first phase, the relay decodes one part, re-encodes it and forwards it in the second phase. Then, at the end of second phase, the destination decodes both parts by using joint typicality decoding. This chapter is an expanded edition of our conference paper [27], where this coding scheme has been first presented. Since information theory may be unfamiliar to the reader, the chapter also covers the main mathematical tools that have been used and the particular information-theoretic vocabulary.

Notations: Upper-case letters, e.g. X, denote random variables. Lower-case letters, e.g. x, stand for outcomes of such random variables, and calligraphic letters, e.g. $\mathcal{X}$, refer to the sets of possible outcomes. $|$ denotes conditionality (for probabilities, entropy or mutual information).

3.1 The relay channel in Information Theory

In order to set up the vocabulary and notations, this chapter serves to recall the main notions of the Shannon's theory: the channel, the entropy, the mutual information and the capacity. We refer the reader to [63, 90] for general introductions on information theory, and to [37, 46, 91] for further detailed analysis.

3.1.1 What is a channel ?

Data transmission is performed over a channel that is supported by an usually imperfect physical medium, typically an electrical wire, an optical fiber or air. Due to noise, interference from other users or inappropriate signal processing, this channel may affect in a random manner the transmitted sequence of bits. The received sequence is thus random, but hopefully has a

distribution which depends on the transmitted sequence. Using such dependence, the receiver aims at recovering the source data.

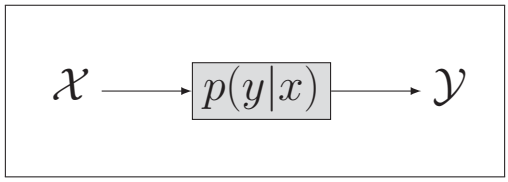


Figure 3.1 A communication channel in information theory

A **channel** is defined in information theory as a system in which the output Y depends probabilistically on its input X. More precisely, the input X and output Y are two random variables, whose values are respectively taken from the sets $\mathcal{X}$ and $\mathcal{Y}$, called **alphabets**. In practice, the alphabet refers to the modulation scheme. Then, the channel is determined by a transition matrix, which gives the conditional probability $p(y|x)$ to receive a given output $y \in \mathcal{Y}$ when $x \in \mathcal{X}$ has been sent. It should be noted that the definition of a channel is related not only to the physical medium supporting the transmission, but also to the modulation scheme. It is denoted as $(\mathcal{X}, p(y|x), \mathcal{Y})$ and is depicted in Figure 3.1.

3.1.2 What is the channel capacity ?

The **entropy** H of a random variable X, distributed according to $p(x)$, indicates how much information can be carried by this random variable. For example, an uncoded QPSK constellation with four equiprobable symbols can carry 2 bits of information, whereas a 64-QAM can carry 6 bits. The entropy is computed as:

$$H(X) = -\sum_{x \in X} p(x) \log_2 \left(p(x) \right).$$

Regarding data transmission, the entropy $H(X)$ of the channel input measures how uncertain the receiver is when it has to decide which $x \in \mathcal{X}$ has been sent by the source. However, this receiver is helped in its decision by observing the channel output $y \in \mathcal{Y}$, such that the

uncertainty is reduced to $H(X|Y)$. The average amount of uncertainty in X that is resolved by the observation of the outcome Y is called the **mutual information** and is defined as

$$I(X,Y) = H(X) - H(X|Y).$$

It refers to the amount of information that is transmitted at each channel use, given the input distribution $p(x)$ and the considered alphabet (or modulation scheme) $\mathcal{X}$. It is measured as a data rate, i.e. in bits per channel use or bits per symbols.

By modifying the structure of the input X and its distribution $p(x)$, i.e. by designing a coding scheme, we can increase the mutual information, and thereby, the data rate that can be transmitted over the considered channel. For example, in a broadcast channel where the input X contains information intended for several users ($x = (x_1, x_2...x_n)$ for n users), the sum rate can be increased by introducing dependence between the codewords X_i intended for each users and by setting $p(x) = p(x_1)p(x_2|x_1)p(x_3|x_1x_2)\dots$ rather than simply having $p(x) = \prod_i p(x_i)$. A data rate is said **achievable** if we can find how to encode data at that rate and decode with arbitrarily low error probability. We will come back to encoding and decoding in Section 3.2.

Now, by taking the right input structure and distribution $p(x)$, the mutual information can be maximized: this is the **channel capacity**, defined as

$$C = \max_{p(x)} \; I(X,Y).$$

The channel capacity can also be understood as the supremum of all achievable rates. Note that the term "capacity" in network engineering refers more to a maximum achievable rate (given a code structure, multiple access technique, scheduling policy and so on) than to the channel capacity as defined in information theory.

3.1.3 The relay channel: definition, achievable rates and capacity

The elementary relay network consists of a source S transmitting data to a destination D and supported by a third node, the relay R. In its most restricted definition, the relay does not have its own data to send and its sole purpose is to assist the source transmission.

Definition 3.1. The relay channel is defined by:

- two channel input alphabets: $\mathcal{X}$ at the source and $\mathcal{X}_r$ at the relay,
- two channel output alphabets: $\mathcal{Y}_r$ at the relay and $\mathcal{Y}$ at the destination,
- a set of distributions $p(y, y_r|x, x_r)$ describing the transition matrix of the channel.

It is depicted in Figure 3.2, where $\mathcal{W}$ and $\tilde{\mathcal{W}}$ respectively stand for the transmitted and decoded message sets, before modulation. A channel is said **memoryless** when the received sequences y and y_r depend only on the inputs x and x_r transmitted at that time and are conditionally independent of previous channel inputs or outputs.

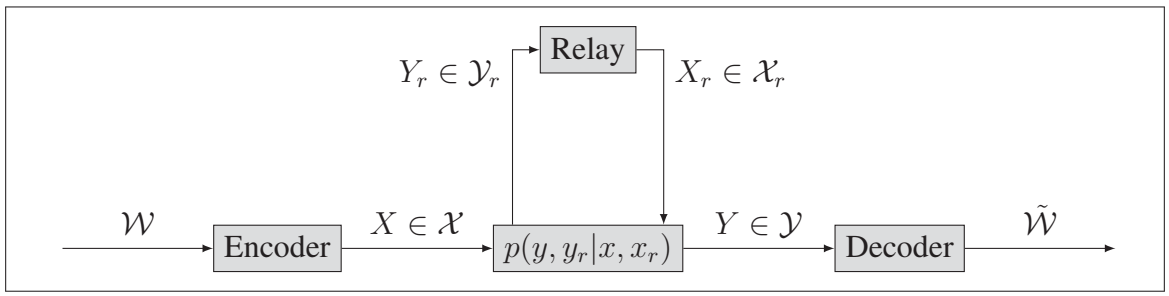


Figure 3.2 Coding for the relay channel

The half-duplex relay channel: In this thesis, we consider memoryless half-duplex relay channel and, without loss of generality, we assume time-division. Thus, the transmission is carried out in two phases of total duration T. The first phase, of duration θT with $\theta \in [0, 1]$, refers to the **broadcast phase**: the source transmits while both the destination and the relay listen. The channel is given by $p(y, y_r|x)$. The second phase, of duration $\bar{\theta}T = (1 - \theta)\, T$, refers to the **multiple access phase**: both the source and the relay transmit to the destination

and the channel is given by $p(y|x, x_r)$. For analysis, we consider a normalized transmission duration, i.e. $T = 1$. This is depicted in Figure 3.3.

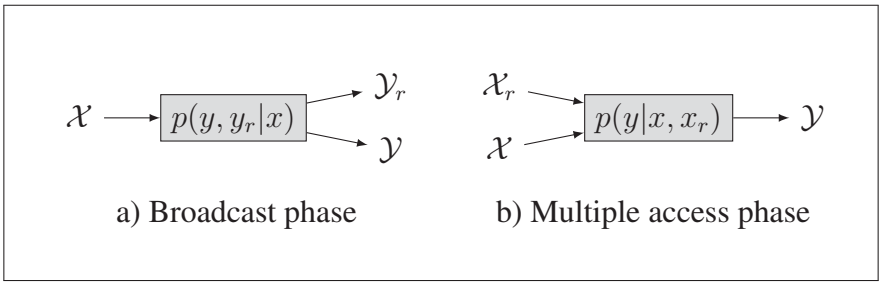


Figure 3.3 The half-duplex relay channel

Main results on the relay channel: The capacity of the relay channel is still not known in general but lower and upper bounds have been derived. Main results are summarized in [37, 48, 92]. A well-known upper-bound for the relay channel capacity is the **cut-set bound**, which directly comes from the max-flow min-cut theorem [48]:

$$C \leq \sup_{p(x,x_r)} \min \left[I(XX_r; Y), I(X; YY_r|X_r) \right]$$

Many achievable lower-bounds have been proposed, such as the Decode-Forward and Partial Decode-Forward bounds or the Compress-Forward bound. However, none of them seems to outperform the others in a general manner. A generalization of decode-forward and compress-forward has been proposed in [93].

Whereas most of presented results in information theory are for full-duplex channels, practical constraints, where the relay cannot send and receive at the same time or in the same frequency band, have been investigated in [24, 39, 94]. Larger relay networks have been widely explored as well, for instance the parallel relay channel [95], the multi-hop relay channel [96, 97, 98], the diamond relay network [99, 100], the relay broadcast channel [49, 50]. Another important reference on this topic is the work of Kramer et al. [92].

3.2 Coding for the relay channel

So far, we have presented the channel and its main properties. We now move on to encoding and decoding for successful data transmission on such channels.

3.2.1 What is a coding scheme ?

In a noisy channel, two different inputs may lead to the same output, such that data cannot be recovered at the destination. For example, we consider the simple noisy channel depicted in Figure 3.4a, with input and output alphabets $\mathcal{X} = \mathcal{Y} = \{s_1, s_2, s_3, s_4\}$ (QPSK modulation) and transition probability $p(Y = s_i|X = s_j)$ all equal to 1/2. If we consider that X takes values in the whole set $\mathcal{X}$, all bits of the considered alphabet (or modulation scheme) are data bits. The bits are said **uncoded**. However in this case, when receiving the symbol s_j, it is impossible to decide whether s_j or $s_{(j-1) \bmod 4}$ has been sent. Now, if we consider only the subset $X \in \{s_1, s_3\}$ rather than the whole alphabet $\mathcal{X}$ as input, data can be recovered without error. In this case, the symbols of the QPSK modulation scheme are **coded bits**, i.e. they contain both **data bits** (1 bit in our example) and **redundancy bits** (since a BPSK would be sufficient to transmit $\{s_1, s_3\}$ in a noiseless channel). The number of data bits per symbol or channel use is the **data rate R** (1 bit per channel use in our example), which accounts for both the modulation rate (2bits/symbol here) and the error correction rate (1/2 here).

Thus, to successfully transmit data, it is essential to identify the channel outputs that can be confused and to find the inputs that are sufficiently "far apart", in the sense that their noisy versions at the channel output are distinguishable. This is the objective of **coding**.

So far, we have considered how to recover data after only one transmission, or channel use. In practical systems, the receiver tries to recover data after n channel uses, by considering the received output $y^n = (y_1, y_2 \dots y_n) \in \mathcal{Y}^n$. Hence, the channel input and output are referred as **sequences** and coding is performed by **blocks**. By analogy with the "alphabet", the input sequences x^n are called **codewords** and the set of all codewords is the **codebook**. To connect information theory with network engineering, one block usually corresponds to one packet, and a typical block length for a wireless channel could be, for example, 1024 coded bits.

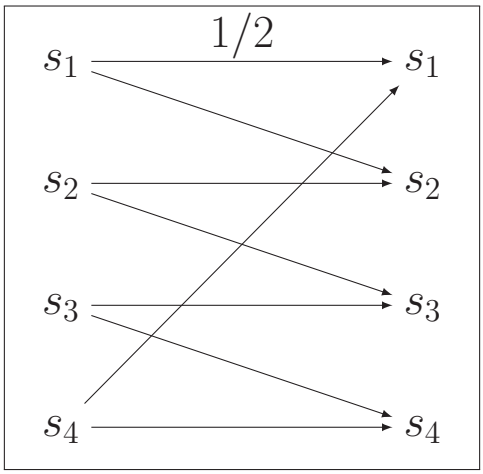


Figure 3.4 A simple noisy channel

A **coding scheme** is thus characterized by the block length n and the number of data sequences M that are encoded in one block, but also by how to generate the codebook, how to map data messages to codewords and how to decode. Now, coming back to the relay channel, we have the following definition.

Definition 3.2. A (M, n)-code for the relay channel $(\mathcal{X}, \mathcal{X}_r, p\left(y, y_r|x, x_r\right), \mathcal{Y}, \mathcal{Y}_r)$ as described in Definition 3.1 consists of:

- a message set $\mathcal{W} = \{1, 2, \ldots, M\}$,
- two encoders (at the source and the relay),
- two decoders (at the relay and the destination).

Each message from the set $\mathcal{W}$ can be represented using $\log_2(M)$ bits and is transmitted over n channel uses. The rate R of such code is defined as

$$R = \frac{\log_2(M)}{n} \quad \text{in bits per channel use.}$$

In the following, we present the coding scheme that has been proposed for the discrete memoryless half-duplex relay channel in our conference paper [27]. We describe the codebook generation, the encoding and decoding process, and finally, we derive the rate that is achieved with such scheme. This coding scheme will be considered as benchmark for energy-efficiency throughout the rest of this thesis.

3.2.2 Proposed coding scheme for the relay channel

As explained earlier, by modifying the structure of the input sequence X^n and its distribution $p(x^n)$, we can increase the data rate that can be transmitted. Thus, we propose the following scheme. To send to the destination a message m of rate R (in bits per channel use), the source divides m into two parts (m_d, m_r), with rates R_d and R_r respectively, where $R_d + R_r = R$. Note that the values for R_r and R_d will be determined in next chapter, as part of the resource allocation optimization. In such message splitting, depicted in Figure 3.5, the message m_d is directly decoded by the destination whereas m_r is intended to be relayed. Contrary to more classical partial decode-forward schemes where only m_r is transmitted during Phase 1, both message parts are sent during both transmission phases.

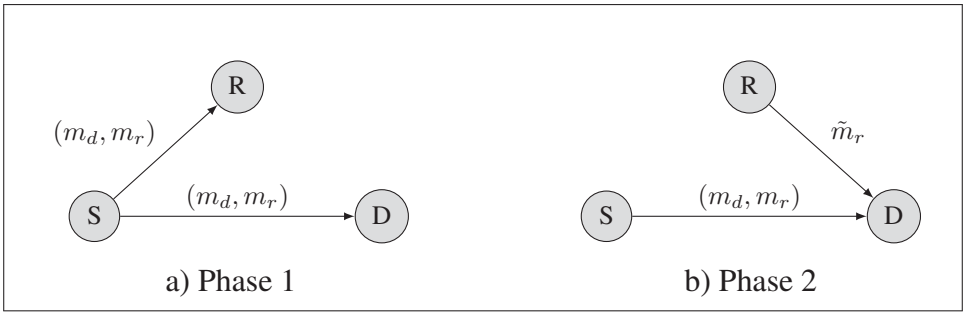


Figure 3.5 A general view of the proposed coding scheme

3.2.2.1 Codebook generation and encoding technique

In this thesis, we consider for analysis random coding techniques as initially formulated in the Shannon's theory, rather than lattice or trellis codes [101]. First, we define the joint probability distribution $p(x, u, x_r)$ and fix $p(x, u, x_r) = p(x_r)p(u|x_r)p(x|u, x_r)$, where X and X_r are two random variables, with outcomes respectively taken from the source and the relay alphabets $\mathcal{X}$ and $\mathcal{X}_r$, and where U is an auxiliary random variable.

Second, the codewords for each message parts are randomly generated as depicted in Figure 3.6. We generate 2^{nR_r} i.i.d. sequences $x_r^n(m_r) \sim \prod_{i=1}^{n} p(x_{r_i})$, where each sequence encodes one message m_r at the relay. Next, for each x_r^n, we generate one sequence $u^n(m_r)$

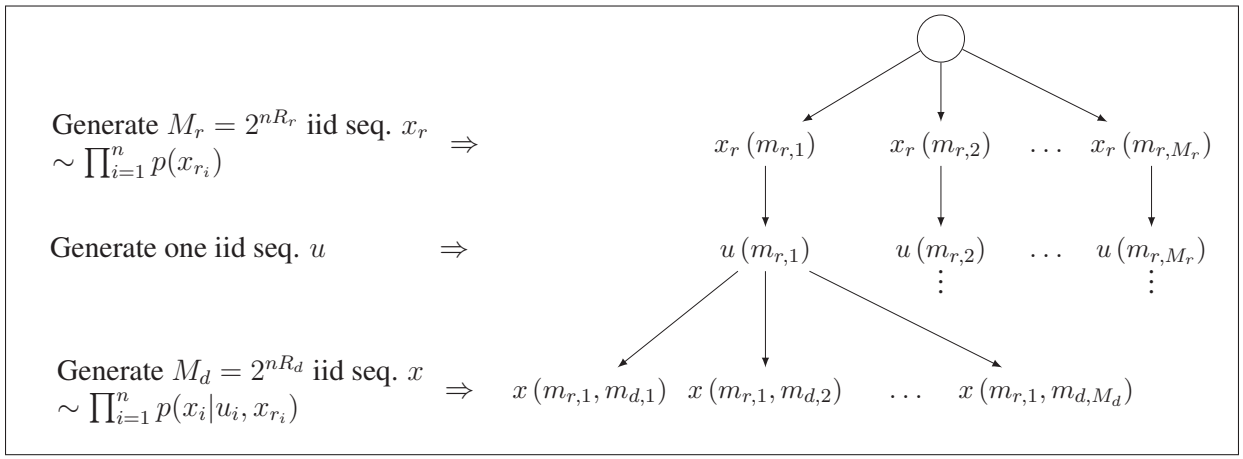


Figure 3.6 Codebook generation

$\sim \prod_{i=1}^{n} p(u_i|x_{r_i})$, that encodes m_r at the source. This ensures correlation between U^n and X_r^n, which encode for the same message part m_r. Finally, for each pair $(u(m_r), x_r(m_r))$, we generate 2^{nR_d} sequences $x^n(m_d, m_r) \sim \prod_{i=1}^{n} p(x_i|u_i, x_{r_i})$. Note that the final rate is equal to $R = R_d + R_r$ as desired. In this procedure, m_r is encoded first because the relay does not decode m_d. If encoded in second, information on m_r would be only contained in X^n, such that the relay would have to decode $U^n(m_d)$ first.

Third, we proceed to mapping and determine the transmission process. To send a message $m = (m_d, m_r)$, the source encoder maps it to the codeword $x^n(m_d, m_r) \in \mathcal{X}^n$. For convenience, we respectively denote $x^{\theta n} = [x_1, x_2 \dots x_{\theta n}]$ and $x^{\bar{\theta} n} = [x_{\theta n+1}, x_{\theta n+2} \dots x_n]$ the codewords sent over the first phase (of length θn) and over the second phase (of length $\bar{\theta} n$). During the first phase, the source sends $x^{\theta n}$, while the relay listens. At the end of the first phase, the relay decodes $\tilde{m}_r$ from the received signal and re-encodes it into the codeword $x_r^n(\tilde{m}_r)$. Then, during the second phase, the source sends $x^{\bar{\theta} n}$, while the relay sends $x_r^{\bar{\theta} n}$.

3.2.2.2 Decoding technique

At the receivers (relay and destination), the optimal decoding strategy to minimize the error probability is the maximum likelihood decoding. In this approach, the receiver picks the a

posteriori most likely message $\tilde{m}$, i.e.

$$p\left(Y^n|X^n(\tilde{m})\right) \geq p\left(Y^n|X^n(m_i)\right) \quad \forall m_i \neq \tilde{m} \in \mathcal{W}.$$

However, this decoding strategy is difficult to analyze. Instead, we consider the joint typicality decoding, which allows simpler analysis of the achievable rate and is asymptotically optimal (when $n \longrightarrow \infty$).

Joint typicality decoding: With regards to a single random variable $X \in \mathcal{X}$ distributed according to $p(x)$, a sequence of length n is said **typical** if the empirical entropy $-\frac{1}{n}\log_2 p(x^n)$ is approaching the true entropy $H(X)$. An important property is that the set of all typical sequences has "nearly" $2^{nH(X)}$ elements, all being "nearly" equiprobable. Considering typical sequences can be understood as assuming "sufficiently random" sequences, by avoiding "very particular" ones, such as all 1's or all 0's. When n becomes large, the probability to pick a "very particular" sequence tends to 0 (we will come back later on this property, called the AEP).

Similarly, we define the typicality set of a pair of random variables (X, Y), or **joint typical set**, as follows [46].

Definition 3.3. The joint typical set $\mathcal{A}_\epsilon^n$ with respect to the probability distribution $p(x, y)$, is the set of all sequences $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ for which, after n channel uses,

$$\begin{cases} |-\frac{1}{n}\log_2 p(x^n) - H(X)| \leq \epsilon, & \text{i.e. } x^n \text{ is typical} \\ |-\frac{1}{n}\log_2 p(y^n) - H(Y)| \leq \epsilon, & \text{i.e. } y^n \text{ is typical} \\ |-\frac{1}{n}\log_2 p(x^n, y^n) - H(X,Y)| \leq \epsilon. \end{cases}$$
$$\text{with} \quad p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$$

Note that there are about $2^{nH(X)}$ typical x^n sequences and about $2^{nH(Y)}$ typical y^n sequences, but there are only about $2^{nH(X,Y)}$ jointly typical (x^n, y^n) sequences, meaning that not all pairs of typical x^n and y^n are jointly typical. **Joint typical decoding** is based on this property. When observing a typical sequence y^n at the channel output, a receiver decides which one between

m_1 or m_2 has been sent by determining which codeword between $x^n(m_1)$ or $x^n(m_2)$ is jointly typical with y^n. This is depicted in Figure 3.7.

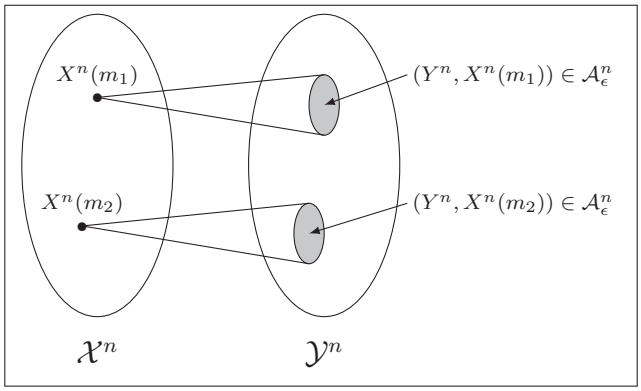


Figure 3.7 Joint typical decoding

Proposed decoding technique: After phase 1, the relay chooses the unique m_r such that

$$\left(Y_r^{\theta n}, U^{\theta n}\left(m_r\right)\right) \in \mathcal{A}_\epsilon^{\theta n}.$$

If no message $m_r \in \mathcal{W}$ or more than one message satisfy this condition, an error is declared.

Then, at the end of phase 2, the destination performs joint typicality decoding using the signal received during both transmission phases. Given the received sequence $\left[Y_1^{\theta n}\ Y_2^{\bar{\theta} n}\right]$, it chooses the unique set (m_d, m_r), such that

$$\begin{aligned} \left(Y_1^{\theta n}, U^{\theta n}\left(m_r\right), X^{\theta n}\left(m_d, m_r\right)\right) \in \mathcal{A}_\epsilon^{\theta n} \quad \text{and} \\ \left(Y_2^{\bar{\theta} n}, U^{\bar{\theta} n}\left(m_r\right), X^{\bar{\theta} n}\left(m_d, m_r\right), X_r^{\bar{\theta} n}\left(m_r\right)\right) \in \mathcal{A}_\epsilon^{\bar{\theta} n} \end{aligned}$$

where both line describes the decoding conditions for the first and the second transmission phases respectively. If no pair (m_d, m_r) or more than one satisfy those conditions, the destination declares an error.

3.3 Achievable rates with the proposed scheme

We first give an insight of the maximal rate that is achieved with the proposed coding scheme by looking once more at Figure 3.7 and by drawing an analogy with sphere packing. If there are few messages m_i (i.e. low data rate), the sets of jointly typical sequences, plotted in grey, are well separated and data can be decoded easily. However in this case, many sequences of $\mathcal{Y}^n$ remain unused, meaning that the data rate can be increased. On the contrary, when the number of messages m_i is too high, the grey sets start to overlap, meaning that there exists sequences y^n which can be generated out of several codewords. To determine the maximal achievable rate, we first list the potential error events. We then compute the error probability and deduce the rate bounds.

3.3.1 Analysis of the error events

Since the message inputs (m_r, m_d) are equiprobable, we assume without loss of generality that $(1,1)$ has been sent. In a general manner, we define the event E_i as given in [46] by:

$$E_i = \{(Y^n, X^n(i)) \in \mathcal{A}_\epsilon^n\},$$

meaning that the i^{th} codeword $X^n(i)$ and Y^n are jointly typical, where Y^n is the sequence that resulted from transmitting the first codeword $X^n(1)$ over the channel. Using such notation, we describe the error events with regards to our coding scheme, as in [27]. At the destination, three types of errors can be defined:

- **Error type A:** $(1,1)$ is not jointly typical with the signal received at the destination, i.e. the received sequence is unrelated to the transmitted sequence. We denote E_0^c such error event, where subscript ^{c} stands for complementary. Separating the two transmission phases, we have $E_0^c = E_1^c \cup E_2^c$ with

$$E_1^c = \left\{\left(Y_1, U^{\theta n}(1), X^{\theta n}(1,1)\right) \notin \mathcal{A}_\epsilon^{\theta n}\right\}$$
$$E_2^c = \left\{\left(Y_2, U^{\bar{\theta} n}(1), X^{\bar{\theta} n}(1,1), X_r^{\bar{\theta} n}(1)\right) \notin \mathcal{A}_\epsilon^{\bar{\theta} n}\right\}$$

- **Error type B:** m_r is decoded erroneously, whatever the decision on m_d. This error event is denoted as $E_r = E_3 \cap E_4$ where, for some $\hat{m}_r \neq 1$ and any $\hat{m}_d$,

$$E_3 = \left\{ \left(Y_1, U^{\theta n}\left(\hat{m}_r\right), X^{\theta n}\left(\hat{m}_d, \hat{m}_r\right) \right) \in \mathcal{A}_\epsilon^{\theta n} \right\}$$
$$E_4 = \left\{ \left(Y_2, U^{\bar{\theta} n}\left(\hat{m}_r\right), X^{\bar{\theta} n}\left(\hat{m}_d, \hat{m}_r\right), X_r^{\bar{\theta} n}\left(\hat{m}_r\right) \right) \in \mathcal{A}_\epsilon^{\bar{\theta} n} \right\}$$

- **Error type C:** the destination does not decode correctly m_d, given that m_r is correct. The error event is referred as $E_d = E_5 \cap E_6$ where, for some $\hat{m}_d \neq 1$,

$$E_5 = \left\{ \left(Y_1, U^{\theta n}\left(1\right), X^{\theta n}\left(\hat{m}_d, 1\right) \right) \in \mathcal{A}_\epsilon^{\theta n} \right\}$$
$$E_6 = \left\{ \left(Y_2, U^{\bar{\theta} n}\left(1\right), X^{\bar{\theta} n}\left(\hat{m}_d, \right), X_r^{\bar{\theta} n}\left(1\right) \right) \in \mathcal{A}_\epsilon^{\bar{\theta} n} \right\}$$

Similarly, we have at the relay:

- **Error type D:** $\hat{m}_r = 1$ is not jointly typical with the received signal. Such error event is denoted as $E_{0,R}^c$, where $E_{0,R}^c = \left\{ \left(Y_r^{\theta n}, U^{\theta n}\left(1\right) \right) \notin \mathcal{A}_\epsilon^{\theta n} \right\}$
- **Error type E:** m_r is decoded erroneously. This error event is denoted as $E_{r,R}$ where, for some $\hat{m}_r \neq 1$, $E_{r,R} = \left\{ \left(Y_r^{\theta n}, U^{\theta n}\left(\hat{m}_r\right) \right) \in \mathcal{A}_\epsilon^{\theta n} \right\}$

3.3.2 Derivation of rate bounds

Thanks to the definition of error events, we can deduce the rate bounds that must be satisfied to reach arbitrarily low error probability. Denoting $\mathbb{P}$ a probability, the overall probability that an error occurs while $(1, 1)$ has been sent is equal to:

$$\mathbb{P}_e \leq \mathbb{P}\left(E_0^c\right) + \mathbb{P}\left(E_{0,R}^c\right) + \mathbb{P}\left(E_d\right) + \mathbb{P}\left(E_r\right) + \mathbb{P}\left(E_{r,R}\right).$$

Error types A and D: We show that error events E_0^c and $E_{0,R}^c$ have arbitrarily low probability by using the Asymptotic Equipartition Property (AEP), formulated by Shannon in his original 1948 paper [45]. It is the equivalent of the law of large numbers. More precisely, for our proof,

we consider the joint AEP, i.e. the AEP for a pair of random variables, which states that

$$-\frac{1}{n}\log p\left(X^n, Y^n\right) \underset{n\to\infty}{\longrightarrow} H(X,Y) \quad \text{almost surely,}$$

for two sets of random variables $X^n = (X_1, X_2 \dots X_n)$ and $Y^n = (Y_1, Y_2 \dots Y_n)$ i.i.d. and distributed according to $p(x, y)$. It is equivalent to say that the probability to get a given realization for (X^n, Y^n), where Y^n is the noisy version of X^n, is close to $2^{-nH(X,Y)}$ with high probability, meaning that $(X^n, Y^n) \in \mathcal{A}_\epsilon^n$ almost surely, given the definition of $\mathcal{A}_\epsilon^n$ in Definition 3.3.

Thus, by the joint AEP, the probability of events E_1, E_2 and $E_{0,R}^c$ tend to zero when the block length n is increasing and $(1, 1)$ (respectively 1) is almost surely jointly typical with the signal received at the destination (respectively at the relay):

$$\mathbb{P}\left(E_0^c\right) \leq \epsilon \quad \text{and} \quad \mathbb{P}\left(E_{0,R}^c\right) \leq \epsilon \quad \text{for some } \epsilon \text{ and } n \text{ sufficiently large.}$$

Other Error types: Then, we focus on the error types B, C and E which constitute the heart of our problem. As presented in [27], we have the following rate bounds:

$$\begin{aligned}
\mathbb{P}(E_r) \to 0 \quad &\text{if} \quad R_d + R_r \leq \theta I\left(Y_1; UX\right) + \bar{\theta} I\left(Y_2; UXX_r\right) \\
\mathbb{P}(E_d) \to 0 \quad &\text{if} \quad R_d \leq \theta I\left(Y_1; X|U\right) + \bar{\theta} I\left(Y_2; X|UX_r\right) \\
\mathbb{P}(E_{r,R}) \to 0 \quad &\text{if} \quad R_r \leq \theta I\left(Y_r; U\right)
\end{aligned}$$

We detail below the proof of one rate bound, the other bounds following the same principles. For instance, we focus on the error event $E_d = E_5 \cap E_6$, i.e. m_d is decoded erroneously given that m_r is correct. As explained in the encoding process, for each m_r, we generated 2^{nR_d} sequences of length nR_d to encode for m_d. Therefore, there are $2^{nR_d} - 1$ codewords that could be erroneously jointly typical with the received Y^n. Denoting $E_{d,i}$ each particular error event

(for each $\hat{m}_d \neq 1$), we have:

$$\begin{aligned}\mathbb{P}(E_d) &= \sum_{i=2}^{2^{nR_d}} \mathbb{P}(E_{d,i}) = \sum_{i=2}^{2^{nR_d}} \mathbb{P}(E_{5,i})\mathbb{P}(E_{6,i}) \\ &\lesssim \sum_{i=1}^{2^{nR_d}} \mathbb{P}(E_{5,i}) \times \mathbb{P}(E_{6,i}) \qquad \text{for } n \text{ large.}\end{aligned}$$

We recall that $E_{5,i} = \left\{ \left(Y_1, U^{\theta n}(1), X^{\theta n}(\hat{m}_{d,i}, 1)\right) \in \mathcal{A}_\epsilon^{\theta n} \right\}$ with $\hat{m}_{d,i}$ the i^{th} codeword. Thus,

$$\mathbb{P}(E_{5,i}) \leq \sum_{(Y_1,U,X)\in\mathcal{A}_\epsilon^{\theta n}} p(U, X, Y_1)$$

From the AEP, we know that $\mathcal{A}_\epsilon^{\theta n}$ has at most $2^{\theta n(H(Y_1,U,X)+\epsilon)}$ elements, with each member having probability "close-to" $2^{-n(H(Y_1,U,X)+\epsilon)}$. However in the case of an error, Y_1, generated out of message $(1,1)$, is independent from the codeword $X(\hat{m}_d, 1)$ with $\hat{m}_d \neq 1$ and we have in reality $p(U, X, Y_1) = p(U, X)p(Y_1|U)$. By the chain rule, we have $H(Y_1, U, X) = H(U, X) + H(Y_1|U, X)$. Therefore,

$$\begin{aligned}\mathbb{P}(E_{5,i}) &\leq 2^{\theta n(H(Y_1,U,X)+\epsilon)} 2^{-\theta n(H(U,X)-\epsilon)} 2^{-\theta n(H(Y_1|U)-\epsilon)} \\ &\leq 2^{\theta n(H(Y_1|U,X)+2\epsilon)} 2^{-\theta n(H(Y_1|U)-\epsilon)} \\ &\leq 2^{-\theta n(I(Y_1;X|U)-3\epsilon)}\end{aligned}$$

A similar analysis is done with $E_{6,i} = \left\{ \left(Y_2, U^{\bar{\theta} n}(1), X^{\bar{\theta} n}(\hat{m}_d,), X_r^{\bar{\theta} n}(1)\right) \in \mathcal{A}_\epsilon^{\bar{\theta} n} \right\}$. In this case, both U and X_r are known since m_r has been decoded correctly, but X is independent from the output Y_2. Once again, using the AEP and the chain rule, we have:

$$\begin{aligned}\mathbb{P}(E_{6,i}) &\leq \sum_{(Y_2,U,X,X_r)\in\mathcal{A}_\epsilon^{\bar{\theta} n}} p(U, X, X_r)\, p(Y_2|U, X_r) \\ &\leq 2^{\bar{\theta} n(H(Y_2,U,X,X_r)+\epsilon)} 2^{-\bar{\theta} n(H(U,X,X_r)-\epsilon)} 2^{-\bar{\theta} n(H(Y_2|U,X_r)-\epsilon)} \\ &\leq 2^{\bar{\theta} n(H(Y_2|U,X,X_r)+2\epsilon)} 2^{-\bar{\theta} n(H(Y_2|U,X_r)-\epsilon)} \\ &\leq 2^{-\bar{\theta} n(I(Y_2;X|UX_r)-3\epsilon)}\end{aligned}$$

From the expressions of $\mathbb{P}(E_{5,i})$ and $\mathbb{P}(E_{6,i})$, we deduce that:

$$\begin{aligned}\mathbb{P}(E_d) \leq \sum_{i=1}^{2^{nR_d}} \mathbb{P}(E_{5,i}) \times \mathbb{P}(E_{6,i}) &\leq \sum_{i=1}^{2^{nR_d}} 2^{-\bar{\theta}n(I(Y_2;X|UX_r)-3\epsilon)} \times 2^{-\theta n(I(Y_1;X|U)-3\epsilon)} \\ &\leq 2^{nR_d} \times 2^{-\bar{\theta}n(I(Y_2;X|UX_r)-3\epsilon)} \times 2^{-\theta n(I(Y_1;X|U)-3\epsilon)} \\ &\leq 2^{3\epsilon n} \times 2^{n\left(R_d-\theta I(Y_1;X|U)-\bar{\theta}I(Y_2;X|UX_r)\right)}\end{aligned}$$

Thus, $\mathbb{P}(E_d) \to 0$ if $R_d \leq \theta I\left(Y_1; X|U\right) + \bar{\theta} I\left(Y_2; X|UX_r\right)$.

Final rate bounds: Given the error analysis detailed above, we deduce that the total source rate $R = R_d + R_r$ should satisfy the following rate bounds [27]:

$$\begin{aligned}R &\leq \theta I\left(Y_1; UX\right) + \bar{\theta} I\left(Y_2; UXX_r\right) \\ R &\leq \theta I\left(Y_r; U\right) + \theta I\left(Y_1; X|U\right) + \bar{\theta} I\left(Y_2; X|UX_r\right)\end{aligned}$$

to be achievable for some joint distribution $p(x_r)p(u|x_r)\ p(x|u,x_r)p(y_r,y|x,x_r)$.

3.4 From Information Theory to practical Gaussian channels

So far, we analyzed the discrete memoryless channel (DMC). We now explain how we can apply such results to the more practical Gaussian case.

Model for the AWGN relay channel: While the DMC is characterized by the set of distributions $p(y, y_r|x, x_r)$, the half-duplex complex Gaussian relay channel is given by:

$$Y_r = h_s X_1 + Z_r\ ; \qquad Y_1 = h_d X_1 + Z_1 \quad \text{and} \quad Y_2 = h_d X_2 + h_r X_r + Z_2, \tag{3.1}$$

where h_d, h_s and h_r respectively stand for the complex gain of the direct link, the source-to-relay link and the relay-to-destination link and where Z_1, Z_2 and Z_r are independent AWGN with same variance N_0.

Proposed coding scheme: A practical system has limited transmit power and, at each transmission phase, each node allocates to each message a portion of its available power. We denote

P_s and P_r the power constraints at the source and the relay respectively. $P_s^{(c)}$ and $P_r^{(c)}$ respectively stand for the power consumption at the source and the relay. Then, η_1 and η_2 refer to the portions of the source power that are respectively allocated to m_d in the first phase (of duration θT) and in the second phase (of duration $\bar{\theta} T$). Similarly, we denote ρ_1 and ρ_2 the portion of the source power for m_r at each phase and ρ_r the portion of the relay power used to forward $\tilde{m}_r$ to the destination. Thus, we have:

$$P_s^{(c)} = \theta\left(\eta_1 + \rho_1\right) P_s + \bar{\theta}\left(\eta_2 + \rho_2\right) P_s \leq \ P_s \quad \text{and} \quad P_r^{(c)} = \bar{\theta}\rho_r P_r \leq \ P_r$$

We recall that, to send a message m, the source divides it into two parts (m_d, m_r), with rates R_d and R_r respectively, and maps it to the codeword $X^n(m_d, m_r)$. Here, the encoding process is based on superposition coding, as depicted in Figure 3.8-a for two QPSK constellations.

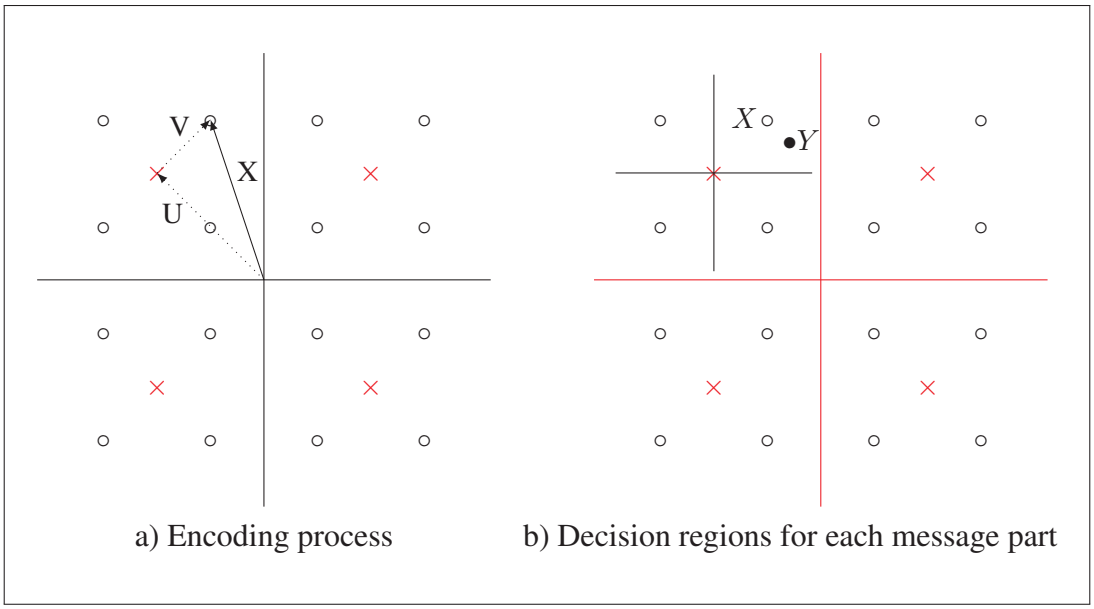


Figure 3.8 Superposition coding of 2 QPSK constellations and ML decoding

Applying the proposed scheme to the Gaussian relay channel, we generate 2^{nR_r} iid sequences $u^n(m_r) \sim \mathcal{N}(0,1)$. Then, we generate 2^{nR_d} sequences $v^n(m_r) \sim \mathcal{N}(0,1)$, independent from U^n. Taking U^n and V^n from Gaussian distributions gives Gaussian channel outputs and maximizes the achievable rate. Considering Gaussian channel inputs rather than practical modu-

lations such as QPSK or QAM defines upper-bounds for the system performance. Moreover, by the central limit theorem, the performance of an OFDM system closely approaches that of Gaussian signaling [90]. Denoting $X^n = \left[X_1^{\theta n}\ X_2^{\bar{\theta} n}\right]$, the transmit signal in the two phases is defined as

$$X_r = \sqrt{\rho_r P_r} U\ ; \qquad X_1 = \sqrt{\rho_1 P_s} U + \sqrt{\eta_1 P_s} V \quad \text{and} \quad X_2 = \sqrt{\rho_2 P_s} U + \sqrt{\eta_2 P_s} V \quad (3.2)$$

Note that the optimal X^n and X_r^n are fully correlated to allow receive beamforming gain at the destination. This scheme covers direct transmission and two-hops relaying as special cases, as well as the maximum-rate scheme proposed in [102], by simply setting to 0 some of the energy portions (η_i / ρ_j) dedicated to the transmission of m_r or m_d during the first or the second phase.

Achievable rate region: Now, we derive the rate bounds for the Gaussian case. We recall that we have for the DMC

$$\begin{aligned} R &\leq \theta I\left(Y_1; UX\right) + \bar{\theta} I\left(Y_2; UXX_r\right) \\ R &\leq \theta I\left(Y_r; U\right) + \theta I\left(Y_1; X|U\right) + \bar{\theta} I\left(Y_2; X|UX_r\right), \end{aligned}$$

and that the entropy of a complex Gaussian random variable, with zero mean and variance P, is given by:

$$\begin{aligned} h(x) &= 2\int_x p(x) \log \frac{1}{p(x)} dx \qquad \text{with} \quad p(x) = \frac{1}{\sqrt{2\pi P}} \exp\left(-\frac{x^2}{2P}\right) \\ &= \log\left(2\pi e P\right). \end{aligned}$$

For deriving rate bounds, we need to compute the variance of the received signals. Given the encoding process in Eq. (3.2) and the Gaussian channel model of Eq. (3.1), we can show that:

$$\begin{aligned} Y_r &\sim \mathcal{N}\left(0, \eta_1 P_s |h_s|^2 + \rho_1 P_s |h_s|^2 + N_0\right) \\ Y_1 &\sim \mathcal{N}\left(0, \eta_1 P_s |h_d|^2 + \rho_1 P_s |h_d|^2 + N_0\right) \\ Y_2 &\sim \mathcal{N}\left(0, \left(\eta_2 + \rho_2\right) P_s |h_d|^2 + \rho_r P_r |h_r|^2 + 2\sqrt{P_s |h_d|^2 P_r |h_r|^2 \rho_2 \rho_r} + N_0\right). \end{aligned}$$

Then, using the properties of the mutual information, we have:

$$\begin{aligned} I\left(Y_1; UX\right) &= H\left(Y_1\right) - H\left(Y_1|UX\right) \\ &= H\left(Y_1\right) - H\left(h_d X_1 + Z_1|UX\right) \\ &= H\left(Y_1\right) - H\left(Z_1\right) \\ &= \log_2\left(2\pi e\left(\eta_1 P_s|h_d|^2 + \rho_1 P_s|h_d|^2 + N_0\right)\right) - \log_2\left(2\pi e N_0\right) \\ &= \log_2\left(1 + \frac{\left(\eta_1 + \rho_1\right) P_s|h_d|^2}{N_0}\right) \end{aligned}$$

The other mutual informations can be computed similarly, from which we get the following rate bounds:

$$R \leq \theta \log_2\left(1 + \frac{\left(\eta_1 + \rho_1\right) P_s|h_d|^2}{N_0}\right) + \bar{\theta} \log_2\left(1 + \frac{\left(\eta_2 + \rho_2\right) P_s|h_d|^2 + \rho_r P_r|h_r|^2 + 2\sqrt{P_s|h_d|^2 P_r|h_r|^2 \rho_2 \rho_r}}{N_0}\right)$$

$$R \leq \theta \log_2\left(1 + \frac{\rho_1 P_s|h_s|^2}{N_0 + \eta_1 P_s|h_s|^2}\right) + \theta \log_2\left(1 + \frac{\eta_1 P_s|h_d|^2}{N_0}\right) + \bar{\theta} \log_2\left(1 + \frac{\eta_2 P_s|h_d|^2}{N_0}\right)$$

At the end of this chapter, we have proposed a new coding scheme for the half-duplex relay channel and deduced its rate bounds for achievability for both the DMC and the Gaussian channel. In next Chapter, we derive the power and rate allocations which are optimal for the energy consumption while maintaining a desired source rate.

CHAPTER 4

ENERGY MINIMIZATION FOR THE HALF-DUPLEX RELAY CHANNEL WITH DECODE-FORWARD RELAYING

Fanny Parzysz[†], Mai Vu[‡], François Gagnon[†]

[†] Département de Génie Électrique, École de Technologie Supérieure,

1100 Notre-Dame West, Montreal, Quebec, (H3C 1K3) Canada

[‡] Department of Electrical and Computer Engineering, Tufts University,

161 College Ave, Medford, MA, (02155) USA

This paper has been published

in *IEEE Transactions on Communications*.[1]

Summary of this Chapter: The Gaussian relay channel constitutes the core of this Chapter, which is an edition of our first manuscript [28]. Based on the coding scheme presented in previous Chapter, we now derive three optimal sets of power allocation, which respectively minimize the network, the relay and the source energy consumption. These optimal power allocations are given in closed-form, which have so far remained implicit for maximum-rate schemes. Moreover, analysis shows that minimizing the network energy consumption at a given rate is not equivalent to maximizing the rate given energy, since it only covers part of all rates achievable by decode-forward. Then, we combine the optimized schemes for network and relay energy consumptions into a generalized one, which then covers all achievable rates. This generalized scheme is not only energy-optimal for the desired source rate but also rate-optimal for the consumed energy. The results also give a detailed understanding of the power consumption regimes and allow a comprehensive description of the optimal message coding and resource allocation for each desired source rate and channel realization. Finally, we simulate the proposed schemes in a realistic environment, considering path-loss and shadowing. Sig-

[1] *Fanny Parzysz (2013, p.2232)*

nificant energy gain can be obtained over both direct and two-hop transmissions, particularly when the source is far from relay and destination.

4.1 Introduction

As a key feature of future wireless systems, relaying has attracted significant attention in recent years and a rich literature has contributed to analyzing the capacity of relay channels. Upper and lower bounds were first established in [48], and derived for AWGN relay channels considering energy-per-bit in [24]. More recently, bounds have been extended to multiple relays in [92] and MIMO channels in [103, 104]. Considering the three-node networks, power allocation maximizing the source rate at a given power has been extensively analyzed in [102, 105].

In practice, however, user applications are mainly associated with a fixed minimum rate, defined as a Quality-of-Service feature [106], and run on power-limited devices. This limitation has spurred new analysis and designs for energy efficiency. For example, bounds on the minimum energy-per-bit have been established in [24] and power allocations for energy efficiency of the relay channel have been derived in [107, 108]. In spite of this interest, energy efficiency analysis is mostly based on power allocations optimized for rate as in [108]. The study of coding and optimization directly for energy remain scarce. One reason is that the two problems of maximizing rate at a given network power and minimizing network power at a given rate are often thought to be equivalent. Nevertheless, neither optimization problem is fully characterized. Power allocations optimized for rate remain implicit without closed-form solutions, while power allocations for energy remain scarce and are mainly based on unrealistic constraints, such as full-duplex transmissions [107] or no individual power constraints [107, 108].

In this paper, we explore energy optimization for the three-node relay channel in Gaussian noise environments, considering Gaussian signaling as typically used in such analysis to define upper-bounds for the system performance [65]. We first propose a comprehensive half-duplex coding scheme based on time division for fading channels with Gaussian noise. We then explicitly optimize its resource allocation for energy efficiency at a desired source rate and individual node power constraints. We consider three objectives: the network, the relay and the source

energy consumption. We next combine the proposed schemes into a generalized one that maximizes the energy efficiency of the Gaussian relay channel and analyze its performance as a function of both the source rate and the mobile user location. The contributions are threefold:

- We show that maximizing the source rate and minimizing the network energy consumption are not equivalent as often believed. Specifically, there are different power-optimized regimes that cannot be uncovered by maximizing rates.
- We highlight the different regimes of coding technique and optimal power allocation (full or partial decode-forward, with or without beamforming), which constitute the lower bound on the capacity of the Gaussian relay channel by using decode-forward.
- We provide the optimal power allocation in closed-form.

