ADVANCES IN COOPERATION TECHNIQUES FOR WIRELESS COMMUNICATION NETWORKS

YINGHAO GUO

NATIONAL UNIVERSITY OF SINGAPORE

2016

ADVANCES IN COOPERATION TECHNIQUES FOR WIRELESS COMMUNICATION NETWORKS

YINGHAO GUO

(*B. Eng. Dalian University of Technology, P. R. China*)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE

2016

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Yinghao Guo

05-08-2016

Acknowledgements

First and foremost, I want to express my sincere gratitude to my Ph.D. supervisor Prof. Rui Zhang for his guidance during the past four years for my study at National University of Singapore. I have benefited tremendously from his pioneering vision in the field of wireless communication, and rigorousness in research. Without his continual advice and encouragement, this thesis would certainly not be possible. The knowledge and experience I gained under his supervision will be an immeasurable wealth for me throughout my future career and life. Next, I want to thank Prof. Lingjie Duan from the Singapore University of Technology and Design. I have learnt substantially from his expertise on wireless networking and game theory. The helpful discussion with him has been a great source of inspiration for me on my research. His passion for research always motivates me to be an avid learner in and beyond my Ph.D. studies.

I thank all the current and past members in the Communication and Networking Lab, including Jie Xu, Hyungsik Ju, Yong Zeng, Suzhi Bi, Liang Liu, Shixin Luo, Xun Zhou, Mohammad Reza, Katayoun Rahbar, Seunghyun Lee, Shuowen Zhang, Reuben Stephen, Hong Xing, Pierrick Abiven, Seungmin Kang and many others. I also want to thank my good friends in China and Singapore for all of their supports and companions.

Last but not the least, I wish to express my sincere gratitude to my families: my parents, Zhonghua Guo and Ying Sun, and my grandmother, Zhenying Hou for their encouragement, support and unconditional love. I am happy and grateful that I am always accompanied by them on this journey.

Table of Contents

Summary

Cooperative communication is an effective technique to improve the reliability as well as energy and spectrum efficiency of wireless communication systems. However, the recent development of other pertinent technologies such as short-range communication (SRC), local device caching, and smart power grid have brought new challenges to the optimal design of cooperative communication, which are not yet solved. This thesis is thus devoted to the investigation of new cooperation techniques under the setups of energy-aware cooperative relaying, cooperative local caching with heterogeneous file preferences, and joint energy and spectrum cooperation among the base stations (BSs) powered by smart grid, respectively.

First, this thesis investigates the energy saving performance of the mobile terminals (MTs) via their uplink cooperative relaying in the wireless cellular network under heterogeneous battery levels. In order to address the incentive issues in such cooperation, a novel pricing-based mechanism is proposed to incentivize the uplink cooperative relaying. We consider two scenarios where the MTs in the network belong to the same entity or different entities, respectively. For the former scenario with full cooperation, the problem is formulated as a deterministic relay selection problem and a simple threshold structure is shown for the optimal solution. For the latter case with partial cooperation, we investigate the pricing and load sharing problem under the uncertainties of the channel conditions and battery levels. An efficient algorithm based on alternating optimization is proposed for solving the problem.

Next, we study the performance trade-offs among different MTs and the effects of their selfish behaviours in cooperative local caching under heterogeneous

file preferences. We practically categorize the MTs into different interest groups according to the MTs' preferences. We define each group's utility as the probability of successfully discovering the file in the neighbouring MTs. By modelling MTs' mobilities in different groups as independent homogeneous Poisson point processes (HPPPs), the MTs' utilities are derived in closed-form. We first consider the fully cooperative case where a centralizer helps all groups to make the caching decisions. The problem is formulated as a weighted-sum utility maximization problem, through which the maximum utility trade-offs of different groups are characterized. We also consider the case of partial cooperation with limited inter-group file-sharing to demonstrate the effects of selfish behaviours in the cooperation, as well as the case of no cooperation to show the gain of inter-group file sharing. The optimal caching distributions for these two cases are also derived.

Finally, we study the BS-side joint energy and spectrum cooperation for reducing the energy costs of the wireless systems exploiting the renewable energy. We propose a joint energy and spectrum cooperation scheme between different cellular systems to address the incentive issues. When the systems are fully cooperative (e.g., belonging to the same operator), we formulate the cooperation design as a convex optimization problem to minimize their weighted-sum energy cost and obtain the optimal solution in closed-form. We also study the partially cooperative scenario where the two systems have their own interests and will cooperate only if their energy costs can be reduced at the same time. The condition for partial cooperation is shown, based on which we also develop a distributed algorithm for the two systems to gradually and simultaneously reduce their costs from the non-cooperative benchmark to the Pareto optimum.