The paper is organized as follows. The new half-duplex coding scheme is analyzed for the Gaussian relay channel in Section 4.2. We optimize the resource allocation for network, relay and source energy efficiency in Section 4.3 and build the generalized scheme in Section 4.4. Section 4.5 presents the performance analysis and Section 4.6 concludes this paper.

4.2 A Comprehensive Half-Duplex Decode-Forward Scheme for Gaussian Channels

4.2.1 Channel model

We consider a half-duplex channel with time division, such that the transmission is carried out in two phases within each code block of normalized length. During the first phase, of duration $\theta \in [0, 1]$, the source transmits while the relay listens. During the second phase, of duration $\bar{\theta} = (1 - \theta)$, both the source and the relay transmit. We consider complex Gaussian channels, where h_d, h_s and h_r respectively stand for the gain of the direct link, the source-to-relay link and the relay-to-destination link, as depicted in Figure 4.1. We assume that channel gain information is known globally at all nodes and consider independent Gaussian noises Z_1, Z_2 and Z_r with variance N. The half-duplex Gaussian relay channel can be written as follows

$$\begin{aligned} &\textit{Phase 1:} \quad Y_r = h_s X_1 + Z_r \quad ; \quad Y_1 = h_d X_1 + Z_1 \\ &\textit{Phase 2:} \quad Y_2 = h_d X_2 + h_r X_r + Z_2 \end{aligned} \tag{4.1}$$

where at each time either Y_r and Y_1 occur together or only Y_2 can occur.

4.2.2 Coding scheme for the Gaussian relay channel

We consider Gaussian signaling and superposition coding as in the half-duplex partial decode-forward scheme presented in our conference version [27] for the discrete memoryless and Gaussian channels. For cooperative transmission, superposition modulation benefits from a renewed interest for potential implementation. As shown in [109], this strategy outperforms other techniques, such as bit-interleaved coded modulation, at even lower receiver complexity.

We describe the coding scheme as follows. To send a message m of rate R to the destination, the source performs message splitting. It divides the initial message into two parts (m_d, m_r), with rates R_d and R_r respectively, where $R_d + R_r = R$. Both are encoded at the source using superposition coding. The message m_d is directly decoded by the destination at the end of the second phase, whereas m_r is intended to be relayed. This scheme is depicted in Figure 4.1.

The source and the relay have individual power constraints P_s and P_r within the same bandwidth. At each phase, each node allocates to each message a portion of its available transmit power. Denote η_1 and η_2 as the portion of source power P_s allocated to m_d in the first and second phase respectively, similarly, ρ_1 and ρ_2 as the portion for m_r. Denote ρ_r as the portion of the relay power P_r used to forward $\tilde{m}_r$ to the destination. We consider transmit power constraint at each node such that

$$\begin{aligned} P_s^{(c)} &= \theta\left(\eta_1 + \rho_1\right) P_s + \bar{\theta}\left(\eta_2 + \rho_2\right) P_s \leq \ P_s \\ P_r^{(c)} &= \bar{\theta}\rho_r P_r \leq \ P_r\,, \end{aligned} \tag{4.2}$$

where superscript (c) refers to the consumed power.

Applying the scheme to the Gaussian relay channel, the transmit signal can be written in two separate phases as the channel in (4.1) as follows

$$\begin{aligned} &\textit{Phase 1:} \quad X_1 = \sqrt{\rho_1 P_s}U + \sqrt{\eta_1 P_s}V \\ &\textit{Phase 2:} \quad X_r = \sqrt{\rho_r P_r}U \quad ; \quad X_2 = \sqrt{\rho_2 P_s}U + \sqrt{\eta_2 P_s}V \end{aligned} \tag{4.3}$$

where $U\,(m_r) \sim \mathcal{N}(0,1)$ and $V\,(m_d) \sim \mathcal{N}(0,1)$ are independent Gaussian signals that carry messages m_r and m_d respectively. Note that U appears in both X_r and X, which allows a beamforming gain at the destination.

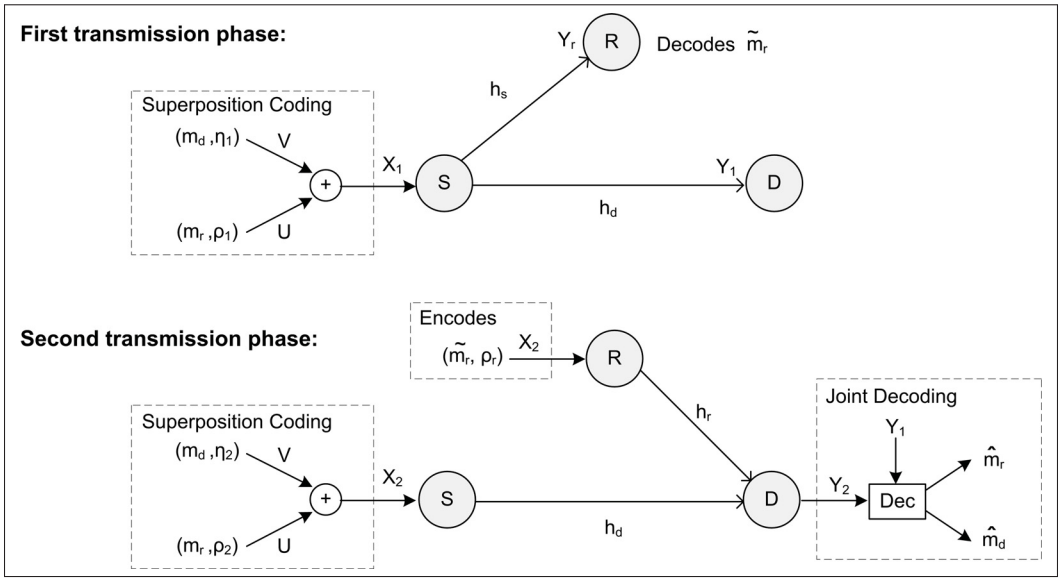


Figure 4.1 A Half-Duplex Coding Scheme for Relay Channels

Joint decoding is considered at the receivers. At the end of Phase 1, the relay decodes m_r, i.e. chooses the unique $\tilde{m}_r$ for which U and Y_r are typical. At the end of Phase 2, the destination jointly decodes (m_d, m_r) by choosing the unique $(\hat{m}_d, \hat{m}_r)$ for which U and X_1 are jointly typical with Y_1, and simultaneously U, X_2 and X_r are jointly typical with Y_2. More details on decoding can be found in our previous work [27]. Note that instead of joint typicality, maximum likelihood decoding can be used to achieve the same rate.

For Gaussian relay channels, all rates satisfying constraints in Eq.(4.4) are achievable.

$$R \leq \theta \log_2 \left(1 + \frac{(\eta_1 + \rho_1) P_s |h_d|^2}{N}\right) + \bar{\theta} \log_2 \left(1 + \frac{(\eta_2 + \rho_2) P_s |h_d|^2 + \rho_r P_r |h_r|^2 + 2\sqrt{P_s |h_d|^2 P_r |h_r|^2 \rho_2 \rho_r}}{N}\right) = I_1 \quad (4.4)$$

$$R \leq \theta \log_2 \left(1 + \frac{\rho_1 P_s |h_s|^2}{N + \eta_1 P_s |h_s|^2}\right) + \theta \log_2 \left(1 + \frac{\eta_1 P_s |h_d|^2}{N}\right) + \bar{\theta} \log_2 \left(1 + \frac{\eta_2 P_s |h_d|^2}{N}\right) = I_2 \quad (4.5)$$

In this equation, $\theta \in [0, 1]$, and power splitting factors $(\eta_1, \rho_1, \eta_2, \rho_2, \rho_r)$ are non-negative and satisfies constraint (4.2). Both rate constraints are results of the decoding at the relay over the first phase and joint decoding at the destination simultaneously over both phases. The first constraint captures the coherent transmission between source and relay in the second phase, whereas the second constraint comes from the decoding of m_r at the relay and decoding of m_d at the destination. We refer to [27] for the general proof in the discrete memoryless channel case. This coding scheme includes as special cases direct transmissions, two-hops relaying, as well as the maximum-rate scheme proposed in [102, 105].

4.2.3 Energy-optimization problem

We propose to optimize the above coding scheme for energy efficiency in the Gaussian relay channel at a given source rate. We consider three cases: the network, the relay and the source power consumptions. Specifically, we look for the optimal set of power allocation $(\rho_1, \eta_1, \rho_2, \eta_2, \rho_r)$ solving the following general problem:

$$\begin{aligned} \min \quad & \omega_s \left[\theta (\eta_1 + \rho_1) P_s + \bar{\theta} (\eta_2 + \rho_2) P_s\right] + \omega_r \bar{\theta} \rho_r P_r \\ \text{s.t.} \quad & I_1 \geq R \quad ; \quad I_2 \geq R \\ & P_s^{(c)} \leq P_s \quad ; \quad P_r^{(c)} \leq P_r \end{aligned} \quad (4.6)$$

where ω_s and ω_r are either 0 or 1, depending on the targeted optimization. The two rate constraints ensure achievability of the source rate, given two individual power constraints.

Optimization for Gaussian signaling in Gaussian channel as formulated in (4.6) provides an upper-bound on the performance of practical systems. It defines the limit of what is possible in practice, as well as gives an insight of how a practical scheme should be structured. Especially, it is found in [47] that the system that emulates most closely the structure of signals defined by information theory also reaches the highest performance. The authors of [47, 110] analyze the maximum rates allowed by information theory for the multiple-access Gaussian channel using cooperation, as well as implementation in a CDMA system. Other analysis is provided in [111] for the interference Gaussian channel in a tri-sectored OFDMA network. In this reference, an interference mitigation scheme with power allocation is proposed for Gaussian signaling and then is simulated for various practical modulations and coding rates. The practical system approaches the channel capacity with Gaussian signaling within 1.2 bit per channel use. These examples of existing works illustrate that optimization for Gaussian signaling is a valid approach for determining optimal power allocation for practical implementation.

4.3 Three energy efficient schemes for the relay channel

In this section, we explore energy optimization for the relay channel considering three objectives: the network, the relay and the source energy consumption. Then, we compare the maximum achievable rate of each optimized scheme.

4.3.1 Network energy optimal set of power allocation (N-EE)

In this optimization, we consider the total source and relay power consumption during both transmission phases. To minimize the total consumption while maintaining a desired source rate, we analyze the following problem ($\omega_s = 1$, $\omega_r = 1$), denoted as N-EE (for Network Energy-Efficient scheme):

$$\begin{aligned} \min \quad & \left[\theta\left(\eta_1+\rho_1\right)P_s + \bar{\theta}\left(\eta_2+\rho_2\right)P_s\right] + \bar{\theta}\rho_r P_r \\ \text{s.t.} \quad & I_1 \geq R \quad ; \quad I_2 \geq R \\ & P_s^{(c)} \leq P_s \quad ; \quad P_r^{(c)} \leq P_r \end{aligned} \tag{4.7}$$

The optimal power allocation depends on the desired rate. If the source rate is higher than a certain threshold $R^{(n)}$, then partial decode-forward is applied (sub-scheme $A^{(n)}$), as detailed in Proposition 4.1. Otherwise, the source does not split its message and uses full decode-forward (sub-scheme $B^{(n)}$), with power allocation as given in Proposition 4.2. Note that the power consumption of N-EE is a continuous function of the source rate, even at $R^{(n)}$. This result is summarized in Algorithm 4.1 and the proof can be found in Appendix 1, p.173.

Algorithm 4.1 Optimal scheme for network energy efficiency, N-EE, solving (4.7)

if $|h_s|^2 \geq |h_d|^2$ **then**

Find $\rho_1^\star \in [0, 1/\theta]$ satisfying $g_1(\rho_1^\star) = 0$ where

$$g_1(\rho_1^\star) = v_d^{1/\bar{\theta}} - 1 + \frac{|h_r|^2}{|h_d|^2}\left(\frac{|h_s|^2}{|h_d|^2} v_s^{1/\bar{\theta}} - 1\right) \tag{4.8}$$

$$\text{with} \quad v_s = \frac{2^R}{\left(1 + \frac{\rho_1^\star P_s |h_s|^2}{N}\right)} \quad \text{and} \quad v_d = \frac{2^R}{\left(1 + \frac{\rho_1^\star P_s |h_d|^2}{N}\right)} = \frac{2^R \frac{|h_s|^2}{|h_d|^2} v_s}{2^R + \left(\frac{|h_s|^2}{|h_d|^2} - 1\right) v_s}$$

else

if $\rho_1^\star$ *exists* **then**

$R^{(n)} \leftarrow \theta \log_2\left(1 + \frac{\rho_1^\star P_s |h_s|^2}{N}\right)$

else

$R^{(n)} \leftarrow \infty$

end

if $R \geq R^{(n)}$ **then**

Use sub-scheme $A^{(n)}$ (partial DF) given in Proposition 4.1;

else

Use sub-scheme $B^{(n)}$ (full DF) given in Proposition 4.2;

end

Use direct transmission or declare outage if not feasible

end

Proposition 4.1. In sub-scheme $A^{(n)}$, the source splits its message into two parts m_r and m_d.

- In Phase 1, the source sends m_r with power $\rho_1^\star P_s$ and rate $R_r = \theta \log_2\left(1 + \frac{\rho_1 P_s |h_s|^2}{N}\right)$. The relay decodes as $\tilde{m}_r$.
- In Phase 2, the relay sends $\tilde{m}_r$ with power $\rho_r^\star P_r$ and the source sends (m_r, m_d) with power $(\rho_2^\star P_s, \eta_2^\star P_s)$. The rate of m_d is $R_d = R - R_r$.

To compute the optimal power allocation set for $A^{(n)}$, $\rho_1^\star$ is first found numerically by solving Eq. (4.8) of Algorithm 4.1. Next, $(\rho_r^\star, \rho_2^\star, \eta_1^\star, \eta_2^\star)$ are deduced from $\rho_1^\star$ as follows:

$$\eta_1^\star = 0 \; ; \quad \rho_r^\star = \frac{N|h_r|^2}{P_r\left(|h_r|^2+|h_d|^2\right)^2}\left(\frac{2^{R/\bar{\theta}}}{\left(1+\frac{\rho_1^\star P_s|h_d|^2}{N}\right)^{\theta/\bar{\theta}}} - \frac{2^{R/\bar{\theta}}}{\left(1+\frac{\rho_1^\star P_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}}\right)$$

$$\eta_2^\star = \left(\frac{2^{R/\bar{\theta}}}{\left(1+\frac{\rho_1^\star P_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}} - 1\right)\frac{N}{P_s|h_d|^2} \; ; \quad \rho_2^\star = \frac{P_r|h_d|^2}{P_s|h_r|^2}\rho_r^\star$$

Regarding the computation of optimal allocation, solving Eq. (4.8) of Algorithm 4.1 in terms of v_s, rather than in terms of $\rho_1^\star$, reduces to solving a polynomial equation and is sufficient to deduce the whole allocation. This shows that computing the optimal allocation is relatively simple, especially in case of equal time division ($\theta = \bar{\theta} = \frac{1}{2}$) which is widely used in practical systems.

As shown in Algorithm 4.1, partial decode-forward is applied as long as $R^{(n)} \leq R \leq R_{\max}^{(n)}$. The lower bound $R^{(n)}$ comes from the non-negativity constraint on $\eta_2^\star$. The upper bound $R_{\max}^{(n)}$ corresponds to the maximum rate that is achievable given the power constraints defined in (4.2). We will come back to this upper bound in Section 4.3.4. Note that, when the direct link is close to 0, the maximum feasible rate for m_d goes to zero. Due to power constraints, outage can occur at some rate $R \leq R^{(n)}$, such that sub-scheme $A^{(n)}$ may not exist. When $R \leq R^{(n)}$, we apply sub-scheme $B^{(n)}$ as discussed next.

Proposition 4.2. In sub-scheme $B^{(n)}$, the source does not split its message m.

- In Phase 1, the source sends m with power $\rho_1^\dagger P_s$ and the relay decodes as $\tilde{m}$.
- In Phase 2, the relay sends $\tilde{m}$ with power $\rho_r^\dagger P_r$ and the source sends m with power $\rho_2^\dagger P_s$.

The optimal power allocation for sub-scheme $B^{(n)}$ is

$$\eta_1^\dagger = \eta_2^\dagger = 0 \; ; \quad \rho_1^\dagger = \left(2^{R/\theta} - 1\right) \frac{N}{P_s |h_s|^2} \; ; \quad \rho_2^\dagger = \frac{P_r |h_d|^2}{P_s |h_r|^2} \rho_r^\dagger$$

$$\rho_r^\dagger = \frac{N |h_r|^2}{P_r \left(|h_r|^2 + |h_d|^2\right)^2} \left(\frac{2^{R/\bar{\theta}}}{\left(1 + \frac{\rho_1^\dagger P_s |h_d|^2}{N}\right)^{\theta/\bar{\theta}}} - 1 \right).$$

$B^{(n)}$ is applied when the source rate is low. In this case, full decode-forward is energy-efficient and the whole message benefits from the receive beamforming gain between the source and the relay.

As an example, Figure 4.2 illustrates the power consumptions of the above two sub-schemes, which differ mainly in the second phase. Thus, the figure plots the set of optimal power allocation and the total energy consumed by the source and the relay during this phase, as a function of the source rate. When the source rate is not achievable given power constraints, an outage occurs, which is depicted by a cut-off in the curve.

4.3.2 Relay energy optimal set of power allocation (R-EE)

Next, we minimize the relay power consumption only and allow the source to consume up to P_s during both phases of the transmission. This scheme is of particular interest in networks where the relay is shared and has to serve many users, or where the relay has its own data to send. Thus, we analyze the following optimization problem, denoted as R-EE (for Relay Energy-Efficient scheme):

$$\min \quad \bar{\theta} \rho_r P_r \quad \text{s.t.} \quad \begin{array}{lll} I_1 \geq R & ; & I_2 \geq R \\ P_s^{(c)} \leq P_s & ; & P_r^{(c)} \leq P_r \end{array} \tag{4.9}$$

Once again, the optimal power allocation depends on the desired rate. Direct transmission (sub-scheme $C^{(r)}$) is used whenever feasible, i.e. as long as the source rate is under the capacity of the direct link, denoted as $R_1^{(r)}$. Note that direct transmissions can be optimal for relay energy but not for source or network energy. Indeed, the optimization problem considered here

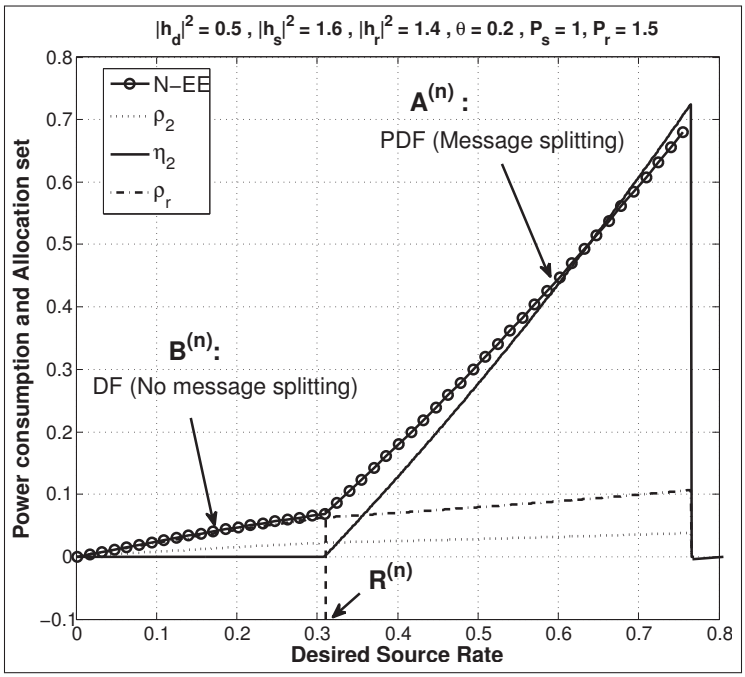


Figure 4.2 Power consumption for N-EE and allocation set during phase 2

is relaxed in terms of the source consumption. This implies that some sub-schemes can be optimal for R-EE but not for N-EE.

If direct transmission is not feasible, the optimal scheme for relay energy uses, similarly to N-EE, either partial decode-forward (sub-scheme $A^{(r)}$) or full decode-forward (sub-scheme $B^{(r)}$), depending on a certain threshold $R_2^{(r)}$. This result is summarized in Algorithm 4.2 and the proof can be found in Appendix 2, p.176.

In the following propositions, we present each sub-scheme and associated power allocation going from low to high rates. A scheme which aims at minimizing the relay energy consumption should always give priority to direct transmissions, such that, if possible, the relay is not used at all. Thus, as long as the direct link is strong enough to support the source rate, R-EE uses sub-scheme $C^{(r)}$.

Algorithm 4.2 Optimal scheme for relay energy efficiency, R-EE, solving (4.9)

Input: $R_1^{(r)} \leftarrow \log_2\left(1+\frac{P_s|h_d|^2}{N}\right)$
if $R \le R_1^{(r)}$ **then**
  Use direct transmission, called sub-scheme $C^{(r)}$, as given in Proposition 4.3;
else
  if $|h_s|^2 \ge |h_d|^2$ **then**
    Find $\rho_1^\star \in [0, 1/\theta]$ satisfying $g_2(\rho_1^\star) = 0$ where

$$g_2(s) = \frac{2^{R/\bar{\theta}}}{\left(1+\frac{sP_s|h_d|^2}{N}\right)^{1/\bar{\theta}}} - 1 + g_3(s)\left(\frac{|h_s|^2}{|h_d|^2}\frac{2^{R/\bar{\theta}}}{\left(1+\frac{sP_s|h_s|^2}{N}\right)^{1/\bar{\theta}}} - 1\right) \tag{4.10}$$

$$\text{and} \quad g_3(s) = \frac{2^{R/\bar{\theta}}}{\left(1+\frac{sP_s|h_d|^2}{N}\right)^{\theta/\bar{\theta}}} - \frac{2^{R/\bar{\theta}}}{\left(1+\frac{sP_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}} \tag{4.11}$$

    if $\rho_1^\star$ *exists* **then**
      $R_2^{(r)} \leftarrow \theta \log_2\left(1+\frac{\rho_1^\star P_s|h_s|^2}{N}\right)$
    else
      $R_2^{(r)} \leftarrow \infty$
    end
    if $R \ge R_2^{(r)}$ **then**
      Use sub-scheme $A^{(r)}$ (partial DF) given in Proposition 4.5;
    else
      Use sub-scheme $B^{(r)}$ (full DF) given in Proposition 4.4;
    end
  else
    Declare outage.
  end
end

Proposition 4.3. Sub-scheme $C^{(r)}$ is direct transmission. The optimal power allocation is as follows.

$$\eta_1^\ddagger = \eta_2^\ddagger = \left(2^R - 1\right)\frac{N}{P_s|h_d|^2}\,; \quad \rho_1^\ddagger = \rho_2^\ddagger = \rho_r^\ddagger = 0.$$

If the source rate increases above the capacity of the direct link, relaying is required and either sub-scheme $A^{(r)}$ or $B^{(r)}$ is applied, as discussed next. These sub-schemes are similar to those in N-EE but with different power allocation. In the case of R-EE, the source consumes all its available power. If $R \leq R_2^{(r)}$ as in Algorithm 4.2, we apply sub-scheme $B^{(r)}$.

Proposition 4.4. In sub-scheme $B^{(r)}$, the source uses full decode-forward.

- In Phase 1, the source sends m with power $\rho_1^\dagger P_s$. The relay decodes as $\tilde{m}$.
- In Phase 2, the relay sends $\tilde{m}$ with power $\rho_r^\dagger P_r$ and the source sends m again, with power $\rho_2^\dagger P_s$.

The optimal power allocation for sub-scheme $B^{(r)}$ is $\eta_1^\dagger = \eta_2^\dagger = 0$ and

$$\rho_1^\dagger = \left(2^{R/\theta} - 1\right) \frac{N}{P_s|h_s|^2} \; ; \quad \rho_2^\dagger = \frac{1 - \rho_1^\dagger \theta}{\bar{\theta}} \; ; \quad \rho_r^\dagger = \frac{N}{P_r|h_r|^2} \left(\sqrt{g_3(\rho_1^\dagger)} - \sqrt{\frac{P_s|h_d|^2 \rho_2^\dagger}{N}} \right)^2 .$$

Note that sub-scheme $B^{(r)}$ only exists if the channel gains are such that $R_1^{(r)} < R_2^{(r)}$. This condition essentially depends on the strength of the direct link. When $R_1^{(r)} \geq R_2^{(r)}$, sub-scheme $A^{(r)}$ is directly applied.

When the source rate increases even further to be above $R_2^{(r)}$, message splitting is required and sub-scheme $A^{(r)}$ is applied.

Proposition 4.5. In sub-scheme $A^{(r)}$, the source splits its message into two parts m_r and m_d.

- In Phase 1, the source sends m_r with power $\rho_1^\star P_s$ and rate $R_r = \theta \log_2 \left(1 + \frac{\rho_1 P_s |h_s|^2}{N}\right)$. The relay decodes as $\tilde{m}_r$
- In Phase 2, the relay sends $\tilde{m}_r$ with power $\rho_r^\star P_r$ and the source sends (m_r, m_d) with power $(\rho_2^\star P_s, \eta_2^\star P_s)$. The rate of m_d is $R_d = R - R_r$.

To compute the optimal power allocation set for sub-scheme $A^{(r)}$, $\rho_1^\star$ is first found numerically by solving Eq. (4.10) of Algorithm 4.2. Next, $g_3(\rho_1^\star)$ is computed using (4.11) and

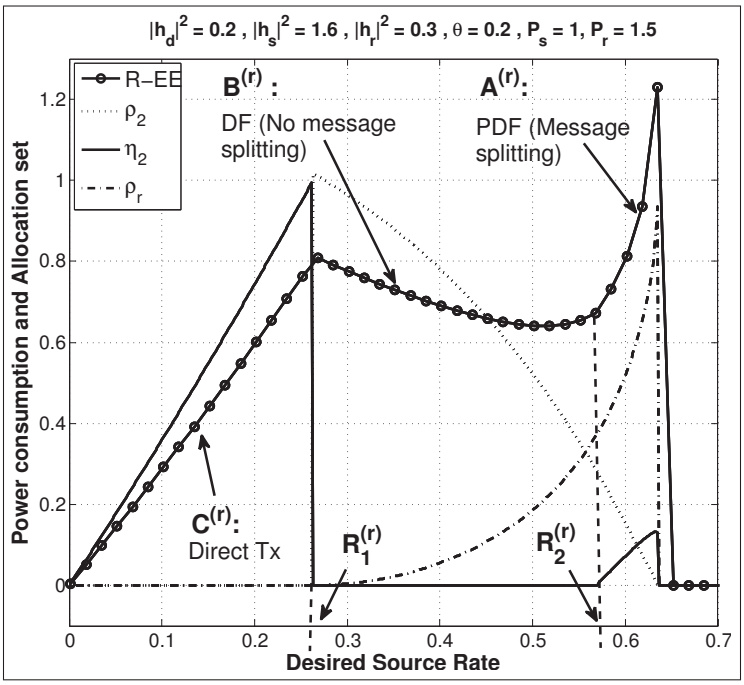


Figure 4.3 Power consumption for R-EE and allocation set during phase 2.

$(\rho_r^\star, \rho_2^\star, \eta_1^\star, \eta_2^\star)$ are deduced from $\rho_1^\star$ and $g_3(\rho_1^\star)$ as follows:

$$\eta_2^\star = \left(\frac{2^{R/\bar{\theta}}}{\left(1 + \frac{\rho_1^\star P_s |h_s|^2}{N}\right)^{\theta/\bar{\theta}}} - 1 \right) \frac{N}{P_s |h_d|^2} \; ; \quad \eta_1^\star = 0 \; ; \quad \rho_2^\star = \frac{(1 - \rho_1^\star \theta)}{\bar{\theta}} - \eta_2^\star$$

$$\rho_r^\star = \frac{N}{P_r |h_r|^2} \left(\sqrt{g_3(\rho_1^\star)} - \sqrt{\frac{P_s |h_d|^2 \rho_2^\star}{N}} \right)^2 .$$

If $\eta_i^\star$ and $\rho_i^\star$ as in Proposition 4.5 do not satisfy the power constraints in (4.2), then the desired source rate cannot be achieved and outage is declared. For example and similarly to $A^{(n)}$, when the direct link is very weak, sub-scheme $A^{(r)}$ may not exist.

Figure 4.3 illustrates the above sub-schemes and plots the total energy consumption during the second phase as a function of the source rate. Somewhat unexpectedly, the energy consumed during this second phase is decreasing in R when applying sub-scheme $B^{(r)}$. Since the destination decodes the message using the received signals during both phases, the optimization

shows that it is relay energy-optimal for the source to allocate more energy during the first phase. Nevertheless, the total energy consumed during both phases is strictly increasing with rate, as expected.

4.3.3 Source energy optimal set of power allocation (S-EE)

As the last problem, we minimize the source power consumption alone, which is relevant to networks in which the relay is not power critical. We thus allow the relay to consume up to P_r during both transmission phases. We analyze the following optimization problem, denoted as S-EE (for Source Energy-Efficient scheme):

$$\min \quad \theta\left(\eta_1+\rho_1\right)P_s+\bar{\theta}\left(\eta_2+\rho_2\right)P_s \quad \text{s.t.} \quad \begin{array}{lll} I_1 \geq R & ; & I_2 \geq R \\ P_s^{(c)} \leq P_s & ; & P_r^{(c)} \leq P_r \end{array} \tag{4.12}$$

The optimal scheme for source energy efficiency is a function of the source rate and is composed of four sub-schemes: two-hop relaying ($D^{(s)}$), decode-forward ($B^{(s)}$), partial decode-forward with spatial beamforming ($A^{(s)}$) and without beamforming($C^{(s)}$). Note that two-hop relaying and partial decode-forward without beamforming are optimal only for source energy but not for relay or network energy. These sub-schemes are applied depending on the source rate compared with two thresholds $R_1^{(s)}$ and $R_2^{(s)}$, which are closely related to the non-negativity constraints on ρ_2 and η_2. This result is summarized in Algorithm 4.3 and the proof can be found in Appendix 3, p.177. Note that the functions g_2 and g_3 of Algorithm 4.3 are the same in Algorithm 4.2.

In the following propositions, we present each sub-scheme going from low to high rates and refer to Figure 4.4 for an illustration. For very low rates, it is energy-efficient for the source to apply two-hop relaying, such that the source only consumes energy during the first phase. However, contrary to simple routing, the destination is listening during both transmission phases. This corresponds to sub-scheme $D^{(s)}$.

Proposition 4.6. Sub-scheme $D^{(s)}$ is two-hop relaying with destination simultaneously decoding over both phases.

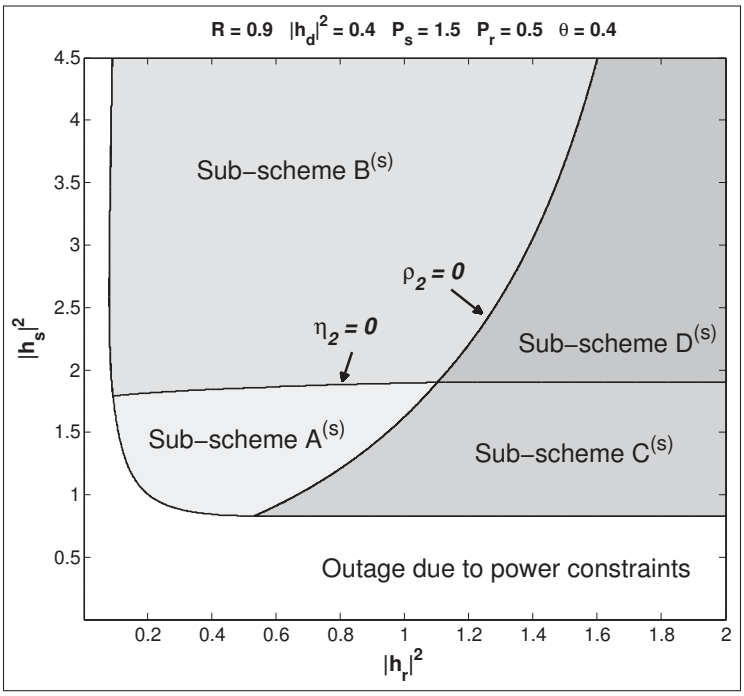


Figure 4.4 Applied sub-schemes in S-EE,
as function of the SR- and RD- links.

- In Phase 1, the source sends m, the relay decodes as $\tilde{m}$.
- In Phase 2, the relay sends $\tilde{m}$ and the source is silent.

The optimal power allocation is as follows.

$$\eta_1^\circ = \eta_2^\circ = \rho_2^\circ = 0 \; ; \; \rho_1^\circ = \left(2^{R/\theta} - 1\right) \frac{N}{P_s |h_s|^2} \; ; \; \rho_r^\circ = \frac{N}{P_r |h_r|^2} g_3(\rho_1^\circ)$$

where g_3 is defined by (4.14) of Algorithm 4.3.

If the rate increases, two-hop relaying is no longer energy-efficient for the source. In this case, either sub-scheme $C^{(s)}$ or $B^{(s)}$ is applied, depending on which of the SR- or RD-link is limiting.

If the SR-link is the bottleneck of the network, the second rate constraint $R \leq I_2$ is the limiting one and the constraint on the relay allocation is relaxed. In this case, $R_C \leq R_B$, as defined in Algorithm 4.3, and sub-scheme $C^{(s)}$ is applied.

Algorithm 4.3 Optimal scheme for source energy efficiency, S-EE, solving (4.12)

if $|h_s|^2 \geq |h_d|^2$ **then**

$$R_B \leftarrow \theta \log_2\left(1+\frac{|h_d|^2}{|h_s|^2}\left(2^{R/\theta}-1\right)\right)+\bar{\theta}\log_2\left(1+\frac{P_r|h_r|^2}{\bar{\theta}N}\right)$$

$$R_C \leftarrow \theta \log_2\left(\frac{|h_s|^2}{|h_d|^2}\right)$$

Find $\rho_1^\star \in [0, 1/\theta]$ satisfying $g_2(\rho_1^\star) = 0$ where

$$g_2(s) = \frac{2^{R/\bar{\theta}}}{\left(1+\frac{sP_s|h_d|^2}{N}\right)^{\theta/\bar{\theta}}} - 1 + g_3(s)\left(\frac{|h_s|^2}{|h_d|^2}\frac{2^{R/\bar{\theta}}}{\left(1+\frac{sP_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}}-1\right) \tag{4.13}$$

$$\text{and} \quad g_3(s) = \frac{2^{R/\bar{\theta}}}{\left(1+\frac{sP_s|h_d|^2}{N}\right)^{\theta/\bar{\theta}}} - \frac{2^{R/\bar{\theta}}}{\left(1+\frac{sP_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}} \tag{4.14}$$

if $R_B \leq R_C$ **then**

$$R_1^{(s)} \leftarrow R_B \quad \text{and} \quad R_2^{(s)} \leftarrow \theta \log_2\left(1+\frac{\rho_1^\star P_s|h_s|^2}{N}\right) \text{ or } \infty \text{ if } \rho_1^\star \text{ does not exist}$$

else

$$R_2^{(s)} \leftarrow \bar{\theta}\log_2\left(\frac{P_r|h_r|^2}{\bar{\theta}N}\right) - \bar{\theta}\log_2\left(\frac{1}{\left(1+\frac{\rho_1^\star P_s|h_d|^2}{N}\right)^{\theta/\bar{\theta}}} - \frac{1}{\left(1+\frac{\rho_1^\star P_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}}\right)$$

or ∞ if $\rho_1^\star$ does not exist and $R_1^{(s)} \leftarrow R_C$

end

case $R \geq R_1^{(s)}$ *and* $R \geq R_2^{(s)}$

Use sub-scheme $A^{(s)}$ (partial DF with beamforming) given in Prop. 4.9;

end

case $R_1^{(s)} \leq R \leq R_2^{(s)}$ *and* $R_B \leq R_C$

Use sub-scheme $B^{(s)}$ (full DF with beamforming) given in Prop. 4.8;

end

case $R_1^{(s)} \leq R \leq R_2^{(s)}$ *and* $R_B \geq R_C$

Use sub-scheme $C^{(s)}$ (partial DF without beamforming) given in Prop. 4.7;

end

case $R \leq R_1^{(s)}$ *and* $R \leq R_2^{(s)}$

Use $D^{(s)}$ (full DF without beamforming) given in Prop. 4.6;

end

else

Use the direct link or declare outage if not feasible.

end

Proposition 4.7. Sub-scheme $C^{(s)}$ uses partial decode-forward without receive beamforming.

- In Phase 1, the source sends m_r and the relay decodes as $\tilde{m}_r$.
- In Phase 2, the source sends m_d, the relay sends $\tilde{m}_r$.

The optimal power allocation for sub-scheme $C^{(s)}$ is such that

$$\eta_1^\ddagger = \rho_2^\ddagger = 0 \; ; \quad \rho_1^\ddagger = \frac{N}{P_s|h_s|^2}\left(2^R\left(\frac{|h_s|^2}{|h_d|^2}\right)^{\bar{\theta}} - 1\right)$$
$$\eta_2^\ddagger = \rho_1^\ddagger + \frac{N}{P_s}\left(\frac{1}{|h_s|^2} - \frac{1}{|h_d|^2}\right) \; ; \quad \rho_r^\ddagger = \frac{N}{P_r|h_r|^2} g_3(\rho_1^\ddagger)$$

where g_3 is defined by (4.14) of Algorithm 4.3.

Sub-scheme $C^{(s)}$ is specific to S-EE. The source transmits m_r to the relay in Phase 1, but does not send it again in the second phase. The RD-link is strong enough to carry all information contained in m_r by itself. Thus, the source saves energy by only sending new information (m_d) during the second phase, and does not require receive beamforming (which is possible if both the source and the relay transmit in the second phase). Also note that, as the RD-link is non limiting, the relay does not necessarily use all its available power to transmit $\tilde{m}_r$.

On the other hand, if the RD-link is the bottleneck of the network ($R_B \leq R_C$), the relay cannot support the transmission of $\tilde{m}_r$ alone, even by transmitting with full power. In this case, the source should repeat the message m_r during Phase 2, such that beamforming gain is obtained. Then, $B^{(s)}$ is applied as follows.

Proposition 4.8. Sub-scheme $B^{(s)}$ uses full decode-forward.

- In Phase 1, the source sends m, the relay decodes as $\tilde{m}$.
- In Phase 2, the relay sends $\tilde{m}$ and the source continues to send m.

The optimal power allocation for sub-scheme $B^{(s)}$ is $\eta_1^\dagger = \eta_2^\dagger = 0$ and

$$\rho_r^\dagger = \frac{1}{\bar{\theta}} \; ; \quad \rho_1^\dagger = \left(2^{R/\theta} - 1\right)\frac{N}{P_s|h_s|^2} \; ; \quad \rho_2^\dagger = \frac{N}{P_s|h_d|^2}\left(\sqrt{g_3(\rho_1^\dagger)} - \sqrt{\frac{P_r|h_r|^2}{\bar{\theta}N}}\right)^2 .$$

Finally, if the source rate is very high, such that $R \geq \max\left\{R_1^{(s)}, R_2^{(s)}\right\}$, partial decode-forward with receive beamforming is required and sub-scheme $A^{(s)}$ is applied.

Proposition 4.9. In sub-scheme $A^{(s)}$, the source splits its message into two parts m_r and m_d.

- In Phase 1, the source sends m_r and the relay decodes as $\tilde{m}_r$.
- In Phase 2, the relay sends $\tilde{m}_r$ and the source sends (m_r, m_d).

To compute the optimal power allocation set, $\rho_1^\star$ is first found numerically by solving (4.13) of Algorithm 4.3. Next, $g_3(\rho_1^\star)$ is computed using (4.14) and $(\rho_r^\star, \rho_2^\star, \eta_1^\star, \eta_2^\star)$ are deduced from $\rho_1^\star$ and $g_3(\rho_1^\star)$ as follows:

$$\eta_2^\star = \left(\frac{2^{R/\bar{\theta}}}{\left(1+\frac{\rho_1^\star P_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}} - 1\right)\frac{N}{P_s|h_d|^2}\;; \quad \rho_r^\star = \frac{1}{\bar{\theta}}$$

$$\eta_1^\star = 0\;; \quad \rho_2^\star = \frac{N}{P_s|h_d|^2}\left(\sqrt{g_3(\rho_1^\star)} - \sqrt{\frac{P_r|h_r|^2}{\bar{\theta}N}}\right)^2$$

This sub-scheme is similar to $A^{(n)}$ and $A^{(r)}$, but with different power allocation. Even if the relay already transmits with full power, the RD-link is still too weak to support the transmission of m_r alone. Thus, the source repeats the message m_r during the second phase to obtain beamforming gain. Moreover, contrary to $B^{(s)}$, the SR-link is also too weak to support the transmission of the whole source message. Thus, part of the source message (m_d) has also to be transmitted via the direct link only.

Existence of the sub-schemes: Considering the optimized scheme as function of the source rate, S-EE is composed of either $(D^{(s)}, B^{(s)}, A^{(s)})$ or $(D^{(s)}, C^{(s)}, A^{(s)})$, in that order, depending on which link is the bottleneck of the network. Depending on the power constraints and channel gains, S-EE can also be composed of fewer than three sub-schemes and outage can occur with any sub-scheme $D^{(s)}$, $C^{(s)}$, $B^{(s)}$ or $A^{(s)}$. Specifically, when the gains of the SD- and SR-links are close, sub-scheme $D^{(s)}$ (two-hop relaying) may not be applied at all. Indeed, in this case and even for low source rates, two-hop relaying under-uses the direct link to transmit information and is thus sub-optimal. Furthermore, given the power constraints, outage can occur before the sequences $(D^{(s)}, B^{(s)}, A^{(s)})$ and $(D^{(s)}, C^{(s)}, A^{(s)})$ are completed, such that

$A^{(s)}$, $B^{(s)}$ or $C^{(s)}$ may not be applied. We next describe the condition of existence of these sub-schemes.

First, consider the sequence $(D^{(s)}, C^{(s)}, A^{(s)})$, i.e. two-hop relaying, full decode-forward without receive beamforming and partial decode-forward with beamforming. When the SR-link is the limiting one, S-EE is composed of only $(D^{(s)}, C^{(s)})$ for most channel realizations. Partial decode-forward with beamforming ($A^{(s)}$) is only applied if the relay power P_r is very low compared to the source power P_s.

Second, we consider the sequence $(D^{(s)}, B^{(s)}, A^{(s)})$, i.e. two-hop relaying, full decode-forward with beamforming and partial decode-forward with beamforming. When the network is severely limited by both the SD- and RD-links, S-EE is composed of only $(D^{(s)}, B^{(s)})$. Recall that in $B^{(s)}$, the relay constraint is already met. Thus, if the direct link is weak given the source power constraint, the scheme rapidly goes into outage and cannot achieve high rates.

Finally, S-EE can also be reduced to only two-hop relaying ($D^{(s)}$) when the direct link is almost null.

4.3.4 Maximum rates achieved by the three optimized schemes

We now analyze the maximum rate achievable for the three optimized schemes. Formal proof can be found in Appendix 4, p.180. First, in N-EE (which minimizes the network consumption), the source power constraint is always met with equality before the relay power constraint. Therefore, at the expense of consuming more total energy, S-EE and R-EE (which respectively minimize the source and the relay consumption) allow the source to reach higher rates than the maximal rate achieved by scheme N-EE, denoted as $R_{\max}^{(n)}$.

Second, analyzing the outage S-EE sub-schemes gives a comprehensive description of the overall maximum achievable rate $R_{\max}$. Outage is declared when the power constraints cannot be met given the source rate, which can occur with any sub-scheme, as explained in Section 4.3.3. For sub-schemes $A^{(s)}$ and $B^{(s)}$, the relay power constraint is already met and S-EE goes in outage as soon as the source power constraint is met. Therefore, in this case, at $R = R_{\max}$,

both source and relay power constraints are met. On the contrary, when sub-schemes $C^{(s)}$ and $D^{(s)}$ are applied, the relay power constraint is not limiting. Thus, S-EE goes in outage only if the source power constraint is met. In this case, the maximum achievable rate $R_{\max}$ can be derived in closed-form. The values of $R_{\max}$ are described in Table 4.1. Figure 4.6 gives an example of $R_{\max}$.

Last, given power constraints and channel gains, S-EE and R-EE schemes reach the same maximum achievable rate $R_{\max}$ in Table 4.1 and have the same optimal power allocation at this specific rate.

Table 4.1 Maximum achievable rate $R_{\max}$

Channel conditions	**Outage with**	$\mathbf{R}_{\max}$
Weak SR-link and low P_r or Weak RD-link	$A^{(s)}$	I_1 in (4.4), with η_1, ρ_1, η_2 and ρ_2 of $A^{(s)}$ and $\rho_r = \frac{1}{\theta}$
Weak SD- and RD-links	$B^{(s)}$	I_1 in (4.4), with η_1, ρ_1, η_2 and ρ_2 of $B^{(s)}$ and $\rho_r = \frac{1}{\theta}$
Weak SR-link	$C^{(s)}$	$\log_2\left(\theta + \frac{\lvert h_s\rvert^2}{\lvert h_d\rvert^2}\bar{\theta} + \frac{P_s\lvert h_s\rvert^2}{N}\right) + \bar{\theta}\log_2\left(\frac{\lvert h_d\rvert^2}{\lvert h_s\rvert^2}\right)$
SD-link close to 0	$D^{(s)}$	$\theta\log_2\left(1 + \frac{P_s\lvert h_s\rvert^2}{\theta N}\right)$

4.4 A generalized scheme for energy efficiency (G-EE)

So far, we have proposed three optimized schemes N-EE, R-EE and S-EE with power allocations that respectively minimize the network, the relay and the source energy consumption. We now propose a generalized scheme that combines these optimal power allocations in a smooth way.

4.4.1 Implementing a generalized scheme (G-EE)

As shown in Section 4.3.4, N-EE sub-schemes are insufficient to describe the whole range of achievable rates and both R-EE and S-EE can increase this range.

Considering power consumption as a function of the source rate, the resource allocation that minimizes the network consumption (N-EE) is extended beyond its maximal achievable rate

$R_{\max}^{(n)}$ by the allocation that minimizes the relay consumption (R-EE) in a continuous and differentiable manner (see Appendix 5, p.182 for the detailed proof). Therefore, we propose a generalized scheme that combines these allocations and allows higher achievable rates. In this generalized scheme, denoted as G-EE, the message splitting and power allocation for minimizing the total energy consumption are applied as long as feasible, given the rate and power constraints. In particular, $B^{(n)}$ (full decode-forward) is used for source rates lower than $R^{(n)}$, and $A^{(n)}$ (partial decode-forward) is used for rates above $R^{(n)}$ and up to $R_{\max}^{(n)}$, which is computed by I_1 in (4.4) with equalized source power constraint (4.2). Once the source rate is infeasible for N-EE, the message splitting and power allocation for minimizing the relay energy consumption alone is applied, by using $A^{(r)}$ (partial decode-forward).

Algorithm 4.4 Generalized optimal scheme for energy efficiency, G-EE

```
if R ≤ R^(n) then
    Apply B^(n) (full DF for network energy-efficiency)
else
    if R^(n) ≤ R ≤ R_max^(n) then
        Apply A^(n) (partial DF for network energy-efficiency)
    else
        if R_max^(n) ≤ R ≤ R_max then
            Apply A^(r) (partial DF for relay energy-efficiency)
        else
            Declare outage
        end
    end
end
```

4.4.2 Comparison between G-EE and rate-optimal schemes

Considering rates as a function of energy, G-EE is also rate-optimal for the consumed energy in the half-duplex relay channel with decode-forward. Indeed, if a source rate achieved with G-EE is not rate-optimal for this energy, then there exists another set of power allocation that consumes the same energy but achieves a better rate R_m. For this higher rate R_m, G-EE consumes more energy that the maximum-rate scheme and is thus suboptimal for energy, which

is contradictory. Thus, for all achievable rates, G-EE is also rate-optimal. Also G-EE achieves all the rates that are achievable by the maximum-rate scheme (see below).

Considering energy as a function of rates and given fixed individual power constraints, the maximum source rate $R_{\max}$ achieved with energy minimization G-EE is equal to the maximum rate achieved by rate maximization. This comes from the fact that both G-EE and the maximum-rate scheme are optimization problems with four constraints (two rate and two power constraints). The two rate constraints are always met with equality in both problems and, with source rate up to $R_{\max}$, G-EE is equivalent to the rate-optimal scheme. When $R \to R_{\max}$, the source power constraint is also met with equality (G-EE applies R-EE). Hence, in the interior neighbourhood of $R_{\max}$, each optimization problem has just one degree of freedom left. When $R \geq R_{\max}$, an outage occurs: either the desired source rate becomes infeasible, or the relay power constraint is also met with equality. Thus, above $R_{\max}$, there is no more degree of freedom for neither problem and $R_{\max}$ is the maximal rate that is achieved for both G-EE and maximum-rate schemes.

Finally, both previous points imply that only the combined scheme G-EE is equivalent to the maximum-rate scheme. Figure 4.6 of next section illustrates this equivalence. As a reference, we consider the rate-optimal scheme proposed in [102, 105]. This scheme is similar to $A^{(n)}$ in coding: the source performs message splitting, sending one part during phase 1, and both during phase 2. Then, the power allocation is numerically optimized to maximize the achievable rate. For the comparison, we compute the maximum capacity achievable with the rate-optimal scheme based on the energy consumed using G-EE (plot with reversed axis).

4.4.3 Discussion

Even though both analysis and simulation show that the generalized scheme G-EE and the maximum-rate scheme proposed in [102, 105] lead to the same set of power allocation, we argue that the proposed scheme G-EE brings two significant advantages.

First, existing maximum-rate schemes proposed in the literature require numerical search of the optimal solution for each channel realization. On the contrary, the G-EE scheme, which is both rate- and energy-optimal, provides the optimal power allocation in closed-form.

Second, this energy efficiency approach reveals more details about sub-schemes and power allocations than the maximum-rate approach. We show that minimizing the network energy consumption is not equivalent to maximizing the network capacity as often believed, since the whole range of source rates achievable with decode-forward cannot be covered. Furthermore, we highlight that several regimes of coding technique and power allocation exist (partial or full decode-forward, with or without beamforming) and each of them corresponds to a specific energy optimization problem (N-EE or R-EE).

4.5 Simulation results

4.5.1 Description of the simulation environment

We simulate the performance of the proposed schemes in a realistic environment and consider both pathloss and shadowing. Shadowing is the most severe random attenuation factor encountered in urban cellular environment [112]. It refers to mid-term channel variations which are due to the multipath propagation generated by obstructions between the transmitter and receiver. Since blocking objects do not vary rapidly in geographical position, size or dielectric properties, shadowing remains slow and predictable. Therefore, in our case, we can reasonably assume that channel gains are constant over the two transmission phases and that nodes can track those variations at each instant.

Figure 4.5 depicts the considered scenario. In a dense urban environment with high surrounding buildings, a mobile user (source S) sends data to a base station (destination D) with the help of a fixed relay station (relay R), such as in LTE or Wimax. The mobile user is located outdoors, at street level. The relay station is below the top of surrounding building and the base station is far above the rooftop. This channel model corresponds to the Stationary Feeder scenario described in the European WINNER project [69]. Pathloss and shadowing effects are given by

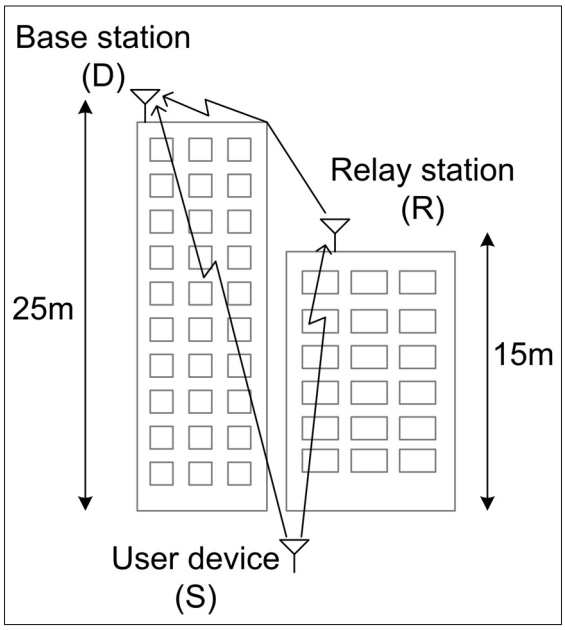


Figure 4.5 Communication in a realistic environment

(4.23) of [69]:

$$\text{PL} = A\log_{10}(d[\text{m}]) + B + C\log_{10}\left(\frac{f_c[\text{GHz}]}{5.0}\right) + \chi$$

where d is the distance between the transmitter and receiver and f_c is the frequency carrier (within the range 2-6GHz). χ refers to shadowing effect and is modelled as a zero-mean log-normal random variable, with standard deviation σ^2. The parameters A, B, C and σ^2 depend on the global location of the transmitters and receivers (street level, rooftop...). Considering the relative positions of nodes and using notations of [69], the SR-link is modelled by scenario B1, the SD-link by scenario C2 and the RD-link by scenario B5f. Table 4.2 gives the distances transmitter-receiver, as well as the pathloss and shadowing parameters for each link. For simulations, we take $f_c = 3.5$GHz, $P_s = 100$mW, $P_r = 200$mW, which refers to a Wimax situation. We also consider $\theta = 0.3$ and $N = 10^{-10}$W.

Table 4.2 Pathloss and shadowing parameters

Link	**d[m]**	**A**	**B**	**C**	σ^2
SR	70	22.7	41.0	20	3
SD	100	26	39	20	4
RD	30	23.5	57.5	23	8

4.5.2 Reference Schemes

As reference schemes, we consider direct transmission, two-hop transmission and the maximum-rate scheme in [102, 105]. Direct transmission is one-phase communications. The source sends the message m with no splitting ($m_d = m$ and $m_r = 0$) and minimum power

$$\eta_D P_s = \frac{\left(2^R - 1\right) N}{|h_d|^2} \ .$$

In the classical two-hop routing, the source first sends the whole message to the relay, again without splitting, and with a minimum power $\rho_{2H}^S P_s$. If the relay can decode the message, it forwards to the destination with a power of $\rho_{2H}^R P_r$, where

$$\rho_{2H}^S P_s = \frac{\left(2^{R/\theta} - 1\right) N}{|h_s|^2} \quad ; \quad \rho_{2H}^R P_r = \frac{\left(2^{R/\bar{\theta}} - 1\right) N}{|h_r|^2} \ .$$

The destination then decodes using only the signal received from the relay (not both signals received from source and relay as in the two-hop sub-scheme $D^{(s)}$).

The maximum-rate scheme in [102, 105] uses partial decode-forward with power allocation optimal for rate. Allocation is computed numerically solving a convex optimization problem. Whereas classical performance simulations plot capacity as a function of energy, we reverse the axes to plot energy as a function of rate and to allow comparison. We then analyze the performance of the proposed schemes in both non-fading and fading environments.

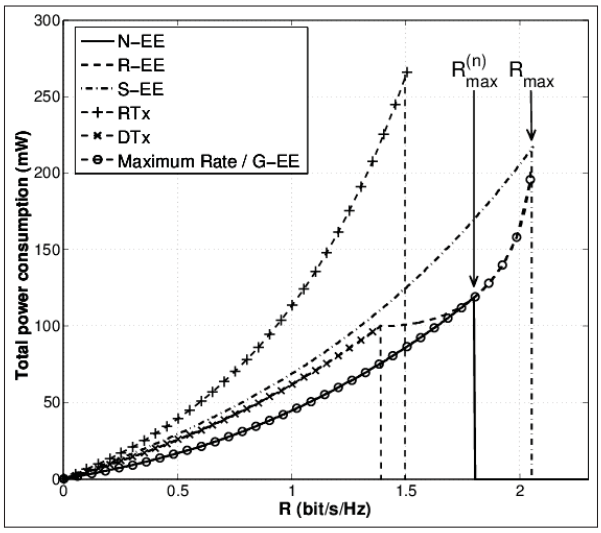


Figure 4.6 Total power consumption
as function of the rate during both phases.

4.5.3 Performance in non shadowing environment

We first consider a non fading environment, with only pathloss (χ is set to zero). This can also be viewed as instantaneous performance.

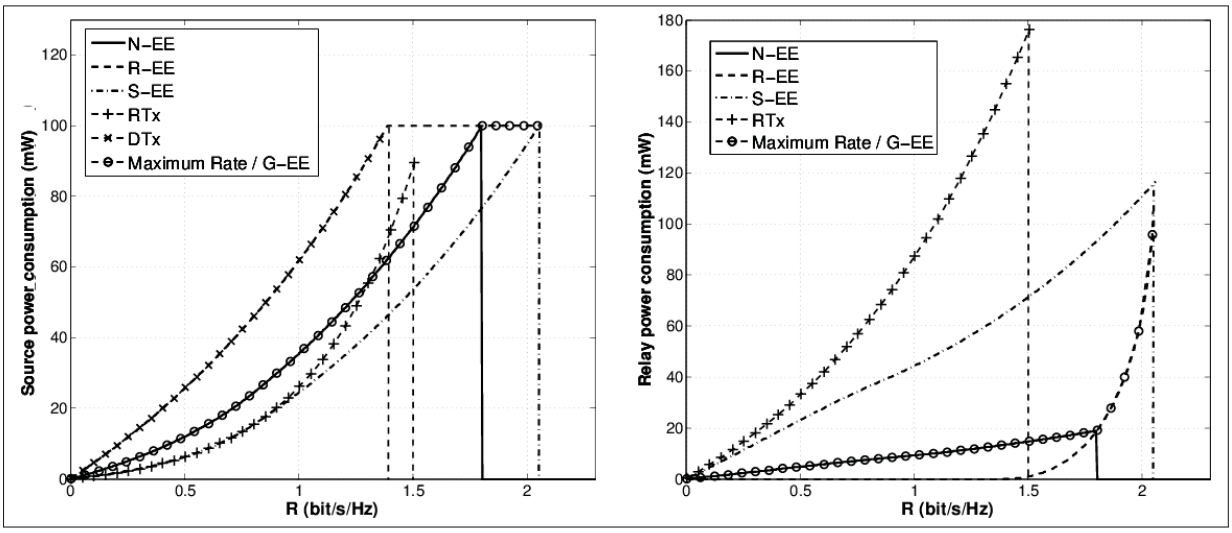


Figure 4.7 Power consumption per node: source (left) and relay (right)

Figures 4.6 and 4.7 respectively show the total power and the per-node power needed to maintain a range of desired source rates. When a source rate is not achievable, an outage occurs, which is depicted by a cut-off in the curve. These figures illustrate the analysis of Section 4.3.4. First, higher rates are achieved by optimizing only one node consumption (either the source as in S-EE, or the relay as in R-EE), rather than the whole network consumption (N-EE). Moreover, both schemes S-EE and R-EE achieve the same maximum source rate. Second, the combined scheme G-EE is also rate-optimal and significantly outperforms both direct and two-hop routing for every achievable source rate. For example, at a rate of 1.35 normalized b/s/Hz, 1.2dB of energy gain is obtained over direct transmissions and 4.5dB over two-hop routing. The gain comes mainly from lessening the relay power consumption. Note that at $R = R_{\max}$, the relay power constraint is not necessarily met with equality if rates beyond $R_{\max}$ are infeasible due to rate constraints. Similarly, R-EE and S-EE provide significant power gain at the relay and the source respectively. Note that maximum-rate scheme is neither source nor relay energy-efficient as it consumes more power at each respective node than the optimal scheme for that node.

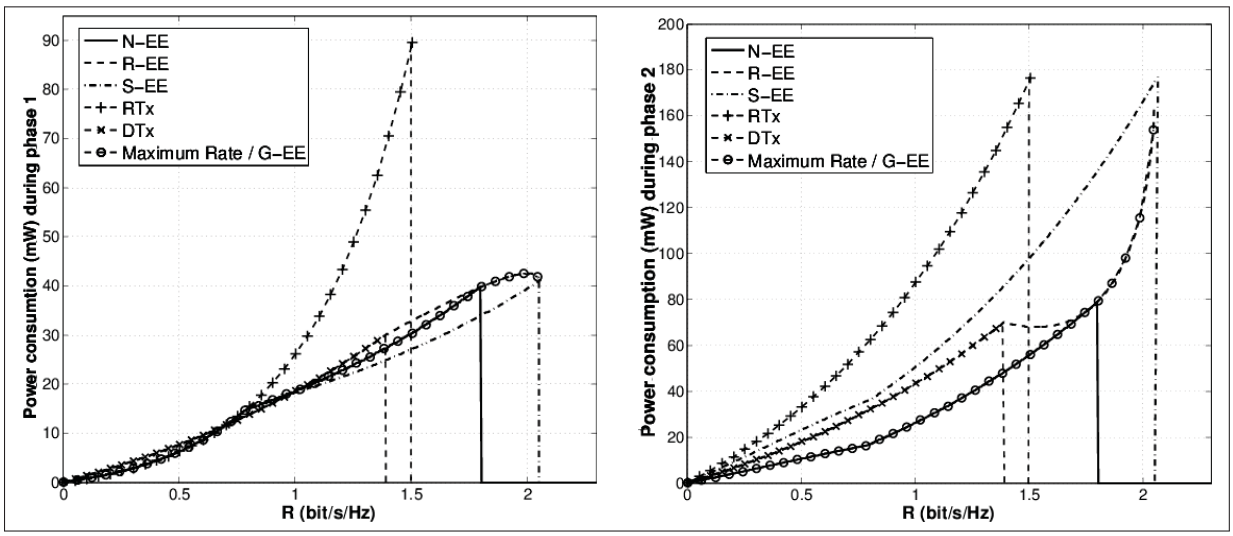


Figure 4.8 Power consumption per phase: phase 1 (left) and phase 2 (right)

Two-hop routing generally suffers from very high instantaneous power consumption. Since the whole message is sent twice, one by the source and one by the relay, within the same time

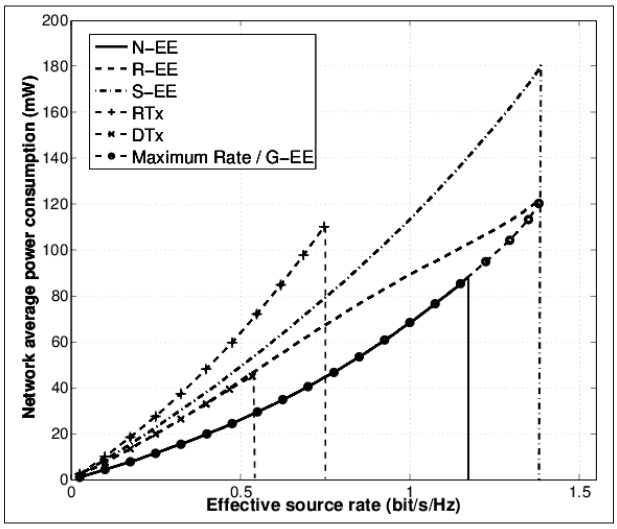


Figure 4.9 Total average energy gain
as a function of the source rate

slot, both transmitters have to increase their transmit power to sustain the desired rate. We plot in Figure 4.8 the power consumed at each phase for the proposed and reference schemes. We see that the proposed schemes allocate power in a smooth manner and reduce transmit power peaks. Contrary to routing, energy is more evenly spread over both phases and over both direct and relaying paths. Moreover, the consumption per phase as function of the source rate is similar to the direct transmission consumption. This can help managing interference in networks which allow both direct and relayed transmissions.

4.5.4 Performance in shadowing environment

We consider in this section both pathloss and shadowing. Due to random shadowing, there exists a non-zero probability that the channel gains are too weak to perform transmission. We plot in Figure 4.9 the average total network energy consumption. The cut-off in the curve depicts the maximum rate above which the outage probability is greater than 0.05. Simulations show that the proposed schemes significantly reduce the average energy consumed, as well as increase the maximal achievable source rate, given the outage requirement.

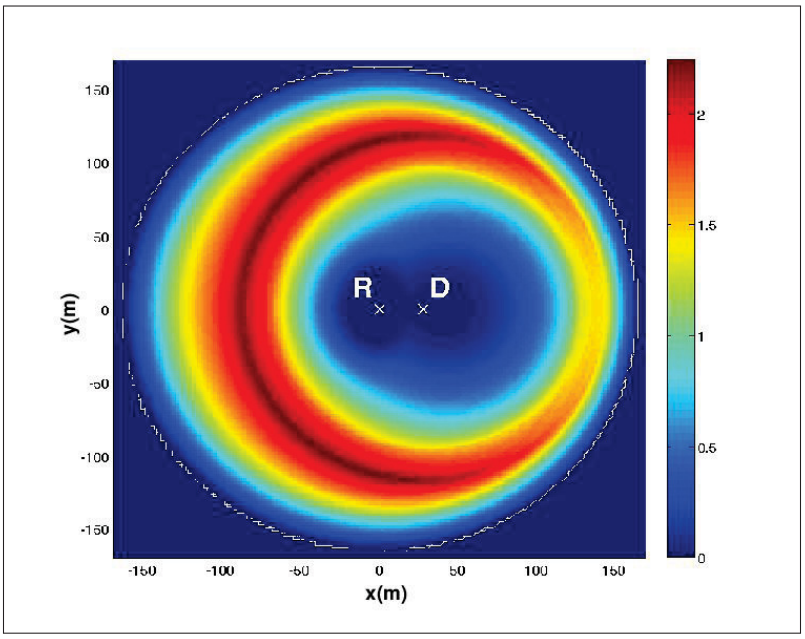


Figure 4.10 Total average energy gain (dB)
as function of the user position

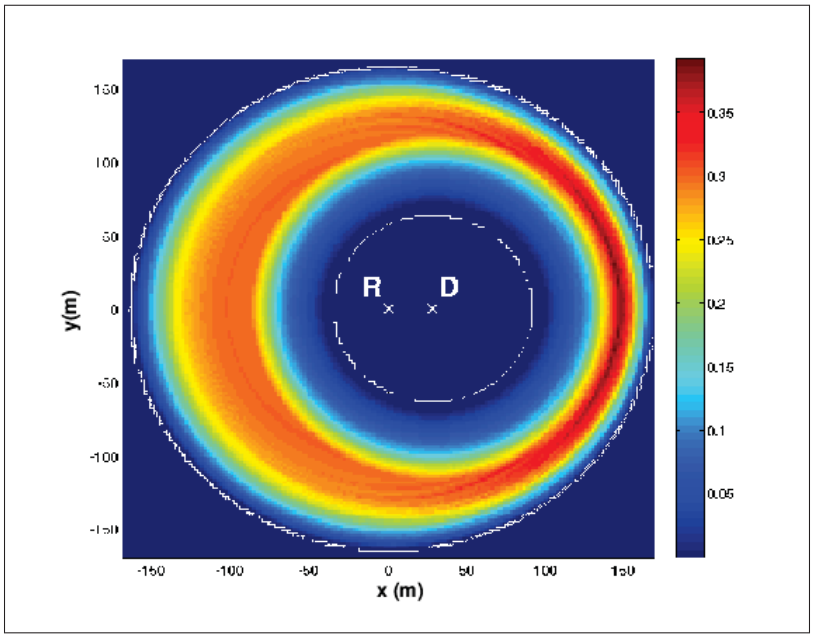


Figure 4.11 Outage reduction (%)
as a function of the mobile user position

Finally, we analyze the utilization region of the proposed schemes and evaluate the energy gain as a function of the mobile user location, similarly to [39]. We consider the environment

and heights of relay and base station as described in Section 4.5.1. The mobile user is free to locate anywhere at street level. As illustrated in Figure 4.10 The relay station is located at the origin $(0, 0)$ and the destination at $(20\sqrt{2}, 0)$ so that $d_{\text{RD}} = 30$m. As reference, we consider a combined scheme which uses either direct transmissions (DTx) or two-hop routing (RTx), depending on the total consumed energy. For simulations, equal time division is assumed, with a source rate of 1bit/s/Hz. Figures 4.10 and 4.11 respectively depict the power gain in dB realized by the proposed generalized scheme G-EE over the reference combined scheme and the outage gain (computed as the difference between outage probabilities of reference and G-EE). Simulation shows that G-EE provides up to 2.2dB of energy gain, as well as decreases the outage probability by up to 37.5%. G-EE is particularly beneficial at cell-edge, for both energy and outage.

4.6 Conclusion

We explore the issue of energy efficiency in relay channels and design a half-duplex coding scheme with optimal resource allocations that maintain a desired source rate and minimize the energy consumption. We consider the network, the relay and the source consumption separately and then combine them into a generalized set of power allocation G-EE which is energy optimal for the relay channel.

We show that minimizing the network energy consumption is not equivalent to maximizing the network capacity as often believed, since the whole range of source rates achievable with Decode-Forward cannot be covered by individual schemes. This equivalence only holds for the proposed generalized scheme G-EE.

Furthermore, this energy efficiency approach remains meaningful since it allows a comprehensive description of the optimal coding (full or partial decode-forward, with or without beamforming) as well as closed-form solutions of the optimal power allocation, which remains implicit in maximum-rate schemes. Reversely, we highlight that a rate-optimal scheme is not energy-efficient for either node.

The generalized scheme G-EE gives a practical algorithm for message coding and associated power allocation to reach energy efficiency and rate optimality for each desired source rate and channel realization. G-EE is particularly beneficial when the source is far from the relay, and up to 2.2dB energy gain can be obtained, as well as a reduction of 37.5% of the outage probability.

CHAPTER 5

IMPACT OF PROPAGATION ENVIRONMENT ON ENERGY-EFFICIENT RELAY PLACEMENT: MODEL AND PERFORMANCE ANALYSIS

Fanny Parzysz[†], Mai Vu[‡], François Gagnon[†]

[†] Département de Génie Électrique, École de Technologie Supérieure,

1100 Notre-Dame West, Montreal, Quebec, (H3C 1K3) Canada

[‡] Department of Electrical and Computer Engineering, Tufts University,

161 College Ave, Medford, MA, (02155) USA

This paper has been published

in *IEEE Transactions on Wireless Communications*.[1]

Summary of this Chapter: Whereas theoretically optimal communication schemes as presented in earlier Chapters may achieve significant energy gain, they may rapidly turn out to be computationally infeasible. Thus, it is essential to make wise use of advanced coding schemes, such that the provided gains overcompensate for their increased complexity. This can be done by carefully choosing the relay location in the network. This Chapter, which is an edition of our second manuscript [29], addresses such issue from a renewed point of view. Existing results for optimal relay placement do not reflect how the choice of the coding scheme and the radio propagation environment, i.e. the node distances, the relay height and the line-of-sight conditions, can impact system performance. The first objective of this Chapter is to propose a geometrical model for energy-efficient relay placement that requires only a small number of characteristic distances. The second objective is to estimate the maximum cell coverage of a relay-aided cell given power constraints, and conversely, the averaged energy consumption given a cell radius. We show that the practical full decode-forward scheme performs close to the energy-optimized partial decode-forward scheme (i.e. the G-EE scheme as presented in

[1] *Fanny Parzysz (2014, p.2214)*

Chapter 4) when the relay is ideally located. However, away from this optimum relay location, performance rapidly degrades and more advanced coding scheme is needed to maintain good performance and allow more freedom in the relay placement. Finally, we define a trade-off between cell coverage and energy efficiency, and show that there exists a relay location for which increasing the cell coverage has a minimal impact on the average energy consumed per unit area.

5.1 Introduction

Relay-aided transmissions promise significant gains in both coverage extension and performance enhancement. Relay-based architecture for cellular network has been considered for next generation wireless systems, such as the LTE-A [113, 114, 115], and is envisioned as part of future heterogeneous networks, along with macro-, pico-, and femtocells. Picocells are operator-deployed and help alleviate coverage dead zones and traffic hot zones, while femtocells are usually consumer-deployed in an unplanned manner to increase in-door coverage. Both are connected to their own wired backhaul connection, in contrary to relays. The use of low power, low complexity relays in lieu of additional wired connected stations avoids expensive backhaul links. Relay nodes are deployed with a more traditional RF planning, with site deployment and maintenance by the operator, thus allowing optimization of their placement within the cell. At the same time, as a result of the tremendous increase of subscribers, jointly with the ever-growing demand for application requirements, energy efficiency becomes a great challenge for future wireless systems. However, the energy impact of the relay location and of the relay coding scheme remains under-explored in the literature.

5.1.1 Motivation and Prior work

Resource allocation and relaying strategies in cooperative systems have mostly targeted improving the cell capacity or spectral efficiency. A rich literature has investigated relay placement from this perspective [59, 72, 116, 117, 118]. For example, authors of [59] analyze different relaying strategies, amplify-forward, decode-forward and compress-forward, and compare their spectral efficiency.

However, the quality of service (QoS) of value-added mobile applications, such as on-demand video, real-time games or video-conferences, is mostly expressed in terms of a minimum rate to be maintained, along with delay and jitter requirements [106]. This perspective has spurred new analysis and designs for cellular networks. For example, in [60], relay placement is optimized to enhance capacity while meeting a traffic demand. From another perspective, references [56, 57, 58] address coverage extension, given a user rate. Analysis has also focused on improving cell topology, such as [119] which proposes an interesting novel dual-relay architecture.

Since user applications mostly run on power-limited devices, reducing the energy consumption, while meeting a rate requirement, has in particular attracted significant attention. Energy minimization for relay-aided cellular networks has ignited new techniques such as relay-based handover strategies [18, 120], sleep mode operations [121], improved resource management [122] and relay-aided load balancing [123]. Nevertheless, little can be found on the specific issue of energy-efficient relay placement and on the performance gains obtained by using advanced coding schemes rather than simpler schemes. Moreover, extending cell coverage implies serving far users and consequently, consuming more energy. This trade-off has neither been investigated.

Another motivation also brought us to analyze energy-efficient relay placement. Although network simulations offer wide possibilities and numerous variable factors for refined investigation, they are often time consuming and may confine the obtained results only to the considered settings. They provide neither generalization to other settings nor performance limits like an analysis based on capacity bounds. Thus, a model is needed to give insights on the relay utilization and helps set basis for future investigations and extensive simulations. Models for heterogeneous networks have attracted significant attention in recent years and various advanced models have been proposed, such as in [89] for femtocell networks. Such models however, are not easy to apply to relay-aided networks, where a cell has to manage both direct and two-hop transmissions and where physical-layer considerations, such as the coding scheme, need to be included. Current works using models for relaying are mostly based on simplified patterns composed of inner and outer circular regions for direct and relayed transmissions, such as in

[62]. In such a model, the radius of the inner region corresponds to the base-station-to-relay distance, which is neither coverage- nor energy-optimized.

Finally, though relay placement has been addressed in the aforementioned works, none of them allow the characterization of the impact of relay placement based on accurate path-loss models. First, geometrical models for relay positioning are incomplete. In [59, 61], the considered model is one-dimensional, and the relay is located on a straight line between the source and the destination. This model cannot provide accurate solutions for actual relay placement in a network where the user is mobile, while the base station and relay are fixed. In [53, 56, 57, 58, 60, 72, 116, 117], two-dimensional models are considered, but the relay height is not taken into account, despite its impact on both path-loss and line-of-sight (LOS) conditions. Three-dimensional models are not fully investigated and relay height has been considered only in [118]. Second, a majority of results on relay placement, such as [53, 56, 57, 58, 59, 60, 61, 72], are for AWGN channels with path-loss exponent, which may be oversimplified as emphasized in [71], or for fading channels with the same path-loss model for all links in the network. More accurate path-loss channel models, as in [116, 117, 118], should be considered.

5.1.2 Main contributions

We introduce in this paper the notion of Relay Efficiency Area and propose a geometrical model for energy-efficient relay placement, for both uplink and downlink transmissions, which allows meaningful performance analyses without the need for excessive simulations. The proposed model consists of only a few characteristic distances, but accurately approaches simulations on the probability of relaying within 3% and the energy consumption within 1%. Then, we build an analysis framework for the performance of relaying based on time-division full decode-forward and provide comparison with the comprehensive partial decode-forward scheme, which is studied with the perspective of rate maximization in [102] and energy minimization in [28].

In the first part, we consider as a baseline the setting with a fixed number of relays at regular locations in each cell and no inter-cell interference, leaving these factors to future extensions.

This baseline model provides a tractable benchmark for performance comparison and analysis, and shows that the classical inner and outer regions as considered in [62] do not reflect the performance gains obtained by relaying.

In the second part, using the proposed geometrical model, we highlight relay configurations which are beneficial for coverage extension and for energy efficiency. Given that a network designer is constrained by the urban topography and by administrative by-laws for station placement, flexibility in the relay location is also a relevant performance criteria. We thus consider performance both at and nearby the optimal relay location. Furthermore, we show that the optimal relay placements for coverage and energy efficiency are different. We thus analyze the trade-off between energy gain and coverage extension and highlight the deployment cost required to reduce the energy consumption. We finally propose a new definition for optimal relay placement regarding the average energy consumed per unit area.

5.1.3 Paper overview

We define our 6-sector urban cell framework and channel path-loss model, in Section 5.2, and the considered coding schemes in Section 5.3. Section 5.4 introduces the characteristic distances that constitute the proposed geometrical model. These distances are derived for both full decode-forward and partial decode-forward in Section 5.5. We then deduce the system performance in Section 5.6. Simulation results are described in Section 5.7. Finally, Section 5.8 concludes this paper.

Notation: B stands for the base station, R for the relay and U for the user. Calligraphic $\mathcal{R}$ refers to the user rate. Upper-case letters denote constant distances, such as R, D or H, while lower-case letters stand for variable distances or angles, such as r or θ. Subscripts $_d$, $_s$ and $_r$ respectively refer to the link from user to base station (direct link), user to relay station and relay to base station. $\mathbb{P}$ is used for probabilities and $\mathbb{E}$ for expectation. Finally, DTx refers to direct transmission, while RTx stands for relay-aided transmission, in the considered coding schemes.

5.2 System model for a 6-sector urban cell

In this section, we present the cell and path-loss models. In particular, we analyze two different path-loss models, depending on the relay being below or above the rooftop.

5.2.1 System model for relay placement analysis

As presented in [124] and depicted in Figure 5.1(a), we consider a 6-sector hexagonal cell. Note that, however, the analysis presented in this paper can be adapted to other cell models, such as circular shape or varying number of sectors. R_{cov} refers to the maximum radius, which includes the coverage extension provided by the relay station. The base station is located at the center of the cell and each sector is provided with a relay station, located at a distance D_r from the base station. The user is positioned at a distance $r \leq \mathsf{R}_{\text{cov}}$ from the base station, with an angle θ, as illustrated in Figure 5.1(b). r_s refers to the distance from the user to relay. We consider a three-dimensional geometrical model. We analyze the impact of the relay height on the system performance. We denote H_B, H_R and H_U the heights of the base station, the relay station and the user respectively.

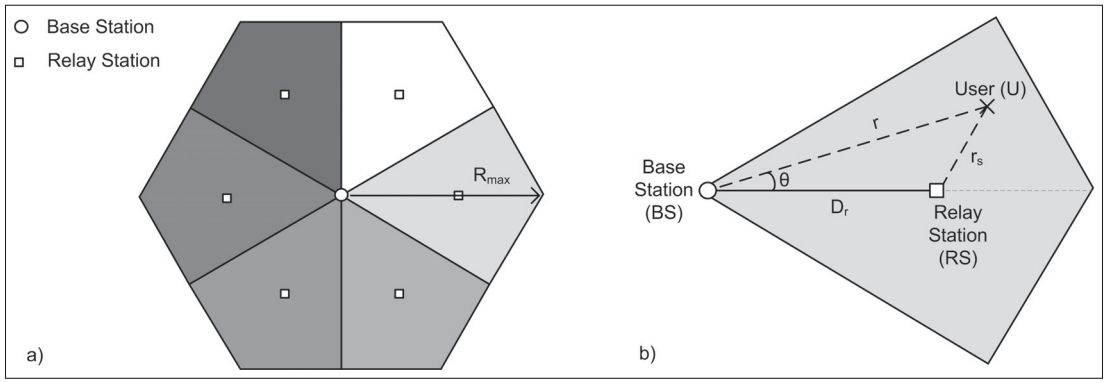


Figure 5.1 System model for a 6-sector urban cell

The multiple access strategy allows orthogonality between users, such that only one user is served for a given time and frequency resource. We assume that the relay operates in the same frequency resource as the user it serves (*in-band* relaying).

5.2.2 Channel model for half-duplex relaying

We consider half-duplex relaying performed in time division. Nevertheless, the following analysis and results are valid with other multiplexing schemes. Without loss of generality, we describe the channel model for uplink transmissions. Here, a transmission of unitary length is carried out in two phases of equal duration. We denote the transmitted codewords X_1 and X_2 for the user at each phase respectively, and X_r for the relay. Y_r and Y_1 respectively are the received signals at the relay and destination at the end of phase 1, Y_2 is the received signal at the destination at the end of phase 2. Thus, the half-duplex relay channel is written as follows:

$$\begin{aligned} Y_r &= h_s X_1 + Z_r\,; \quad Y_1 = h_d X_1 + Z_1 \\ Y_2 &= h_d X_2 + h_r X_r + Z_2 \end{aligned} \tag{5.1}$$

where Z_r, Z_1 and Z_2 are independent additive white Gaussian Noises (AWGN) with equal variance N. We respectively denote h_d, h_s and h_r the complex channel from user to base station (direct link), from user to relay and from relay to base station. For the channel coefficients $|h_i|^2 = \frac{1}{\gamma_i}$, we use the path-loss models as proposed in the WINNER II model [69]:

$$\gamma_i[dB] = A_i \log_{10}(d) + B_i + C_i \log_{10}\left(\frac{f_c}{5}\right) + D_i \log_{10}\left((\mathsf{H}_{\text{Tx}} - 1)(\mathsf{H}_{\text{Rx}} - 1)\right) \tag{5.2}$$

where H_{Tx} and H_{Rx} are the respective heights (in meters) of transmitter and receiver, d is the distance (in meters) between them and f_c the carrier frequency (in GHz). The parameters A_i, B_i, C_i and D_i are constants dependent on the global location of the transmitter and receiver (street level, rooftop...). For convenience, we will use the following notation

$$|h_i|^2 = \frac{1}{\gamma_i} = \frac{1}{K_i d^{A_i/10}} \quad \text{with} \quad K_i = 10^{B_i/10}\left(\frac{f_c}{5}\right)^{C_i/10}\left((\mathsf{H}_{\text{Tx}} - 1)(\mathsf{H}_{\text{Rx}} - 1)\right)^{D_i/10}. \tag{5.3}$$

Since the relay height is allowed to vary from below to above rooftop, the propagation environment for h_s and h_r changes notably. However, the WINNER II project does not provide a continuous path-loss model as a function of the relay height. We will therefore consider the two

following situations: 1) *Vicinity relay:* the relay station is located below rooftop, implying that the user-to-relay link is generally strong but at the cost of a weaker relay-to-base station link; 2) *Base-station-like relay:* the relay station is located above the rooftop, like a base station, the h_r link is very strong and can be assumed as free space (FS). Nevertheless, the user experiences the same path-loss model to access the relay as the base station. We refer the reader to Appendix 1, p.185 for further details on the path-loss model. Note that this paper only considers path-loss. The analysis and the proposed model can thus be understood as the mean coverage and relaying area over a sufficient period of time such that fading and shadowing are averaged out. Future investigations on this topic will include the shadowing effect.

In this channel model, we assume reasonably low inter-cell interference, as a starting point to create a baseline model. Indeed, techniques such as Fractional Frequency Reuse (FFR), scheduling based on inter-base-station coordination and directional antennas between the relay and the base station, have been spurred for next cellular OFDMA systems to suppress or avoid interference from neighbouring cells [125].

5.3 Relaying schemes and Power Allocations for Energy Efficiency

We present the three reference coding schemes for half-duplex time-division relaying which are used for performance analysis. Once again, and without loss of generality, we describe the schemes for uplink transmissions. We consider Gaussian signaling rather than practical modulations and coding. Gaussian signaling is not only suitable for theoretical investigation (by providing upper bound on practical system performance), but is also appropriated for analysis of practical systems. Indeed, practical systems such as LTE are OFDM-based and data is transmitted over numerous parallel frequency channels such that the output is well approximated by a complex Gaussian distribution, as developed in [125, 126].

For each coding scheme, the power allocation minimizes the energy consumption over the normalized transmission duration, while maintaining a given user rate $\mathcal{R}$. The user and the relay have individual power constraints over the two transmission phases within the same bandwidth, respectively denoted as $P_U^{(\max)}$ and $P_R^{(\max)}$. The power constraint of the base station is set to

$P_B^{(\max)}$. At each phase, each node allocates to each transmitted codeword a portion of its available power. We consider *in-band* relaying, where the relay and the user operate in the same frequency band.

5.3.1 Direct Transmission (DTx)

As the first reference, we consider direct transmissions from user to base station. For fair comparison, direct transmission scheme is performed over the same duration as the relaying scheme, and not the half duration as generally done in the literature. By doing this, relaying does not create excess delay, nor uses more resource (time and/or frequency) than direct transmissions. The user rate $\mathcal{R}$ is feasible as long as it is below the channel capacity, i.e.

$$\mathcal{R} \leq \log_2 \left(1 + \frac{P_U^{(\max)} |h_d|^2}{N} \right).$$

The minimum required energy is equal to $\left(2^{\mathcal{R}} - 1\right) \frac{N}{|h_d|^2}$. Outage occurs when the user power constraint $P_U^{(\max)}$ is not satisfied.

5.3.2 Relay-aided Transmission (RTx)

For relay-aided uplink transmissions, two schemes are considered: full-decode forward and partial decode-forward. For each, we use the power allocation that minimizes total energy consumption, i.e. consumption of the user and the relay station together. We furthermore assume that the user chooses to use either direct transmission (DTx) or relayed transmission (RTx) depending on which one consumes less energy.

5.3.2.1 Full decode-forward scheme (Full-DF)

During Phase 1, the user sends its message to the base station with rate $2\mathcal{R}$ and power P_U. Then, the relay decodes the message, re-encodes it with rate $2\mathcal{R}$ and forwards it to the destination during Phase 2 with power P_R.

In the literature, two main decoding techniques are generally considered i) *Two-hop relaying*, as used in practical system, in which the base station only considers the signal received from the relay and thus the relay is merely a repeater; ii) *Repetition-coded full decode-forward*, as defined in [39], in which the user repeats its message in Phase 2. Then, the base station uses MRC and combines the signals received from both the user and relay. Rate bounds and energy consumption of both schemes can be similarly written using a parameter α, with $\alpha = 0$ for two-hop relaying and $\alpha = 1$ for repetition-coded full decode-forward, as follows:

$$\begin{cases} \mathcal{R} & \leq \frac{1}{2}\log_2\left(1+\frac{P_U|h_s|^2}{N}\right) \\ \mathcal{R} & \leq \frac{1}{2}\log_2\left(1+\alpha\frac{P_U|h_d|^2}{N}+\frac{P_R|h_r|^2}{N}\right) \end{cases} \tag{5.4}$$

where the two lines describe the decoding constraints at the relay and the base station respectively. The power allocation set (P_U, P_R) satisfies the following constraints:

$$\frac{1}{2}P_U \leq P_U^{(\max)} \quad \text{and} \quad \frac{1}{2}P_R \leq P_R^{(\max)}. \tag{5.5}$$

The allocation set which minimizes the energy consumption is

$$P_U = \left(2^{2\mathcal{R}}-1\right)\frac{N}{|h_s|^2} \quad \text{and} \quad P_R = \left(2^{2\mathcal{R}}-1\right)\left(1-\alpha\frac{|h_d|^2}{|h_s|^2}\right)\frac{N}{|h_r|^2}. \tag{5.6}$$

However, we can show that, in an urban cellular network considering path-loss only, combining the signals received from both the user and relay station as done in repetition-coded full decode-forward only brings little energy gain compared to two-hop relaying. This small gain is present only when the direct link is strong but, in this case, direct transmission is most of the time more energy-efficient than relaying. Subsequently, we will consider two-hop relaying and repetition-coded full decode-forward as two variants of the same reference scheme, denoted as full decode-forward (Full-DF).

5.3.2.2 Energy-optimized partial decode-forward scheme (EO-PDF)

As the second relay-aided scheme, we consider the partial decode-forward scheme optimized for energy proposed in our previous work [28]. In the EO-PDF scheme, only part of the initial

message is relayed, the rest being sent via the direct link. The scheme is based on rate splitting and superposition coding at transmitter, and joint decoding at receivers. It uses receive beamforming between the relay and the user in the second transmission phase. This scheme arguably requires complex implementation and fine synchronization, however, from the aspect of information theory, it is both energy- and rate-optimal as discussed in [28], and thus, gives the theoretical upper-bound of the performance that can be achieved with decode-forward based relaying.

The optimal joint rate and power allocation of this coding scheme consists of three sub-schemes, which depend on the desired user rate:

1) If the desired user rate is below a certain threshold $\mathcal{R}^\star$,

- In Phase 1, the user sends m with power P_U and the relay decodes as $\tilde{m}$.
- In Phase 2, the relay sends $\tilde{m}$ with power P_R and the user sends m with power P_U.

The allocation (P_U, P_R) minimizes the total energy consumption of the user and the relay.

2) If the desired user rate is above $\mathcal{R}^\star$, the source splits its message into two parts m_r and m_d, of rates $\mathcal{R}^{(r)}$ and $\mathcal{R}^{(d)}$ respectively, with $\mathcal{R} = \mathcal{R}^{(d)} + \mathcal{R}^{(r)}$.

- In Phase 1, the source sends m_r with power $P_U^{(1)}$. The relay decodes as $\tilde{m}_r$.
- In Phase 2, the relay sends $\tilde{m}_r$ with power P_R and the source sends (m_r, m_d) with power $(P_U^{(2)}, P_U^{(3)})$, using superposition coding. The base station jointly decodes m_r and m_d.

The power and rate allocation $\left(\mathcal{R}^{(r)}, \mathcal{R}^{(d)}, P_U^{(1)}, P_U^{(2)}, \ P_U^{(3)}, P_R\right)$ minimizes the total energy consumption, with $\frac{1}{2}\left(P_U^{(1)} + P_U^{(2)} + P_U^{(3)}\right) \leq P_U^{(\max)}$.

3) If this second sub-scheme is in outage due to very high user rate, the same coding scheme is used, i.e. rate splitting, superposition coding and joint decoding, but with other power allocation. Here, $\left(\mathcal{R}^{(r)}, \mathcal{R}^{(d)}, P_U^{(1)}, P_U^{(2)}, P_U^{(3)}, P_R\right)$ minimizes the energy consumption of the relay only.

We refer the reader to [28] for the detailed power and rate allocation, as well as for the expression of the threshold $\mathcal{R}^\star$.

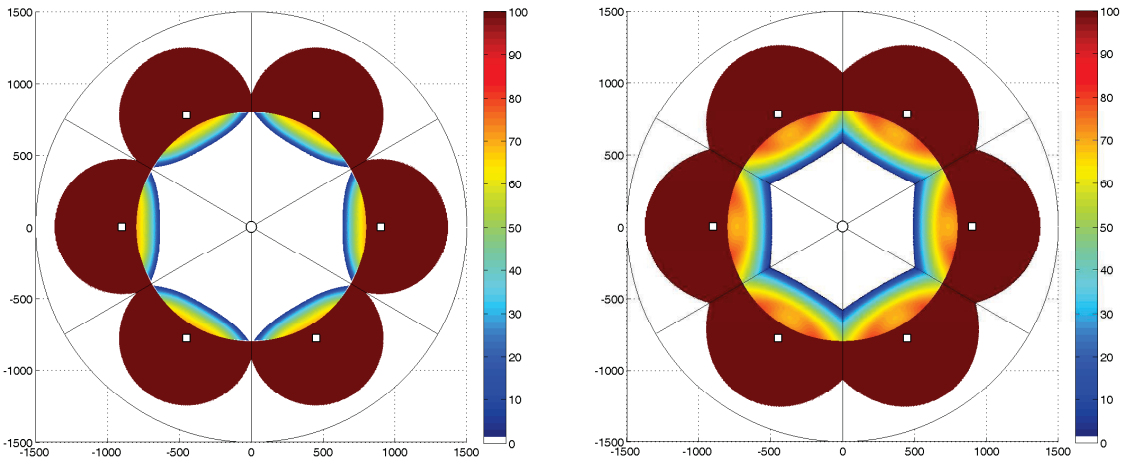
5.4 Characterization of efficient relay placement

In this section, we define the relaying efficiency, both in terms of energy consumption and coverage extension. To do so, we introduce the notion of Relay Efficiency Area (REA), and build our geometrical model based on this REA for both uplink and downlink transmissions. Note that we will consider Cartesian coordinates (x, y) and polar coordinates (r, θ) for the user position.

5.4.1 On the necessity to model the relaying efficiency

As a first insight of the potential energy gain brought by relaying, we plot in Figure 5.2 the energy gain in percentage (i.e. $100\frac{E_{\text{DTx}}-E_{\text{RTx}}}{E_{\text{DTx}}}$) that is obtained by using either full decode-forward (Full-DF) or the energy-optimized partial decode-forward scheme (EO-PDF) over direct uplink transmission, as a function of the user location in the 6-sector cell. The gain is set to 100% when direct transmission is not feasible, i.e. when relaying enhances the cell coverage. As shown in the figure, relaying can significantly reduce the energy consumption. Considering the cell-edge of conventional systems using direct transmission only, an average of 46% of energy gain is obtained with full decode-forward and 70% with EO-PDF. In addition to energy savings, relaying also increases the radius of the corresponding hexagonal cell by 33% for full decode-forward and by 56% for EO-PDF.

Figure 5.2 shows that conventional relaying schemes, such as two-hop relaying or repetition-coded full decode-forward, can provide notable performance enhancement, especially in terms of coverage extension. Nevertheless, those schemes are clearly suboptimal compared to the potential gains of decode-forward, upper-bounded by the EO-PDF scheme. It is then necessary to estimate the gains that implementing advanced coding schemes can bring over simpler schemes, both in terms of energy savings and coverage extension. Furthermore, the efficiency of the relay obviously depends on its location in the cell, as well as its propagation environment. Estimating the network performance cannot be reasonably done by simulating all possible relay network configurations and user locations. Thus, a model is needed to give insight of relay utilization and energy consumption.



(a) using Full decode-forward (Full-DF) (b) performance upper-bound (EO-PDF)

Figure 5.2 Energy gain compared to DTx
($\mathcal{R} = 3$bit/s/Hz, vicinity relay, LOS conditions)

5.4.2 Analytical model for Relay Efficiency Area (REA)

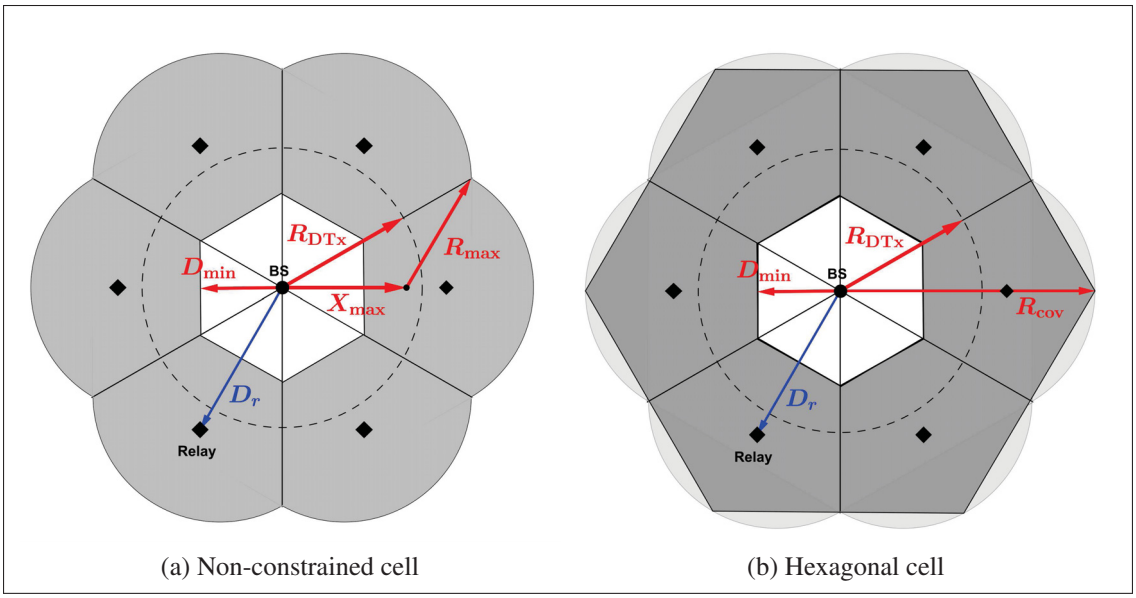
A relay-aided cell is characterized by the probability for a user to be served by the relay station and by the energy saved by using the relay. Those two performance criteria determine if a relay location is beneficial, and thus, characterize relay efficiency. To investigate relay efficiency, we introduce here the concept of *Relay Efficiency Area (REA)* for a cellular network served by relay stations. It is defined as the set of all user locations for which relaying is more energy-efficient than DTx or for which DTx is not feasible. Therefore, the REA covers both performance improvement and coverage extension.

The simulated REA's of Full-DF and the EO-PDF scheme are illustrated in Figure 5.2, and its related model in Figure 5.3. In Figure 5.2, the REA corresponds to the cell area where the energy gain is strictly positive (non-white area). When the user is close to the base station, the h_d link is very strong and DTx consumes less energy than relaying. Otherwise, relaying is more energy-efficient or allows successful transmission while DTx is in outage.