List of Tables

List of Figures

List of Abbreviations

AC	Alternating Current
AWGN	Additive White Gaussian Noise
BS	Base Station
CoMP	Coordinated Multiple Point
CSCG	Cyclic Symmetric Complex Gaussian
CRAN	Cloud Radio Access Network
CT	Cooperative Transmission
DC	Direct Current
DT	Direct Transmission
D2D	Device-to-Device
EPC	Evolved Packet Core
FemtoCaching	Femto-cell Caching
HPPP	Homogeneous Poisson Point Process
i.i.d.	Independent and Identically Distributed
ISM	Industrial, Scientific and Medical
KKT	Karush-Kuhn-Tucker
LP	Linear Programming
LTE	Long Term Evolution
MAC	Mediem Access Control
MT	Mobile Terminal
M2M	Machine to Machine
MIMO	Multiple-input Multiple-output
OFDMA	Orthogonal Frequency Division Multiplexing Access

List of Abbreviations

PA	Power Amplifier
PMF	Probability Mass Function
ProSe	Proximity Service
PSD	Power Spectral Density
PU	Primary User
QoS	Quality of Service
RV	Random Variable
SC-FDMA	Single-carrier Frequency Division Multiple Access
SNR	Signal to Noise Ratio
SRC	Short-range Communications
SU	Secondary User
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
W/O	Without
3GPP	3rd Generation Partnership Project
5G	5th Generation

List of Symbols

Throughout this thesis, scalars are denoted by lower-case letters, and vectors are denoted by bold-face lower-case letters. Sets are denoted by calligraphic letters.

$\mathbb{E}[\cdot]$	Statistical Expectation Opeartor
$\mathbb{P}(\cdot)$	Probability Operator
$\lvert\mathcal{X}\rvert$	Cardinality of Set $\mathcal{X}$
$\mathcal{X} \setminus \mathcal{Y}$	Relative Complement of $\mathcal{Y}$ in $\mathcal{X}$ as $\{x \in \mathcal{X} \,\vert\, x \notin \mathcal{Y}\}$
$\emptyset$	Empty Set
$\sim$	"Distributed as"
$\triangleq$	"Defined as"

Chapter 1

Introduction

The wireless data traffic has experienced an exponential increase during the last decade due to the advancement of smartphones and mobile Internet. It is predicted by Cisco that the global data volume will exceed 24.3 exabytes (i.e., 24.3×10^{18} bytes) per month by 2019 [1]. In order to cope with such a challenge, many promising communication techniques have been proposed and among them, *cooperative communication* [2] is particularly appealing. The key idea of cooperative communication is to exploit different forms of diversities (in channel condition, energy availability, rate requirement, etc.) among different mobile terminals (MTs) or base stations (BSs) in the wireless network, which is also known as *cooperative diversity* in the literature. For example, on the BS side, different BSs may serve different numbers of MTs or have different channel conditions with their served MTs; while on the MT side, different MTs may have different battery levels, channel conditions, rate requirements, etc. By exploiting these complementarities, the transmission reliability, as well as the spectrum and energy efficiency of wireless communication systems can be improved. Due to the above advantages, cooperative communication has been not only the focus of the research community of wireless communication, but also included in various wireless standards, such as the cooperative relaying [3] and coordinated multiple point (CoMP) transmission [4] in the new releases of the Long Term Evolution (LTE) in [5] and [6], respectively.

Moreover, there are other pertinent technologies being integrated into the cooperative communication system, which bring new opportunities as well as challenges to its optimal design and practical implementation. Examples of such new technologies include the short-range communications (SRC) among the MTs,

local caching in the BSs and MTs, smart grid and renewable energy powered BSs, etc. This thesis thus aims to advance the studies of cooperative communication techniques from these new perspectives. In particular, we make an in-depth investigation of the advanced cooperation techniques in wireless communication under the three setups of energy-aware cooperative relaying, cooperative local caching with heterogeneous file preferences, and joint energy and spectrum cooperation among BSs powered by smart grid, respectively.

The organization of the rest of this chapter is as follows. We first provide an overview of the emerging technologies in wireless communication in Section 1.1. We then introduce the challenges as well as motivations for the design of new cooperation techniques under these setups in Section 1.2. Finally, we present an overview and the main contributions of this thesis in Section 1.3.