We propose to characterize the REA of each coding scheme using four characteristic distances, as depicted in Figure 5.3a: $\mathsf{D}_{\min}$ and $\mathsf{R}_{\mathrm{DTx}}$ for the inner bound, $\mathsf{X}_{\max}$ and $\mathsf{R}_{\max}$ for the outer bound. The inner bound is modelled by combining the strait line of equation $x = \mathsf{D}_{\min}$ and a portion of the circle of radius $\mathsf{R}_{\mathrm{DTx}}$ and centred at the base station, which refers to the minimum cell radius over which direct transmission is infeasible given outage requirement. The outer bound is characterized by $\mathsf{X}_{\max}$ and $\mathsf{R}_{\max}$. It is modelled by a portion of the circle of radius $\mathsf{R}_{\max}$ and centred at $(\mathsf{X}_{\max}, 0)$. The Relay Efficiency Area is thus modelled as follows:

$$\mathrm{REA} = \left\{(r,\theta) \ \in \ \left((\mathsf{D}_{\min} \leq r\cos(\theta)) \ \cup \ \bar{\mathcal{C}}\left(0, \mathsf{R}_{\mathrm{DTx}}\right)\right) \cap \ \left(\mathcal{C}\left(\mathsf{X}_{\max}, \mathsf{R}_{\max}\right)\right)\right\} \tag{5.7}$$

where (r,θ) is the user location. $\mathcal{C}\left(\mathsf{X},\mathsf{R}\right)$ stands for the portion of the disk centred at $(\mathsf{X},0)$ with radius R and delimited by the sector bounds. $\bar{\mathcal{C}}$ stands for the exterior of this disk. Note that this model for REA is independent of the user distribution and of the cell shape. These factors are only considered for the computation of the performance metrics, e.g. the coverage radius, the probability of relaying and the average energy consumption.



(a) Non-constrained cell (b) Hexagonal cell

Figure 5.3 A model for Relay Efficiency Area (REA)

Using this first characterization of REA, we derive the model for hexagonal cells and define the maximum cell radius R_{cov} that guarantees coverage for all users located inside the hexagonal cell without overlap, as depicted in Figure 5.3b. Three characteristic distances are sufficient to define this hexagonal model for REA: $\mathsf{D}_{\min}$, R_{DTx} and R_{cov}, where

$$\mathsf{R}_{\text{cov}} = \mathsf{X}_{\max} + \min\left\{\mathsf{R}_{\max}, \sqrt{\frac{(2\mathsf{R}_{\max})^2 - (\mathsf{X}_{\max})^2}{3}}\right\}. \tag{5.8}$$

This expression of R_{cov} insures no coverage hole at sector-edge, contrary to the coverage radius defined in [56, 57] . As we will show later, $\mathsf{R}_{\max}$ is function of the transmit power. The area of a cell sector can then be expressed as follows.

$$\mathcal{A}_{\text{sector}} = \frac{\sqrt{3}}{4}\mathsf{R}_{\text{cov}}^2. \tag{5.9}$$

These characteristic distances, as defined above, allow the direct computation of the overall probability of relaying a communication, given that the user is randomly located in the cell, as well as the average consumed energy to successfully transmit data. In the following sections, we apply this model to compute these characteristic distances for each coding scheme and investigate the probability of relaying in the cell and the total average energy consumption.

5.5 Analysis of Relay Efficiency Area for decode-forward schemes

We now compute the four characteristic distances $\mathsf{D}_{\min}$, R_{DTx}, $\mathsf{X}_{\max}$ and $\mathsf{R}_{\max}$ for Full-DF and EO-PDF, as defined in Section 5.3, and for both uplink and downlink transmissions. These characteristic distances are derived from the channel conditions for transmission feasibility and for energy efficiency.

5.5.1 Common conditions for relaying

First, we focus on R_{DTx}, the maximum radius over which DTx is infeasible. We have

$$\mathsf{R}_{\text{DTx}} = \left(\frac{P_i}{K_d N \left(2^{\mathcal{R}} - 1\right)} \right)^{10/A_d} \tag{5.10}$$

where A_d and K_d are given by the path-loss model and defined in Eq. (5.2) and (5.3), and where $P_i = P_U^{(\max)}$ for uplink and $P_i = P_B^{(\max)}$ for downlink.

Second, we define D_{RTx} for the uplink. It refers to the relaying condition $|h_d|^2 \leq |h_s|^2$ for which relaying can outperform DTx. At the limit, we get

$$|h_d|^2 = |h_s|^2 \quad \Leftrightarrow \quad \left(\frac{K_d}{K_s} \right)^{\frac{20}{A_s}} r^{2\frac{A_d}{A_s}} - r^2 + 2\mathsf{D}_r \cos(\theta)\, r - \mathsf{D}_r^2 = 0 \tag{5.11}$$

where D_r is the distance between the relay and base station. Given the cell sector bounds and regarding Cartesian coordinates, we can show that the derivative of x with respect to y is small, such that Eq. (5.11) can be approximated by a straight line. Thus, in the proposed model, we approximate the relaying condition $|h_d|^2 \leq |h_s|^2$ by $x \geq \mathsf{D}_{\text{RTx}}$, where D_{RTx} is defined as the smallest positive solution of r in Eq. (5.11) with $\theta = 0$, i.e.

$$\left(\frac{K_d}{K_s} \right)^{\frac{20}{A_s}} r^{2\frac{A_d}{A_s}} = (r - \mathsf{D}_r)^2 \tag{5.12}$$

In particular, when h_s and h_d have the same path-loss exponent, as it is the case when both h_s and h_d are in LOS, we get:

$$\mathsf{D}_{\text{RTx}} = \frac{\mathsf{D}_r}{\left(\frac{K_d}{K_s} \right)^{\frac{10}{A_s}} + 1}. \tag{5.13}$$

This relaying condition must be satisfied for both Full-DF and EO-PDF. In practice however, this condition is not limiting for Full-DF since, for this coding scheme, transmission is over two phases of half-duration and data is thus sent at twice the desired rate. Actually, the user decides if relay-aid is necessary based on energy consumption, as we will see in next paragraph. On the contrary, the relaying condition $|h_d|^2 \leq |h_s|^2$ is relevant for the EO-PDF scheme since only

part of the message is relayed, leading to sufficient energy gain to compensate for the shorter transmission duration.

With regards to downlink, this condition becomes $|h_d|^2 \leq |h_r|^2$, which is always satisfied since the relay-to-base-station link is very strong.

5.5.2 Conditions for energy efficiency and outage of full decode-forward (Full-DF)

5.5.2.1 Energy-efficiency of full decode-forward

First, we focus on energy efficiency and we derive the conditions for which the energy consumption using DTx (denoted E_{DTx}) is greater than the consumption using full decode-forward (denoted E_{DF}). This energy-efficient condition can be seen as the dual of the throughput-oriented condition $\log_2\left(1+\frac{P_s|h_d|^2}{N}\right) \leq \frac{1}{2}\log_2\left(1+\frac{2P_s|h_s|^2}{N}\right)$. By using Eq. (5.6), we get

$$
\begin{aligned}
& E_{\text{DTx}} \geq E_{\text{DF}} && (5.14)\\
\Leftrightarrow \quad & \left(2^{\mathcal{R}}-1\right)\frac{N}{|h_d|^2} \geq \left(2^{2\mathcal{R}}-1\right)\frac{N}{2|h_s|^2} + \left(2^{2\mathcal{R}}-1\right)\frac{N}{2|h_r|^2}\left(1-\alpha\frac{|h_d|^2}{|h_s|^2}\right) \\
\Leftrightarrow \quad & \alpha K_s K_r \mathsf{D}_r^{A_r/10} \leq K_d r^{A_d/10}\left[K_s - \left(\frac{2}{2^{\mathcal{R}}+1}K_d r^{A_d/10} - K_r \mathsf{D}_r^{A_r/10}\right)\frac{1}{r_s^{A_s/10}}\right] \\
\text{where} \quad & r_s^2 = \mathsf{D}_r^2 + r^2 - 2\mathsf{D}_r r\cos(\theta).
\end{aligned}
$$

Note that Eq. (5.14) is symmetric in h_s and h_r, meaning that it is valid for both uplink and downlink. We can also show that a solution of Eq. (5.14) for $\alpha = 0$ is almost solution for $\alpha = 1$. This implies that the bounds for two-hop relaying can approximate the ones of repetition-coded full DF. Simulations in Section 5.7 show that the REA of repetition-coded full DF is only increased by less than 1% compared to two-hop relaying.

The solution of Eq. (5.14) can be well approximated by the intersection of the half-plane defined by $x \geq \mathsf{D}_{\text{DF}}^{(e)}$ (inner bound) with the disk $\mathcal{C}\left(\mathsf{X}_{\text{DF}}^{(e)}, \mathsf{R}_{\text{DF}}^{(e)}\right)$ (outer bound), where superscript $^{(e)}$ stands for energy efficiency. $\mathsf{X}_{\text{DF}}^{(e)}$ and $\mathsf{R}_{\text{DF}}^{(e)}$ can be computed easily as follows. Noting that there exists only one circle with its center on the x-axis and going through any two given points

of the plane, using geometrical principles, it is therefore sufficient to find the two user locations for which $E_{\text{DTx}} = E_{\text{DF}}$ with $\theta = 0$ and $\theta = \frac{\pi}{6}$ to deduce $\mathsf{X}_{\text{DF}}^{(e)}$ and $\mathsf{R}_{\text{DF}}^{(e)}$. We propose to reject all relay positions for which $\mathsf{R}_{\text{cov}} < \mathsf{R}_{\text{DTx}}$, i.e. for which $\mathcal{C}\left(\mathsf{X}_{\text{DF}}^{(e)}, \mathsf{R}_{\text{DF}}^{(e)}\right)$ is strictly included in $\mathcal{C}\left(0, \mathsf{R}_{\text{DTx}}\right)$, i.e. the area covered by the base station alone. Consequently, $\mathcal{C}\left(\mathsf{X}_{\text{DF}}^{(e)}, \mathsf{R}_{\text{DF}}^{(e)}\right)$ is only used to determine the acceptance or rejection of a relay configuration, but does not appear in the expression of the REA itself, as defined in Eq. (5.7).

5.5.2.2 Outage of full decode-forward

Next, we focus on the outage condition. Assuming the relay power is sufficient, outage occurs when the transmission between the relay and user is not feasible, i.e. when the channel h_s is too weak, given the power constraint and user rate. Note that direct transmission is not feasible either in this case. Relayed transmission is in outage for all users positions outside the disk centred at the relay station and of radius:

$$\mathsf{R}_{\text{DF}}^{(o)} = \left(\frac{2P_i}{K_s N\left(2^{2\mathcal{R}} - 1\right)}\right)^{\frac{10}{A_s}}. \tag{5.15}$$

where $^{(o)}$ stands for outage and where $P_i = P_U^{(\max)}$ for uplink and $P_i = P_R^{(\max)}$ for downlink.

5.5.2.3 Characteristic distances for full decode-forward

Based on the previous analysis, we can deduce the characteristic distances for full decode-forward as follows:

$$\begin{cases} \mathsf{D}_{\min}^{(\text{DF})} & = \max\left\{\mathsf{D}_{\text{RTx}}, \mathsf{D}_{\text{DF}}^{(e)}\right\} \quad \text{with} \quad \mathsf{D}_{\text{DF}}^{(e)} \leq \mathsf{R}_{\text{DTx}} \\ \mathsf{X}_{\max}^{(\text{DF})} & = \mathsf{D}_r \ ; \quad \mathsf{R}_{\max}^{(\text{DF})} = \mathsf{R}_{\text{DF}}^{(o)}. \end{cases} \tag{5.16}$$

where D_{RTx} is given by Eq. (5.11), $\mathsf{D}_{\text{DF}}^{(e)}$ by Eq. (5.14) and $\mathsf{R}_{\text{DF}}^{(o)}$ by Eq. (5.15). These values define the Relay Efficiency Area of full decode-forward $\text{REA}^{(\text{DF})}$, as defined in Eq. (5.7). Note that REA of Full-DF regarding uplink is strictly included in the REA for downlink. $\mathsf{D}_{\min}^{(\text{DF})}$, which indicates whether the user should use DTx or RTx is the same, but the coverage radius is extended for downlink since the relay station disposes of more power than the user.

5.5.3 Conditions for energy efficiency and outage of the EO-PDF scheme

The computation of the characteristic distances for the EO-PDF scheme follows the same basis as for full decode-forward.

5.5.3.1 Energy-efficiency of EO-PDF

We recall that the inner-bound is modelled by the combination of the strait line of equation $x = \mathsf{D}_{\min}$ and a portion of the circle centred at the base station and of radius R_{DTx}. $\mathsf{D}_{\min}$ usually depends on both the relaying condition defined in Eq. (5.11) and on the energy-efficiency condition $E_{\text{DTx}} \geq E_{\text{EO}}$.

With regards to uplink, simulations show that there can exist user locations for which direct transmissions is more energy-efficient than the EO-PDF scheme. Indeed, in DTx, messages are sent once over the two phases of transmission, with rate $\mathcal{R}$, whereas in the EO-PDF scheme, messages are sent twice, over each phase, with rate $2\mathcal{R}$. For some specific user locations, the energy gain obtained by EO-PDF does not compensate the loss due to transmitting at a rate of $2\mathcal{R}$ during each phase instead of $\mathcal{R}$ during both phases, and thus, direct transmission becomes more energy-efficient. Nevertheless, simulations show that these the area of user locations is negligible for relay-aided cellular networks. First, they occur only for uncommon relay configurations, specifically when the relay is very close to base station and with very low user rates. Moreover, the area of concern remains small compared to the cell size and the energy gain obtained by DTx over EO-PDF is admittedly positive but almost zero. For these reasons, we can consider that $\mathsf{D}_{\min}$ reduces to the relaying condition $|h_d|^2 \leq |h_s|^2$, as described in Eq. (5.11), and is not limited by the energy-efficiency constraint as it is the case for full decode-forward.

Considering downlink transmission, and as said earlier, the relaying condition $|h_d|^2 \leq |h_r|^2$ is always satisfied and D_{RTx} for downlink is large compared to the cell radius. The inner-bound $\mathsf{D}_{\min}$ thus only depends on $\mathsf{D}_{\text{EO}}^{(e)}$, given by the energy-efficiency condition $E_{\text{DTx}} \geq E_{\text{EO}}$. $\mathsf{D}_{\text{EO}}^{(e)}$ is computed numerically as the user-to-base-station distance for which $E_{\text{DTx}} = E_{\text{EO}}$ with $\theta = 0$.

5.5.3.2 Outage of EO-PDF

The outer bound for the EO-PDF scheme is modelled by a portion of circle centred at $(\mathsf{X}_{\mathrm{EO}}^{(o)}, 0)$ and of radius $\mathsf{R}_{\mathrm{EO}}^{(o)}$. Both are determined by the outage condition and are computed similarly to $\mathsf{X}_{\mathrm{DF}}^{(e)}$ and $\mathsf{R}_{\mathrm{DF}}^{(e)}$ for full decode-forward in Section 5.5.2.1, using the coordinates of two user locations at the outage limit. However, generally, $\mathsf{X}_{\mathrm{EO}}^{(o)} \neq \mathsf{D}_r$ and $\mathsf{R}_{\mathrm{EO}}^{(o)} \neq \mathsf{R}_{\mathrm{DF}}^{(o)}$.

5.5.3.3 Characteristic distances for the EO-PDF scheme

Based on the previous analysis, we can deduce the characteristic distances for the EO-PDF scheme as follows:

$$\begin{cases} \mathsf{D}_{\min}^{(\mathrm{EO})} = \max\left\{\mathsf{D}_{\mathrm{RTx}}, \mathsf{D}_{\mathrm{EO}}^{(e)}\right\} \ ; \\ \mathsf{X}_{\max}^{(\mathrm{EO})} = \mathsf{X}_{\mathrm{EO}}^{(o)} \ ; \quad \mathsf{R}_{\max}^{(\mathrm{EO})} = \mathsf{R}_{\mathrm{EO}}^{(o)}. \end{cases} \tag{5.17}$$

where $\mathsf{D}_{\mathrm{RTx}}$ is given by Eq. (5.11). These values define the Relay Efficiency Area of the EO-PDF scheme $\mathrm{REA}^{(\mathrm{EO})}$, as defined in Eq. (5.7).

Note that for both Full-DF and EO-PDF, we have $\mathsf{D}_{\min} \leq \mathsf{D}_r$. This shows that models based on inner and outer regions as in [62] do not reflect the performance gains obtained by relaying and that the cell area for energy-efficient relaying is much wider than the classical outer region.

5.6 Estimated probability of relaying and average energy consumption

So far, we have introduced the Relay Efficiency Area and derived characteristic distances for both Full-DF and EO-PDF. Based on this model, we now derive the estimated probability of relaying and average energy consumption of the cell.

5.6.1 Probability of Relaying

Here, we analyze the probability of relaying a communication. We assume that the user is randomly located in the cell and that the location distribution is uniform. Such distribution allows us to derive a simple expression of the probability of relaying.

The probability of relaying, denoted $\mathbb{P}_{\text{RTx}}$, is obtained by the ratio of relaying area vs. total area and is expressed as follows for hexagonal cells.

$$\mathbb{P}_{\text{RTx}} = 1 - \frac{8}{\sqrt{3}\mathsf{R}_{\text{cov}}^2}\left(\frac{\mathsf{D}_{\min}^2}{2}\tan(\varphi) + \frac{\mathsf{R}_{\text{DTx}}^2}{2}(\phi-\varphi) + \frac{3}{8}\mathsf{R}_{\text{cov}}^2\tan\left(\frac{\pi}{6}-\phi\right)\right) \tag{5.18}$$

where we assume $\mathsf{R}_{\text{DTx}} \leq \mathsf{R}_{\text{cov}}$. φ and ϕ are as in Eq. (5.19), with $\mathsf{X} = \frac{3}{4}\mathsf{R}_{\text{cov}} + \frac{1}{4}\sqrt{4\mathsf{R}_{\text{DTx}}^2 - 3\mathsf{R}_{\text{cov}}^2}$.

$$\begin{cases} \varphi = \phi = \frac{\pi}{6} & \text{if } \mathsf{D}_{\min} \leq \min\left(\frac{\sqrt{3}}{2}\mathsf{R}_{\text{DTx}}, \frac{3}{4}\mathsf{R}_{\text{cov}}\right) \\ \varphi = \arccos\left(\frac{\mathsf{D}_{\min}}{\mathsf{R}_{\text{DTx}}}\right); \phi = \frac{\pi}{6} & \text{if } \frac{\sqrt{3}}{2}\mathsf{R}_{\text{DTx}} \leq \frac{3}{4}\mathsf{R}_{\text{cov}} \quad \text{and } \frac{\sqrt{3}}{2}\mathsf{R}_{\text{DTx}} \leq \mathsf{D}_{\min} \\ \varphi = \phi = \arctan\left(\sqrt{3}\left(\frac{\mathsf{R}_{\text{cov}}}{\mathsf{D}_{\min}} - 1\right)\right) & \text{if } \frac{3}{4}\mathsf{R}_{\text{cov}} \leq \frac{\sqrt{3}}{2}\mathsf{R}_{\text{DTx}} \quad \text{and } \frac{3}{4}\mathsf{R}_{\text{cov}} \leq \mathsf{D}_{\min} \leq \mathsf{X} \\ \varphi = \arccos\left(\frac{\mathsf{D}_{\min}}{\mathsf{R}_{\text{DTx}}}\right); \phi = \frac{\pi}{6} - \arccos\left(\frac{\sqrt{3}\mathsf{R}_{\text{cov}}}{2\mathsf{R}_{\text{DTx}}}\right) & \text{if } \frac{3}{4}\mathsf{R}_{\text{cov}} \leq \frac{\sqrt{3}}{2}\mathsf{R}_{\text{DTx}} \quad \text{and } \frac{3}{4}\mathsf{R}_{\text{cov}} \leq \mathsf{X} \leq \mathsf{D}_{\min} \end{cases} \tag{5.19}$$

Conditions on ϕ and φ are defined based on geometrical principles of a cell sector and depend on whether $\mathsf{D}_{\min} \leq x$ or $(x, y) \in \bar{\mathcal{C}}(0, \mathsf{R}_{\text{DTx}})$, as defined in (5.7), which is the limiting condition for the REA. The complete proof can be found in Appendix 2, p.185. Note that, for circular shape cells, we have $\phi = \frac{\pi}{6}$ and the constraints regarding $\frac{3}{4}\mathsf{R}_{\text{cov}}$ are relaxed.

5.6.2 Average energy consumption per user transmission

We now focus on the energy consumption, averaged over all user locations. The expressions for downlink and uplink are similar, and we thus only focus on the uplink case. As before, we assume that the user is randomly located in the cell and that its distribution is uniform. We respectively denote $\mathbb{E}[E_{\text{DTx}}]$, $\mathbb{E}[E_U]$ and $\mathbb{E}[E_R]$ the average energy consumed by the user when DTx is used, by the user when RTx is used and by the relay station. The total average energy consumption, referred as $\mathbb{E}[E]$, can thus be computed as follows.

$$\mathbb{E}[E] = (1 - \mathbb{P}_{\text{RTx}})\,\mathbb{E}[E_{\text{DTx}}] + \mathbb{P}_{\text{RTx}}\left(\mathbb{E}[E_U] + \mathbb{E}[E_R]\right). \tag{5.20}$$

This expression reflects both the energy consumed by the user only when DTx is used, and the energy consumed by both the user and the relay station when RTx is employed.

We first obtain $\mathbb{E}\left[E_{\text{DTx}}\right]$, the average energy consumed by a user who is not located in the REA. It depends on $\mathsf{D}_{\min}$, R_{DTx} and R_{cov} as follows:

$$\mathbb{E}\left[E_{\text{DTx}}\right] = \iint_{(x,y)\in\overline{\text{REA}}} \left(2^{\mathcal{R}}-1\right)\frac{N}{|h_d|^2}\mathbb{P}\left((x,y)\in\overline{\text{REA}}\right)dxdy \tag{5.21}$$

$$\begin{aligned}\mathbb{E}\left[E_{\text{DTx}}\right] =& \frac{2\left(2^{\mathcal{R}}-1\right)NK_d}{\mathcal{A}_{\text{sector}}\left(1-\mathbb{P}_{\text{RTx}}\right)}\left(\frac{A_d}{10}+2\right)\left[\mathsf{D}_{\min}^{A_d/10+2}\int_0^{\varphi}\left(\frac{1}{\cos(\theta)}\right)^{A_d/10+2}d\theta\right.\\ &\left.+\left(\mathsf{R}_{\text{DTx}}\right)^{A_d/10+2}(\phi-\varphi)+\left(\sqrt{3}\mathsf{R}_{\text{cov}}\right)^{A_d/10+2}\int_{\phi}^{\frac{\pi}{6}}\left(\frac{1}{\sin(\theta)+\sqrt{3}\cos(\theta)}\right)^{A_d/10+2}d\theta\right]\end{aligned}$$

where φ and ϕ are as defined in Eq. (5.19).

Second, the average energy consumed by the relay depends on the coding scheme. Although this energy cannot be derived in closed-form for the EO-PDF scheme, we provide in next paragraph the expression of the average energy for the full decode-forward scheme.

5.6.3 Average energy consumed by the Full-DF scheme

We analyze the average total energy consumption when the user is located in the REA, using full decode-forward. Recall that, as exemplified by simulations, repetition-coded full decode-forward only brings a few percent of energy savings compared to two-hop relaying. Therefore, to compute the average energy consumption, we focus on the two-hop relaying scheme, as defined in Section 5.3.2.1. The average energy is denoted as $\mathbb{E}\left[E_U^{(\text{DF})}\right]$ for the user consumption and $\mathbb{E}\left[E_R^{(\text{DF})}\right]$ for the relay consumption. Both depend on $\mathsf{D}_{\min}$, R_{DTx} and R_{cov}.

Similarly to $\mathbb{E}\left[E_{\text{DTx}}\right]$ in Eq. (5.21), we have

$$\mathbb{E}\left[E_U^{(\text{DF})}\right] = \iint\limits_{(x,y)\in\text{REA}} \left(2^{2\mathcal{R}}-1\right)\frac{N}{2|h_s|^2}\mathbb{P}\left((x,y)\in\text{REA}\right)dxdy \tag{5.22}$$

$$= \frac{\left(2^{2\mathcal{R}}-1\right)NK_s}{\mathcal{A}_{\text{sector}}\mathbb{P}_{\text{RTx}}}\left[\int_0^{\varphi}\int_{\frac{D_{\min}}{\cos(\theta)}}^{r_{\max}(\theta)} r_s^{A_s/10} r dr d\theta + \int_{\varphi}^{\phi}\int_{\mathsf{R}_{\text{DTx}}}^{r_{\max}(\theta)} r_s^{A_s/10} r dr d\theta\right]$$

$$\text{where} \quad \begin{cases} r_s & = \sqrt{\mathsf{D}_r^2 + r^2 - r\mathsf{D}_r\cos(\theta)} \\ r_{\max}(\theta) & = \frac{\sqrt{3}\mathsf{R}_{\text{cov}}}{\sin(\theta)+\sqrt{3}\cos(\theta)} \end{cases}$$

In this, $r_{\max}(\theta)$ describes the cell bound. φ and ϕ are as defined in Eq. (5.19). Note that, in most of cases, we have $A_s = 40$, for which the integral $\int r_s^{A_s/10} r dr$ becomes a polynomial and is equal to $\int \left(\mathsf{D}_r^2 + r^2 - \mathsf{D}_r\cos(\theta)r\right)^2 r dr$, which can be computed easily.

We compute the average relay energy consumption $\mathbb{E}\left[E_R^{(\text{DF})}\right]$ as follows.

$$\mathbb{E}\left[E_R^{(\text{DF})}\right] = \iint\limits_{(x,y)\in\text{REA}} \left(2^{2\mathcal{R}}-1\right)\frac{N}{2|h_r|^2}\mathbb{P}\left((x,y)\in\text{REA}\right)dxdy$$

$$= \left(2^{2\mathcal{R}}-1\right)\frac{N}{2}K_r\mathsf{D}_r^{A_r/10} \tag{5.23}$$

where $\mathcal{A}_{\text{sector}}$ and $\mathbb{P}_{\text{RTx}}$ are as defined in Eq. (5.9) and Eq. (5.18) respectively. These expressions for the user and relay consumptions conclude the analysis of Relay Efficiency Area.

5.7 Simulation results and energy-efficient relay location

In this section, we simulate the performance obtained by relay-aided transmissions and highlight energy-efficient relay configurations. For simulation, if not specified, we use the parameters of Table 5.1. Also note that we do not consider configurations where the direct link h_d is LOS but the user-to-relay link h_s is NLOS. In this case, DTx is performed due to the bad quality of the relaying path. We also assume that the relay remains inside the cell, i.e. $\mathsf{D}_r \leq \mathsf{R}_{\text{cov}}$.

Table 5.1 Default simulation parameters

Carrier frequency	f_c=2.6GHz
Noise	$N = 5.10^{-13}$
Max power	$P_U^{(\max)}$=500mW $P_R^{(\max)}$=1W
Node heights	H_B=30m, H_U=1.5m 10m $\leq \mathsf{H}_R \leq$ 30m
User rate	$\mathcal{R}$= 3bits/s/Hz
Other	h_d / h_s / h_r LOS

5.7.1 Validation of the model

We first validate the REA model, i.e. the proposed characteristic distances R_{DTx}, R_{cov} and $\mathsf{D}_{\min}$. An error occurs for a given user position if relaying is decided according to the REA model while DTx would be in reality more energy-efficient, or reversely, if DTx is decided rather than relaying. The validity of the proposed model mostly depends on the approximation of $\mathsf{D}_{\min}$, which embraces the relaying condition, as written in Eq. (5.11), and the energy-efficiency condition, i.e. $E_{\text{DTx}} \geq E_{\text{RTx}}$.

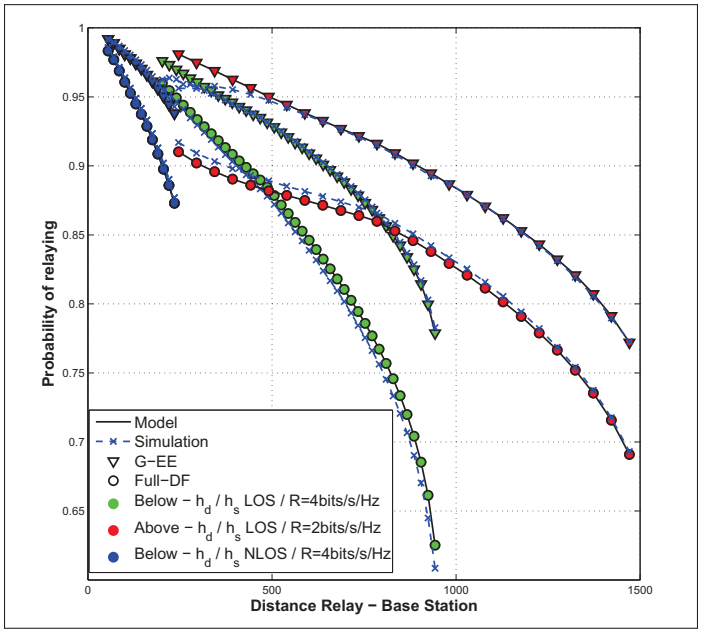


Figure 5.4 Validation of the relaying probability $\mathbb{P}_{\text{RTx}}$

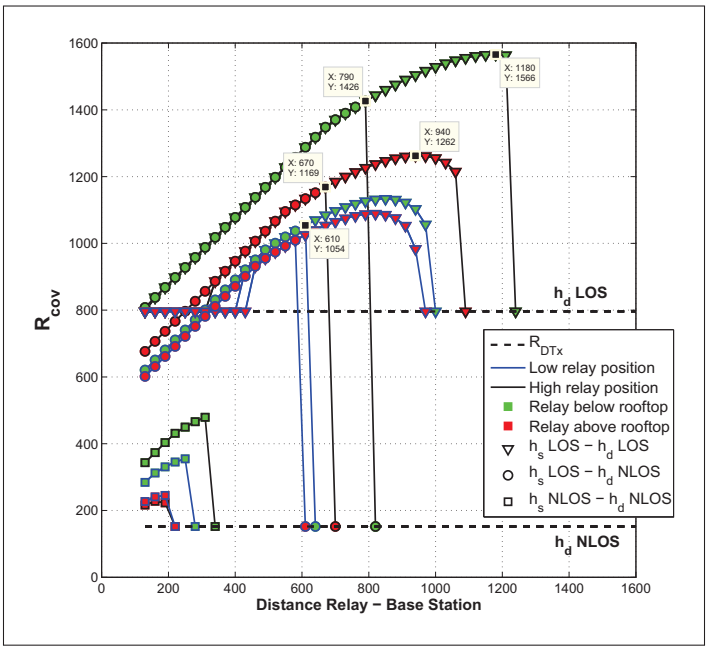
We plot in Figure 5.4 the probability of relaying as a function of the relay to base station distance D_r. This probability is evaluated both by simulation and computation of Eq. (5.18). We consider both relay coding schemes in various propagation environments. We see that the proposed characteristic distances lead to a probability of using the relay within 3% of the simulated probability. Thus, wrong decision occurs only for 3% of all the possible user positions in the cell, those errors being mostly located around the strait line $x = \mathsf{D}_{\min}$. Note that the REA model approaches the simulated consumption within 1%. Indeed, in the neighbourhood of $x = \mathsf{D}_{\min}$ where most of erroneous decisions occurs, distances from user to relay and user to base station do not vary significantly enough to imply notable differences in the energy consumption for either RTx or DTx.

5.7.2 Performance analysis on coverage extension and energy efficiency using Full-DF

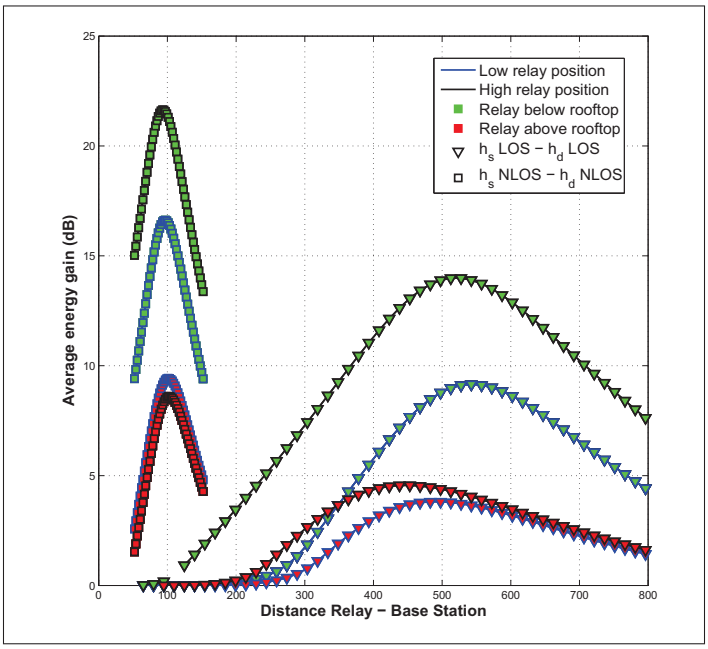
In this subsection, we analyze separately the coverage extension allowed by relaying given power constraints, and the minimal energy consumed for a given cell coverage. We first base our analysis on uplink transmissions for Full-DF, since this scheme is widely considered for relaying in future cellular standard such as LTE. As performance criteria, we consider the coverage and energy gains when the relay is optimally located, and analyze how much performance is degraded when the relay is away from this optimal location.

We plot in Figure 5.5 the maximal cell radius and the energy gain over DTx as functions of the relay-to-base-station distance D_r for various relay heights and LOS conditions. We consider the energy gain (in dB) using Full-DF compared to DTx, given that the cell coverage is not extended ($\mathsf{R}_{\text{cov}} = \mathsf{R}_{\text{DTx}}$). We refer to *low relay location* when $\mathsf{H}_R = 10$m (resp. $\mathsf{H}_R = 20$m) and the relay is below (resp. above) rooftop. Similarly, *high relay location* refers to $\mathsf{H}_R = 20$m (resp. $\mathsf{H}_R = 30$m) when the relay is below (resp. above) rooftop. The cut-off in the curves denotes that, above the corresponding D_r, the relay cannot cover the whole cell surface.

Result 5.1. If located below the rooftop, the relay station offers not only better performance when optimally located, but it also allows a wider range of relay locations for which performance is good, even if not optimal.



(a) Maximal coverage extension



(b) Energy gain over DTx (in dB) with $R_{cov} = R_{DTx}$

Figure 5.5 Impact of the environment using Full-DF

Since the relay-to-base-station link h_r is not the limiting one, it is beneficial to locate the relay such that the user-to-relay link h_s is stronger. Regarding energy, while a base-station-

like relay (i.e. above rooftop) hardly offers 5dB gain compared to DTx, a vicinity relay (i.e. below rooftop) achieves this gain for almost all relay locations and offers up to 14dB gain at optimal location. This significant gain mostly comes from cell-edge. While all cell-edge users are transmitting with almost full power using DTx, relaying allows them to achieve similar performance to a closer user. Regarding coverage, we can illustrate Result 1 using Figure 5.6. In this, the rectangles refer to the range of D_r for which the coverage is over 1200m. If the relay is above rooftop, like a base station, we have $R_{\text{cov}} \geq 1200m$ for $\mathsf{D}_r \in [680m, 1075m]$. However, for a vicinity relay, this range is extended by more than 200m, which offers much more flexibility for a designer to locate a relay station adequately. In Figure 5.6, the maximal coverage extension and the corresponding optimal D_r are respectively plotted with a green line and a green diamond. Note that the vicinity relay outperforms the base-station-like relay when optimally positioned and allows to increase the maximal coverage by 300m.

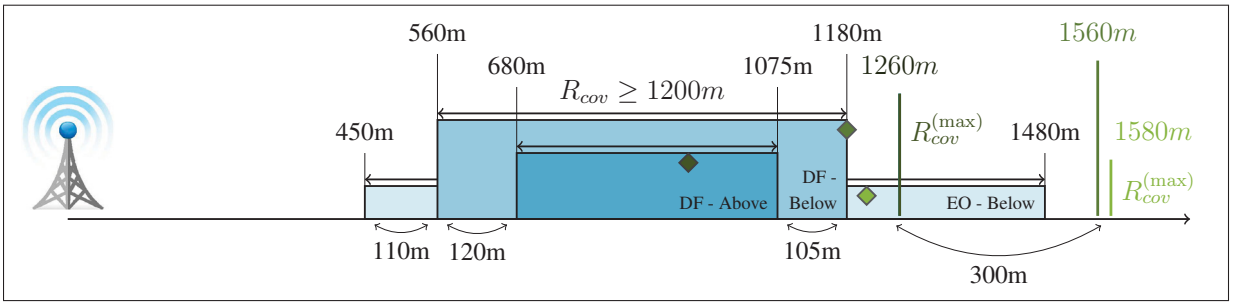


Figure 5.6 Impact of the relay height on the cell coverage (dark blue: with Full-DF, relay above rooftop, medium blue: with Full-DF, relay below rooftop, light blue: with EO-PDF, relay below rooftop)

Result 5.2. The system performance is strongly impacted by the effective relay height using a vicinity relay compared to a base-station-like relay.

Even if a vicinity relay still outperforms a base-station-like relay when being in a low position, decreasing the relay height from 20m (high relay position) to 10m (low position) causes a loss of more than 400m in the coverage and a 5dB energy loss, considering LOS condition (green triangles), as shown in Figure 5.5. Thus, with respect to both coverage and energy efficiency, the relay should be placed at the highest location possible as long as it remains below rooftop,

which allows higher quality and also more probable LOS condition for the user-to-relay link than a base-station-like relay.

Result 5.3. The energy-optimal distance D_r is different from the coverage-optimal distance.

Regarding coverage, it is beneficial to locate the relay far from the base station. On the contrary, regarding energy consumption, the relay should be placed around the middle of the cell, so as to decrease the value for $\mathsf{D}_{\min}$ and consequently, to increase the probability of energy-efficient relaying. For example, the coverage-optimal D_r is 1180m for a vicinity relay with LOS conditions (green triangles) while the energy-optimal D_r is only 520m.

Result 5.4. The loss of LOS of h_d mostly affects the range for beneficial distances D_r, and limitedly impacts the maximum cell coverage, especially when relay is positioned in a high location but still below rooftop.

This is illustrated in Figure 5.7, considering a vicinity relay. Indeed, with LOS for the direct link, a user located in the middle of the cell can use DTx, allowing the relay station to be far from the base station. With h_d NLOS, middle-cell users then require relaying to perform the transmission at same rate, meaning that the relay cannot be away from the cell center.

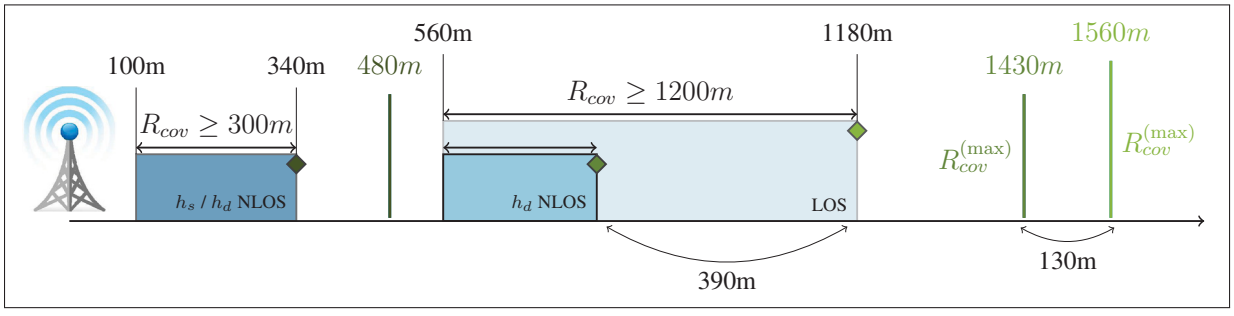


Figure 5.7 Impact of the LOS conditions on the cell coverage using Full-DF (dark blue: h_s/h_d both NLOS, medium blue: only h_d NLOS, light blue: h_s/h_d both LOS)

When both h_s and h_d are NLOS, performance are significantly degraded, but same comments as above can be made, i.e. vicinity relays outperform base-station-like relays but are more sensitive to the effective relay height.

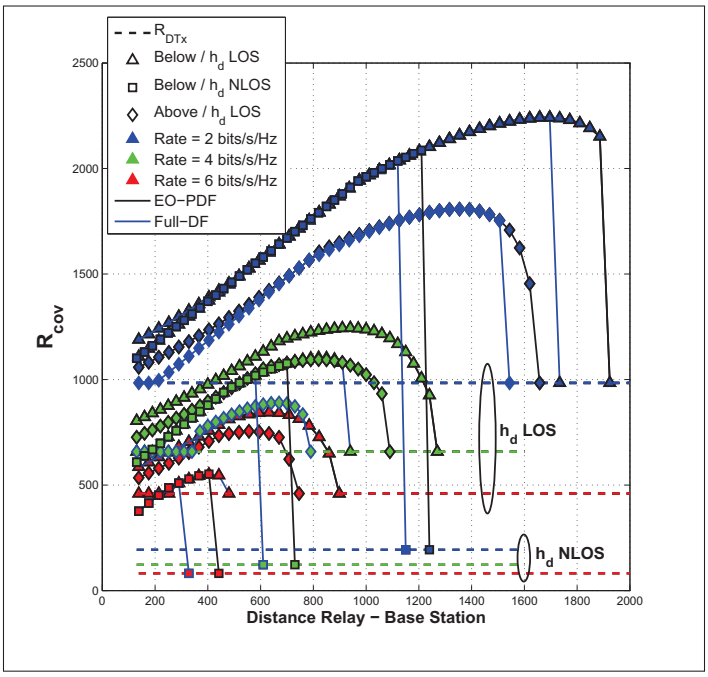
5.7.3 Performance evaluation with optimized coding scheme

So far, we have analyzed the system performance for the practical Full-DF scheme. We now compare the results with the upper-bound performance obtained by using the energy-optimal scheme EO-PDF. We plot in Figure 5.8 the maximum coverage radius and the average total energy consumed by each scheme as functions of the relay-to-base-station distance. We consider various environment conditions and several user rates. The energy consumption is computed with $\mathsf{R}_{\text{cov}} = \mathsf{R}_{\text{cov}}^{(\text{DF})}$. Also note that it consists of two parts due to the hexagonal cell constraint since the coverage radius is written as a minimum of two terms, as shown in Eq. (5.8). The first part of the curve corresponds to $\mathsf{X}_{\max} + \mathsf{R}_{\max}$ being the minimum.

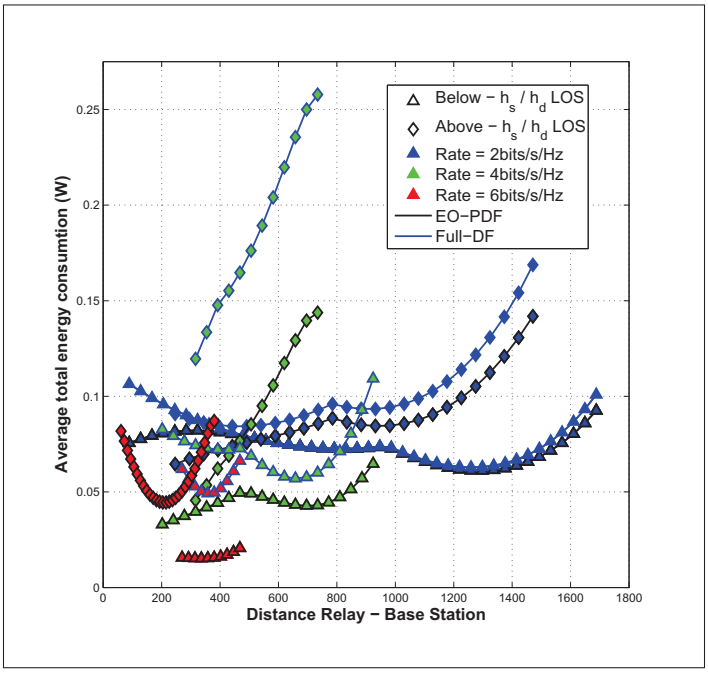
Result 5.5. With regards to both coverage and energy, the EO-PDF performs significantly better than Full-DF for high rates and under weaker channel conditions, i.e. when the direct link is NLOS (plotted with squares) or when the relay is above rooftop (diamonds).

Given good channel conditions (plotted with triangles), i.e. vicinity relay and LOS conditions, Full-DF performs close to the EO-PDF scheme for low rates ($\mathcal{R} \leq 2$bit/s/Hz), with regards to both coverage and energy. However, performances of Full-DF are rapidly degraded when the user rate increases. As illustrated in Figure 5.6 for $\mathcal{R} = 3$bit/s/Hz, the maximum cell coverage obtained with Full-DF remains close to the coverage given by the EO-PDF, but the range of beneficial relay locations for which $R_{\text{cov}} \geq 1200m$ is far smaller. Also note that high rates ($\mathcal{R} = 6$bit/s/Hz) are only achievable with base-station-like relay (i.e. above rooftop) by using the EO-PDF, and with vicinity relays, the coverage using Full-DF is only 65% of the one obtained using EO-PDF.

Same comments can be made with regards to energy efficiency. For example, when $\mathsf{D}_r = 650$m with $\mathcal{R} = 4$bits/s/Hz, the energy loss of Full-DF is around 1.2dB compared to EO-PDF. This loss is relatively low compared to the increase in complexity necessary to implement EO-PDF. However, the performance of Full-DF is significantly degraded when the relay is not optimally positioned. If the relay is closer to the base station ($\mathsf{D}_r \leq 400$m), between 2 and 4dB energy loss is observed. For $\mathcal{R} = 6$bits/s/Hz, this loss attains 6dB.



(a) Maximal coverage extension



(b) Energy consumption with $R_{cov} = R_{cov}^{(DF)}$

Figure 5.8 Comparison between Full-DF and EO-PDF

5.7.4 Analysis of downlink performance

So far, we have analyzed the performance of relaying for uplink transmissions, we know investigate downlink. Due to the symmetry of the scheme, Full-DF has the same performance for both. We thus focus only on the EO-PDF scheme. In Figure 5.9, we plot the average total energy consumed by the EO-PDF scheme given the cell coverage obtain for uplink ($\mathsf{R}_{\text{cov}} = \mathsf{R}_{\text{cov}}^{(\text{EO,UL})}$).

Result 5.6. Vicinity relays still outperforms base-station-like relays. Interestingly, EO-PDF is more energy-efficient for uplink transmissions than for downlink.

This is due to equal time division which is suboptimal, especially when the relay is located above rooftop. In this case, the link from base station to relay station is very strong. Thus, the EO-PDF scheme tends to relay the whole message, using limited rate splitting ($m \sim m_r$), and consumes most of energy in the second phase to take advantage of the beamforming gain between the base station and the relay. On the contrary, locating the relay station below rooftop allows more balance between h_s and h_r. This implies better use of rate splitting which is more suitable for equal time division.

5.7.5 Trade-off between coverage extension and energy efficiency

Here, we refine the analysis of the energy consumption vs. the covered area. Indeed, increasing the coverage radius implies that far-users are included within the cell. Those far-users consuming a lot of energy due to distance, the energy efficiency is thus decreased. However, extending the cell coverage reduces the number of base stations required to cover a given area, which is cost-efficient for the network deployment. Consequently, a network designer may opt for better energy efficiency regardless of the coverage, or may prefer to insure a larger coverage even if more energy is consumed in average. This constitutes a fundamental trade-off on relay deployment.

We thus propose to simulate performance based on the average energy consumption per unit area. As defined in [3], this is the ratio of the average energy consumption over the coverage area $\frac{\mathbb{E}[E]}{\mathcal{A}_{\text{sector}}}$, with $\mathbb{E}[E]$ and $\mathcal{A}_{\text{sector}}$ as expressed in Eq. (5.20) and (5.9) respectively. We plot in

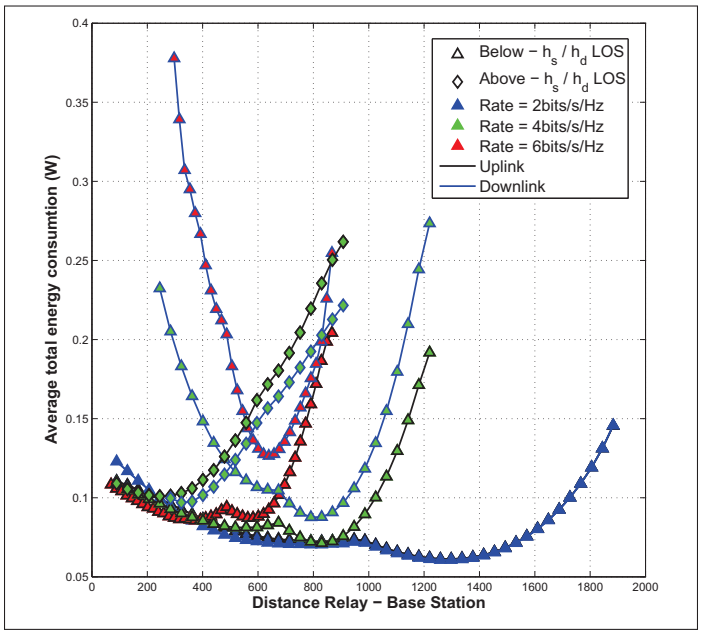
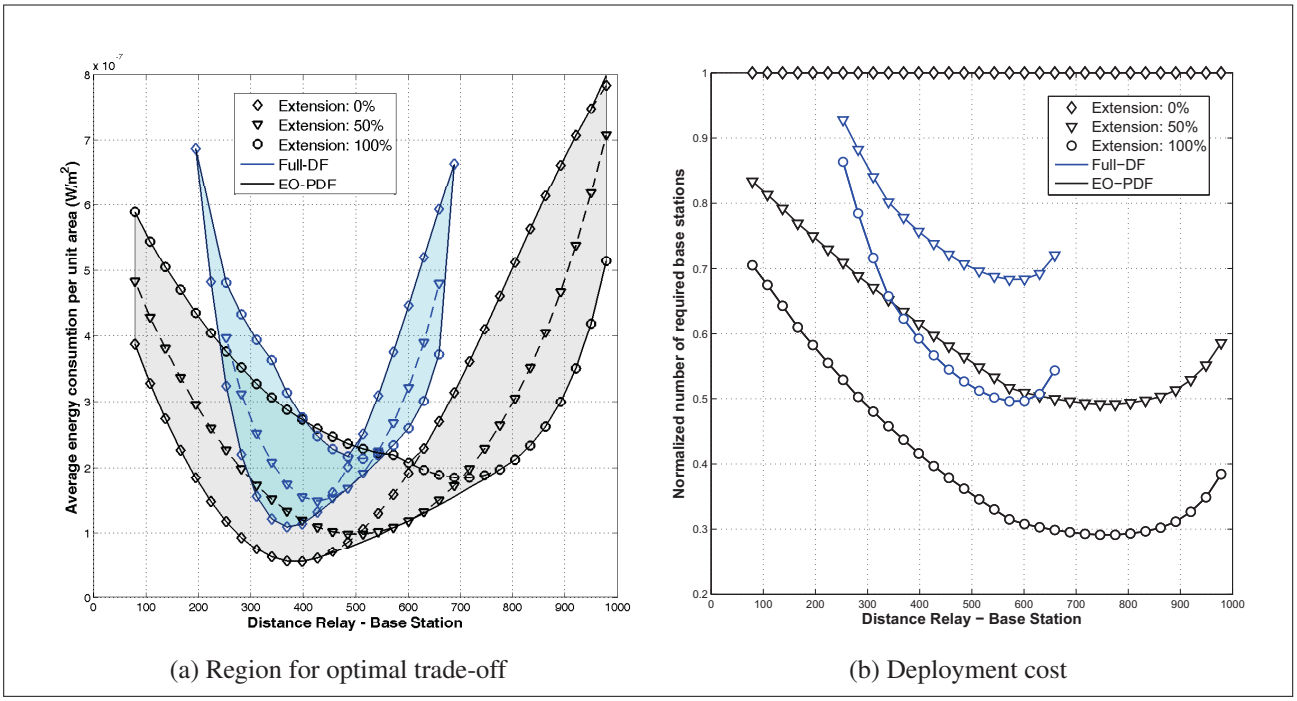


Figure 5.9 Downlink and uplink performance using EO-PDF

Figure 5.10a the average energy consumption per unit area considering several coverage radius, expressed as $R_{\text{cov}} = R_{\text{DTx}} + \beta\left(R_{\text{cov}}^{(\max)} - R_{\text{DTx}}\right)$, where β refers to the coverage extension in percentage and $R_{\text{cov}}^{(\max)}$ to the maximum coverage radius as plotted in Figure 5.8a. By plotting all curves given by $\beta \in [0, 100\%]$, we obtain the shaded area, which corresponds to the region of optimal trade-off. For any set $(\frac{\mathbb{E}[E]}{\mathcal{A}_{\text{sector}}}, \mathsf{D}_r)$ taken in this region, there exists a coverage R_{cov} for which this energy per unit area is optimal with this relay-to-base-station distance D_r. In Figure 5.10b, we plot the ratio of the area covered by direct transmission over the extended area $\frac{\mathcal{A}_{\text{DTx}}}{\mathcal{A}_{\text{RTx}}}$. It refers to the normalized number of required base stations, i.e. the deployment cost, and is directly linked to the maximal coverage extension, as plotted in Figure 5.8a. For example, at D_r =600m, using EO-PDF with maximal coverage extension requires only 30% of the base stations needed to cover the same area without cell extension.

The lower bound of the shaded area in Figure 5.10a gives the minimal energy consumed per unit area as a function of D_r and does not refer to a single coverage extension β. We observe once again that the EO-PDF scheme offers better performances than Full-DF both at the optimal distance $\mathsf{D}_r = 375$m and around this optimum. However, when maximum cover-



(a) Region for optimal trade-off

(b) Deployment cost

Figure 5.10 Trade-off between coverage extension and energy consumption, with $\mathcal{R} = 5$bits/s/Hz

age extension ($\beta = 100\%$) is considered, Full-DF outperforms the energy-optimized scheme EO-PDF for some values of D_r. As explained earlier, far-users that consume large amount of energy are included within the cell by using EO-PDF, thus illustrating the need for defining a trade-off between energy and coverage.

Result 5.7. The minimal average energy consumption per unit area is achieved when the cell coverage is not extended, i.e. when only energy efficiency is considered.

This minimum is attained with $\mathsf{D}_r = 375$m for both Full-DF and EO-PDF. Up to this distance D_r, no coverage extension ($\beta = 0\%$) is more energy-efficient than extended cell. Above this distance, coverage extension has to be considered for better energy efficiency. Note that, when the relay is far from the base station (over 800m for EO-PDF and 600m for Full-DF), reducing the coverage extension is no more energy-efficient.

Result 5.8. The minimal deployment cost, i.e. the minimal required number of base stations, is achieved when the relay is far from the base station, but shows in return severe loss regarding the energy consumed per unit area.

Between the energy-optimal relay location ($\mathsf{D}_r = 375$m) and the cost-optimal location (D_r =750m for EO-PDF and 575m for Full-DF), the energy loss, compared to the lower-bound of the shaded area, is equal to 5.2dB for EO-PDF and 3.3dB for Full-DF.

Result 5.9. There exists a relay-to-base-station distance D_r for which extending or reducing the cell coverage has minimum impact on the energy consumed per unit area.

This impact is illustrated in Figure 5.10a by the thickness of the shaded area at a given D_r. For example, at the energy-optimal D_r =375m, extending the cell coverage from R_{DTx} (Extension: 0%) to $R_{\text{cov}}^{(\text{max})}$ (Extension: 100%) induces a significant energy loss of 7dB for EO-PDF. This loss is reduced to 2.5dB when D_r =600m, distance for which the deployment cost is only 31% of the cost of non-extended cells.

To summarize results, energy-efficient relay placement should be jointly considered with cell coverage extension, and a moderate extension provides better performance in terms of energy. However, the optimal trade-off is, to our mind, neither the energy-optimal nor the cost-optimal relay location, but rather the location for which modifying the cell coverage has minimum impact on the energy consumed per unit area. This is particularly suitable for heterogeneous networks. Indeed, the deployment of unplanned nodes, especially at cell edge, implies that users are served by those nodes rather than by the base station, which virtually reduces the coverage radius of the macrocell. Relay deployment cannot be considered to be energy-efficient only for a given coverage and showing up a 7dB loss as soon as a pico- or femtocell is added, or reversely, switched off. This optimal trade-off is also relevant in networks where dynamic resource allocations and sleep mode operations enable the change of cell coverage as a function of the cell load and desired user rates.

5.8 Conclusion

We have investigated relay placement for noise-limited urban cells and considered both cell coverage and energy efficiency. Using an approach complementary to designing algorithms for optimal placement existing in the literature, we analyzed how the propagation environ-

ment, user rate and relay coding scheme affect the choice of the relay location and the network performance. To do so, we proposed a geometrical model for evaluating energy efficiency and coverage extension, and highlighted that a trade-off exists between them.

Two options can be deduced from this work to efficiently deploy relays in a cell. On the one hand, a network designer may choose to keep terminals simple and use two-hop routing. In this case, attention has to be paid on carefully positioning the relay station, so as to provide close to optimal coverage or energy gain. On the other hand, this designer may be constrained by the cell topology and restricted in potential relay locations. Then, more complex coding schemes, such as partial decode-forward, should be considered to fully benefit from the relay station. With regards to heterogeneous networks, the relay should be located such that increasing the cell coverage has a minimal impact on energy efficiency.

CHAPTER 6

MODELING AND ANALYSIS OF ENERGY EFFICIENCY AND INTERFERENCE FOR CELLULAR RELAY DEPLOYMENT

Fanny Parzysz[†], Mai Vu[‡], François Gagnon[†]

[†] Département de Génie Électrique, École de Technologie Supérieure,

1100 Notre-Dame West, Montreal, Quebec, (H3C 1K3) Canada

[‡] Department of Electrical and Computer Engineering, Tufts University,

161 College Ave, Medford, MA, (02155) USA

This paper has been submitted for publication

in *IEEE Journal on Selected Areas in Communications*.[1]

Summary of this Chapter: By relying on a wireless backhaul link, relay stations enhance the performance of cellular networks at low infrastructure cost, but at the same time, they can aggravate the interference issue. In this Chapter, which edits our third journal paper [30], we analyze for several relay coding schemes the maximum energy gain provided by a relay, taking into account the additional relay-generated interference to neighboring cells. First, we define spatial areas for relaying efficiency in log-normal shadowing environments and propose three easily-computable and tractable models. These models allow the prediction of 1) the probability of energy-efficient relaying, 2) the spatial distribution of energy consumption within a cell and 3) the average interference generated by relays. Second, we define a new performance metric that jointly captures both aspects of energy and interference, and characterize the optimal number and location of relays. These results are obtainable with significantly lower complexity and execution time when applying the proposed models as compared to system simulations. We highlight that energy-efficient relay deployment does not necessarily lead to interference reduction and conversely, an interference-aware deployment is suboptimal in the

[1] *Fanny Parzysz (2015)*

energy consumption. We then propose a map showing the optimal utilization of relay coding schemes across a cell. This map combines two-hop relaying and energy-optimized partial decode-forward as a function of their respective circuitry consumption. Such a combination not only alleviates the interference issue, but also leads to a reduction in the number of relays required for the same performance.

6.1 Introduction

In the urge to limit the energy consumption of cellular networks while maintaining service quality and ubiquitous access, relaying is a flexible and economical solution to enhance performance, eliminate coverage dead zones or alleviate traffic hot zones. It is envisioned as part of next-generation cellular networks, along with pico- and femtocells [127]. Unlike small cells, relay stations are not connected to the core network through a wireline backhaul connection but have to rely on wireless transmission to access the base station. This offers significant infrastructure cost reduction and deployment flexibility but, at the same time, can aggravate the interference issue. Exploring energy-optimized relaying jointly with interference reduction and choice of coding scheme opens new perspectives for efficient relay deployment.

6.1.1 Motivation and Prior work

As highlighted in [2, Chapter 6], the relay location within a cell can significantly affect the system performance. Its optimization is necessary to guarantee maximal gains within reasonable deployment cost and to avoid poor relay locations with negligible gains. Substantial efforts have been paid in optimizing the relay location with regards to capacity [128, 129, 130], coverage [56, 57] and energy [62, 131]. The serving area of relay stations has generally been envisaged as a small circular zone around each relay [132] or covers the cell edge exclusively as in [62]. However, such serving areas are neither energy- nor capacity-optimized. An analysis of the serving area has been provided in [133] for capacity and in [29] for energy, via the newly-defined relay efficiency area. These works, however, account only for the transmit energy but not circuitry consumption, and consider only path-loss but not shadowing, which is a significant cause of signal degradation.

On the one side, analyzing only the useful transmit power, i.e. the power radiated by the antenna, allows fair characterization of the network upper-bound performance, as done in [56, 57, 128, 129, 130]. However, it is generally not sufficient for an energy-efficient analysis. Depending on the considered technology and hardware quality, the energy dissipated in circuitry for site cooling, network maintenance and signal processing (even if no transmission is operated), may dominate the overall consumption [2, 3] such that the performance limits predicted by theory may not be realized. This is particularly true when the network does not operate at full load.

On the other side, considering the transmit energy alone provides a basis for complementary analysis of multi-cell networks, as it is directly related to inter-cell interference (ICI), another great challenge of cellular networks. Unlike base stations, relay stations are generally equipped with omni-directional antennas and can drastically increase the interference to a neighboring cell. Several questions remain open: Does energy-efficient relay deployment necessarily lead to interference reduction? Does deploying few relay stations far from the cell edge but serving a large part of the cell generate less interference than deploying many relay stations transmitting at low power but potentially closer to cell edge? Answering these questions necessitates the understanding of whether there exists a trade-off between energy-efficient and interference-aware relay deployment in cellular networks.

The ICI constraint has been investigated in the context of the relay placement for capacity and coverage enhancement in [56, 57, 130, 134], in which several models have been proposed. In [56, 57, 134], relays are located according to a predefined pattern, e.g. around a circle centered at their serving base station. In this case, interference appears to be very pessimistic, since each neighboring node (among the base stations and relays of each cell) is assumed to interfere with the reference user in each time slot with full transmit power. An existing model for interference evaluation allows for interference from nearby stations to be computed in closed-form, while afar interferers are modeled as a continuum rather than a discrete set [130]. While such models are particularly suitable for capacity maximization or coverage extension, they are not useful in analyzing or optimizing energy consumption because of the assumption that all

stations transmit at the maximum power while the actual transmitted power can be much less. Understanding how energy consumption is distributed across the cell is essential for an energy approach and, to the best of our knowledge, no interference analysis has been performed for energy-optimized relay-assisted networks.

6.1.2 Main contributions and Paper overview

Our objective is to investigate how the deployment of decode-forward-based (DF) relay stations, specifically their number and location, can reduce the overall energy consumption of a cell and how it affects at the same time the performance of neighboring cells, due to the additional interference. First, we propose spatial definitions of relaying efficiency, namely the Efficiency Areas, as the cell area for which a given performance requirement is satisfied. For shadowing environments, we define:

- the Relay Efficiency Area (REA), inside which a user has at least a probability $\mathbb{P}_T$ to be served by the relay; this REA extends the model in [29] to include shadowing.
- the Energy Efficiency Area (EEA), for which the energy consumption does not exceed E_T;

For the energy consumption, we account for (a) the transmit energy, (b) the additional energy consumed by the relay station for signal processing (decoding and re-encoding), (c) the energy loss at the RF amplifier and (d) the transmission-independent energy offset dissipated for network maintenance and site cooling, also called idle energy.

Simulating the network performance offers wide possibilities but is time-consuming and can rapidly turns out to be unfeasible in shadowing environments. Moreover, they provide neither generalization to other simulation settings nor performance limits like an analysis based on capacity bounds. Therefore, we propose in this paper easily-computable models for relaying probability (REA) and energy consumption (EEA), which allow meaningful performance analysis without requiring extensive simulations. Such models have wide application and offer valuable support for load balancing, resource management or base-station switch off.

Second, based on the proposed models for REA and EEA, we define a new framework for interference analysis. We characterize the average relay-generated interference to a given neighboring user and propose a model for it. Next, we define a new performance metric, denoted as Γ, which balances the energy gain provided by relays and the additional interference they generate.

Third, we apply the proposed models to investigate efficient relay deployment. With an energy-efficiency perspective, we characterize how the circuitry energy consumption, dissipated for signal processing and network maintenance, affects the network performance, as a function of the cell radius and the number of relays. We also highlight that energy-efficient relay deployment does not necessarily lead to interference reduction and conversely, that an interference-aware relay deployment is suboptimal in terms of energy reduction.

Finally, most approaches to efficient relaying are limited to the analysis of the simple two-hop relaying scheme. This scheme provides some performance enhancement but fails to capture the true potential of relays. Thus, in addition to two-hop relaying, we consider two energy-optimized partial decode-forward schemes of [28], which respectively minimize the network consumption and the relay consumption only. The second scheme is particularly relevant for analyzing relay-generated interference. Given the increased circuitry consumption of such optimized schemes, we draw a map showing the cell areas for the optimal use of each scheme that maximizes the energy-to-interference ratio Γ. The resulting map not only alleviates the interference issue, but also allows the reduction of the number of relays necessary to reach the same performance.

This paper is organized as follows. The cell configuration and channel model are described in Section 6.2. The reference coding schemes and related model for energy consumption are described in Section 6.3. The models for REA and EEA are proposed and analyzed respectively in Sections 6.4 and 6.5. The new framework for interference analysis is characterized in Section 6.6. The application of these model in energy- and interference-aware relay deployment is covered in Section 6.7 and Section 6.8 concludes this paper.

6.2 System Model for a Relay-Aided Cell

We present in this section the main assumptions on the cell environment and channel model. We use the following notation: BS stands for the base station, RS for the relay station and U for the user. Calligraphic $\mathcal{R}$ refers to the user rate. Upper-case letters denote constant distances, such as D, R for radius, or H for height. Lower-case letters stand for variable distances or angles, e.g. the mobile user coordinates (x, y) or (r, θ). Subscripts $_d$, $_b$ and $_r$ respectively refer to the link from BS to user (direct link), from BS to relay and from relay to user. $\mathbb{P}$ is used for probabilities and $\mathbb{E}$ for expectation. Finally, DTx refers to direct transmission, while RTx stands for relay-aided transmission, in the considered coding scheme.

6.2.1 Cell topology

We consider an hexagonal cell with edge distance D_b, as depicted in Figure 6.1. It consists of 3 sectorized base stations BS_i, $i \in \{1, 3\}$, located above surrounding buildings. We assume a typical radiation pattern for each base station and consider the antenna gain as given in [135]. Each 120°-sector i is served by N_r relay stations, all equipped with omni-directional antennas. A given relay is at a distance of D_r from its assigned BS. Finally, we assume a mobile user positioned at (x, y). It is associated with the closest base station and the relay that provides the highest energy gain. We do not consider techniques for relay selection or multi-relays schemes but the proposed analysis can be adapted to such configurations. We define the maximal coverage by the outage requirement $\mathbb{P}_{\text{out}}$ that has to be maintained throughout the whole cell.

6.2.2 Description of the relay channel

We consider half-duplex relaying performed, without loss of generality, in time division. We assume that the relay operates on the same frequency resource as the user it serves (namely *in-band* relaying in LTE-systems). The multiple access strategy allows orthogonality between users within the cell, such that only one user is served for a given time and frequency resource. A downlink transmission is carried out in two phases of equal duration. We assume complex Gaussian channels with independent additive white Gaussian Noises (AWGN) with equal vari-

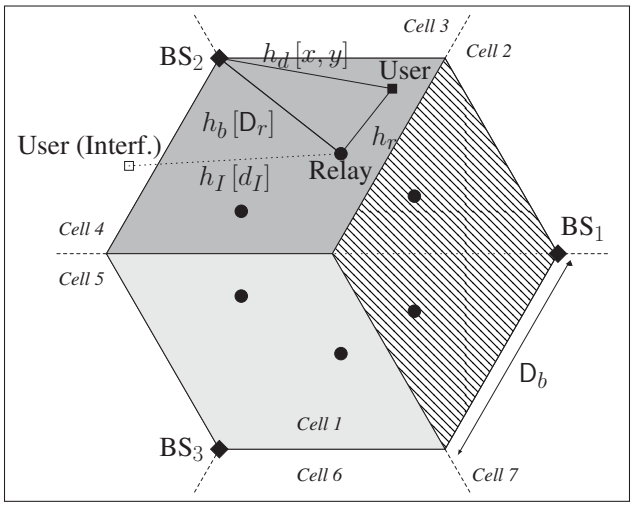


Figure 6.1 System model for a hexagonal cell aided by 6 relays

ance N on all links. We respectively denote h_d, h_b and h_r the channel coefficients from base station to user (direct link), from base station to relay (wireless backhaul link) and from relay to user. In addition, h_I refers to the channel from the interfering relay to a user in a neighboring cell.

We assume that the transmitted signal is degraded by both path-loss and shadowing. Note that we do not consider small-scale fading, since fading coefficients are rarely known for power allocation and terminals are designed to be sufficiently robust against such small-scale parameters, especially for relatively low mobility environments. Hence, this work is based on a per-block capacity analysis and can be understood as the mean relaying probability and energy consumption over a sufficient period of time such that small-scale fading is averaged out.

The BS, RS and user have different heights, moving neighborhoods and transmission ranges, such that different links have different properties, particularly in terms of path-loss. To fit this heterogeneity, we consider the channel model proposed in the WINNER II project [69], where the path-loss of link $k \in \{d, b, r, I\}$, denoted γ_k, is specified by four parameters A_k, B_k, C_k and D_k depending on the global location of the transmitter and receiver (street level, rooftop...). The shadowing coefficient s_k is modelled as a log-normal random variable of variance σ_k^2. We

assume that all s_k's are mutually independent. Channel gains are given by:

$$|h_k|^2 = \frac{s_k}{\gamma_k} = \frac{s_k}{K_k d^{\alpha_k}} \tag{6.1}$$

$$\text{with} \quad \begin{cases} \alpha_k = A_k/10 \\ 10\log_{10}(K_k) \quad = B_k + C_k \log_{10}\left(\frac{f_c}{5}\right) + D_k \log_{10}((\mathsf{H}_{\text{Tx}} - 1)(\mathsf{H}_{\text{Rx}} - 1)) \end{cases}$$

where H_{Tx} and H_{Rx} are the respective heights (in meters) of transmitter and receiver, d is the distance (in meters) between them and f_c the carrier frequency (in GHz).

6.3 Coding schemes and Models for Energy Consumption

We discuss the relaying schemes considered for analysis and their respective overall energy consumption. To simplify notation and focus on energy, we consider *normalized transmissions of unitary length*, thus setting up a direct relation between energy and power.

6.3.1 On the overall energy consumption

The transmit energy, i.e. the energy radiated at the output of the RF antenna front-end, is referred as $E^{(\mathsf{RF})}$ and varies as a function of the channel propagation condition. We assume individual energy constraints over the two transmission phases for the BS and RS, denoted as $\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}$ and $\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}$ respectively (with $^{(\mathsf{m})}$ for maximum). Most literature only accounts for this useful transmit energy, which is a fair assumption for capacity or coverage analysis. However, it is generally not sufficient for an energy analysis.

First, significant energy is dissipated in circuit electronics for data transmission, especially in the RF amplifier which usually performs with considerably low efficiency. We consider the simplified, yet meaningful, approach of [136], where the amplifier inefficiency is assumed linear in the transmit energy $E^{(\mathsf{RF})}$ and is characterized by a multiplicative coefficient, denoted η_R for the relay and η_B for the base station.

Second, we account for the circuitry consumption related to a transmission, i.e. to signal processing at the encoder and decoder. As in [2, Chapter 6], we model this consumption by the

following energy offsets: $\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})}$ for the BS transmission, $\mathsf{E}_{\mathsf{R}}^{(\mathsf{Rx})}$ and $\mathsf{E}_{\mathsf{R}}^{(\mathsf{Tx})}$ for the RS reception and transmission, and $\mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})}$ for the reception at the mobile user. Such energy offsets mostly depend on the quality of electronics and on the complexity of the signal processing performed at terminals, notably given by the considered coding scheme. Using simple relaying schemes can help decrease this energy consumption, but at the cost of potentially degraded network performance.

Third, in addition to the energy related to the transmission itself, we consider the transmission-independent consumption, also referred as the idle energy. This is the energy dissipated for site cooling, network maintenance and additional signaling, which is consumed whether or not data is transmitted. As proposed in [136], it is modelled by an offset, consumed at each station and denoted as $\mathsf{E}_{\mathsf{R}}^{(\mathsf{idle})}$ for relays, and $\mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})}$ for base stations.

6.3.2 Coding schemes considered for analysis

We now present the reference coding schemes and express for each scheme the overall energy consumption. In this work, we consider Gaussian signaling which accurately approximates OFDM-based communications [126], as used in practical systems including LTE and WiMAX. We consider downlink transmissions and assume that channel coefficients are known at transmitters to achieve the best possible performance. The power allocation aims at minimizing the energy consumption over the normalized transmission duration, given a fixed user rate $\mathcal{R}$.

6.3.2.1 Direct transmission (DTx)

Data is transmitted directly from the BS to the user over the two transmission phases, and not just only during the first one as done in most literature. This two-phase transmission allows fair performance comparison since, in this case, both direct and relay-aided transmissions have the same delay and consume the same time resource. The energy consumption is given by:

$$\mathsf{E}_{\mathsf{DTx}} = \eta_B E_B^{(\mathsf{RF})} + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})}\right) + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})} + N_r \mathsf{E}_{\mathsf{R}}^{(\mathsf{idle})}\right). \tag{6.2}$$

Without loss of generality in this analysis, we assume equal weights for the user, RS and BS consumption offsets. Setting $N_r = 0$ gives the consumption of a reference scenario, where no relay station is deployed. However, when $N_r > 0$, the energy consumption of DTx should account for the idle energy $N_r \mathsf{E}_{\mathsf{R}}^{(\mathsf{idle})}$ dissipated at relay stations, even if DTx does not actually use those relays.

6.3.2.2 Two-hop relaying (2Hop)

Two-hop relaying is the simplest decode-forward scheme and thus, gives lower-bounds of the performance that can be achieved with DF-based relaying. In this scheme, the user sends its message to the relay station with rate $2\mathcal{R}$ during Phase 1. Then, the relay decodes the message, re-encodes it with rate $2\mathcal{R}$ and forwards it to the destination during Phase 2. The base station finally decodes using only the signal received from the relay station (the direct link is ignored). The energy consumption is expressed as:

$$\begin{aligned}\mathsf{E}_{\mathsf{2Hop}} =& \eta_B E_B^{(\mathsf{RF})} + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{Rx})}\right) + \eta_R E_R^{(\mathsf{RF})} \\ &+ \left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})}\right) + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})} + N_r \mathsf{E}_{\mathsf{R}}^{(\mathsf{idle})}\right). \end{aligned} \tag{6.3}$$

We denote $\mathsf{E}_{\mathsf{2Hop}}^{(\mathsf{dsp})} = \mathsf{E}_{\mathsf{R}}^{(\mathsf{Rx})} + \mathsf{E}_{\mathsf{T}}^{(\mathsf{Rx})}$ the additional energy dissipated at the RS for decoding and re-encoding, in comparison with the energy offset of DTx described in Eq. (6.2).

6.3.2.3 Optimized partial decode-forward schemes

Finally, we consider for simulations the partial decode-forward (PDF) schemes optimized for energy proposed in [28]. In this, only part of the initial message is relayed, the rest being sent via the direct link. The base station divides the data message into two parts m_r and m_d using rate splitting. During the first transmission phase, m_r is broadcast to both the relay station and the user. At the end of this phase, only the relay decodes m_r and then re-encodes it. During the second phase, the relay sends $\tilde{m_r}$ and the base station jointly sends (m_r, m_d) using superposition coding. At the end of phase 2, the user jointly decodes m_r and m_d to recover

the initial message. Several power allocations have been proposed in [28]. In this work, we consider the two following schemes:

- Energy-optimized PDF (EO-PDF), which minimizes the total transmit energy consumption and is both energy- and rate-optimal (referred to as G-EE in [28]).
- Interference-at-Relay PDF (IR-PDF), which minimizes the energy transmitted by the relay only, thus minimizing the relay-generated interference (namely R-EE in [28]).

We refer the reader to [28] for the detailed power and rate allocation for m_r and m_d. Such optimized schemes arguably require complex implementation and fine synchronization but they provide theoretical upper-bounds of the performance achievable with DF-based relaying.

Little information is available on the circuitry energy consumed by such coding techniques (rate splitting, superposition coding, joint decoding...). To account for the increased complexity of these schemes, we consider an additional energy offset $\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)}$ to the overall consumption as follows:

$$\begin{aligned}\mathsf{E}_{\mathsf{EO}} =&\eta_B E_B^{(\mathsf{RF})} + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{Rx})}\right) + \eta_R E_R^{(\mathsf{RF})} \\ &+ \left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})}\right) + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})} + N_r \mathsf{E}_{\mathsf{R}}^{(\mathsf{idle})}\right) + \mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)}. \end{aligned} \tag{6.4}$$

In the performance analysis of Section 6.7, we will consider several values for $\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)}$. As for two-hop relaying, we denote $\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp})} = \mathsf{E}_{\mathsf{R}}^{(\mathsf{Rx})} + \mathsf{E}_{\mathsf{T}}^{(\mathsf{Rx})} + \mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)}$, representing the additional energy dissipated in circuitry to perform this relaying scheme.

6.4 Relaying Probability and Relay Efficiency Area with shadowing

We first characterize the relaying efficiency in shadowing environments and define the Relay Efficiency Area (REA) as the cell area where RTx has a larger probability to save energy compared to DTx. We analyze how to compute such area and propose a simplified model for fast and accurate performance evaluation. Such analysis is described in detail since it will be used as basis for the energy and interference analysis, presented in Sections 6.5 and 6.6.

6.4.1 A new definition of the Relay Efficiency Area

For a given channel realization (h_d, h_b, h_r), the area covered by a relay-aided cell can be divided into two geographical regions, depending on whether DTx or RTx should be performed. The wider is the cell area served by the relay, the more efficient can a relay station be considered. As highlighted in the introduction, various models are considered in the literature to characterize the serving area of a relay, but none are energy-optimized or account for shadowing. Here, we extend the definition of the pathloss-only Relay Efficiency Area (REA) in [29] to account also for shadowing environment. We define the REA as the set of all user locations for which relaying is *statistically* more energy-efficient than DTx or is necessary to satisfy a given *outage requirement* $\mathbb{P}_{\text{out}}$. Both coverage extension and energy gains are included in the REA definition.

Definition 6.1. The Relay Efficiency Area (REA) of a network in shadowing environment is defined by the pair $(\mathcal{A}_{\mathsf{R}}, \mathbb{P}_T)$. Any mobile user M located at (x, y) within the geographical area $\mathcal{A}_{\mathsf{R}}$ is served by the relay station with at least the probability $\mathbb{P}_T$, either because RTx is more energy-efficient or because DTx is not feasible, i.e. $\mathcal{A}_{\mathsf{R}} = \{M\,(x,y) \quad \text{s.t.} \quad \mathbb{P}_T \leq \mathbb{P}_{\text{RTx}}\}$, where $\mathbb{P}_{\text{RTx}}\,(x,y)$ is the probability for RTx at user $M\,(x,y)$.

The REA depends on the user rate, the relaying coding scheme and the channel radio propagation. As an example, the REA for two-hop relaying is illustrated in Figure 6.2 for various $\mathbb{P}_T$ and $\mathsf{E}_{\mathsf{2Hop}}^{(\mathsf{dsp})}$. The dotted lines refer to the model for the REA proposed in Section 6.4.2. Similarly, we can define $\mathbb{P}_{\text{DTx}}\,(x,y)$ as the probability for DTx at user $M\,(x,y)$, either because DTx is more energy-efficient or because RTx is not feasible. Then, the outage condition is expressed as $1 - \mathbb{P}_{\text{out}} \leq \mathbb{P}_{\text{RTx}}\,(x,y) + \mathbb{P}_{\text{DTx}}\,(x,y)$ for user $M\,(x,y)$.

6.4.2 Characterization of the relaying probability $\mathbb{P}_{\mathbf{RTx}}\,(x,y)$

For readability, the relaying probability is analyzed for downlink two-hop relaying only, but is valid for other scenarios, as discussed in Section 6.4.5. We first focus on the case $\mathsf{E}^{(\mathsf{dsp})} = 0$.

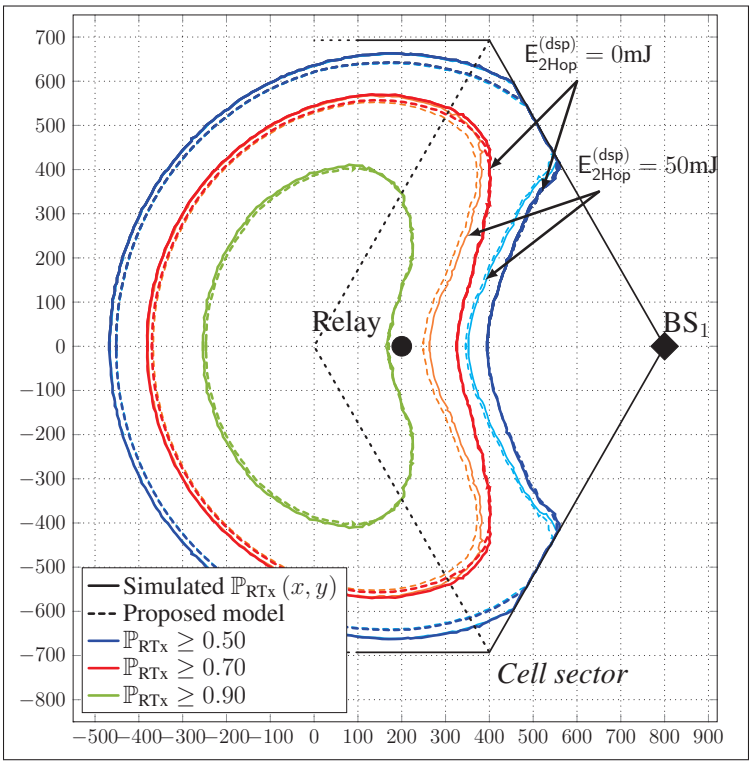


Figure 6.2 Relay Efficiency Area: simulation and model
(Two-hop relaying, $N_r = 1$, $\mathsf{D}_r = 700$m)

6.4.2.1 Notation for the energy consumption

We define E_d as the RF transmit energy consumed by the BS when DTx is used. Similarly, E_b and E_r stand for the energy consumed by the BS and the relay respectively when RTx is used. Considering Gaussian signaling, we deduce the energy consumption based on capacity formulas. To send data to user $M(x, y)$, we have:

$$E_d = \left(2^{\mathcal{R}} - 1\right) N \frac{\gamma_d [x, y]}{s_d} \; ; \; E_b = \left(2^{2\mathcal{R}} - 1\right) N \frac{\gamma_b [\mathsf{D}_r]}{2 s_b} \; ; \; E_r = \left(2^{2\mathcal{R}} - 1\right) N \frac{\gamma_r [x, y]}{2 s_r}, \quad (6.5)$$

where N is the variance of the AWGN, s_k is the shadowing coefficient and γ_k the path-loss, as given in Eq. (6.1). Then, we use superscript $^{(0)}$ to refer to the non-shadowing case, when only path-loss is considered (i.e. the variance of the shadowing coefficient s_k is $\sigma_k^2 = 0$). Assuming

log-normal shadowing environment, we have

$$E_k = \frac{E_k^{(0)}}{s_k} \sim \log \mathcal{N}\left(\mu_k, \sigma_k^2\right) \quad \text{with} \;\; \mu_k = \ln\left(E_k^{(0)}\right). \tag{6.6}$$

6.4.2.2 Probability for energy-efficient relaying

For a given channel realization, a transmission is relayed if DTx is not feasible or if RTx is more energy-efficient. This implies:

$$\mathbb{P}_{\text{RTx}}(x, y) = \mathbb{P}_{\text{CR}}(x, y) + \mathbb{P}_{\text{ER}}(x, y) \tag{6.7}$$

where $\mathbb{P}_{\text{CR}}$ and $\mathbb{P}_{\text{ER}}$ are as defined below. As we will see in the next section, these probabilities are necessary to define a model for the energy consumption and interference.

Definition 6.2. When the user-to-BS link is too weak and data cannot be sent using DTx given the channel realization, relaying is performed to extend the cell coverage and maintain the outage requirement. This is referred to as the Coverage Condition for Relaying (CR), which occurs with probability $\mathbb{P}_{\text{CR}}$. Given the energy constraints $\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}$ at BS and $\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}$ at RS, we have

$$\mathcal{C}_{\text{CR}} = \left\{ E_d > \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \;\; \cap \;\; E_b \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \;\; \cap \;\; E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \right\}.$$

Since energy consumption in each scenario is often independent, $\mathbb{P}_{\text{CR}}$ can generally be computed as $\mathbb{P}_{\text{CR}} = \mathbb{P}\left(E_d > \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}\right) \mathbb{P}\left(E_b \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}\right) \mathbb{P}\left(E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right)$. Similarly, we define the Coverage Condition for Direct transmission (CD) for which data can be sent only via the direct link. It occurs with probability $\mathbb{P}_{\text{CD}}$ and is generally computed in closed-form as for $\mathbb{P}_{\text{CR}}$.

Definition 6.3. When both DTx and RTx are feasible, relaying is performed if it is more energy-efficient. This defines the Energy-Efficient Condition for Relaying (ER), expressed as:

$$\mathcal{C}_{\text{ER}} = \left\{ E_k \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \;\; \cap \;\; E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \;\; \cap \;\; E_b + E_r \leq E_d \right\}.$$

with $k \in \{d, b\}$. The ER-condition has probability $\mathbb{P}_{\text{ER}}$.

Similarly, we define the Energy-Efficient Condition for Direct transmission (ED) for which DTx is more energy-efficient. It occurs with probability $\mathbb{P}_{\text{ED}}$:

$$\mathbb{P}_{\text{ED}} = \mathbb{P}\left(E_{\{d,b\}} \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \ \cap \ E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) - \mathbb{P}_{\text{ER}}. \tag{6.8}$$

Thus, the outage condition at user $M\,(x, y)$ is given by

$$1 - \mathbb{P}_{\text{out}} \leq \ \mathbb{P}_{\text{ER}} + \mathbb{P}_{\text{CR}} + \mathbb{P}_{\text{ED}} + \mathbb{P}_{\text{CD}} \quad \leq \ \mathbb{P}\left(E_{\{d,b\}} \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \ \cap \ E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) + \mathbb{P}_{\text{CR}} + \mathbb{P}_{\text{CD}}$$

Due to power constraints, $\mathbb{P}_{\text{ER}}$ is obtainable by a triple integral over s_d, s_b and s_r, as in Eq. (A III-1) of Appendix A. Thus, contrary to $\mathbb{P}_{\text{CR}}$, it is not separable and may not exist in closed form. Even tough numerical computation can be envisaged using mathematical software, this approach rapidly becomes unsuitable for large networks, even considering the simple two-hop scheme.

6.4.3 Proposed model for the REA

To address the key issue of computation, we propose a closed-form lower bound for $\mathbb{P}_{\text{ER}}$.

Lemma 6.1. The probability $\mathbb{P}_{\text{ER}}$ for energy-efficient relaying is lower-bounded by the sum $\mathbb{P}_{\text{low}} = \mathbb{P}_{\text{low}}^{(1)} + \mathbb{P}_{\text{low}}^{(2)}$ where

$$\begin{aligned}\mathbb{P}_{\text{low}}^{(1)} =& \max\Big(0, \mathbb{P}\left(E_{b+r} \leq E_d\right) \ - \Big[\ \mathbb{P}\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) \mathbb{P}\left(E_{b+r} \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) \\ &+\mathbb{P}\left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d\right)\Big]\Big) \end{aligned} \tag{6.9}$$

$$\mathbb{P}_{\text{low}}^{(2)} = \mathbb{P}\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}\right) \mathbb{P}\left(E_{b+r} \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) \tag{6.10}$$

Contrary to $\mathbb{P}_{\text{ER}}$, the lower-bound $\mathbb{P}_{\text{low}}$ consists of elementary probabilities that can be individually computed or have closed-form approximations.

Proof. See Appendix 1, p.187.

Also note that, using $\mathbb{P}_{\text{low}}$, we get an upper-bound for $\mathbb{P}_{\text{ED}}$ as

$$\mathbb{P}_{\text{ED}} \leq \mathbb{P}\left(E_{\{d,b\}} \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \ \cap \ E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) - \mathbb{P}_{\text{low}}. \tag{6.11}$$

Model 6.1 (REA). The Relay Efficiency Area, as characterized in Definition 6.1, is modeled in log-normal outdoors propagation environments by the pair $(\widehat{\mathcal{A}_{\mathsf{R}}}, \mathbb{P}_T)$, where $\widehat{\mathcal{A}_{\mathsf{R}}} = \{M(x,y) \text{ s.t. } \mathbb{P}_T \leq \mathbb{P}_{\text{low}}(x,y) + \mathbb{P}_{\text{CR}}(x,y)\}$. $\mathbb{P}_{\text{low}}$ and $\mathbb{P}_{\text{CR}}$ are given by Lemma 6.1 and Definition 6.2 respectively.

Any mobile user $M(x,y)$ circulating within $(\widehat{\mathcal{A}_{\mathsf{R}}}, \mathbb{P}_T)$ is also within $(\mathcal{A}_{\mathsf{R}}, \mathbb{P}_T)$, such that it has at least the probability $\mathbb{P}_T$ to save energy via relaying.

6.4.4 Accounting for the circuitry consumption

Up to now, we have focused on the case $\mathsf{E}^{(\mathsf{dsp})} = 0$, i.e. accounting only for the transmit energy. These results, however, can be generalized to show the overall energy consumption. In this case, data is relayed if the overall energy consumed using RTx is less than the consumption using DTx. As defined in Section 6.3.2, the overall consumption includes the RF amplifier efficiency (given by η_R and η_B) and the additional energy dissipated for decoding and re-encoding at the relay station (given by $\mathsf{E}^{(\mathsf{dsp})}$). The energy-efficient condition for relaying becomes:

$$\begin{aligned} \mathcal{C}_{\text{ER}}^{\circ} = & \left\{ E_b^{\circ} + E_r^{\circ} + \mathsf{E}^{(\mathsf{dsp})} \leq E_d^{\circ} \quad \cap \quad E_{\{d,b\}}^{\circ} \leq \eta_B \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \quad \cap \quad E_r^{\circ} \leq \eta_R \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \right\} \\ & \text{with} \quad E_k^{\circ} \sim \log \mathcal{N}\left(\mu_k + \ln(\eta_k), \sigma_k^2\right) \end{aligned}$$

Model 6.2 (REA, extended). The REA accounting for the overall energy consumption is modelled by $(\widehat{\mathcal{A}_{\mathsf{R}}^{\circ}}, \mathbb{P}_T)$, with $\widehat{\mathcal{A}_{\mathsf{R}}^{\circ}} = \{M(x,y) \quad \text{s.t.} \quad \mathbb{P}_T \leq \mathbb{P}_{\text{low}}^{\circ}(x,y) + \mathbb{P}_{\text{CR}}^{\circ}(x,y)\}$. $\mathbb{P}_{\text{low}}^{\circ}$ and $\mathbb{P}_{\text{CR}}^{\circ}$ are given by Lemma 6.1 and Definition 6.2 respectively, but with the following replacement:

$$\begin{aligned} E_k \longrightarrow \ & E_k^{\circ}, \quad k \in \{d,b,r\} \quad ; \quad E_{b+r} \longrightarrow \ E_b^{\circ} + E_r^{\circ} + \mathsf{E}^{(\mathsf{dsp})} \\ \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \longrightarrow \ & \eta_B \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \quad ; \quad \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \longrightarrow \ \eta_R \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} + \mathsf{E}^{(\mathsf{dsp})} \end{aligned} \tag{6.12}$$

This model is illustrated in Figure 6.2 for $\mathsf{E}_{\mathsf{2Hop}}^{(\mathsf{dsp})} = 50\text{mJ}$.

6.4.5 Discussion

6.4.5.1 Model validity

The proposed model has been validated under several outdoors environment settings and for both uplink and downlink transmissions. For the uplink, we generally have $\mathsf{E}_{\mathsf{U}}^{(\mathsf{m})} \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}$ for the user and relay energy constraints, implying that $\mathbb{P}_{\text{low}}^{(2)} = 0$. This model can be also used for any DF schemes for which the consumed energies E_d, E_r and E_b are log-normally distributed, such as the repetition-coded full DF scheme in [39] where the user decodes data using maximum ratio combining on the signal received from the BS and RS during both phases.

6.4.5.2 Model utilization

The primary use of the REA is to compute the spatial distribution of transmit energy consumption, as described in the next section. Thanks to the proposed model, many relay configurations and propagation environments can be analyzed in a reasonable time, which even allows finding the optimal relay location by exhaustive search. Moreover, this model can help decide between DTx and RTx when only statistics of the channel realizations are known at the transmitters (partial CSIT). In this case, the proposed model specifies for each user location the path which has the highest probability to save energy.

Next, the concept of Efficiency Area has wide application since it is based on the network geometry and ensures a minimum performance. Such a framework allows the deployment of relay stations in an efficient manner, simply by locating relays such that hotspots or cell regions with poor performance are included within the corresponding Efficiency Area. This concept can be particularly applicable to non-uniform random user locations, e.g. a hotspot-type distribution. More specifically, the probability for energy-efficient relaying offers valuable support for network resource management, relay selection, load-balancing, scheduling or BS switch off. For example, switch-off can be decided if the probability to reduce energy consump-

tion via relaying is above a given threshold $\mathbb{P}_T$, i.e. if the considered cell area is included in $(\mathcal{A}_{\mathsf{R}}, \mathbb{P}_T)$.

6.5 Energy Consumption and Energy Efficiency Area with shadowing

We analyze the spatial distribution of the energy consumption within the cell and define the Energy Efficiency Area (EEA) as the cell area for which the average transmit energy consumption does not exceed E_T. The EEA is related to the maximum energy necessary to transmit data at a given rate. Such a metric is relevant for performance analysis and encompasses fairness between the served users. Indeed, a relay station cannot be energy-efficient for users located in a certain area of the cell only, while showing unacceptably high energy consumption elsewhere.

6.5.1 Definition of the EEA

The average transmit energy, denoted as $\mathbb{E}\left[E^{(\mathsf{RF})}\right]$, consumed to send data to a given user $M(x,y)$, is equal to the sum of the energy consumed by BS only when DTx is used, and by both BS and RS when RTx is used. It is expressed as follows, with $E_{b+r} = E_b + E_r$:

$$\begin{aligned}\mathbb{E}\left[E^{(\mathsf{RF})}\right] =& \mathbb{P}_{\text{CR}}\,\mathbb{E}\left[E_{b+r}\,|\,\mathcal{C}_{\text{CR}}\right] + \mathbb{P}_{\text{ER}}\,\mathbb{E}\left[E_{b+r}\,|\,\mathcal{C}_{\text{ER}}\right] \\ &+ \mathbb{P}_{\text{CD}}\,\mathbb{E}\left[E_d\,|\,\mathcal{C}_{\text{CD}}\right] + \mathbb{P}_{\text{ED}}\,\mathbb{E}\left[E_d\,|\,\mathcal{C}_{\text{ED}}\right]. \end{aligned} \tag{6.13}$$

Definition 6.4. The Energy Efficiency Area (EEA) of a network in shadowing environment defines the energy range across the cell. It is characterized by the pair $(\mathcal{A}_{\mathsf{E}}, \mathsf{E}_T)$: the average transmit energy consumed by any mobile user $M(x,y)$ within $\mathcal{A}_{\mathsf{E}}$ does not exceed E_T, i.e.

$$\mathcal{A}_{\mathsf{E}} = \left\{ M(x,y) \quad \text{s.t.} \quad \mathbb{E}\left[E^{(\mathsf{RF})}\right] \leq \mathsf{E}_T \right\}.$$

Figure 6.3 illustrates $\mathcal{A}_{\mathsf{E}}$ and its corresponding model for various values of E_T. Also note that the areas for relaying probability $\mathcal{A}_{\mathsf{R}}$ and for energy $\mathcal{A}_{\mathsf{E}}$ are distinct. Next, we describe the proposed model for the EEA.

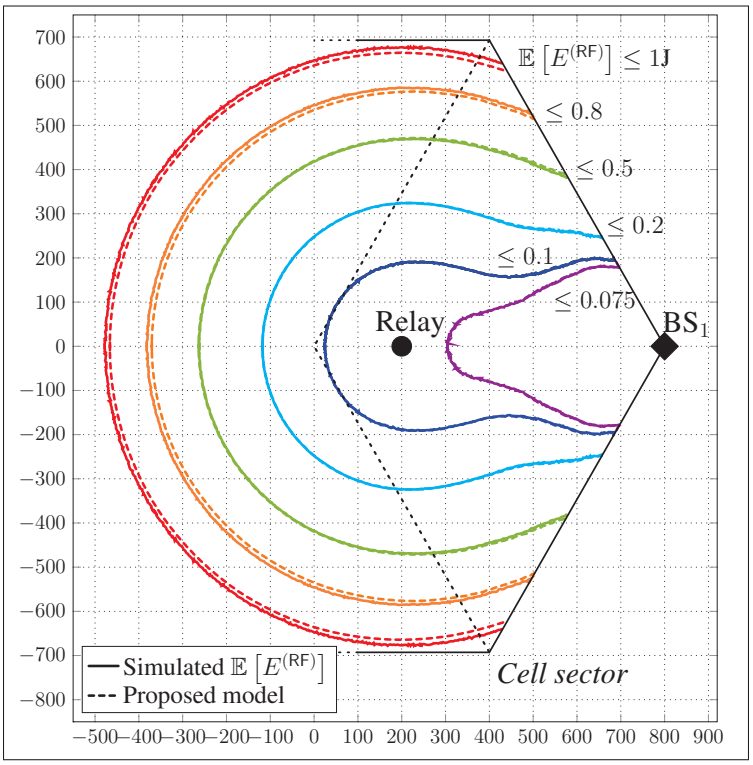


Figure 6.3 Energy Efficiency Area: simulation and model
(Two-hop relaying, $N_r = 1$, $\mathsf{D}_r = 700$m)

6.5.2 Analysis of the EEA with shadowing

We focus on the computation of $\mathbb{E}\left[E^{(\mathsf{RF})}\right]$ in Eq. (6.13) for log-normal shadowing environments. For the coverage conditions CD and CR, both expectations are computed in closed-form:

$$\mathbb{E}\left[E_{s+r} \,|\, \mathcal{C}_{\mathrm{CR}}\right] = \sum_{j \in \{r,b\}} g(j, \mathsf{E}_{\mathsf{j}}^{(\mathsf{m})}) \quad \text{and} \quad \mathbb{E}\left[E_d \,|\, \mathcal{C}_{\mathrm{CD}}\right] = g(d, \mathsf{E}_{\mathsf{b}}^{(\mathsf{m})}) \tag{6.14}$$

$$\text{where} \quad g(k, \mathsf{E}) = \exp\left(\mu_k + \frac{\sigma_k^2}{2}\right) \frac{\Phi\left(-\sigma_k + \frac{\ln(\mathsf{E}) - \mu_k}{\sigma_k}\right)}{\Phi\left(\frac{\ln(\mathsf{E}) - \mu_k}{\sigma_k}\right)}. \tag{6.15}$$

and where Φ is the cumulative distribution function of the standard normal distribution. On the contrary, the conditional expectations $\mathbb{E}\left[E_d \,|\, \mathcal{C}_{\mathrm{ED}}\right]$ and $\mathbb{E}\left[E_s + E_r \,|\, \mathcal{C}_{\mathrm{ER}}\right]$ for energy efficiency

can only be expressed in an integral form. We thus propose to bound both of them, as done for the probabilities $\mathbb{P}_{\text{ER}}$ and $\mathbb{P}_{\text{ED}}$ in the previous section.

Lemma 6.2. When RTx is more energy-efficient than DTx, the average consumed energy $\mathbb{P}_{\text{ER}}\mathbb{E}\left[E_{b+r} \,|\, \mathcal{C}_{\text{ER}}\right]$ in log-normal shadowing environments is lower-bounded by $\mathsf{E}_{\text{low}}^{(\text{ER})}$, where

$$\mathsf{E}_{\text{low}}^{(\text{ER})} = \exp\left(\mu_{b+r} + \frac{\sigma_{b+r}^2}{2}\right)\mathbb{P}_{\text{low}}^{(1,\text{ER})} + g(b+r, \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})})\mathbb{P}_{\text{low}}^{(2)}.$$

Here, $\mathbb{P}_{\text{low}}^{(1,\text{ER})}$ is computed similarly to $\mathbb{P}_{\text{low}}^{(1)}$ but considering the scaled distribution $\exp(\sigma_{b+r}^2)E_{b+r}$ rather than E_{b+r}. We recall that $\mathbb{P}_{\text{low}}^{(1)}$ and $\mathbb{P}_{\text{low}}^{(2)}$ are given by Eq. (6.9) and (6.10) respectively.

Proof. See Appendix 2, p.189.

Lemma 6.3. When DTx is more energy-efficient than RTx, the average energy consumption is equal to $\mathbb{P}_{\text{ED}}\mathbb{E}\left[E_d \,|\, \mathcal{C}_{\text{ED}}\right]$ and is upper-bounded by

$$\begin{aligned}\mathsf{E}_{\text{up}}^{(\text{ED})} = g(d, \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})})\mathbb{P}\left(E_d \le \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}\right)\mathbb{P}\left(E_b \le \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \,\cap\, E_r \le \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) - \exp\left(\mu_d + \frac{\sigma_d^2}{2}\right)\mathbb{P}_{\text{low}}^{(1,\text{ED})} \\ - \exp\left(\mu_d + \frac{\sigma_d^2}{2}\right)\frac{\Phi\left(-\sigma_d + \frac{\ln\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right)-\mu_d}{\sigma_d}\right) - \Phi\left(-\sigma_d + \frac{\ln\left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}\right)-\mu_d}{\sigma_d}\right)}{\Phi\left(\frac{\ln\left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}\right)-\mu_d}{\sigma_d}\right) - \Phi\left(\frac{\ln\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right)-\mu_d}{\sigma_d}\right)}\mathbb{P}_{\text{low}}^{(2)} \qquad (6.16)\end{aligned}$$

$\mathbb{P}_{\text{low}}^{(1,\text{ED})}$ is computed similarly to $\mathbb{P}_{\text{low}}^{(1)}$ but with the scaled distribution $\exp(\sigma_d^2)E_d$ instead of E_d.

Proof. The proof is similar to that of Lemma 6.2.

We now apply Lemmas 6.2 and 6.3 to deduce a model for the RF energy consumption.

Model 6.3 (EEA). We model the Energy Efficiency Area (EEA) of a relay-aided network in shadowing environment by the pair $(\widehat{\mathcal{A}_{\mathsf{E}}}, \mathsf{E}_T)$. The average transmit energy to any mobile user $M(x,y)$ within $\widehat{\mathcal{A}_{\mathsf{E}}}$ can be assumed to be below E_T, i.e.

$$\widehat{\mathcal{A}_{\mathsf{E}}} = \left\{M(x,y) \quad \text{s.t.} \quad \mathbb{E}\left[\widehat{E^{(\mathsf{RF})}}\right] \le \mathsf{E}_T\right\}$$

where $\mathbb{E}\left[\widehat{E^{(\mathsf{RF})}}\right]$ is given by Eq. (6.13) but with the bounds $\mathsf{E}_{\text{low}}^{(\text{ER})}$ and $\mathsf{E}_{\text{up}}^{(\text{ED})}$.

Simulations show that $\mathbb{E}\left[\widehat{E^{(\mathsf{RF})}}\right]$ can be considered as a tight upper-bound for $\mathbb{E}\left[E^{(\mathsf{RF})}\right]$ and any user located with the model area $\widehat{\mathcal{A}_{\mathsf{E}}}$ can be assumed to be within $\mathcal{A}_{\mathsf{E}}$ as well. The proposed model for the EEA can be extended to account for the circuitry energy consumption by using the same variable replacement as for the REA, given in Eq. (6.12).

6.6 A new framework for joint analysis of relay-generated ICI and energy

The assumption that all interferers are transmitting at maximum power is not realistic within an energy-efficiency context, and we propose to characterize the impact of a RS on the interference imposed on a neighboring user, given the *actual* relay energy consumption. Such an analysis provides helpful support for interference management techniques as well. Indeed, techniques to mitigate interference from neighboring cells, as considered for next generation cellular OFDMA systems [85, 86], require a fine understanding of the interference profile over time and frequency. This is particularly true in a network where both DTx and RTx occur simultaneously.

6.6.1 Approximation of the relay-generated interference

Based on the previous analysis for probability of energy-efficient relaying and overall energy consumption, we compute the average interference received at a mobile user located in another cell, at distance d_I from the interfering relay, as illustrated in Figure 6.1. To isolate the impact of a relay station, we do not consider other source of interference. This assumption is fair given that relays are usually equipped with omnidirectional antennas, as opposed to base stations.

We denote $M_I(x_I, y_I)$ as a given neighboring user, $\gamma_I = K_I d_I^{\alpha_I}$ as the path-loss between the relay station and this user, and s_I as the corresponding shadowing coefficient. The average interference received at $M_I(x_I, y_I)$ is expressed as

$$I(x_I, y_I) = \mathbb{E}_{s_I}\left[\mathbb{E}\left[E_r^{(\mathsf{RF})}\right]\frac{s_I}{\gamma_I}\right] = \mathbb{E}\left[E_r^{(\mathsf{RF})}\right]\frac{\exp\left(\sigma_I^2/2\right)}{K_I d_I^{\alpha_I}} \tag{6.17}$$

where $\mathbb{E}\left[E_r^{(\mathsf{RF})}\right]$ is the energy radiated by the relay, averaged over s_d, s_s and s_r. It is given by

$$\mathbb{E}\left[E_r^{(\mathsf{RF})}\right] = \mathbb{P}_{\text{CR}}\,\mathbb{E}\left[E_r^{(\mathsf{RF})} \,|\, \mathcal{C}_{\text{CR}}\right] + \mathbb{P}_{\text{ER}}\,\mathbb{E}\left[E_r^{(\mathsf{RF})} \,|\, \mathcal{C}_{\text{ER}}\right]. \tag{6.18}$$

First, the energy radiated by the relay when DTx is not feasible, due to fading, is computed as:

$$\mathbb{P}_{\text{CR}}\,\mathbb{E}\left[E_r^{(\mathsf{RF})} \,|\, \mathcal{C}_{\text{CR}}\right] = \mathbb{P}_{\text{CR}} g(r, \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}) \tag{6.19}$$

with g given in Eq. (6.15). Second, using a proof similar to Lemma 6.2, the energy radiated by the relay station when RTx is more energy-efficient than DTx, is lower-bounded as follows:

$$\mathbb{P}_{\text{ER}}\mathbb{E}\left[E_r^{(\mathsf{RF})} \,|\, \mathcal{C}_{\text{ER}}\right] \geq \mathsf{E}_{\text{low}}^{(\text{I})} = \exp\left(\mu_r + \frac{\sigma_r^2}{2}\right)\left(\mathbb{P}_{\text{low}}^{(1,\text{I})} + \mathbb{P}_{\text{low}}^{(2,\text{I})}\right), \tag{6.20}$$

where $\mathbb{P}_{\text{low}}^{(1,\text{I})}$ and $\mathbb{P}_{\text{low}}^{(2,\text{I})}$ are computed similarly to $\mathbb{P}_{\text{low}}^{(1)}$ and $\mathbb{P}_{\text{low}}^{(2)}$ in Eq. (6.9) and (6.10), but considering the scaled distribution $\exp(\sigma_r^2)E_r$ instead of E_r.

Model 6.4 (**Approximation for** $I(x_I, y_I)$)**.** The relay-generated interference at a given neighboring user $M_I(x_I, y_I)$, located at a distance d_I from the relay station, is lower-bounded by

$$\widehat{I}(x_I, y_I) = \mathbb{E}\left[\widehat{E_r^{(\mathsf{RF})}}\right]\frac{\exp\left(\sigma_I^2/2\right)}{K_I d_I^{\alpha_I}} = \left(\mathbb{P}_{\text{CR}} g(r, \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}) + \mathsf{E}_{\text{low}}^{(\text{I})}\right)\frac{\exp\left(\sigma_I^2/2\right)}{K_I d_I^{\alpha_I}},$$

where g is given by Eq. (6.15) and $\mathsf{E}_{\text{low}}^{(\text{I})}$ by Eq. (6.20).

One can argue that upper-bounding the interference $\widehat{I}(x_I, y_I)$ would be more suitable for performance analysis. However, we highlight that the proposed lower-bound is tight, as shown in Section 6.7.1. It is thereby accurate enough to investigate interference-aware relay deployment. The proposed model for the ICI can be extended to account for the circuitry energy consumption by using the variable replacement of Eq. (6.12), similarly to REA and EEA.

6.6.2 A new metric for analyzing energy and interference

A relay station can provide significant *energy gain* and coverage extension for the cell it serves. But, at the same time, it is an additional source of interference, implying that neighboring cells experience an *energy loss* to maintain the same data rate for their own users. In consequence, a relay deployment is efficient if the achieved energy gain, referred as υ_{Gain}, is higher than the resulted energy loss, referred as υ_{Loss}. We propose to use their ratio as a metric to jointly capture the aspects of energy and interference.

To evaluate the energy gain υ_{Gain}, we consider a user $M(x,y)$ served by BS_1. We compare the energy $\mathsf{E}_1^{(\mathsf{N_r}=0)}$ consumed to send data to this user when BS_1 is not supported by relay stations ($N_r = 0$) and the energy $\mathsf{E}_1^{(\mathsf{N_r})}$ consumed when BS_1 is supported by N_r relay stations. We have:

$$\begin{cases} \mathsf{E}_1^{(\mathsf{N_r}=0)}[x,y] = & \eta_B E^{(\mathsf{RF})}[x,y] + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})}\right) + \mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})} \\ \mathsf{E}_1^{(\mathsf{N_r})}[x,y] = & \eta_B E_B^{(\mathsf{RF})}[x,y] + \eta_R E_R^{(\mathsf{RF})}[x,y] + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})} + \mathsf{E}^{(\mathsf{dsp})}\right) \\ & + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})} + N_r \mathsf{E}_{\mathsf{R}}^{(\mathsf{idle})}\right). \end{cases}$$

Here, the various energy offsets accounted in $\mathsf{E}_{\mathsf{idle}}$ are described in Section 6.3.2. We recall that we can focus only on Sector 1 since the three base stations BS_1, BS_2 and BS_3 within the considered cell are orthogonal and do use the same resource (in time and frequency).

Similarly, to evaluate the energy loss υ_{Loss}, we consider a user $M(x,y)$, which is located in a neighboring cell $i \neq 1$ and performs DTx. We compare the energy $\mathsf{E}_{\mathsf{i}}^{(\mathsf{N_r}=0)}$ consumed by the neighboring BS_i to send data to this user when BS_1 is not supported by relay stations ($N_r = 0$) and the energy $\mathsf{E}_{\mathsf{i}}^{(\mathsf{N_r})}$, consumed to maintain the same rate, when BS_1 is supported by N_r relay stations generating interference. To isolate the impact of the N_r relays, we assume an ideal network, without any other source of interference. Denoting $I(x,y)$ the interference received at $M(x,y)$ as defined in Eq. (6.17) of previous interference analysis, we have:

$$\forall i \neq 1 \qquad \begin{cases} \mathsf{E}_{\mathsf{i}}^{(\mathsf{N_r}=0)}[x,y] = \eta_B E^{(\mathsf{RF})}[x,y] + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})}\right) + \mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})}. \\ \mathsf{E}_{\mathsf{i}}^{(\mathsf{N_r})}[x,y] = \eta_B E^{(\mathsf{RF})}[x,y]\left(1 + \frac{2I(x,y)}{N}\right) + \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})}\right) + \mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})}. \end{cases}$$

Remark: In addition to energy gain, relay stations also provide coverage extension. To account for such extension and compute υ_{Gain} and υ_{Loss}, we do not consider a power constraint for DTx.

Definition 6.5. To capture both the energy and interference aspects, we define the ratio Γ as

$$\Gamma = \frac{\upsilon_{\text{Gain}}}{\upsilon_{\text{Loss}}} \quad \text{with} \quad \upsilon_{\text{Gain}} = \mathbb{E}\left[\frac{\mathsf{E}_1^{(\mathsf{N_r}=0)} - \mathsf{E}_1^{(\mathsf{N_r})}}{\mathsf{E}_1^{(\mathsf{N_r}=0)}}\right] \text{ and } \quad \upsilon_{\text{Loss}} = \mathbb{E}\left[\frac{\sum\limits_{i\neq 1}\mathsf{E}_\mathsf{i}^{(\mathsf{N_r}=0)} - \sum\limits_{i\neq 1}\mathsf{E}_\mathsf{i}^{(\mathsf{N_r})}}{\sum\limits_{i\neq 1}\mathsf{E}_\mathsf{i}^{(\mathsf{N_r}=0)}}\right].$$

Here, υ_{Gain} is averaged over all users served by BS_1 and υ_{Loss} is averaged over all users located in the neighboring cells 2 to 7, as depicted in Figure 6.1. If $\Gamma > 1$, the considered relay configuration is efficient, if $0 < \Gamma < 1$, the relay stations result in more energy loss for neighboring cells than they actually provide energy gain in their own cell. If $\Gamma < 0$, relaying does not provide any energy gain, due to the circuitry consumption.

6.7 Performance analysis for Energy- and ICI-efficient Relay Deployment

In this section, we first validate the proposed models for relaying probability, energy consumption and interference. Then, we jointly analyze the energy consumption and the generated ICI using two-hop relaying. In the last subsection, the impact of the relay coding scheme on the network performance is explored. If not specified, we consider the simulation parameters of Table 6.1, taken from [2, 3, 136, 137]. For the channel gains, the direct link h_d and interference link h_I are modelled by scenario C2 of the WINNER II project [69], the RS-to-user link h_r by scenario B1, the BS-to-RS link h_b by B5c. We recall that we consider normalized transmissions of unitary length, setting up a direct relation between energy and power.

6.7.1 Models validation

For validation of the proposed model for REA, we account for all users $M(x, y)$ located in the simulated $(\mathcal{A}_{\text{R}}, \mathbb{P}_T)$ but not declared in $\left(\widehat{\mathcal{A}_{\text{R}}}, \mathbb{P}_T\right)$, for some given threshold $\mathbb{P}_T$ (or reversely, $M(x, y)$ is declared inside while it is actually outside). This means that, for such user, the effective relaying probability, obtained by Monte-Carlo simulations, is s.t. $\mathbb{P}_{\text{RTx}} \geq \mathbb{P}_T$ but the proposed lower-bound gives $\mathbb{P}_{\text{low}}(x, y) + \mathbb{P}_{\text{CR}}(x, y) \leq \mathbb{P}_T$ (or reversely). We define the error

Table 6.1 Simulation parameters

Energy	
$\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}$	1J
$\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}$	500mJ
$\mathsf{E}_{\mathsf{B}}^{(\mathsf{idle})}$	25mJ
$\mathsf{E}_{\mathsf{R}}^{(\mathsf{idle})}$	10mJ
$\mathsf{E}_{\mathsf{2Hop}}^{(\mathsf{dsp})}$	0-50mJ
$\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)}$	0-50mJ
η_B	2.66
η_R	3.1
$\mathsf{E}_{\mathsf{B}}^{(\mathsf{Tx})} + \mathsf{E}_{\mathsf{U}}^{(\mathsf{Rx})}$	90mJ

Channel	
N	-93dBm
σ_d	6dB
σ_b	3dB
σ_r	4dB
σ_I	6dB
Normalized Tx (1s)	

Others	
$\mathbb{P}_{\text{out}}$	0.02
$\mathcal{R}$	3bit/ch.use
f_c	2.6GHz
H_{b}	30m
H_{r}	20m
H_{u}	1.5m

ratio ζ_{R} as the proportion of such erroneous users, i.e.

$$\zeta_{\mathsf{R}} = \frac{\iint \mathbb{1}_{\mathcal{E}_{\mathsf{R}}} dxdy}{\iint \mathbb{1}_{\mathcal{A}_{\mathsf{R}}} dxdy} \quad \text{with} \quad \mathcal{E}_{\mathsf{R}} = \left\{ M(x,y) \in \mathcal{A}_{\mathsf{R}} \cap M(x,y) \notin \widehat{\mathcal{A}_{\mathsf{R}}} \right\} \cup \left\{ M(x,y) \notin \mathcal{A}_{\mathsf{R}} \cap M(x,y) \in \widehat{\mathcal{A}_{\mathsf{R}}} \right\}$$

Similarly, the error ratio ζ_{E} for the EEA refers to the proportion of erroneous users, for which $\mathbb{E}\left[E^{(\mathsf{RF})}\right] \leq \mathsf{E}_T$ and $\mathbb{E}\left[\widehat{E^{(\mathsf{RF})}}\right] \geq \mathsf{E}_T$ (or reversely), for some given threshold E_T. For the interference analysis, we focus on the approximation of the average energy radiated by the relay and define the error ratio ζ_{I} as the proportion of users for which $\mathbb{E}\left[E_r^{(\mathsf{RF})}\right] \leq \mathsf{E}_{T,r}$ and $\mathbb{E}\left[\widehat{E_r^{(\mathsf{RF})}}\right] \geq \mathsf{E}_{T,r}$ (or reversely), for some given threshold $\mathsf{E}_{T,r}$.

We plot in Figure 6.4(a) (resp. b and c) the error ratio ζ_{R} (resp. ζ_{E} and ζ_{I}) obtained for a wide range of $\mathbb{P}_T$ (resp. E_T and $\mathsf{E}_{T,r}$) and several $\mathsf{E}^{(\mathsf{dsp})}$ (over a unitary time unit). More precisely, the plotted ζ_{X} stands for the ratio averaged over various RS-to-BS distances ($\mathsf{D}_b \in [600, 1000]$m), various user rates ($\mathcal{R} \in [2,4]$bits/ch. use) and the two outdoors propagation environments described in [29, Appendix A]. For the purpose of validation, we consider a wide cell coverage by fixing the outage requirement $\mathbb{P}_{\text{out}}$ to 0.1, which is very large for encoded data.

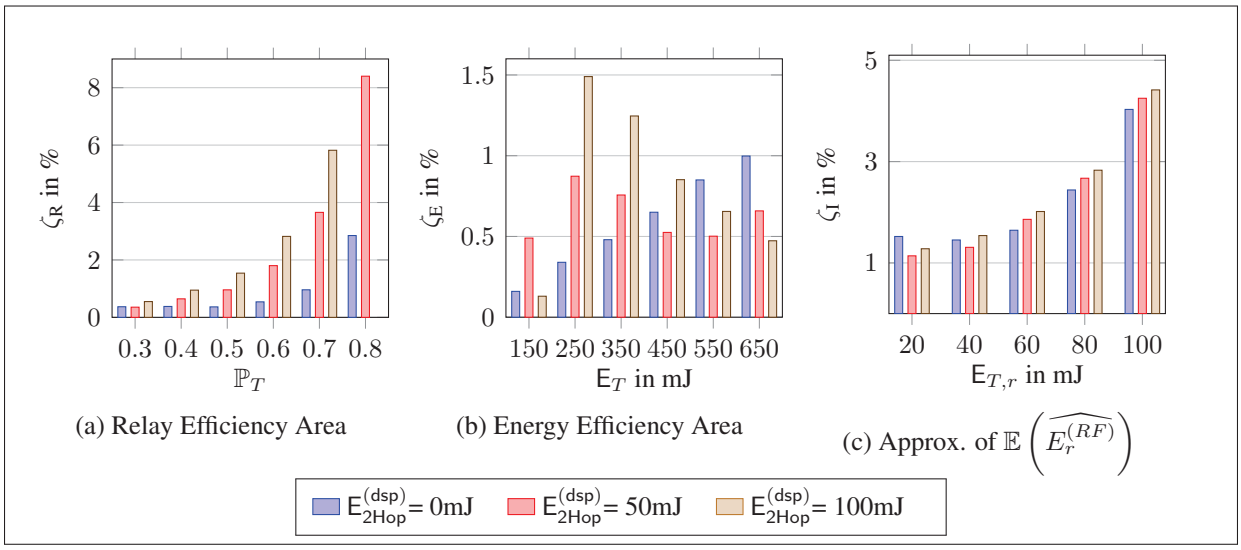


Figure 6.4 Validation of the proposed models: error ratio ζ for various simulation settings

From Figure 6.4(a), we observe that the error ratio for the REA does not exceed 3% when the circuitry consumption is not considered ($\mathsf{E}^{(\mathsf{dsp})} = 0$). Although the model for $\mathsf{E}^{(\mathsf{dsp})} > 0$ is less accurate, such error increase does not impact at all the proposed model for energy consumption, as illustrated in Figure 6.4(b). Indeed, when an error occurs and a user $M(x, y)$ is wrongly declared in $\widehat{\mathcal{A}_{\mathsf{R}}}$ while it is not (or reversely), we have $E_b + E_r \simeq E_d$. For the EEA, the error ration does not exceed 1.5% and, for the interference approximation, it is below 5%, as plotted in Figure 6.4(c). Furthermore, error are mostly located at cell edge and, by considering restricted cell coverage ($\mathbb{P}_{\text{out}} = 0.02$), as for the rest of this paper, the error ratio for the ICI falls under 2.5%.

The proposed models for the REA, EEA and ICI thus provide very efficient frameworks for performance analysis with regards to both accuracy and savings in the simulation time. Non-model based simulations were also performed taking 50 000 samples for the channel gains of each link and required several hours for a single relay configuration. By comparison, model-based simulations were completed in less than 3 seconds. Subsequently, we will refer to "simulations" for "model-based simulations".

6.7.2 On the minimal energy consumption per unit area

As a first step, we do not consider the impact of relays in terms of interference and analyze the relay performance regarding the EEA only, in the shadowing model used in the WINNER project. The Joule-per-bit metric has been widely used for energy efficiency analysis. Yet, in practice, a large part of the network is primarily providing coverage and does not operate at full load, even at peak traffic hours. Due to the energy dissipated in circuitry to maintain the network operational, the energy efficiency can be particularly poor under low-traffic loads and restricted coverage [3]. To capture the aspect of the cell coverage, we consider the maximal energy consumption $\mathsf{E}_{\max}$ that is required to send data at rate $\mathcal{R}$ to any user located within the cell sector, i.e. $\mathcal{A}_{\text{sector}} \in (\mathcal{A}_{\mathsf{E}}, \mathsf{E}_{\max})$, and divide it by the sector area $\mathcal{A}_{\text{sector}}$. It is expressed in Joule-per-square-meter and denoted as Ψ:

$$\Psi = \frac{\mathsf{E}_{\max} + \mathsf{E}_{\mathsf{idle}}}{\mathcal{A}_{\text{sector}}} \quad \text{where} \quad \mathcal{A}_{\text{sector}} = \frac{\sqrt{3}}{2}\mathsf{D}_b^2. \tag{6.21}$$

Here, E_T accounts for the transmit energy $\mathbb{E}\left[E^{(\mathsf{RF})}\right]$ as in Eq. (6.13), the RF amplifier coefficients η_R and η_B and the energy $\mathsf{E}^{(\mathsf{dsp})}$ dissipated at the relay for decoding and re-encoding. $\mathsf{E}_{\mathsf{idle}}$ refers to the other energy offsets described in Section 6.3.2, for the considered coding scheme. In Figure 6.5, we plot the minimal feasible energy per unit area Ψ as a function of the cell radius D_b, considering different values for N_r and $\mathsf{E}^{(\mathsf{dsp})}$. To do so, for each considered set $\left(\mathsf{D}_b, N_r, \mathsf{E}^{(\mathsf{dsp})}\right)$, we find the location for the N_r relays which minimizes Ψ. Note that proceeding this way would not have been possible in a reasonable time without using the proposed models.

Remark: For comparison purpose, Figure 6.5 also plots the value for Ψ achieved when the relay location maximizes the metric Γ (namely "Optimal for Γ" in the figure). We will come back to this point in the following subsection.

Result 6.1. The energy offset $\mathsf{E}^{(\mathsf{dsp})}$, consumed for decoding and re-encoding at the relay station has severe impact on the cell energy efficiency, except for very large cell radius, where the overall energy is dominated by the RF transmit consumption.

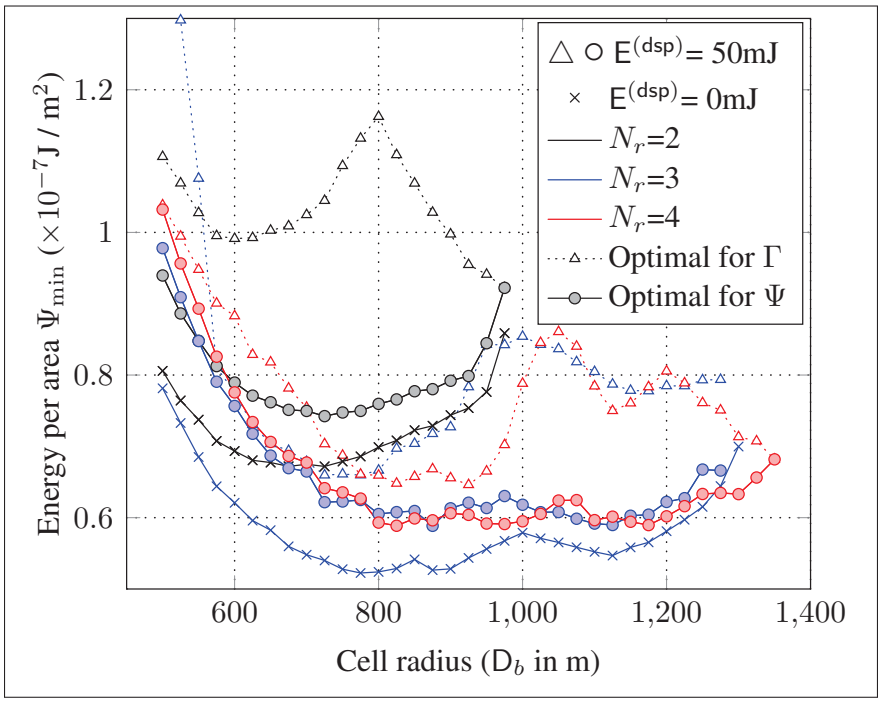


Figure 6.5 Minimal energy consumption per unit area $\Psi_{\min}$

As example, at D_b =700m, increasing $E^{(dsp)}$ from 0 to 50mW leads to a degradation of $\Psi_{\min}$ of 11% with N_r=2 and of 22% with N_r=3.

Result 6.2. Up to N_r=3, increasing the number of relays per cell allows significant energy reduction. Passed this limit, the gain provided by additional relays is not sufficient to compensate for the idle energy $N_r E_R^{(idle)}$, dissipated to maintain the network operational.

With N_r=2, the minimal value for $\Psi_{\min}$ is equal to 0.74e-7J/m^2, while adding one more RS allows a gain of 28%, $\Psi_{\min}$ then reaching 0.58e-7J/m^2. However, there is no significant performance gain between the cases N_r=3 and N_r=4.

We recall that we have considered as performance metric the maximal energy necessary to transmit data at a given rate to any user of the cell sector, as given by the EEA, rather than simply the average energy consumption of the cell. Simulations show that the relay configurations optimal for Ψ and for the average do not match. When optimized for the average consumption, the relay configuration results in a severe energy increase at cell edge (from 10% to 25%), meaning that there exists a major performance gap between users of the cell center, with strong

channel, and cell-edge users, with weak channel. Thus, the average optimization criteria does not provide fairness as does our proposed Energy Efficiency Area and related metric Ψ.

6.7.3 A new energy-interference trade-off on relay deployment

We now investigate the network performance in terms of both energy and interference, by using the Γ-metric. In Figure 6.6, we plot the maximal feasible $\Gamma_{\max}$ as a function of the cell radius D_b, for different values of N_r and $\mathsf{E}^{(\mathsf{dsp})}$. To do so, for each considered considered set $\left(\mathsf{D}_b, N_r, \mathsf{E}^{(\mathsf{dsp})}\right)$, we find the location for the N_r relays which maximizes Γ.

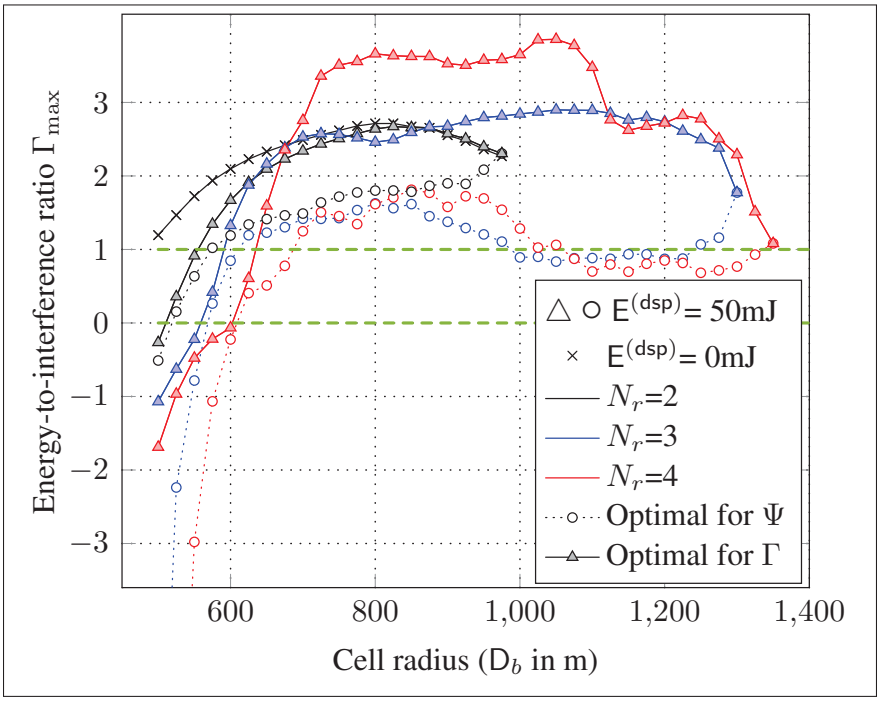


Figure 6.6 Maximal energy-to-interference ratio $\Gamma_{\max}$

First, for a given number of relays, we observe that $\Gamma_{\max}$ is poor for small-size cells ($\mathsf{D}_b \leq$600m) and is significantly affected by $\mathsf{E}^{(\mathsf{dsp})}$. For example, at $\mathsf{D}_b \leq$550m, $\Gamma_{\max}$ is divided by 2 when $\mathsf{E}^{(\mathsf{dsp})}$ is increased from 0 to 50mW. On the contrary, for wider cells, the overall energy consumption is dominated by the RF transmit energy and the interference issue is relaxed due to distance. The impact of $\mathsf{E}^{(\mathsf{dsp})}$ is minor and $\Gamma_{\max}$ increases. For the sake of clarity, the case

$\mathsf{E}^{(\mathsf{dsp})}$=0mJ has been plotted for N_r=2 only, but results do not change for larger number of relays.

Result 6.3. Except for very short cell radius, the energy offset $\mathsf{E}^{(\mathsf{dsp})}$, consumed for decoding and re-encoding at the relay station, has little impact on the energy-to-interference ratio Γ.

Second, as depicted in Figure 6.6, increasing the number of relays generally improves $\Gamma_{\max}$. We have shown in the previous subsection that relay configurations with N_r=3 and N_r=4 provide around the same minimal energy consumption per unit area $\Psi_{\min}$. On the contrary, when accounting for the interference generated by relays, the case N_r=4 largely outperforms N_r=3 and, for example, $\Gamma_{\max}$ is increased from 2.46 to 3.66 at D_b=800m. While the transmit energy gain achieved by increasing the number of relays from 3 to 4 is just enough to compensate for the additional idle energy $\mathsf{E}_{\mathsf{R}}^{(\mathsf{idle})}$, dissipated whether or not data is transmitted (for cooling and network maintenance) and affecting both υ_{Gain} and $\Psi_{\min}$, it is largely beneficial for the neighboring cells (reduced υ_{Loss}).

Result 6.4. Accounting for the interference generated by relays, deploying many relay stations potentially closer to cell edge but transmitting at lower power is more efficient than deploying few relay stations far from cell edge but serving a large part of the cell.

We now compare the performance achieved when the relay configuration is optimized either for Ψ or Γ and propose guidelines for efficient relay deployment. To do so, Figure 6.5 also depicts the value for Ψ obtained when the relay location is optimized for Γ (namely, "Optimal for Γ" in the figure) and reversely, Figure 6.6 depicts the value for Γ achieved by a location optimized for Ψ ("Optimal for Ψ").

A first important remark is that both relay configurations are essentially distinct and provide notably different performance. To illustrate the gap between such deployment options, we plot in Figure 6.7 the relay configurations optimal for Γ and for Ψ with $N_r = 2$ and various D_b. Moreover, we observe from Figure 6.6 that the value for Γ achieved with an energy-efficient relay deployment (optimal for Ψ) is below 1 for 1000m $\leq \mathsf{D}_b$ and $N_r \geq 3$, meaning that the network performance is actually degraded.

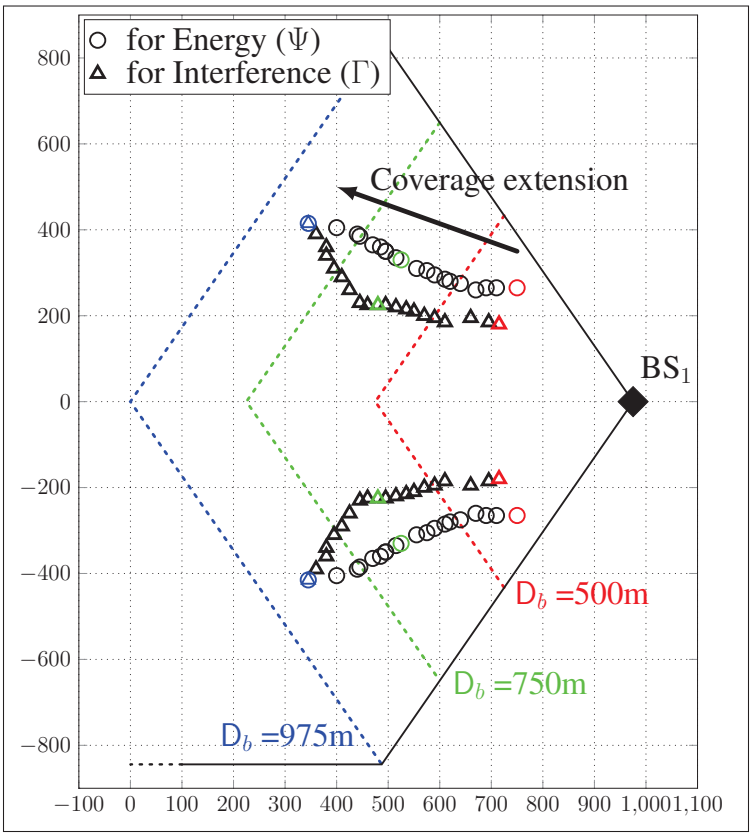


Figure 6.7 Optimal relay positions
(N_r = 2, $\mathsf{E}_{\mathsf{2Hop}}^{(\mathsf{dsp})}$ = 50mJ)

Result 6.5. Energy-efficient relay deployment does not necessarily lead to interference reduction and reversely, an interference-aware location is suboptimal for the cell energy consumption.

Based on the above results, we propose a guideline for efficient relay deployment regarding both Ψ and Γ. First, for short cell radius ($550 \leq \mathsf{D}_b \leq 700$m), deploying two relay stations per sector located to minimize the energy consumption per unit area Ψ can be considered as the best option, Γ remains above 1, meaning that the overall network performance is not degraded.

Second, for wider cell size ($700\text{m} \leq \mathsf{D}_b$), deploying four relay stations per sector provides the optimal results for both Γ and Ψ. However, current cellular networks are already reaching saturation and negotiating new site agreement for antenna deployment is getting ever harder for cellular operators. Thus, we argue that considering $N_r = 3$ may actually be the best practical

choice. When 700m $\leq \mathsf{D}_b \leq$ 1000m, optimizing the relay deployment for Γ (resp. Ψ) does not degrade too much the performance in Ψ (resp. Γ), such that both deployment options can be considered. However, for 700m $\leq \mathsf{D}_b$, a deployment optimized for Ψ should not be considered since the overall network performance is degraded ($\Gamma < 1$).

6.7.4 Impact of the relay coding scheme

Up to now, we have shown that a relay deployment can be energy-efficient at the scale of a single cell (measured by Ψ) but without necessarily being efficient at a larger scale (measured by Γ). We now investigate the performance achieved by the energy-optimized relaying schemes described in Section 6.3.2. Regarding the Γ-metric, EO-PDF maximizes the energy gain υ_{Gain}, while decreasing at the same time the energy loss υ_{Loss} experienced by neighboring cells. As detailed in [28], the energy transmitted by EO-PDF is more uniformly spread over the two transmission phases and over both the direct and relaying links, reducing at same time the power peaks causing high interference. The IR-PDF scheme minimized the use of the relay station by transmitting the most data possible via the direct link. The energy loss υ_{Loss} is minimized, but in return, the energy gain υ_{Gain} is reduced. We recall that the circuitry consumption of such partial DF schemes is expressed as $\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp})} = \mathsf{E}_{\mathsf{2Hop}}^{(\mathsf{dsp})} + \mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)}$, where $\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)}$ is an additional offset accounting for their increased complexity.

6.7.4.1 Objective and simulation settings

For analysis, we consider a cell sector aided by two RS only ($N_r = 2$), with low, medium and maximal cell radius ($\mathsf{D}_b \in \{600, 800, 975\text{m}\}$). Such configuration provides suboptimal performance in both Ψ and Γ, compared to a configuration with more RS, but it offers valuable infrastructure cost reduction and deployment simplicity for a cellular operator. We consider as performance basis an energy-efficient relay deployment where both RS are located to minimize the energy per unit area Ψ consumed by two-hop relaying.

To investigate how optimized relaying schemes can alleviate the interference issue, we derive the optimal utilization of coding schemes within the cell sector. To do so, we compare for each user location the performance achieved by two-hop relaying, EO-PDF and IR-PDF and select

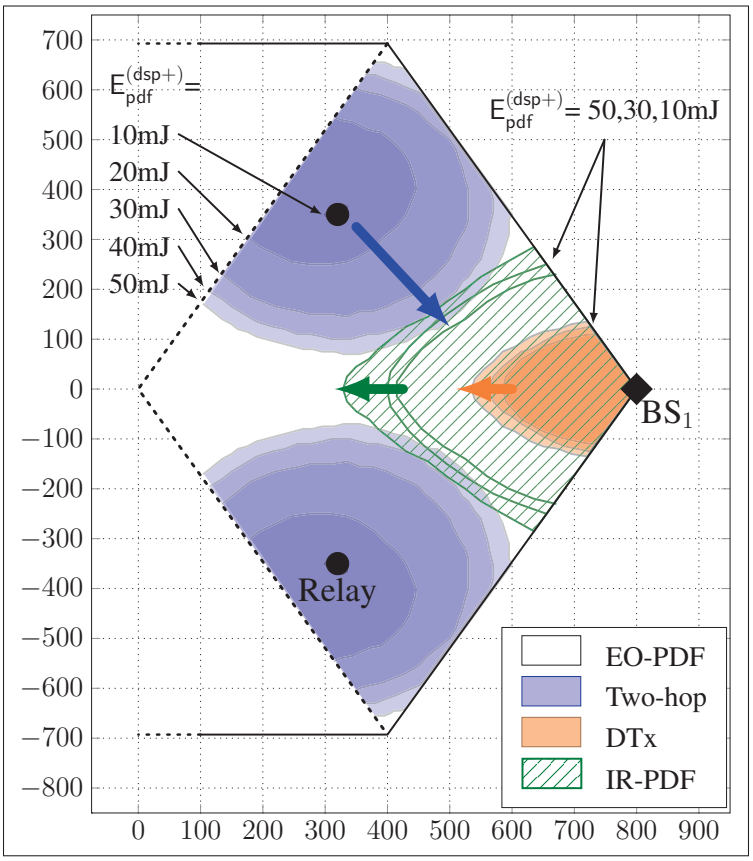


Figure 6.8 Map of the cell sector

for each the one that increases Γ. Proceeding in this way, we design a spatially-optimized utilization of coding schemes and draw a map showing the cell areas where to use each coding scheme to maximize Γ. Such map is illustrated in Figure 6.8 for D_b=800m and $E_{pdf}^{(dsp+)} \in [0, 50]$mJ.

6.7.4.2 Spatial analysis

With a medium cell radius (D_b=800m), when the additional circuitry consumption $E_{pdf}^{(dsp+)}$ is below 10mJ, EO-PDF outperforms two-hop relaying in almost the whole cell area, as illustrated in Figure 6.8. For higher values of $E_{pdf}^{(dsp+)}$, EO-PDF does not provide sufficient reduction in the transmit energy around the RS and cannot compensate for the dissipated energy $E_{pdf}^{(dsp+)}$. However, even with $E_{pdf}^{(dsp+)}$=50mJ (i.e. EO-PDF consumed twice as much energy as two-hop relaying to process data), EO-PDF still outperforms two-hop relaying when the user-to-relay

link is weaker. In larger cells (D_b=975m), EO-PDF outperforms two-hop relaying for any user location and any value of $E_{pdf}^{(dsp+)} \in [0, 50]$mJ. The additional circuitry consumption $E_{pdf}^{(dsp+)}$ has only marginal effect on Γ. In smaller cells (D_b=600m), the overall energy consumption is dominated by the circuitry consumption. Except from the case $E_{pdf}^{(dsp)} = E_{2Hop}^{(dsp)}$, the EO-PDF scheme improves the cell performance only if used very far from the relay stations.

Also note that the IR-PDF scheme can only reach the same performance as other coding schemes but without outperforming them. The corresponding cell areas are plotted in green.

6.7.4.3 Coding schemes and relay deployment

In the following, we denote "combination EO-PDF / 2Hop" as the spatially-optimized utilization of coding schemes, as previously described and illustrated in Figure 6.8 and plot in Figure 6.9, the maximal energy-to-interference ratio $\Gamma_{\max}$ achieved by this combination and by two-hop relaying only, with $N_r = 2$ and $N_r = 3$.

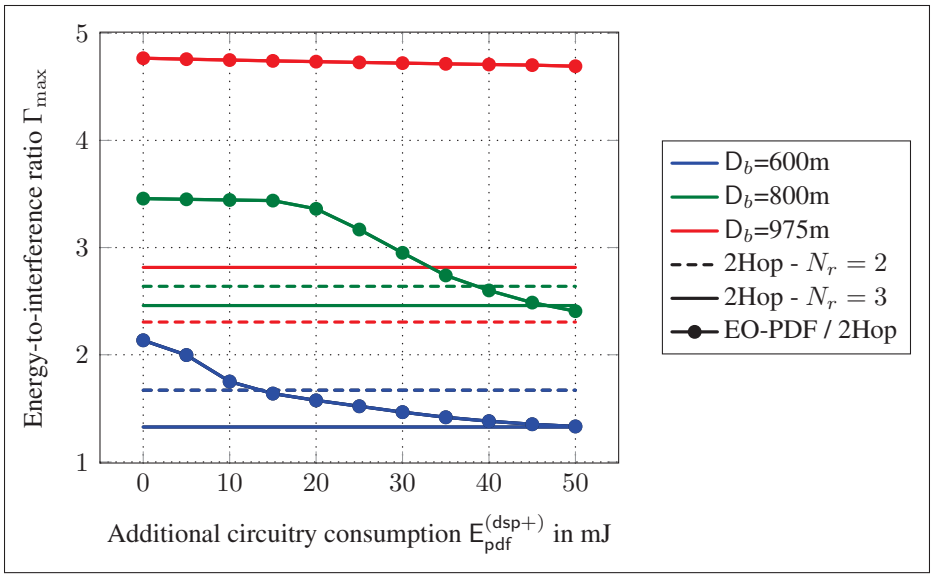


Figure 6.9 Maximal energy-to-interference ratio $\Gamma_{\max}$

Remark: For some data points, two-hop relaying only outperforms the combination EO-PDF / 2Hop. Indeed, such combination is used with an energy-efficient deployment (optimized for

Ψ), which is sub-optimal for Γ. On the contrary, the plotted performance of two-hop relaying is obtained with an interference-aware deployment (optimized for Γ) and can be understood as the maximum Γ reached by two-hop relaying in the best feasible relay configuration.

Result 6.6. In a small-size cell, the energy-optimized partial decode-forward scheme (EO-PDF) is severely affected by its increased circuitry consumption such that the optimal combination EO-PDF / 2Hop does not provide much performance enhancement when $\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)} \geq$30mJ.

For example, when D_b=600m and $\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)} =$50mJ, EO-PDF / 2Hop achieves a ratio Γ of 1.33, which outperforms two-hop relaying used with same relay location (Γ=1.19) but does not reach the performance of two-hop relaying when used with an interference-aware location (Γ=1.67).

Result 6.7. When the additional circuitry consumption $\mathsf{E}_{\mathsf{pdf}}^{(\mathsf{dsp}+)}$ is low, or when the cell radius is wide, the combination EO-PDF / 2Hop approaches and even outperforms the maximal $\Gamma_{\max}$ achieved by two-hop relaying, even with a relay location suboptimal for Γ. It also allows a reduction of the number of relays per sector for the same, or even better, ratio Γ.

As observed in Figure 6.9, the maximal $\Gamma_{\max}$ achieved by the combination EO-PDF / 2Hop is higher than two-hop relaying with $N_r = 3$, for almost any cell size and any additional circuitry consumption. It even outperforms the case $N_r = 4$ for wide coverage extension (D_b=975m).

6.8 Conclusion

We have highlighted a new trade-off on relay deployment for cellular networks that balances system energy efficiency and performance loss experienced by neighbouring users due to the additional interference generated by relays. To this end, we first formulated a spatial definition of the relay efficiency and proposed three tractable models allowing meaningful analysis without requiring time-consuming simulations. Next, we analyzed the joint impact of the circuitry consumption, the location and number of relays as well as the relaying coding scheme on the network performance. By significantly reducing the transmit power peaks, energy-optimized coding schemes alleviate the interference issue, and by performing well even with suboptimal relay location or reduced number of relays, they offer valuable deployment flexibility.

CHAPTER 7

DISCUSSION AND RECOMMENDATIONS

This research work has already been widely discussed within each journal paper. This last chapter reviews its main highlights and possible extensions, in particular, the implementation of the proposed coding schemes and the issue of the network saturation.

7.1 On the practical implementation of energy-optimal relaying scheme

In the first journal paper, lower-bounds on the energy consumption for half-duplex decode-forward relaying have been characterized. Optimal resource allocations have been proposed to maintain a desired source rate and minimize the energy consumption of: a) the considered three-node relay network (N-EE), b) the relay alone (R-EE) and c) the source alone (S-EE).

From a theoretical aspect, an important result is that minimizing the network energy consumption is not equivalent to maximizing the network capacity as it is often believed. The whole range of source rates achievable with decode-forward is not covered by any of the individual schemes (N-EE, R-EE and S-EE), but only by their combination. The proposed energy-efficiency approach has a certain benefit over the maximum-rate approach since it leads to the closed-form solution of the optimal power allocation and not solely to a solution on the form "$\underset{x}{\text{argmax}}\ f(x)$". Moreover, it allows a comprehensive description of the optimal coding (full or partial decode-forward, with or without beamforming etc.).

Yet, the major concern raised by the reviewers of this first journal paper regards the practical aspect of the proposed analysis, principally the performance of such scheme considering practical binary alphabets and modulations (e.g. M-QAM, QPSK), with common channel decoder (e.g. turbo codes), and the possibility to implement such complex coding scheme. We next discuss such issue, based on four key aspects: 1) the gap between the results obtained with Gaussian signaling and finite modulations, 2) the use of finite block length codes, 3) the requirements for CSIT acquisition and 4) the practical techniques for rate splitting and superposition coding.

On the use of Gaussian signaling: Considering a Gaussian distribution for the channel input, as done in this thesis, provides benchmark for implementation of any proposed cooperative strategy and gives insight on how to structure a practical scheme. As stated in the paper by Sendonaris, Erkip and Aazhang, which won the best paper award on IEEE Transactions on Communications ([47], p. 1929), "within a given transmission framework, the system that most closely emulated the signal structure of the information-theoretic capacity-maximizing system, also had the highest throughput." Furthermore, from the central limit theorem, Gaussian signaling accurately approximates OFDM-based communications, as used in practical systems such as LTE or WiMAX. This implies that proposed theoretical results can be fairly applied to practical networks.

On the finite block length: One drawback of classical information-theoretical tools is that the deduced bounds on maximum rate or minimum energy are achieved using codes with infinite, or at least very long, block length. The block length is critical when designing practical error correction codes. A long processing duration may be computationally infeasible and definitely causes excessive latency, which is a particularly severe issue for real-time applications. With too short processing, the decoding process may not be able to recover data and the rate bounds, as defined by information theory, cannot be guaranteed. The maximal rate achievable at a given block length and error probability has been investigated in [138] and references therein. Such topic constitutes an interesting theoretical extension to the work proposed in this thesis.

CSIT acquisition and feedback errors: Throughout this thesis, the different power allocations that have been considered all require channel station information (CSI) at the transmitter. For example, in OFDM systems, CSI is estimated via the detection of reference symbols spread over the time and frequency OFDM-structure. The main advantage of an approach based on coherent systems is the simplicity of implementation. Non-coherent strategies relax the constraint of CSI acquisition but in return, require complex algorithms for equivalent performance.

The CSI acquisition is mostly limited by how fast the channel condition is varying. While instantaneous CSI can be accurately estimated in slow fading environment, fast fading systems undergo imperfect channel estimation or outdated CSI feedback. This is particularly challeng-

ing in relaying networks. While the next-link quality can be reasonably tracked, the channel state of distant links may be hardly known in time. For example, it is harder for the source to track accurately the variation of the link from the relay to the destination. In our conference paper [32], we have explored such topic and have proposed distributed power allocations, assuming only local channel knowledge and long-term statistics for the distant link. Other line of research may also consider incremental redundancy, which does not require channel state information at transmitter.

Another challenge of CSI acquisition is related to the overhead induced by reference symbols. The required resource does not carry any data and consumes additional energy, resulting in performance loss. Whereas the trade-off between the CSI accuracy and the system spectral efficiency has been widely investigated [139, 140, 141, 142], its impact on energy efficiency remains an open research problem. Some preliminary work on this topic can be found in [81] for MIMO systems and [26] for CoMP. With regards to relay networks, most techniques for relay selection and cooperation require large amount of feedback, meaning that the energy loss due to CSI acquisition is potentially severe.

Techniques for rate splitting and superposition modulation: Superposition modulation (SM) belongs to the family of multi-level modulations, which has been proposed for single-source single-destination transmissions by Imai and Hirakawa in as soon as 1977 [143], even before trellis-coded modulations (1982) or bit-interleaved coded modulation (1998) [109]. Whereas information theory has largely motivated its use, it is only recently that SM is envisaged for implementation.

SM is non-bijective, contrary to other modulation schemes such as PSK or QAM. It is widely believed that it induces higher peak-to-average power ratio, that it is more sensitive to hardware non-idealities (e.g. carrier frequency offset or phase noise...) and above all, that its detection, demodulation and decoding are more complex than for bijective modulations. Yet, it is actually quite the opposite and main hindrances of SM turned out to be "myths", as expressed in [109]. Superposition coding outperforms QAM modulation at even lower receiver complexity and provides a natural framework for multi-user transmission, as well as MIMO and OFDM tech-

nologies [109]. Further analysis on the theory and practice of SM has been provided in [144], with a particular focus on error correction. Finally, while the high peak-to-average power remains a challenge, it must be noted that it is inherent to any Gaussian-like distribution, and not only to SM. Such issue has been investigated for example in [145].

Finally, some recent work has focused on the practical implementation of SM for multi-user networks. In [146], the performance of various superposition modulation schemes, in association with Reed-Solomon codes, is analyzed for relay networks and in [147], an experimental set up for SM in broadcast channels has been designed using off-the-shelf single-user coding and decoding techniques, and tested on the open-source GNU Radio platform, with the Universal Software Radio Peripheral (USRP) hardware board as RF front-end. All these recent contributions make us believe that superposition coding is no more out-of-reach regarding practical implementation and that it is a promising technique for future cellular networks, notably for cooperative scenarios.

In our opinions, the main hindrance for implementing the coding scheme proposed in our first journal with current modulation techniques paper lies in the rate splitting and mapping challenge. Optimizing the constellation size and picking the right constellation point, together with the right error correction rate, is a **discrete** problem. Even if based on typical QAM modulations, efficiently exploiting the available constellation space, i.e. identifying the different clusters of the constellation points which will be used for each layer of the superposition coding, remains a serious challenge. This identification must account for the Euclidean distances between the constellation points within each cluster but also from one cluster to another one. Then, mapping bits to each constellation point should maintain uniform and minimal bit error rate. Such process is illustrated in [146]. Designing a general method for finding the best constellation size, cluster sets and error correction rate (or equivalently, an adaptive modulation coding (AMC) scheme for superposition modulation) is definitely a pertinent extension to the research work of this thesis.

7.2 The new challenge of relay deployment in urban environments

The second and third journal papers have proposed a geometry-based performance analysis of relay-aided cellular networks. The concept of efficiency area has been defined as a new optimization framework and gives the geographical distribution of the probability for energy-efficient relaying and of the energy consumption, depending on the relays number and location, the propagation environment and the user rate. To avoid extensive time-consuming simulations, mathematical models of the defined efficiency areas have been proposed for fast and accurate performance evaluation.

More particularly, the second journal paper put emphasis on the radio propagation environment, notably the relay height and LOS conditions. The major concern raised by the Reviewers on this work has regarded the performance of relaying when accounting for shadowing and interference. Those two aspects have been fully addressed in the third journal paper. First, improved geometrical models have been proposed to include shadowing, but also the energy dissipated in circuitry for relaying. Second, pointing out that classical interference analysis (as in capacity-optimized systems) cannot be applied to energy-oriented networks, a new performance metric has been proposed to balance the energy gain provided by relay stations within a cell and the additional interference received in the neighbouring cells, based on the *actual* energy consumption and not on a given fixed power.

The proposed efficiency areas and performance framework have wide application and offer valuable support for network resource management, relay selection, load-balancing, scheduling or base station switch off. In this thesis, we have focused in particular on the relay deployment and have highlighted new trade-offs on energy consumption, coverage extension, deployment cost and interference reduction. Most importantly, we have shown that there exists a relay-to-BS distance for which the coverage can be increased or decreased without harming the energy-efficiency, and that relay deployment for energy reduction does not necessarily lead to interference reduction.

In industrial tracks of the last IEEE Vehicular Technology Conference (VTC-Fall 2014), two important issues have been raised in relation with the research work of this thesis. First, the statistical description of the wireless propagation channels, including the WINNER project as used in this thesis, may not be accurate enough for relay analysis and new campaigns of channel measurements have been launched to improve existing models. Second, it is becoming harder and harder to add new stations in networks already very densely equipped. We next discuss further both aspects.

Improved statistical description of the wireless channels: While the COST-231 [67] and the WINNER models [69] have been widely used so far for performance analysis, some operators and equipment manufacturers, including Orange Labs and Huawei Technologies, are questioning the accuracy of such channel descriptions and claim that the considered scenarios do not fit the actual network utilization. The majority of communications happen from indoors to outdoors and more and more transmissions come from very high buildings, especially professional calls. As highlighted in our second journal paper, the relay height significantly affects the system performance. Yet, channel models do not sufficiently account for the node height and none of them proposes statistical description for high building, from street level to rooftop. The notion of three-dimensional coverage has also to be defined. Such improved channel description will definitely offer new perspectives on the work presented in this thesis.

The challenge of cell densification: The cell densification remains quite a hot topic in the research community [148]. The literature generally proposes performance enhancement for future networks and considers to this end that macro base station can be supported by a dozen of micro and pico base stations, and by one or two hundreds home-deployed access points. However, in practice, the urban environment is already saturated and adding new stations causes severe performance degradation, due to interference and non-appropriate radio resource management. Moreover, getting a new site agreement for operator-managed stations is a serious issue, particularly in Europe. As a matter of comparison, approximately 13.000 wireless antenna sites are deployed in Canada, while the United Kingdom has around 52.000 sites in an area less than a fifth of Canada's surface area [149].

As shown in this thesis, the use of energy-optimal coding schemes, rather than the simple two-hop relaying scheme, can bring part of the answer to the network saturation. Not only it provides more flexibility in the relay deployment and allows to be less restricted in potential relay locations but also, it allows to reduce the number of relay stations necessary to satisfy the same performance requirement. A relevant line of research to extend the work presented in this thesis encompasses unconventional relaying strategies, notably cognitive relays, multiple-relay multiple-hop systems and two-way relaying.

Finally, the results obtained in this thesis are associated with network planning. Another opportune extension to address such challenge is to explore energy-efficient adaptive relaying. This includes mobile relaying and self-organizing relay networks [150], where the relay selection and path, the antenna tilts, the cell radius, etc. are dynamically adjusted based on the real-time user traffic, requested QoS and propagation conditions.

GENERAL CONCLUSION

This thesis has contributed to a new understanding of energy efficiency in relay-assisted cellular networks and, as far of our knowledge, no other research work has bridged the relay geographical deployment to the considered coding scheme.

From a theoretical perspective, we have characterized the minimal energy consumption achieved by decode-forward relaying and developed the concept of Efficiency Area as a novel spatial approach to analyze the relaying efficiency in various scenarios, facilitate relay deployment and support a wide range of other energy-efficient techniques. To address the key issue of time-consuming simulations in log-normal shadowing environments, we have proposed easily-computable and tractable models for the probability of energy-efficient relaying, the energy consumption and the interference generated by relay stations. Finally, we have defined a new performance metric which jointly captures both aspects of energy and interference in energy-optimized relay-assisted cellular networks.

From a practical perspective, we have designed a novel half-duplex relaying coding scheme, together with energy-optimized power allocations, which respectively minimize the network consumption, the consumption of the relay only and the consumption of the source only. Such coding scheme and power allocations have then been used as reference for upper-bound performance throughout the rest of this thesis. Furthermore, we have developed a system-level model for the energy consumption, accounting for the RF transmit energy, the additional circuitry consumption for encoding and decoding at the relay station and finally, the idle energy consumed at each deployed station, even if not transmitting.

We have fulfilled our objective to investigate energy-efficient relay deployment and analyze the gap between the simple but practical two-hop relaying scheme versus the energy-optimized partial decode-forward. Using an approach complementary to designing algorithms for optimal placement, we have characterized the impact of the radio propagation environment and the interference generated by relays on the network efficiency. Through our performance analysis, we have refined trade-offs on energy-efficient relaying.

Most importantly, this thesis spotlights that the use of energy-optimized partial decode-forward schemes should be definitely considered as a key to improve energy efficiency in relay-aided cellular networks and reduce constraints on relay deployment at the same time. They provide robustness to channel impairment and harsh radio environments, achieve satisfying performance even with suboptimal relay location, alleviate the interference issue for neighbourhood cells and allow a reduction in the number of relays necessary to reach same performance.

APPENDIX I

APPENDIX OF FIRST JOURNAL PAPER

1 Proof of Algorithm 4.1 : Power allocation for the network energy efficiency

1.1 General problem setting

Denote $X = (\eta_1, \rho_1, \eta_2, \rho_2, \rho_r)^T$. Consider optimization problem (4.6). We look for the X that minimizes the objective function f, such that

$$f(X) = \omega_s \left[\theta \left(\eta_1 + \rho_1\right) P_s + \bar{\theta} \left(\eta_2 + \rho_2\right) P_s\right] + \omega_r \bar{\theta} \rho_r P_r$$

given the two rate constraints $c_1(X) = I_1 - R$ and $c_2(X) = I_2 - R$, as defined in (4.4). Here, we first ignore the power constraints. The Lagrangian can then be formed as

$$\mathcal{L} = f(X) - \lambda_1 c_1(X) - \lambda_2 c_2(X)$$

The KKT conditions are expressed by equations (A I-1) - (A I-5) and by (A I-6).

$$\frac{\partial \mathcal{L}}{\partial \eta_1} = \omega_s \theta P_s - \frac{\lambda_1 \theta \frac{P_s |h_d|^2}{N}}{\left(1 + \frac{(\eta_1+\rho_1) P_s |h_d|^2}{N}\right)} - \frac{\lambda_2 \theta P_s}{N} \left(\frac{|h_s|^2}{\left(1 + \frac{(\eta_1+\rho_1) P_s |h_s|^2}{N}\right)} - \frac{|h_s|^2}{\left(1 + \frac{\eta_1 P_s |h_s|^2}{N}\right)} + \frac{|h_d|^2}{\left(1 + \frac{\eta_1 P_s |h_d|^2}{N}\right)} \right) = 0 \quad \text{(A I-1)}$$

$$\frac{\partial \mathcal{L}}{\partial \rho_1} = \omega_s \theta P_s - \lambda_1 \frac{\theta \frac{P_s |h_d|^2}{N}}{\left(1 + \frac{(\eta_1+\rho_1) P_s |h_d|^2}{N}\right)} - \lambda_2 \frac{\theta P_s}{N} \frac{|h_s|^2}{\left(1 + \frac{(\eta_1+\rho_1) P_s |h_s|^2}{N}\right)} = 0 \quad \text{(A I-2)}$$

$$\frac{\partial \mathcal{L}}{\partial \eta_2} = \omega_s \bar{\theta} P_s - \lambda_1 \bar{\theta} \frac{\frac{P_s|h_d|^2}{N}}{\left(1 + \frac{(\eta_2+\rho_2)P_s|h_d|^2+|h_r|^2\rho_r P_r+2\sqrt{P_s|h_d|^2 P_r|h_r|^2\rho_2\rho_r}}{N}\right)} - \lambda_2 \bar{\theta} \frac{\frac{P_s|h_d|^2}{N}}{\left(1 + \frac{\eta_2 P_s|h_d|^2}{N}\right)} = 0 \quad \text{(A I-3)}$$

$$\frac{\partial \mathcal{L}}{\partial \rho_2} = \omega_s \bar{\theta} P_s - \lambda_1 \bar{\theta} \frac{\frac{P_s|h_d|^2}{N}\left(1 + \sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}\right)}{\left(1 + \frac{(\eta_2+\rho_2)P_s|h_d|^2+|h_r|^2\rho_r P_r+2\sqrt{P_s|h_d|^2 P_r|h_r|^2\rho_2\rho_r}}{N}\right)} = 0 \quad \text{(A I-4)}$$

$$\frac{\partial \mathcal{L}}{\partial \rho_r} = \omega_r \bar{\theta} P_r - \lambda_1 \bar{\theta} \frac{\frac{P_r|h_r|^2}{N}\left(1 + \sqrt{\frac{P_s|h_d|^2\rho_2}{P_r|h_r|^2\rho_r}}\right)}{\left(1 + \frac{(\eta_2+\rho_2)P_s|h_d|^2+|h_r|^2\rho_r P_r+2\sqrt{P_s|h_d|^2 P_r|h_r|^2\rho_2\rho_r}}{N}\right)} = 0 \quad \text{(A I-5)}$$

$$\lambda_i c_i(X) = 0 \; ; \quad c_i(X) \geq 0 \; ; \quad \lambda_i \geq 0 \; ; \quad i = 1, 2 \quad \text{(A I-6)}$$

1.2 Transmission during the first phase

Note that $\frac{\partial \mathcal{L}}{\partial \eta_1}$ and $\frac{\partial \mathcal{L}}{\partial \rho_1}$ cannot be equal to zero simultaneously, unless $|h_d|^2 = |h_s|^2$. Therefore, we have to relax either ρ_1 or η_1. Considering $|h_d|^2 < |h_s|^2$ for which decode-forward is useful, we get $\frac{\partial \mathcal{L}}{\partial \eta_1} > \frac{\partial \mathcal{L}}{\partial \rho_1}$. Relaxing η_1 gives $\frac{\partial \mathcal{L}}{\partial \eta_1} > 0$ when $\frac{\partial \mathcal{L}}{\partial \rho_1} = 0$, and η_1 should be minimized. Since the objective function and the first constraint $c_1(X)$ only depend on the sum $\eta_1 + \rho_1$, setting $\eta_1 = 0$ weakens the second constraint $c_2(X)$ and minimizes f. Thus, the source only sends m_r during the first phase.

The following notations will be used for the rest of this appendix and next ones:

$$\Gamma_i = 1 + \frac{\rho_1 P_s|h_i|^2}{N} \quad ; \quad G_i = \frac{2^{R/\bar{\theta}}}{\Gamma_i^{\theta/\bar{\theta}}} \quad \text{where } i \in \{s, d\}$$

$$\Gamma_r = 1 + \frac{(\eta_2 + \rho_2) P_s|h_d|^2 + |h_r|^2\rho_r P_r + 2\sqrt{P_s|h_d|^2 P_r|h_r|^2\rho_2\rho_r}}{N}$$

1.3 Optimal allocation set for network energy efficiency

In this case, both ω_s and ω_r are equal to 1. First, from (A I-4) and (A I-5), we get:

$$\lambda_1 = \frac{\Gamma_r N}{|h_d|^2}\left(1+\sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}\right)^{-1} = \frac{\Gamma_r N}{|h_r|^2}\left(1+\sqrt{\frac{P_s|h_d|^2\rho_2}{P_r|h_r|^2\rho_r}}\right)^{-1}$$

$$\text{and} \quad \frac{|h_d|^2}{|h_r|^2} = \sqrt{\frac{P_s|h_d|^2\rho_2}{P_r|h_r|^2\rho_r}} \;\Leftrightarrow\; \rho_2 = \frac{P_r|h_d|^2}{P_s|h_r|^2}\rho_r. \tag{A I-7}$$

As $\lambda_1 > 0$, the first constraint is active and $G_d = \Gamma_r$ gives

$$\frac{2^{R/\bar{\theta}}}{\left(1+\frac{\rho_1 P_s|h_d|^2}{N}\right)^{\theta/\bar{\theta}}} = 1 + \frac{\eta_2 P_s|h_d|^2}{N} + \frac{\left(\sqrt{P_s|h_d|^2\rho_2}+\sqrt{P_r|h_r|^2\rho_r}\right)^2}{N} \tag{A I-8}$$

With (A I-7), we get

$$G_d = 1 + \frac{\eta_2 P_s|h_d|^2}{N} + P_r\rho_r|h_r|^2\left(\frac{|h_d|^2}{|h_r|^2}+1\right)^2 \tag{A I-9}$$

From (A I-3), $\lambda_2 = \left(1-\frac{1}{1+\sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}}\right)\left(\frac{N}{|h_d|^2}+\eta_2 P_s\right) > 0$. Thus, the second constraint is also active and we get

$$\frac{\eta_2 P_s|h_d|^2}{N} = G_s - 1 \tag{A I-10}$$

From (A I-7), (A I-9) and (A I-10), we write ρ_2, ρ_r and η_2 as functions of ρ_1. Replacing the Lagrangian multipliers λ_1 and λ_2 in (A I-2), we prove that ρ_1 solves $g_1(\rho_1) = 0$, where g_1 is defined by (4.8) of Algorithm 4.1. Note that g_1 is decreasing in ρ_1, increasing in R and that $g_1(0) = \frac{|h_s|^2|h_r|^2+|h_d|^4}{|h_d|^2|h_r|^2+|h_d|^4}2^{R/\bar{\theta}} - 1 > 0$ for all R. If, for all $\rho_1 \in [0, 1/\theta]$, $g_1(\rho_1) > 0$, then the desired source rate is infeasible with the allocation as defined for $A^{(n)}$ and sub-scheme $B^{(n)}$ should be considered instead. We can also check that $\rho_r > 0$ and $\rho_2 > 0$ since $|h_s|^2 > |h_d|^2$ by assumption, and $\eta_2 > 0$ as long as $R \geq \theta\log_2\left(1+\frac{\rho_1 P_s|h_s|^2}{N}\right) = R^{(n)}$ On the contrary, when $R \leq R^{(n)}$, η_2 is negative and there is no solution for (A I-10). Thus, we relax η_2 and rewrite the optimization. Following the same steps as above, we can show that constraints are still active, and that (A I-2), (A I-4) and (A I-5) still hold. This leads to Proposition 4.2. If the power constraints are not satisfied, then the desired source rate is infeasible.

2 Proof of Algorithm 4.2 : Power allocation for relay energy efficiency

The proof of Algorithm 4.2 and Propositions 4.3, 4.5 and 4.4 follows the same main steps as in Appendix 1. Here, $\omega_s = 0$ and $\omega_r = 1$. Analysis in Appendices 1.1 and 1.2 still holds. To minimize the energy consumption of the relay, direct transmissions should be performed as long as the direct link is not in outage, or equivalently as long as $R \leq \log_2\left(1+\frac{P_s|h_d|^2}{N}\right)$ (Proposition 4.3).

The relay is needed to perform the transmission as soon as the source meets its power constraint and cannot send more data without going in outage. Then,

$$\rho_2 = \frac{1-\rho_1\theta}{\bar{\theta}} - \eta_2. \tag{A I-11}$$

First, note that Eq. $\frac{\partial \mathcal{L}}{\partial \rho_r} = 0$ (A I-5) is still valid. Thus $\lambda_1 = \Gamma_r \left[\frac{|h_r|^2}{N\ln 2}\left(1+\sqrt{\frac{P_s|h_d|^2\rho_2}{P_r|h_r|^2\rho_r}}\right)\right]^{-1}$. Second, with (A I-11), the Lagrangian equations now satisfy

$$\frac{\partial \mathcal{L}}{\partial \eta_2} = \lambda_1 \frac{\bar{\theta}P_s|h_d|^2}{N\Gamma_d \ln 2} - \lambda_1\bar{\theta}\frac{P_s|h_d|^2}{N\ln 2\Gamma_r} + \lambda_2\frac{\bar{\theta}P_s|h_s|^2}{N\Gamma_s\ln 2} - \lambda_2\bar{\theta}\frac{P_s|h_d|^2}{N\left(1+\frac{\eta_2 P_s|h_d|^2}{N}\right)\ln 2} = 0 \tag{A I-12}$$

$$\frac{\partial \mathcal{L}}{\partial \rho_2} = \lambda_1 \frac{\bar{\theta}P_s|h_d|^2}{N\Gamma_d \ln 2} - \lambda_1\bar{\theta}\frac{P_s|h_d|^2\left(1+\sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}\right)}{N\ln 2\Gamma_r} + \lambda_2\frac{\bar{\theta}P_s|h_s|^2}{N\Gamma_s\ln 2} = 0 \tag{A I-13}$$

Subtracting (A I-12) from (A I-13), we get

$$\lambda_2 = \lambda_1\sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}\left(1+\frac{\eta_2 P_s|h_d|^2}{N}\right)\Gamma_r^{-1} = \frac{N}{|h_r|^2}\sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}\frac{\left(1+\frac{\eta_2 P_s|h_d|^2}{N}\right)}{\left(1+\sqrt{\frac{P_s|h_d|^2\rho_2}{P_r|h_r|^2\rho_r}}\right)} > 0$$

Both constraints are active: Eq. (A I-8) $G_d = \Gamma_r$ and Eq. (A I-10) $\frac{\eta_2 P_s|h_d|^2}{N} = G_s - 1$ still hold. From (A I-10), η_2 is expressed as a function of ρ_1. From (A I-8), (A I-11) and the expression of η_2, ρ_r can be also written as a function of ρ_1.

Finally, substituting the expressions of λ_1 and λ_2 in (A I-8), (A I-10) into (A I-13), we prove that ρ_1 should satisfy $g_2(\rho_1) = 0$, as defined by (4.10) of Algorithm 4.2, with

$$g_3(\rho_1) = \sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}} = \frac{2^{R/\bar{\theta}}}{\left(1+\frac{\rho_1 P_s|h_d|^2}{N}\right)^{\theta/\bar{\theta}}} - \frac{2^{R/\bar{\theta}}}{\left(1+\frac{\rho_1 P_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}}$$

Now, we focus on the existence of this power allocation set. First, to ensure $\eta_2 > 0$, R must be greater than $\theta \log_2\left(1+\frac{\rho_1 Ps|h_s|^2}{N}\right) = R_2^{(r)}$. Moreover, ρ_r is positive as long as $R > \theta \log_2\left(1+\frac{\rho_1 P_s|h_d|^2}{N}\right) + \bar{\theta}\log_2\left(1+\frac{(\eta_2+\rho_2)P_s|h_d|^2}{N}\right) = \log_2\left(1+\frac{P_s|h_d|^2}{N}\right) = R_1^{(r)}$. Finally, the existence of ρ_1, and consequently of ρ_2, follows the same steps as in Appendix 1. This concludes the proof of Proposition 4.5.

When $R < \theta \log_2\left(1+\frac{\rho_1 Ps|h_s|^2}{N}\right)$, η_2 is negative and there is no solution for (A I-10). In this case, we relax η_2 and rewrite the optimization problem. As above, we can show that rate constraints are still active, and that (A I-5) and (A I-13) still hold. This concludes the proof of Proposition 4.4 and Algorithm 4.2.

3 Proof of Algorithm 4.3 : Power allocation for source energy efficiency

The proof of Algorithm 4.3 and Propositions 4.7, 4.8 and 4.9 follows the same steps as in Appendix 1. Here, $\omega_s = 1$ and $\omega_r = 0$. Analysis of Lagrangian in 1.1 and 1.2 still holds. However, the constraint on the relay consumption is relaxed here, such that it can consume up to its maximal power.

First, let's consider that the source rate is high and that all the available relay power is required ($\rho_r = \frac{1}{\bar{\theta}}$). This means that the first rate constraint, which depends on ρ_r will necessarily be tight. From Eq. $\frac{\partial \mathcal{L}}{\partial \rho_1} = 0$ (A I-2) and Eq. $\frac{\partial \mathcal{L}}{\partial \rho_2} = 0$ (A I-4), we can prove that both Lagrangian

multipliers are positive, meaning that both constraints are active, such that

$$\lambda_1 = \frac{\Gamma_r}{\frac{|h_d|^2}{N}\left(1+\sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}\right)} = G_d\frac{N}{|h_d|^2\left(1+\sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}\right)}$$

$$\lambda_2 = \left(1-\frac{\lambda_1|h_d|^2}{N\ln 2\Gamma_d}\right)\frac{N\ln 2\Gamma_s}{|h_s|^2} = \left(1-\frac{G_d}{\left(1+\sqrt{\frac{P_r|h_r|^2\rho_r}{P_s|h_d|^2\rho_2}}\right)}\right)\frac{N\ln 2\Gamma_s}{|h_s|^2}$$

Plugging in these expressions of λ_1 and λ_2, as well as Eq. $G_d = \Gamma_r$ (A I-8) and Eq. $\frac{\eta_2 P_s|h_d|^2}{N} = G_s - 1$ (A I-10) into (A I-4), we prove that ρ_1 should satisfy $g_2(\rho_1) = 0$, as defined by (4.13) of Algorithm 4.3. The existence of this power allocation set follows the same steps as for Appendices 1 and 2. This leads to sub-scheme $A^{(s)}$ and Proposition 4.9. Note that the power allocation is similar to sub-scheme $A^{(r)}$. We now have to insure the non-negativity of this solution, in particular $\eta_2 \geq 0$ and $\rho_2 \geq 0$.

On the one hand, we consider the non-negativity constraint on η_2. η_2 is positive following sub-scheme $A^{(s)}$ as long as $R \geq R_{2,B}^{(s)}$ with $R_{2,B}^{(s)} = \theta\log_2\left(1+\frac{\rho_1^\star P_s|h_s|^2}{N}\right)$. If $R \leq R_{2,B}^{(s)}$, η_2 is relaxed and set to 0. Next, the optimization follows the same steps as for $B^{(n)}$ or $B^{(r)}$, but with $\rho_r = \frac{1}{\theta}$. This leads to sub-scheme $B^{(s)}$ and Proposition 4.8. Note that, following $B^{(s)}$, ρ_2 is positive as long as $R \geq R_{1,B}^{(s)}$, with $R_{1,B}^{(s)} = R_B$ as defined in Algorithm 4.3. This means that the relay requires all its power to forward the source data such that the first rate constraint c_1 is active. However, if the rate decreases under the threshold $R_{1,B}^{(s)}$, the relay may not need all its available power and the first rate constraint becomes inactive. ρ_2 is relaxed and set to 0. Finally, even if the relay consumption is not part of the optimization problem here, we can nevertheless reduce it to the minimum required, such that $R = R_1^{(s)}$ (however, c_1 is still inactive in the sense of Lagrangian optimization). Thus, we deduce the expression of $\rho_r < 1/\bar{\theta}$. The optimization now leads to sub-scheme $D^{(s)}$ and Proposition 4.6.

On the other hand and following sub-scheme $A^{(s)}$ again, we consider the non-negativity constraint on ρ_2 (with $\eta_2 > 0$). ρ_2 is positive following sub-scheme $A^{(s)}$ as long as $R \geq R_{2,C}^{(s)}$

$$\text{with} \quad R_{2,C}^{(s)} = \bar{\theta}\log_2\left(\frac{P_r|h_r|^2}{\bar{\theta}N}\right) - \bar{\theta}\log_2\left(\frac{1}{\left(1+\frac{\rho_1^\star P_s|h_d|^2}{N}\right)^{\theta/\bar{\theta}}} - \frac{1}{\left(1+\frac{\rho_1^\star P_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}}\right).$$

If $R \leq R_{2,C}^{(s)}$, ρ_2 is relaxed and set to 0. This also means that the relay is able to forward the source data without requiring the beamforming gain between the source and the relay. Therefore, the first rate constraint c_1 becomes inactive and $\lambda_1 = 0$. From Eq. $\frac{\partial \mathcal{L}}{\partial \rho_1} = 0$ (A I-2) and Eq. $\frac{\partial \mathcal{L}}{\partial \eta_2} = 0$ (A I-3), we get $\frac{|h_s|^2}{1+\frac{\rho_1 P_s|h_s|^2}{N}} = \frac{|h_d|^2}{1+\frac{\eta_2 P_s|h_d|^2}{N}}$, such that $\eta_2 = \rho_1 + \frac{N}{P_s}\left(\frac{1}{|h_s|^2} - \frac{1}{|h_d|^2}\right)$. As the second constraint c_2 is active, Eq. $\frac{\eta_2 P_s|h_d|^2}{N} = G_s - 1$ (A I-10) is still valid and we get $\eta_2 = \frac{N}{P_s|h_d|^2}\left(\frac{2^{R/\bar{\theta}}}{\left(1+\frac{\rho_1 P_s|h_s|^2}{N}\right)^{\theta/\bar{\theta}}} - 1\right)$. By equalizing both expressions of η_2, we deduce ρ_1. Next, even if the relay consumption is not considered here, we reduce it to its minimum required value. The optimization leads to sub-scheme $C^{(s)}$ and Proposition 4.7. Now, following $C^{(s)}$, η_2 is positive as long as $R \geq R_{1,C}^{(s)}$, with $R_{1,C}^{(s)} = R_C$ as defined in Algorithm 4.3. If the rate decreases under the threshold $R_{1,C}^{(s)}$, η_2 is relaxed and set to 0, which leads to $D^{(s)}$ as previously defined.

So far, depending on non-negativity constraints, we defined two possible sequences of sub-schemes: $(A^{(s)}, B^{(s)}, D^{(s)})$ and $(A^{(s)}, C^{(s)}, D^{(s)})$. We now define how each of them should be used. First, $A^{(s)}$ is applied when $R \geq \max\left\{R_{2,B}^{(s)}, R_{2,C}^{(s)}\right\} = R_2^{(s)}$. If $R_{2,C}^{(s)} \leq R_{2,B}^{(s)}$, the non-negativity constraint on η_2 is stronger than the constraint on ρ_2 and the sequence $(A^{(s)}, B^{(s)})$ is applied. Otherwise, $(A^{(s)}, C^{(s)})$ is applied. Second, $D^{(s)}$ is such that $\eta_2 = 0$ and $\rho_2 = 0$. This sub-scheme is applied as long as $R \leq \min\left\{R_{1,B}^{(s)}, R_{1,C}^{(s)}\right\} = R_1^{(s)}$. Thus, if $R_{1,B}^{(s)} \leq R_{1,C}^{(s)}$, the sequence $(B^{(s)}, D^{(s)})$ is applied, otherwise $(C^{(s)}, D^{(s)})$ is applied.

Finally, we show that upper and lower rate bounds for applying either $B^{(s)}$ or $C^{(s)}$ are consistent. Considering the source consumption as a function of the source rate, sub-schemes are continuous. Moreover, η_2 and ρ_2 are increasing functions of the source rate. Thus, if $R_{1,B}^{(s)} \leq R_{1,C}^{(s)}$ and sequence $(D^{(s)}, B^{(s)})$ is applied for low source rates, we get $\rho_2 > 0$ for any

$R > R_{1,B}^{(s)}$. In this case, we necessarily have $R > R_{2,C}^{(s)}$ which implies $R_{2,C}^{(s)} \leq R_{2,B}^{(s)}$. Therefore, only sequence $(A^{(s)}, B^{(s)}, D^{(s)})$ can occur. Similarly, we can show that only $(A^{(s)}, C^{(s)}, D^{(s)})$ can occur if $R_{1,B}^{(s)} \geq R_{1,C}^{(s)}$. This completes the proof of Algorithm 4.3.

4 Maximal achievable source rate

4.1 R-EE and S-EE achieve higher rates than N-EE

A source rate R is achievable for N-EE (resp. R-EE) as long as there exists a $\rho_1^\star$ satisfying $g_1(\rho_1^\star) = 0$ (resp. $g_2(\rho_1^\star) = 0$). Recall that $g_1(0) > 0$, $g_2(0) > 0$ and that both functions are decreasing (see Appendix 1). Since $\forall x,\ g_2(x) < g_1(x)$, there exists a range of ρ_1 for which $g_1(\rho_1) > 0$ but $g_2(\rho_1) \leq 0$. Therefore, for this range, we can find a solution for $g_2(\rho_1^\star) = 0$ but not for $g_1(\rho_1^\star) = 0$. The corresponding range of source rates is thus achieved with R-EE, but not with N-EE. The proof is similar for S-EE.

4.2 Derivation of $R_{\max}$ with sub-schemes $D^{(s)}$ and $C^{(s)}$

If sub-scheme $D^{(s)}$ is applied when S-EE goes in outage, the source power constraint is met, such that $(\rho_1^\circ + \eta_1^\circ)\theta P_s + (\rho_2^\circ + \eta_2^\circ)\bar{\theta} P_s = P_s$. This leads to $\rho_1^\circ = \frac{1}{\theta}$ and $R_{\max} = \theta \log_2\left(1 + \frac{P_s|h_s|^2}{\theta N}\right)$.

Similarly, if sub-scheme $C^{(s)}$ is applied when S-EE goes in outage, $(\rho_1^\ddagger + \eta_1^\ddagger)\theta P_s + (\rho_2^\ddagger + \eta_2^\ddagger)\bar{\theta} P_s = P_s$. Thus,

$$\theta\rho_1^\ddagger + \bar{\theta}\left(\rho_1^\ddagger + \frac{N}{P_s}\left(\frac{1}{|h_s|^2} - \frac{1}{|h_d|^2}\right)\right) = 1$$
$$\rho_1^\ddagger = 1 + \frac{\bar{\theta}N}{P_s}\left(\frac{1}{|h_d|2} - \frac{1}{|h_s|2}\right) = \frac{N}{P_s|h_s|^2}\left(2^R\left(\frac{|h_s|^2}{|h_d|^2}\right)^{\bar{\theta}} - 1\right),$$

from which we deduce $R_{\max}$ as in Section 4.3.4.

4.3 S-EE and R-EE achieve the same maximum rate $R_{\max}$

We want to prove that the power allocation of S-EE for $R = R_{\max}$ corresponds to the allocation of R-EE at this specific rate such that both achieve the same maximum rate $R_{\max}$. Also note that as long as $|h_s|^2 > |h_d|^2$, the maximum rate achieved by R-EE using $C^{(r)}$ (direct transmission) is strictly lower to the maximum rate achieved by R-EE using $B^{(r)}$ or $A^{(r)}$. Thus, when R-EE declares outage, either sub-scheme $B^{(r)}$ or $A^{(r)}$ is applied. Now, we successively consider the four possible cases for S-EE.

First, assume that $A^{(s)}$ is applied when S-EE goes in outage at $R = R_{\max}$. This means that $R \geq \theta \log_2\left(1 + \frac{\rho_1 P_s |h_s|^2}{N}\right)$ where ρ_1 solves $g_2(\rho_1) = 0$. Therefore, R-EE uses $A^{(r)}$ at this rate and $\rho_1^{(r)} = \rho_1^{(s)}$. Now, note that $\eta_2^{(r)} = \eta_2^{(s)}$, since both solve the same expression. For R-EE, $\rho_2^{(r)} = \frac{1-\rho_1^{(r)}\theta}{\bar{\theta}}$ and S-EE declares outage when the source power constraint is met. Thus, $\rho_2^{(r)} = \rho_2^{(s)}$ at $R = R_{\max}$. Noting that ρ_2 and ρ_r satisfy the same equation in both R-EE and S-EE ($\Gamma_r = G_d$), we get $\rho_r^{(r)} = \rho_r^{(s)} = 1/\bar{\theta}$, which means that the relay constraint is also achieved and R-EE goes in outage for this same source rate.

Second, assume that $B^{(s)}$ is used when S-EE goes in outage. This means that

$$R \leq \theta \log_2\left(1 + \frac{\rho_1 P_s |h_s|^2}{N}\right)$$

where ρ_1 solves $g_2(\rho_1) = 0$ and that R-EE uses $B^{(r)}$ when outage occurs. Similarly to the above analysis, we can show that the power allocation of S-EE for $R = R_{\max}$ corresponds to the allocation of R-EE at this specific rate.

Third, we assume that $C^{(s)}$ is used when S-EE goes in outage. In this case, we have, by concavity,

$$\begin{aligned} R_{\max} &= \log_2\left(\theta\left(1+\frac{P_s|h_s|^2}{N\theta}\right)+\bar{\theta}\frac{|h_s|^2}{|h_d|^2}\right)-\bar{\theta}\log_2\left(\frac{|h_s|^2}{|h_d|^2}\right) \\ &\geq \theta\log_2\left(1+\frac{P_s|h_s|^2}{N\theta}\right)+\bar{\theta}\log_2\left(\frac{|h_s|^2}{|h_d|^2}\right)-\bar{\theta}\log_2\left(\frac{|h_s|^2}{|h_d|^2}\right) \\ &= \theta\log_2\left(1+\frac{P_s|h_s|^2}{N\theta}\right) \end{aligned}$$

Thus, for all $\rho_1 \leq \frac{1}{\theta}$ solving $g_2(\rho_1)=0$, we get $R_{\max} \geq \theta\log_2\left(1+\frac{\rho_1 P_s|h_s|^2}{N}\right)$ such that at $R=R_{\max}$, R-EE uses $A^{(r)}$. Note that S-EE declares outage using $C^{(s)}$ if the source rate is met, such that $\eta_2^{(s)}=\frac{\left(1-\rho_1^{(s)}\theta\right)}{\bar{\theta}}$. Thus, we can show that the allocation of S-EE is also a solution for $A^{(r)}$ at this specific source rate. Similar analysis can be done for $D^{(s)}$.

Thus, for all cases, the power allocation of S-EE for $R=R_{\max}$ corresponds to the allocation of R-EE at this specific rate and both S-EE and R-EE achieve the same maximum rate $R_{\max}$.

5 Analysis of scheme G-EE: Continuity and Differentiability

First, we show the continuity of both schemes N-EE and R-EE at $R=R_{\max}^{(n)}$. Note that η_1 and η_2 solve the same equations for both sub-schemes $A^{(n)}$ and $A^{(r)}$. Then, since the first rate constraint is active, we get $g_3(\rho_1^\star)=\sqrt{\frac{P_r|h_r|^2\rho_r^\star}{P_s|h_d|^2\rho_2^\star}}$. In $A^{(n)}$, $g_3(\rho_1^\star)=\frac{|h_r|^2}{|h_d|^2}$, such that $\rho_1^\star$ and $\rho_r^\star$ of $A^{(n)}$ are also optimal for $A^{(r)}$. Finally, at $R=R_{\max}^{(n)}$, the source power constraint is met, such that $\rho_2^\star=\frac{\left(1-\rho_1^\star\theta\right)}{\bar{\theta}}-\eta_2^\star$ for both sub-schemes. Therefore, the optimal power allocation set for $A^{(n)}$ is also optimal for $A^{(r)}$ at this particular rate and schemes N-EE and R-EE are continuous.

Second, note that parameters ρ_1, η_2, ρ_2 and ρ_r are not differentiable themselves at $R=R_{\max}^{(n)}$, but that the energy of G-EE (the weighted sum of ρ_1, η_2, ρ_2 and ρ_r) is differentiable. To prove it, let's consider the left and right derivatives of G-EE. The left derivative (resp. the right derivative) corresponds to the derivative of the total energy consumed by N-EE (resp. R-EE) at $R=R_{\max}^{(n)}$. If both semi-derivatives are equal, then G-EE is differentiable at this source rate. Let's write both as functions of $\frac{\partial\rho_1}{\partial R}$, $\frac{\partial G_s}{\partial R}$ and $\frac{\partial G_d}{\partial R}$. Also note that $\frac{\partial G_i}{\partial R}=\frac{\ln 2}{\bar{\theta}}G_i-\frac{\theta P_s|h_s|^2}{\bar{\theta}N_0}\frac{G_i}{\Gamma_i}\frac{\partial\rho_1}{\partial R}$.

Let's consider the left derivative (scheme N-EE). In this case, we have:

$$\begin{aligned}\frac{\partial E_{\text{N-EE}}}{\partial R} &= \theta P_s \frac{\partial \rho_1^{(n)}}{\partial R} + \bar{\theta} P_s \left(\frac{\partial \eta_2^{(n)}}{\partial R} + \frac{\partial \rho_2^{(n)}}{\partial R} \right) + \bar{\theta} P_r \frac{\partial \rho_r^{(n)}}{\partial R} \\ &= \theta P_s \frac{\partial \rho_1^{(n)}}{\partial R} + \left(\frac{N\bar{\theta}}{|h_d|^2} - \frac{N\bar{\theta}}{|h_d|^2 + |h_r|^2} \right) \frac{\partial G_s^{(n)}}{\partial R} + \frac{N\bar{\theta}}{|h_d|^2 + |h_r|^2} \frac{\partial G_d^{(n)}}{\partial R}\end{aligned}$$

Since the source power constraint is met with equality, the right derivative (scheme R-EE) is

$$\begin{aligned}\frac{\partial E_{\text{R-EE}}}{\partial R} = \bar{\theta} P_r \frac{\partial \rho_r^{(r)}}{\partial R} = \frac{\bar{\theta} N}{|h_r|^2} \Bigg(& \frac{P_s |h_d|^2}{N} \frac{\partial \rho_2^{(r)}}{\partial R} + \frac{\partial G_d^{(r)}}{\partial R} - \frac{\partial G_d^{(r)}}{\partial R} \\ & - \sqrt{\frac{P_s |h_d|^2}{N}} \left(\sqrt{\frac{(G_d - G_s)}{\rho_2}} \frac{\partial \rho_2^{(r)}}{\partial R} + \sqrt{\frac{\rho_2}{(G_d - G_s)}} \left(\frac{\partial G_d^{(r)}}{\partial R} - \frac{\partial G_d^{(r)}}{\partial R} \right) \right) \Bigg)\end{aligned}$$

Since $\frac{\partial \rho_2^{(r)}}{\partial R} = -\frac{\theta}{\bar{\theta}} \frac{\partial \rho_1^{(r)}}{\partial R} - \frac{N}{P_s |h_d|^2} \frac{\partial G_s^{(r)}}{\partial R}$, then

$$\begin{aligned}\frac{\partial E_{\text{R-EE}}}{\partial R} = & \frac{\bar{\theta} N}{|h_r|^2} \left(\frac{P_s |h_d|^2}{N} - \sqrt{\frac{P_s |h_d|^2}{N} \frac{(G_d - G_s)}{\rho_2}} \right) \frac{\partial \rho_1^{(r)}}{\partial R} \\ & + \frac{\bar{\theta} N}{|h_r|^2} \left(1 - \sqrt{\frac{P_s |h_d|^2}{N} \frac{\rho_2}{(G_d - G_s)}} \right) \frac{\partial G_d^{(r)}}{\partial R} \\ & - \frac{\bar{\theta} N}{|h_r|^2} \left(\frac{N}{P_s |h_d|^2} \left(\frac{P_s |h_d|^2}{N} - \sqrt{\frac{P_s |h_d|^2}{N} \frac{(G_d - G_s)}{\rho_2}} \right) \right. \\ & \left. + \left(1 - \sqrt{\frac{P_s |h_d|^2}{N} \frac{\rho_2}{(G_d - G_s)}} \right) \right) \frac{\partial G_s^{(r)}}{\partial R}\end{aligned}$$

Now, recall that both N-EE and R-EE are continuous at this particular rate. Thus, $\rho_2 = \frac{P_r |h_d|^2}{P_s |h_r|^2} \rho_r = \frac{N |h_d|^2 (G_d - G_s)}{P_s (|h_r|^2 + |h_d|^2)^2}$ and the right derivative of G-EE is also such that

$$\frac{\partial E_{\text{R-EE}}}{\partial R} = \theta P_s \frac{\partial \rho_1^{(r)}}{\partial R} + \left(\frac{N\bar{\theta}}{|h_d|^2} - \frac{N\bar{\theta}}{|h_d|^2 + |h_r|^2} \right) \frac{\partial G_s^{(r)}}{\partial R} + \frac{N\bar{\theta}}{|h_d|^2 + |h_r|^2} \frac{\partial G_d^{(r)}}{\partial R}$$

which is similar to the left derivative $\frac{\partial E_{\text{N-EE}}}{\partial R}$. Now, define the constant C (that may depend on channel gains or on the source rate) such that $\frac{\partial \rho_1^{(r)}}{\partial R} = \frac{\partial \rho_1^{(n)}}{\partial R} + C$. We get $\frac{\partial G_i^{(r)}}{\partial R} = \frac{\partial G_i^{(n)}}{\partial R} -$

$C\frac{\theta P_s|h_i|^2}{\bar{\theta}N}\frac{G_i}{\Gamma_i}$ and

$$\begin{aligned}\frac{\partial E_{\text{R-EE}}}{\partial R} &= \frac{\partial E_{\text{N-EE}}}{\partial R} - C\theta P_s\left(\left(1-\frac{|h_d|^2}{|h_d|^2+|h_r|^2}\right)\frac{|h_s|^2}{|h_d|^2}\frac{G_s}{\Gamma_s}+\frac{|h_d|^2}{|h_d|^2+|h_r|^2}\frac{G_d}{\Gamma_d}-1\right)\\ &= \frac{\partial E_{\text{N-EE}}}{\partial R} - C\theta P_s\left(1+\frac{|h_r|^2}{|h_d|^2}\right)g_1(\rho_1)\end{aligned}$$

Since $g_1(\rho_1)=0$, then $\frac{\partial E_{\text{N-EE}}}{\partial R}=\frac{\partial E_{\text{R-EE}}}{\partial R}$ at $R=R_{\max}^{(n)}$, which completes the proof.

APPENDIX II

APPENDIX OF SECOND JOURNAL PAPER

1 Considered path-loss models

In this paper, we propose to model the path-loss depending on the relay height and base our model on the WINNER II project [69]. We consider an urban environment where the user is located outdoor at street level and the base station is clearly above surrounding buildings. Both the base station and the relay station are fixed. As recommended by 802.16j/m and 3GPP specifications [151], the relay station establishes a high quality link with the base station and LOS is assumed for this channel. Since the relay height is allowed to vary from below to above rooftop, the propagation environment for h_r and h_s changes notably. However, the WINNER II project does not provide a continuous model as a function of the relay height. We will therefore consider the two situations summarized in Table II-1. Furthermore, we assume that the terminals are sufficiently robust against small-scale environment parameters, such as multipath components.

Tableau-A II-1 Considered WINNER II scenarios

	Vicinity Relay	**Base-station-like Relay**
h_s	B1 (LOS / NLOS)	C2 (LOS / NLOS)
h_r	B5c LOS	Free space (FS)
h_d	C2 (LOS / NLOS)	

2 Proof of Probability of Relaying in Eq. (5.18)

The expression of Eq. (5.18) is obtained using geometrical principles. $\mathbb{P}_{\text{RTx}}$ is equal to the surface of the REA divided by the surface of the cell. We can reduce the analysis to one cell sector and rather focus on the surface of $\overline{\text{REA}}$, where DTx is more energy-efficient than RTx. We denote $\mathcal{L}$ the strait line of equation $x = \mathsf{D}_{\min}$ and $\mathcal{C}$ the circle centred at the base station and of radius R_{DTx}. This is illustrated in Figure II-1.

First, we focus on the angles φ and ϕ. When $\frac{\sqrt{3}}{2}\mathsf{R}_{\text{DTx}} \leq \frac{3}{4}\mathsf{R}_{\text{cov}}$, the cell-edge is never limiting and $\overline{\text{REA}}$ is upper-bound by either $\mathcal{L}$ or $\mathcal{C}$ (left part of the hexagonal cell sector). This gives $\phi = \frac{\pi}{6}$. Furthermore, if $\mathsf{D}_{\min} \leq \cos\left(\frac{\pi}{6}\right)\mathsf{R}_{\text{DTx}}$, the set of all user positions within the sector for which $x < \mathsf{D}_{\min}$ is strictly included in the sphere $\mathcal{C}\left(0, \mathsf{R}_{\text{DTx}}\right)$. Therefore, $\overline{\text{REA}}$ is upper-bounded by $\mathcal{L}$ only and $\varphi = \phi$. Otherwise, both $\mathcal{L}$ and $\mathcal{C}$ upper-bound $\overline{\text{REA}}$ and intersect at $(\mathsf{R}_{\text{DTx}}, \varphi)$, with $\varphi = \arccos\left(\frac{\mathsf{D}_{\min}}{\mathsf{R}_{\text{DTx}}}\right)$. Second, when $\frac{3}{4}\mathsf{R}_{\text{cov}} \leq \frac{\sqrt{3}}{2}\mathsf{R}_{\text{DTx}}$ (right part of the hexagonal cell sector), the cell edge can be limiting. We denote X the x-coordinate of the intersection of $\mathcal{C}$ with the cell edge. If $\mathsf{D}_{\min} \leq \frac{3}{4}\mathsf{R}_{\text{cov}}$, as before, $\overline{\text{REA}}$ is upper-bounded by $\mathcal{L}$ only and $\varphi = \phi$. If $\frac{3}{4}\mathsf{R}_{\text{cov}} \leq \mathsf{D}_{\min} \leq \mathsf{X}$, $\overline{\text{REA}}$ is upper-bounded by both $\mathcal{L}$ and the cell edge, which gives $\varphi = \phi$. Finally, if $\mathsf{X} \leq \mathsf{D}_{\min}$, $\overline{\text{REA}}$ is upper-bounded by $\mathcal{L}$, $\mathcal{C}$ and the cell edge, and angles are computed using geometrical properties. This gives the expressions of φ and ϕ as given in Eq. (5.19).

We can now deduce the surface of $\overline{\text{REA}}$. To do so, we decompose this surface into elementary geometrical shapes, whose surface can be easily computed. For $\theta \in [0, \varphi]$, $\overline{\text{REA}}$ is upper-bounded by $\mathcal{L}$ and its surface reduces to a triangle for this range of angles. For $\theta \in [\varphi, \phi]$, $\overline{\text{REA}}$ is upper-bounded by $\mathcal{C}$. The related surface here is a portion of sphere. Finally, when $\phi < \frac{\pi}{6}$, $\overline{\text{REA}}$ is upper-bounded by the cell edge for $\theta \in \left[\phi, \frac{\pi}{6}\right]$. This also corresponds to the surface of a triangle. Using, this decomposition, we get Eq. (5.18).

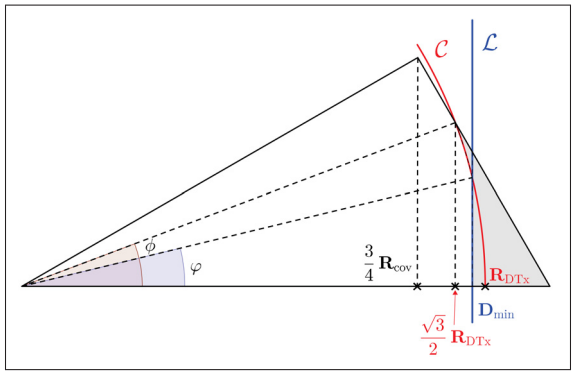


Figure-A II-1 Illustration of the Relay Efficiency Area for a half cell sector

APPENDIX III

APPENDIX OF THIRD JOURNAL PAPER

1 Proof of Lemma 6.1: Lower-bound for $\mathbb{P}_{\text{ER}}$

The probability $\mathbb{P}_{\text{ER}}$ is expressed as follows:

$$\mathbb{P}_{\text{ER}} = \int_0^{\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}} \int_0^{\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}} \int_0^{\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}} \mathbb{1}\left(E_b + E_r \leq E_d\right) f(E_d, E_b, E_r) dE_d dE_b dE_r \tag{A III-1}$$

with $f(E_d, E_b, E_r)$ the p.d.f. of the energy consumptions E_s, E_r and E_d.

Among the wide possibilities for lower bounds, we aim at removing the power constraints which set conditions on $E_b + E_r \leq E_d$. Hence, we decompose $\mathbb{P}_{\text{ER}}$ into elementary probabilities that discard the triple integral. We get $\mathbb{P}_{\text{ER}} = \mathbb{P}^{(1)} + \mathbb{P}^{(2)}$ with

$$\begin{aligned}
\mathbb{P}^{(1)} &= \mathbb{P}\left(E_b + E_r \leq E_d \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) \\
\mathbb{P}^{(2)} &= \mathbb{P}\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \ \cap \ E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \ \cap \ E_b + E_r \leq E_d\right).
\end{aligned}$$

First, we find a lower bound for $\mathbb{P}^{(1)}$. We have:

$$\begin{aligned}
\mathbb{P}\left(E_b + E_r \leq E_d\right) = \ & \mathbb{P}\left(E_b + E_r \leq E_d \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) \\
& + \mathbb{P}\left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d\right) \mathbb{P}\left(E_b \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \ \cap \ E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) && \text{(a)} \\
& + \mathbb{P}\left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \ \cap \ \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \leq E_b \ \cup \ \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_r\right) \ \cap \ E_b + E_r \leq E_d\right) && \text{(b)} \\
& + \mathbb{P}\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) \mathbb{P}\left(E_b + E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right) && \text{(c)} \\
& + \mathbb{P}\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \ \cap \ \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_b + E_r \leq E_d\right) && \text{(d)}
\end{aligned}$$

Regarding the probabilities of lines (a) and (c), the condition $E_b + E_r \leq E_d$ necessarily holds given energy constraints and these two probabilities can be readily computed in closed-form. Second, the probabilities of lines (b) and (d), denoted $\mathbb{P}^{(b)}$ and $\mathbb{P}^{(d)}$ respectively, can only be

expressed in integral form, but are respectively upper-bounded by

$$\mathbb{P}^{(b)} \leq \mathbb{P}\left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \cap \left(\mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \leq E_s \cup \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_r\right)\right)$$
$$\mathbb{P}^{(d)} \leq \mathbb{P}\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \cap \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_s + E_r \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right)$$

Plugging these upper-bounds into the expression for $\mathbb{P}\left(E_b + E_r \leq E_d\right)$, we obtain the lower-bound $\mathbb{P}_{\text{low}}^{(1)}$ given in Eq. (6.9). Next, we have:

$$\mathbb{P}^{(2)} = \mathbb{P}\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \cap E_b + E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right)$$
$$+ \mathbb{P}\left(\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_b + E_r \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} \cap E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right)$$

Note that the energy E_r consumed by BS to transmit data to RS is generally low thanks to strong channel conditions. Thereby, the probability in second line approaches 0 and $\mathbb{P}^{(2)}$ can be tightly lower-bounded by $\mathbb{P}_{\text{low}}^{(2)}$ of Eq. (6.10). We now show that $\mathbb{P}_{\text{low}} = \mathbb{P}_{\text{low}}^{(1)} + \mathbb{P}_{\text{low}}^{(2)}$ can be approximated by a closed-form expression. First, $\mathbb{P}_{\text{low}}$, and thus $\mathbb{P}_{\text{CR}}$ and $\mathbb{P}_{\text{CD}}$, are computed using the cumulative distribution function Φ of the standard normal distribution:

$$\mathbb{P}\left(E_k \leq \mathsf{E}_{\mathsf{k}}^{(\mathsf{m})}\right) = \Phi\left(\frac{\ln\left(\mathsf{E}_{\mathsf{k}}^{(\mathsf{m})}\right) - \mu_k}{\sigma_k}\right) \tag{A III-2}$$

where μ_k and σ_k are given in Eq. (6.6). Even if Φ is written as an integral, it is widely available in scientific tools through well-known tables, such that its computation does not imply much complexity and can be considered as closed-form. In addition, $\mathbb{P}_{\text{low}}$ requires the computation of $\mathbb{P}\left(E_b + E_r \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})} + \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right)$ and $\mathbb{P}\left(E_b + E_r \leq E_d\right) = \mathbb{P}\left(E_b^{(0)}\frac{s_d}{s_b} + E_r^{(0)}\frac{s_d}{s_r} \leq E_d^{(0)}\right)$, both of which involve the sum of two log-normal random variables. Such distributions do not have a closed-form expression, but have been extensively explored in the literature [152, 153]. In this work, we consider the Fenton-Wilkinson approach and approximate these sum distributions by log-normal random variables. $E_b + E_r$ is approximated by $E_{b+r} \sim \log\mathcal{N}\left(\mu_{b+r}, \sigma_{b+r}^2\right)$, where μ_{b+r} and σ_{b+r}^2 are computed as given in [153, Eq. (9-12)]. Similar computation can be

performed for $E_b^{(0)}\frac{s_d}{s_b} + E_r^{(0)}\frac{s_d}{s_r}$, taking into account the correlation coefficient between $E_b^{(0)}\frac{s_d}{s_b}$ and $E_r^{(0)}\frac{s_d}{s_r}$.

Consequently, using the Fenton-Wilkinson approach, we have decomposed $\mathbb{P}_{\text{low}}(x,y)$ into elementary probabilities that can be computed in closed-form using Eq. (A III-2), thus avoiding the computation of a triple integral for each possible user location $M(x,y)$ within the cell.

2 Proof of Lemma 6.2: Bound for $\mathbb{P}_{\mathbf{ER}}\mathbb{E}(E_{b+r}\,|\,\mathcal{C}_{\mathbf{ER}})$

Here, we use the decomposition for $\mathbb{P}_{\text{low}}$ that has been proposed in Lemma 6.1:

$$\begin{aligned}
&\mathbb{P}_{\text{ER}}\mathbb{E}\left[E_{b+r}\,|\,\mathcal{C}_{\text{ER}}\right] \leq\ E_1 + E_2 \\
\text{with}\quad &E_1 = \mathbb{P}^{(1)}\mathbb{E}\left[E_{b+r}\,|\,E_b + E_r \leq E_d \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right] \\
\text{and}\quad &E_2 = \mathbb{P}^{(2)}\mathbb{E}\left[E_{b+r}\,|\,\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})} \leq E_d \leq \mathsf{E}_{\mathsf{B}}^{(\mathsf{m})}\ \cap\ E_b + E_r \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right] \\
&\qquad =\ g(b+r,\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})})\mathbb{P}^{(2)}\ \geq\ g(b+r,\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})})\mathbb{P}_{\text{low}}^{(2)}
\end{aligned}$$

where g is given by Eq. (6.15). With regards to E_1, we come back the integral form. Denoting f_k the probability density function of E_k, we get:

$$\begin{aligned}
E_1 &= \int_0^{\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}}\int_0^{\mathsf{E}_{\mathsf{d}}} (E_s + E_r)\, f_{s+r}\,(E_s + E_r)\, f_d\,(E_d)\, d(Es + Er) dE_d \\
&\simeq \int_0^{\mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}} \exp\left(\mu_{s+r} + \frac{\sigma_{s+r}^2}{2}\right)\Phi\left(-\sigma_{s+r} + \frac{\ln(E_d) - \mu_{s+r}}{\sigma_{s+r}}\right) f_d\,(E_d)\, dE_d \\
&= \exp\left(\mu_{s+r} + \frac{\sigma_{s+r}^2}{2}\right)\mathbb{P}\left(\exp\left(\sigma_{s+r}^2\right)(E_b + E_r) \leq E_d \leq \mathsf{E}_{\mathsf{R}}^{(\mathsf{m})}\right)
\end{aligned}$$

BIBLIOGRAPHY

[1] Nosratinia, A., T. Hunter and A. Hedayat. October 2004. "Cooperative communication in wireless networks". *IEEE Communications Magazine*, vol. 42, n° 10, p. 74-80.

[2] Hossain, E., V. K. Bhargava and G. P. Fettweis, 2012. *Green radio communication networks*. Cambridge University Press.

[3] Correia, L. M., D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer and L. Van Der Perre. 2010. "Challenges and enabling technologies for energy aware mobile radio networks". *IEEE Communications Magazine*, vol. 48, n° 11, p. 66-72.

[4] Hasan, Z., H. Boostanimehr and V. K. Bhargava. Fourth quarter 2011. "Green Cellular Networks: A Survey, Some Research Issues and Challenges". *IEEE Communications Surveys & Tutorials*, vol. 13, n° 4, p. 524-540.

[5] Han, C., T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C.-X. Wang, A. L. Tuan, M. Nakhai, Z. Jiayi and L. Hanzo. 2011. "Green radio: radio techniques to enable energy-efficient wireless networks". *IEEE Communications Magazine*, vol. 49, n° 6, p. 46-54.

[6] Gódor, I. June 2012. *Final Report on Green Network Technologies*. Technical Report INFSO-ICT-247733 EARTH - Deliverable D3.3.

[7] Consortium GreenTouch. 2010. "Challenge & Opportunity". In *Website of GreenTouch*. Online. <http://www.greentouch.org/>. Accessed March, 2014.

[8] Feng, D., C. Jiang, G. Lim, L. J. Cimini Jr, G. Feng and G. Y. Li. 2013. "A survey of energy-efficient wireless communications". *IEEE Communications Surveys & Tutorials*, vol. 15, n° 1, p. 167-178.

[9] Chen, Y., S. Zhang, S. Xu and G. Y. Li. 2011. "Fundamental trade-offs on green wireless networks". *IEEE Communications Magazine*, vol. 49, n° 6, p. 30-37.

[10] Antoniadis, A., N. Barcelo, M. Nugent, K. Pruhs and M. Scquizzato. 2014. "Energy-efficient Circuit Design". In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*. p. 303-312. ACM.

[11] Cui, S., A. J. Goldsmith and A. Bahai. September 2005. "Energy-constrained modulation optimization". *IEEE Transactions on Wireless Communications*, vol. 4, n° 5, p. 2349-2360.

[12] Miao, G. and J. Zhang. December 2011. "On Optimal Energy-Efficient Multi-User MIMO". In *IEEE Global Telecommunications Conference (GLOBECOM)*. p. 1-6.

[13] Li, G. Y., Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen and S. Xu. December 2011. "Energy-efficient wireless communications: tutorial, survey, and open issues". *IEEE Wireless Communications*, vol. 18, n° 6, p. 28-35.

[14] Ngo, H. Q., E. G. Larsson and T. L. Marzetta. April 2013. "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems". *IEEE Transactions on Communications*, vol. 61, n° 4, p. 1436-1449.

[15] Rao, J. B. and A. O. Fapojuwo. First quarter 2014. "A survey of energy efficient resource management techniques for multicell cellular networks". *IEEE Communications Surveys & Tutorials*, vol. 16, n° 1, p. 154-180.

[16] Miao, G., N. Himayat, G. Y. Li, A. T. Koc and S. Talwar. 2009. "Interference-aware energy-efficient power optimization". In *IEEE International Conference on Communications (ICC)*. p. 1-5.

[17] Marsch, P. and G. P. Fettweis, 2011. *Coordinated Multi-Point in Mobile Communications: from theory to practice*. Cambridge University Press.

[18] Yang, W.-H., Y.-C. Wang, Y.-C. Tseng and B.-S. Lin. 2009. "An Energy-Efficient Handover Scheme with Geographic Mobility Awareness in WiMAX-WiFi Integrated Networks". In *IEEE Wireless Communications and Networking Conference (WCNC)*. p. 1-6.

[19] Niu, Z., Y. Wu, J. Gong and Z. Yang. November 2010. "Cell zooming for cost-efficient green cellular networks". *IEEE Communications Magazine*, vol. 48, n° 11, p. 74-79.

[20] Xiong, C., G. Y. Li, S. Zhang, Y. Chen and S. Xu. November 2011. "Energy- and Spectral-Efficiency Tradeoff in Downlink OFDMA Networks". *IEEE Transactions on Wireless Communications*, vol. 10, n° 11, p. 3874-3886.

[21] Héliot, F., M. A. Imran and R. Tafazolli. May 2012. "On the Energy Efficiency-Spectral Efficiency Trade-off over the MIMO Rayleigh Fading Channel". *IEEE Transactions on Communications*, vol. 60, n° 5, p. 1345-1356.

[22] Onireti, O., F. Héliot and M. A. Imran. February 2012. "On the Energy Efficiency-Spectral Efficiency Trade-Off in the Uplink of CoMP System". *IEEE Transactions on Wireless Communications*, vol. 11, n° 2, p. 556-561.

[23] Hong, X., J. Wang, C.-X. Wang and J. Shi. July 2014. "Cognitive radio in 5G: a perspective on energy-spectral efficiency trade-off". *IEEE Communications Magazine*, vol. 52, n° 7, p. 46-53.

[24] El Gamal, A., M. Mohseni and S. Zahedi. April 2006. "Bounds on capacity and minimum energy-per-bit for AWGN relay channels". *IEEE Transactions on Information Theory*, vol. 52, n° 4, p. 1545-1561.

[25] Jain, A., S. Kulkarni and S. Verdú. March 2010. "Minimum Energy per Bit for Wideband Wireless Multicasting: Performance of Decode-and-Forward". In *IEEE INFOCOM*. p. 1-9.

[26] Fehske, A. J., P. Marsch and G. P. Fettweis. December 2010. "Bit per Joule efficiency of cooperating base stations in cellular networks". In *IEEE Global Telecommunications Workshops (GLOBECOM Workshops)*. p. 1406-1411.

[27] Parzysz, F., M. Vu and F. Gagnon. October 2011. "A half-duplex relay coding scheme optimized for energy efficiency". In *IEEE Information Theory Workshop (ITW)*. p. 306 -310.

[28] Parzysz, F., M. Vu and F. Gagnon. 2013. "Energy Minimization for the Half-Duplex Relay Channel with Decode-Forward Relaying". *IEEE Transaction on Communications*, vol. 61, n° 6, p. 2232-2247.

[29] Parzysz, F., M. Vu and F. Gagnon. April 2014. "Impact of Propagation Environment on Energy-Efficient Relay Placement: Model and Performance Analysis". *IEEE Transactions on Wireless Communications*, vol. 13, n° 4, p. 2214-2228.

[30] Parzysz, F., M. Vu and F. Gagnon. 2015. "A Joint Analysis of Energy and Interference for Optimized Relay Deployment in Cellular Networks". *Submitted to IEEE Journal on Selected Areas in Communications, Series on Green Communications and Networking*.

[31] Parzysz, F., M. Vu and F. Gagnon. May 2011. "Energy-efficient schemes for on-demand relaying". In *34th IEEE Sarnoff Symposium*. p. 1-6.

[32] Parzysz, F., M. Vu and F. Gagnon. June 2012. "Optimal distributed coding schemes for energy efficiency in the fading relay channel". In *IEEE International Conference on Communications (ICC)*. p. 2207-2212.

[33] Parzysz, F., M. Vu and F. Gagnon. Sep. 2014. "Trade-offs on Energy-Efficient Relay Deployment in Cellular Networks". In *IEEE Vehicular Technology Conference Fall (VTC-Fall)*.

[34] Debaillie, B., D. van den Broek, C. Lavin, B. van Liempd, E. Klumperink, C. Palacios, J. Craninckx and A. Parssinen. Sept 2014. "Analog/RF Solutions Enabling Compact Full-Duplex Radios". *IEEE Journal on Selected Areas in Communications*, vol. 32, n° 9, p. 1662-1673.

[35] Chen, D. and J. Laneman. July 2006. "Modulation and demodulation for cooperative diversity in wireless systems". *IEEE Transactions on Wireless Communications*, vol. 5, n° 7, p. 1785-1794.

[36] Annavajjala, R., A. Maaref and J. Zhang. May 2010. "Demodulate-and-Forward Relaying with Higher Order Modulations: Impact of Channel State Uncertainty". In *IEEE International Conference on Communications (ICC)*. p. 1-5.

[37] El Gamal, A. and Y.-H. Kim, 2010. *Lecture Notes on Network Information Theory*. Cambridge University Press, Online. <http://202.114.89.42/resource/pdf/6073.pdf>.

[38] Khafagy, M., A. Ismail, M. Alouini and S. Aissa. June 2013. "On the Outage Performance of Full-Duplex Selective Decode-and-Forward Relaying". *IEEE Communications Letters*, vol. 17, n° 6, p. 1180-1183.

[39] Laneman, J. N., D. N. Tse and G. W. Wornell. December 2004. "Cooperative diversity in wireless networks: Efficient protocols and outage behavior". *IEEE Transactions on Information Theory*, vol. 50, n° 12, p. 3062 - 3080.

[40] Ikki, S. and M. Ahmed. February 2011. "Performance analysis of incremental-relaying cooperative-diversity networks over rayleigh fading channels". *IET Communications*, vol. 5, n° 3, p. 337-349.

[41] Beres, E. and R. Adve. January 2008. "Selection cooperation in multi-source cooperative networks". *IEEE Transactions on Wireless Communications*, vol. 7, n° 1, p. 118-127.

[42] Yang, Y., H. Hu, J. Xu and G. Mao. 2009. "Relay technologies for WiMax and LTE-advanced mobile systems". *IEEE Communications Magazine*, vol. 47, n° 10, p. 100–105.

[43] Banerjee, S. and A. Misra. 2002. "Minimum Energy Paths for Reliable Communication in Multi-hop Wireless Networks". In *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking &Amp; Computing*. p. 146–156. ACM.

[44] Lee, D., S. Zhou and Z. Niu. December 2011. "Multi-Hop Relay Network for Base Station Energy Saving and Its Performance Evaluation". In *IEEE Global Telecommunications Conference (GLOBECOM)*. p. 1-5.

[45] Shannon, C. E. 1948. "A mathematical theory of communication". *The Bell System Technical Journal*.

[46] Cover, T. M. and J. A. Thomas, 2012. *Elements of Information Theory*. 2nd edition, John Wiley & Sons.

[47] Sendonaris, A., E. Erkip and B. Aazhang. November 2003. "User cooperation diversity. Part I. System description". *IEEE Transactions on Communications*, vol. 51, n° 11, p. 1927-1938.

[48] Cover, T. and A. El Gamal. September 1979. "Capacity theorems for the relay channel". *IEEE Transactions on Information Theory*, vol. 25, n° 5, p. 572-584.

[49] Liang, Y. and V. V. Veeravalli. 2007. "Cooperative relay broadcast channels". *IEEE Transactions on Information Theory*, vol. 53, n° 3, p. 900-928.

[50] Liang, Y. and G. Kramer. October 2007. "Rate Regions for Relay Broadcast Channels". *IEEE Transactions on Information Theory*, vol. 53, n° 10, p. 3517-3535.

[51] Lei, H., X. fan Wang and P. Chong. Sept 2010. "Opportunistic Relay Selection in Future Green Multihop Cellular Networks". In *IEEE Vehicular Technology Conference (VTC-Fall)*. p. 1-5.

[52] Brandão, A. September 2010. "On the Energy Consumption of Relay Networks". In *IEEE Vehicular Technology Conference (VTC-Fall)*. p. 1-5.

[53] Yang, D., X. Fang and G. Xue. December 2011. "Near-Optimal Relay Station Placement for Power Minimization in WiMAX Networks". In *IEEE Global Telecommunications Conference (GLOBECOM)*. p. 1-5.

[54] Wu, G. and G. Feng. June 2012. "Energy-efficient relay deployment in next generation cellular networks". In *IEEE International Conference on Communications (ICC)*. p. 5757-5761.

[55] Li, H., G. Koudouridis, G. Hedby and T. Wu. June 2012. "Cooperative relay design for energy efficient cell capacity improvements". In *IEEE International Conference on Communications (ICC)*. p. 6018-6023.

[56] Joshi, G. and A. Karandikar. 2011. "Optimal relay placement for cellular coverage extension". In *IEEE National Conference on Communications (NCC)*. p. 1-5.

[57] Khakurel, S., M. Mehta and A. Karandikar. 2012. "Optimal relay placement for coverage extension in LTE-A cellular systems". In *IEEE National Conference on Communications (NCC)*. p. 1-5.

[58] Youssef, R., M. Helard, M. Crussiere and J. Helard. April 2012. "Effect of relaying on coverage in realistic deployment scenarios". In *International Conference on Telecommunications (ICT)*. p. 1-5.

[59] Lin, B., P.-H. Ho, L.-L. Xie and X. Shen. 2007. "Optimal relay station placement in IEEE 802.16 j networks". In *Proceedings of the 2007 international conference on Wireless communications and mobile computing*. p. 25-30. ACM.

[60] Lin, B., P.-H. Ho, L.-L. Xie, X. Shen and J. Tapolcai. February 2010. "Optimal Relay Station Placement in Broadband Wireless Access Networks". *IEEE Transactions on Mobile Computing*, vol. 9, n° 2, p. 259 -269.

[61] Yang, C., W. Wang, S. Zhao and M. Peng. march 2010. "Location optimization for decode-and-forward opportunistic cooperative networks". In *The 5th Ann. ICST Wireless Internet Conf. (WICON)*. p. 1-5.

[62] Chandwani, G., S. N. Datta and S. Chakrabarti. 2010. "Relay assisted cellular system for energy minimization". In *IEEE Annual India Conference (INDICON)*. p. 1-4.

[63] Proakis, J. G., 2001. *Digital Communications*. McGraw-Hill series in electrical and computer engineering : communications and signal processing.

[64] Goldsmith, A., 2005. *Wireless communications*. Cambridge university press.

[65] Tse, D. and P. Viswanath, 2005. *Fundamentals of wireless communication*. Cambridge university press.

[66] Gallager, R., 2008. *Principles of Digital Communication*. Cambridge University Press.

[67] Damosso, E. and L. M. Correia. 1999. *Digital Mobile Radio Towards Future Generation Systems Communications. COST 231 Final Report*. Technical Report INFSO-ICT-247733 EARTH - Deliverable D3.3.

[68] Howard, H., 2006. *Spatial channel model for multiple input multiple output (MIMO) simulations*. 3GPP2 Spacial Model Ad-hoc Group3GPP TR 25.996.

[69] Kyösti, P., J. Meinilä, L. Hentilä, X. Zhao, T. Jämsä, C. Schneider, M. Narandzić, M. Milojević, A. Hong, J. Ylitalo, V.-M. Holappa, M. Alatossava, R. Bultitude, Y. de Jong and T. Rautiainen. September 2007. *WINNER II Channel Models, Part I Channel Models*. Technical Report IST-4-027756 WINNER II, D1.1.2 V1.1 - Public.

[70] Senarath, G., W. Tong, P. Zhu, H. Zhang, D. Steer, D. Yu, M. Naden and D. Kitchener, February 2007. *Multi-hop Relay System Evaluation Methodology (Channel Model and Performance Metric)*. IEEE 802.16j-06-013r3.

[71] Wang, C.-X., X. Hong, X. Ge, X. Cheng, G. Zhang and J. Thompson. February 2010. "Cooperative MIMO channel models: A survey". *IEEE Communications Magazine*, vol. 48, n° 2, p. 80-87.

[72] Lu, H.-C., W. Liao and F.-S. Lin. January 2011. "Relay Station Placement Strategy in IEEE 802.16j WiMAX Networks". *IEEE Transactions on Communications*, vol. 59, n° 1, p. 151 -158.

[73] Canadian Wireless Telecommunications Association. April 2013. "Connecting Canadians: Wireless Antenna Tower Siting in Canada". Retrieved from <http://cwta.ca/wordpress/wp-content/uploads/2011/08/Connecting.pdf>. Accessed March, 2014.

[74] Health Canada, 2009. *Safety Code 6: Health Canada's Radiofrequency Exposure Guidelines*. Online. <http://www.hc-sc.gc.ca/ewh-semt/pubs/radiation/radio_guide-lignes_direct/index-eng.php>. Accessed March, 2014.

[75] Industry Canada, August 2014. *Guide to Assist Land-use Authorities in Developing Antenna System Siting Protocols*. Online. <https://www.ic.gc.ca/eic/site/smt-gst.nsf/vwapj/LUA-e.pdf/\protect\T1\textdollarfile/LUA-e.pdf>. Accessed March, 2014.

[76] Hoydis, J., M. Kobayashi and M. Debbah. March 2011. "Green Small-Cell Networks". *IEEE Vehicular Technology Magazine*, vol. 6, n° 1, p. 37-43.

[77] Irmer, R., H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele and V. Jungnickel. February 2011. "Coordinated multipoint: Concepts, performance, and field trial results". *IEEE Communications Magazine*, vol. 49, n° 2, p. 102-111.

[78] Mondal, B., E. Visotsky, T. Thomas, X. Wang and A. Ghosh. September 2012. "Performance of downlink comp in LTE under practical constraints". In *IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. p. 2049-2054.

[79] Cili, G., H. Yanikomeroglu and F. Yu. September 2013. "Energy Efficiency and Capacity Evaluation of LTE-Advanced Downlink CoMP Schemes Subject to Channel Estimation Errors and System Delay". In *IEEE Vehicular Technology Conference (VTC-Fall)*. p. 1-5.

[80] Hossain, E., D. Kim and V. Bhargava, 2011. *Cooperative Cellular Wireless Networks*. Cambridge University Press.

[81] Kim, H., C.-B. Chae, G. De Veciana and R. W. Heath. August 2009. "A cross-layer approach to energy efficiency for adaptive MIMO systems exploiting spare capacity". *IEEE Transactions on Wireless Communications*, vol. 8, n° 8, p. 4264-4275.

[82] Cui, S., A. J. Goldsmith and A. Bahai. August 2004. "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks". *IEEE Journal on Selected Areas in Communications*, vol. 22, n° 6, p. 1089-1098.

[83] Wu, D., L. Zhou, Y. Cai, R. Hu and Y. Qian. November 2014. "Energy-Aware Dynamic Cooperative Strategy Selection for Relay-Assisted Cellular Networks: An Evolutionary Game Approach". *IEEE Transactions on Vehicular Technology*, vol. 63, n° 9, p. 4659-4669.

[84] Madan, R., N. Mehta, A. Molisch and J. Zhang. August 2008. "Energy-Efficient Cooperative Relaying over Fading Channels with Simple Relay Selection". *IEEE Transactions on Wireless Communications*, vol. 7, n° 8, p. 3013-3025.

[85] Ghosh, A., J. Zhang, J. G. Andrews and R. Muhamed, 2010. *Fundamentals of LTE*. Pearson Education.

[86] Himayat, N., S. Talwar, A. Rao and R. Soni. 2010. "Interference management for 4G cellular standards [WIMAX/LTE UPDATE]". *IEEE Communication Magazine*, vol. 48, n° 8, p. 86-92.

[87] Shafiul Alam, A., L. Dooley and A. Poulton. December 2012. "Energy efficient relay-assisted cellular network model using base station Switching". In *IEEE Globecom Workshops (GC Wkshps)*. p. 1155-1160.

[88] Oh, E., K. Son and B. Krishnamachari. May 2013. "Dynamic Base Station Switching-On/Off Strategies for Green Cellular Networks". *IEEE Transactions on Wireless Communications*, vol. 12, n° 5, p. 2126-2136.

[89] Dhillon, H. S., R. K. Ganti, F. Baccelli and J. G. Andrews. 2012. "Modeling and analysis of K-tier downlink heterogeneous cellular networks". *IEEE Journal on Selected Areas in Communications*, vol. 30, n° 3, p. 550-560.

[90] Kurzweil, J., 2000. *An introduction to digital communications*. John Wiley & Sons.

[91] Gallager, R. G., 1970. *Information theory and reliable communication*. Springer-Verlag, Courses and lectures - International Centre for Mechanical Sciences.

[92] Kramer, G., M. Gastpar and P. Gupta. September 2005. "Cooperative Strategies and Capacity Theorems for Relay Networks". *IEEE Transactions on Information Theory*, vol. 51, n° 9, p. 3037-3063.

[93] Luo, K., R. H. Gohary and H. Yanikomeroglu. October 2011. "On the generalization of decode-and-forward and compress-and-forward for Gaussian relay channels". In *IEEE Information Theory Workshop (ITW)*. p. 623-627.

[94] El Gamal, A. and S. Zahedi. May 2005. "Capacity of a class of relay channels with orthogonal components". *IEEE Transactions on Information Theory*, vol. 51, n° 5, p. 1815-1817.

[95] Liang, Y., V. V. Veeravalli and H. V. Poor. 2007. "Resource allocation for wireless fading relay channels: Max-min solution". *IEEE Transactions on Information Theory*, vol. 53, n° 10, p. 3432-3453.

[96] XIE, L.-L. and P. Kumar. 2005. "An achievable rate for the multiple-level relay channel". *IEEE Transactions on Information Theory*, vol. 51, n° 4, p. 1348-1358.

[97] Razaghi, P. and W. Yu. June 2007. "A Structured Generalization of Decode-and-Forward Strategies for Multiple-Relay Networks". In *IEEE International Symposium on Information Theory (ISIT)*. p. 271-275.

[98] Ghabeli, L. and M. R. Aref. June 2010. "On achievable rates for relay networks based on partial decode-and-forward". In *IEEE International Symposium on Information Theory (ISIT)*. p. 659-663.

[99] Zamani, M. and A. K. Khandani. July 2011. "On the maximum achievable rates in the decode-forward diamond channel". In *IEEE International Symposium on Information Theory (ISIT)*. p. 1534-1538.

[100] Nazaroglu, C., A. Ozgur and C. Fragouli. July 2011. "Wireless network simplification: The Gaussian N-relay diamond network". In *IEEE International Symposium on Information Theory (ISIT)*. p. 2472-2476.

[101] Song, Y. and N. Devroye. 2013. "Lattice codes for the Gaussian relay channel: Decode-and-Forward and Compress-and-Forward". *IEEE Transactions on Information Theory*, vol. 59, n° 8, p. 4927–4948.

[102] Host-Madsen, A. 2002. "On the capacity of wireless relaying". *IEEE Vehicular Technology Conference (VTC-Fall)*, vol. 3, p. 1333-1337.

[103] Wang, B., J. Zhang and A. Host-Madsen. January 2005. "On the capacity of MIMO relay channels". *IEEE Transactions on Information Theory*, vol. 51, n° 1, p. 29 -43.

[104] Simoens, S., O. Muñoz, J. Vidal and A. Del Coso. July 2008. "Capacity bounds for Gaussian MIMO relay channel with channel state information". In *The 9th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. p. 441 -445.

[105] Host-Madsen, A. and J. Zhang. June 2005. "Capacity bounds and power allocation for wireless relay channels". *IEEE Transactions on Information Theory*, vol. 51, n° 6, p. 2020 -2040.

[106] Alasti, M., B. Neekzad, J. Hui and R. Vannithamby. May 2010. "Quality of service in WiMAX and LTE networks [Topics in Wireless Communications]". *IEEE Communications Magazine*, vol. 48, n° 5, p. 104 -111.

[107] Dawy, Z. and H. Kamoun. 2004. "The general Gaussian relay channel: Analysis and insights". *The 5th International Conference on Source and Channel Coding (ITG SCC)*, vol. 181, p. 469.

[108] Gómez-Vilardebó, J., A. I. Pérez-Neira and M. Nájar. 2010. "Energy efficient communications over the AWGN relay channel". *IEEE Transactions on Wireless Communications*, vol. 9, n° 1, p. 32-37.

[109] Hoeher, P. A. and T. Wo. December 2011. "Superposition modulation: myths and facts". *IEEE Communications Magazine*, vol. 49, n° 12, p. 110 -116.

[110] Sendonaris, A., E. Erkip and B. Aazhang. November 2003. "User cooperation diversity. Part II. Implementation aspects and performance analysis". *IEEE Transactions on Communications*, vol. 51, n° 11, p. 1939 - 1948.

[111] Zhou, G., G. Bauch, J. Berkmann and W. Xu. June 2012. "Multi-layer rate splitting scheme for interference mitigation in tri-sectored wireless networks". In *IEEE International Conference on Communications (ICC)*.

[112] Ergen, M., 2009. *Mobile Broadband - Including WiMAX and LTE*. ed. 1st.

[113] Ghosh, A., R. Ratasuk, B. Mondal, N. Mangalvedhe and T. Thomas. 2010. "LTE-advanced: next-generation wireless broadband technology [Invited Paper]". *IEEE Wireless Communications*, vol. 17, n° 3, p. 10-22.

[114] Loa, K., C.-C. Wu, S.-T. Sheu, Y. Yuan, M. Chion, D. Huo and L. Xu. 2010. "IMT-advanced relay standards [WiMAX/LTE Update]". *IEEE Communications Magazine*, vol. 48, n° 8, p. 40-48.

[115] Damnjanovic, A., J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song and D. Malladi. 2011. "A survey on 3GPP heterogeneous networks". *IEEE Wireless Communications*, vol. 18, n° 3, p. 10-21.

[116] Wang, L.-C., W.-S. Su, J.-H. Huang, A. Chen and C.-J. Chang. April 2008. "Optimal Relay Location in Multi-Hop Cellular Systems". In *IEEE Wireless Communications and Networking Conference (WCNC)*. p. 1306 -1310.

[117] Huang, J.-H., L.-C. Wang, C.-J. Chang and W.-S. Su. 2010. "Design of optimal relay location in two-hop cellular systems". *Wireless Networks*, vol. 16, n° 8, p. 2179–2189.

[118] Sambale, K. and B. Walke. 2012. "Decode-and-forward relay placement for maximum cell spectral efficiency". In *The 18th European Wireless Conference*. p. 1-6.

[119] Lin, B., P.-H. Ho, L.-L. Xie and X. Shen. 2008. "Relay station placement in IEEE 802.16 j dual-relay MMR networks". In *IEEE International Conference on Communications (ICC)*. p. 3437–3441. IEEE.

[120] Cho, S., E. W. Jang and J. M. Cioffi. 2009. "Handover in multihop cellular networks". *IEEE Communications Magazine*, vol. 47, n° 7, p. 64-73.

[121] Wang, R., J. S. Thompson and H. Haas. 2010. "A novel time-domain sleep mode design for energy-efficient LTE". In *The 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*. p. 1-4.

[122] Salem, M., A. Adinoyi, M. Rahman, H. Yanikomeroglu, D. Falconer, Y.-D. Kim, E. Kim and Y.-C. Cheong. 2010. "An Overview of Radio Resource Management in Relay-Enhanced OFDMA-Based Networks". *IEEE Communications Surveys & Tutorials*, vol. 12, n° 3, p. 422-438.

[123] Kolios, P., V. Friderikos and K. Papadaki. 2010. "Load Balancing via Store-Carry and Forward Relaying in Cellular Networks". In *IEEE Global Telecommunications Conference (GLOBECOM)*. p. 1-6.

[124] Peters, S. W., A. Y. Panah, K. T. Truong and R. W. Heath. 2009. "Relay architectures for 3GPP LTE-Advanced". *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, n° 1, p. 618-787.

[125] Ghosh, A., J. Zhang, J. G. Andrews and R. Muhamed, 2010. *Fundamentals of LTE*. Pearson Education.

[126] Clark, A., P. J. Smith and D. P. Taylor. 2007. "Instantaneous capacity of OFDM on Rayleigh-fading channels". *IEEE Transactions on Information Theory*, vol. 53, n° 1, p. 355–361.

[127] Aliu, O., A. Imran, M. Imran and B. Evans. First Quarter 2013. "A Survey of Self Organisation in Future Cellular Networks". *IEEE Communications & Surveys Tutorials*, vol. 15, n° 1, p. 336-361.

[128] Islam, M., Z. Dziong, K. Sohraby, M. Daneshmand and R. Jana. Feb. 2012. "Capacity-optimal relay and base station placement in wireless networks". In *International Conference on Information Networking (ICOIN)*. p. 358-363.

[129] Zolotukhin, M., V. Hytonen, T. Hamalainen and A. Garnaev. 2012. "Optimal Relays Deployment for 802.16j Networks". In *Mobile Networks and Management*. p. 31-45. Springer Berlin Heidelberg.

[130] Minelli, M., M. Ma, M. Coupechoux, J.-M. Kelif, M. Sigelle and P. Godlewski. February 2014. "Optimal Relay Placement in Cellular Networks". *IEEE Transaction on Wireless Communications*, vol. 13, n° 2, p. 998-1009.

[131] Wu, G. and G. Feng. June 2012. "Energy-efficient relay deployment in next generation cellular networks". In *IEEE Int. Conf. on Comm. (ICC)*. p. 5757-5761.

[132] Peng, J., P. Hong and K. Xue. January 2015. "Energy-Aware Cellular Deployment Strategy Under Coverage Performance Constraints". *IEEE Transaction on Wireless Communications*, vol. 14, n° 1, p. 69-80.

[133] Minelli, M., M. Ma, M. Coupechoux and P. Godlewski. February 2014. "A geometrical approach for power optimization in relay-based cellular networks". In *20th National Conference on Communications (NCC)*. p. 1-5.

[134] Hamdi, A., M. El-Khamy and M. El-Sharkawy. April 2012. "Optimized dual relay deployment for LTE-Advanced cellular systems". In *IEEE Wireless Communications and Networking Conference (WCNC)*. p. 2869-2873.

[135] Kathrein Scala Division, March 2013. *Professional antennas & filters for mobile communications 700-3800Mhz: 65° triple band panel antenna - 800 10692V01*. Online. <http://www.kathrein-scala.com/catalog/700-3800_C35.pdf>.

[136] Andreev, S., P. Gonchukov, N. Himayat, Y. Koucheryavy and A. Turlikov. 2012. "Energy efficient communications for future broadband cellular networks". *Computer Communication*, vol. 35, p. 1662 - 1671.

[137] Auer, G., V. Giannini, I. Gódor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, C. Desset and O. Blume. 2011. "Cellular energy efficiency evaluation framework". In *IEEE Vehicular Technology Conference (VTC Spring)*. p. 1-6. IEEE.

[138] Polyanskiy, Y., V. H. Poor and S. Verdú. 2010. "Channel coding rate in the finite blocklength regime". *IEEE Transactions on Information Theory*, vol. 56, n° 5, p. 2307–2359.

[139] Øien, G. E., H. Holm and K. J. Hole. May 2004. "Impact of channel prediction on adaptive coded modulation performance in Rayleigh fading". *IEEE Transactions onVehicular Technology*, vol. 53, n° 3, p. 758-769.

[140] Ye, S., R. S. Blum and L. J. Cimini. November 2006. "Adaptive OFDM Systems With Imperfect Channel State Information". *IEEE Transactions on Wireless Communications*, vol. 5, n° 11, p. 3255-3265.

[141] Caire, G., N. Jindal and S. Shamai. November 2007. "On the Required Accuracy of Transmitter Channel State Information in Multiple Antenna Broadcast Channels". In *IEEE Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*. p. 287-291.

[142] Ramprashad, S. A., G. Caire and H. C. Papadopoulos. November 2009. "Cellular and Network MIMO architectures: MU-MIMO spectral efficiency and costs of channel state information". In *IEEE Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*. p. 1811-1818.

[143] Imai, H. and S. Hirakawa. 1977. "A new multilevel coding method using error-correcting codes". *IEEE Transactions on Information Theory*, vol. 23, n° 3, p. 371-377.

[144] Zhang, R. and L. Hanzo. 2011. "A unified treatment of superposition coding aided communications: Theory and practice". *IEEE Communications Surveys & Tutorials*, vol. 13, n° 3, p. 503-520.

[145] Tong, J., L. Ping and X. Ma. June 2006. "Superposition Coding with Peak-Power Limitation.". In *IEEE International Conference on Communications (ICC)*. p. 1718-1723.

[146] Yang, Z., L. Cai, Y. Luo and J. Pan. 2012. "Topology-aware modulation and error-correction coding for cooperative networks". *IEEE Journal on Selected Areas in Communications*, vol. 30, n° 2, p. 379-387.

[147] Vanka, S., S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou and M. Haenggi. 2012. "Superposition coding strategies: design and experimental evaluation". *IEEE Transactions on Wireless Communications*, vol. 11, n° 7, p. 2628-2639.

[148] Bhushan, N., J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel and S. Geirhofer. 2014. "Network densification: the dominant theme for wireless evolution into 5G". *IEEE Communications Magazine*, vol. 52, n° 2, p. 82-89.

[149] (CWTA), T. C. W. T. A. April 2013. Connecting Canadians: Wireless Antenna Towers Siting in Canada. [Online] Available: http://cwta.ca/wordpress/wp-content/uploads/2013/04/Tower-Booklet_E_2013.pdf.

[150] Mumtaz, S., D. Yang, V. Monteiro, C. Politis and J. Rodriguez. 2013. "Self organized energy efficient position aided relays in LTE-A". *Elsevier, Physical Communication*, vol. 7, n° 0, p. 30-43.

[151] Srinivasan, R. and S. Hamiti, 2010. *IEEE 802.16m System Description Document*. IEEE 802.16 Broadband Wireless Access Working Group.

[152] Gao, X., H. Xu and D. Ye. 2009. "Asymptotic Behavior of Tail Density for Sum of Correlated Lognormal Variables". *International Journal of Mathematics and Mathematical Sciences*, vol. 2009.

[153] Abu-Dayya, A. and N. Beaulieu. February 1994. "Outage probabilities in the presence of correlated lognormal interferers". *IEEE Transaction on Vehicular Technology*, vol. 43, n° 1, p. 164-173.

[154] Ng, D. W. K. and R. Schober. 2011. "Resource allocation and scheduling in multi-cell OFDMA systems with decode-and-forward relaying". *IEEE Transactions on Wireless Communications*, vol. 10, n° 7, p. 2246-2258.

[155] Sklar, B. July 1997. "Rayleigh fading channels in mobile digital communication systems .I. Characterization". *IEEE Communications Magazine*, vol. 35, n° 7, p. 90 -100.

[156] Younis, M. and K. Akkaya. 2008. "Strategies and techniques for node placement in wireless sensor networks: A survey". *Ad Hoc Networks*, vol. 6, n° 4, p. 621-655.

[157] Zuari, L., A. Conti and V. Tralli. September 2009. "Effects of relay position and power allocation in space-time coded cooperative wireless systems". In *The 6th International Symposium on Wireless Communication Systems (ISWCS)*. p. 700 -704.

[158] Vu, M. and A. Paulraj. September 2007. "On the capacity of MIMO wireless channels with dynamic CSIT". *IEEE Journal on Selected Areas in Communications*, vol. 25, n° 7, p. 1269 -1283.

[159] Cho, S., J. Kim and J.-H. Kim. 2008. "Relay assisted soft handover in multihop cellular networks". In *The 2nd International Conference on Ubiquitous Information Management and Communications*. p. 136-139. ACM.

[160] Hui, H., S. Zhu and G. Li. April 2009. "Distributed Power Allocation Schemes for Amplify-and-Forward Networks". In *IEEE Wireless Communications and Networking Conference (WCNC)*. p. 1 -6.

[161] Chen, M., S. Serbetli and A. Yener. February 2008. "Distributed power allocation strategies for parallel relay networks". *IEEE Transactions on Wireless Communications*, vol. 7, n° 2, p. 552-561.

[162] Sesia, S., I. Toufik and M. Baker, 2011. *LTE - The Umts Long Term Evolution: From Theory to Practice*. John Wiley & Sons.

[163] Alouini, M.-S. and A. J. Goldsmith. 1999. "Area spectral efficiency of cellular mobile radio systems". *IEEE Transactions on Vehicular Technology*, vol. 48, n° 4, p. 1047-1066.