1.1 Emerging Technologies in Wireless Communication

In this section, we first introduce some emerging technologies in wireless communication. Their interplays with cooperative communication and the resulting new challenges will also be discussed.

1.1.1 Short-range Communications

SRC is a class of technologies designed for peer-to-peer communication between MTs in close proximity to complement the long-range wireless cellular communications enabled by the BSs [7]. One example of wireless cellular network overlaid with SRC is illustrated in Fig. 1.1. With SRC, MTs in the wireless cellular network can communicate directly with each other (denoted by the solid red arrow) without traversing the BSs (denoted by the green dashed arrow). During the past decades, various SRC technologies have been introduced in the wireless industry, a comparison of which is provided in Table 1.1. Among these technologies,

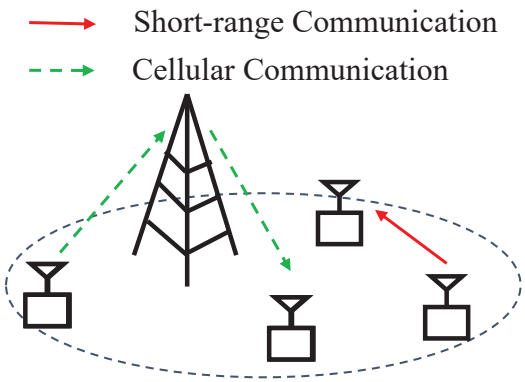


Figure 1.1: Wireless cellular network overlaid with SRC for the MTs to communicate in short range.

Table 1.1: Comparison between different SRC technologies, source [10].

	LTE D2D	Wi-Fi Direct	NFC	ZigBee	Bluetooth
Standardization	3GPP LTE-A	IEEE 802.11	ISO 13157	IEEE 802.15.4	Bluetooth SIG
Working Frequency	LTE Licensed Bands	2.4/5 GHz	13.56 MHz	869/915 MHz 2.4 GHz	2.4 GHz
Max Transmission Distance	1000 m	200 m	0.2 m	10-100 m	10-100 m
Max Data Rate	1 Gbps	500 Mbps	424 kbps	250 kbps	24 Mbps
Applications	Traffic Off-loading, Public Safety, Context Sharing, Local Advertising, Cellular Relay	Context Sharing, Group Gaming, Device Connection	Contact-less Payment, Bluetooth and Wi-Fi Connection	Home Entertainment & Control, Environmental Monitoring	Object Exchange, Peripherals Connection
Infrastructure	in licensed bands	in un-licensed bands			

Bluetooth [8] and Zigbee [9] are the most widely adopted standards in the wireless cellular network. Bluetooth operates on the unlicensed 2.4 GHz industrial, scientific, and medical (ISM) band. It is primarily used for the connection of devices, such as the headsets for the cell phone, audio streaming for the speakers, etc. Zigbee also operates on the ISM band. However, it is mainly designed for the deployment of the wireless mesh network with low cost, low power, and low communication speed.

More recently, a new technology named LTE Device-to-Device (D2D) communication [11, 12] has been proposed for the future fifth-generation (5G) wireless network [13]. The main features of D2D communication are also shown in Table 1.1. Different from the other SRC technologies, D2D communication

can operate on the licensed band of the wireless cellular network under the coordination of the BSs [12]. In this way, the mutual interferences between the MTs with D2D communication and those with the cellular communications can be reduced. Moreover, unlike other SRC technologies that are based on "best effort" communication, D2D communication is required to satisfy a set of quality of service (QoS) rules for the MTs mandated by the telecom operators [14]. From the table, we can also see that D2D communication provides much higher data rate and larger transmission range than the other standards. This will be particularly suitable for various functionalities in the current wireless network with demands for high data rates, such as cooperative relaying between the MTs or data traffic off-loading from the BSs. Finally, because the integration of the D2D communication into the wireless cellular network does not require major changes to the infrastructure, D2D communication can offer an economical and graceful upgrade for the current wireless systems [11].

Due to the above advantages of D2D communication, in the new release of the LTE-Advanced, standards for the D2D communication have been specified in full details as Proximity Service (ProSe) [14], where different functionalities such as cooperative relaying, peer discovery, etc., have been defined. An initial trial of Qualcomm's LTE-Direct [15] has also been done by Deutsch Telecom on the time division duplex (TDD) band of LTE network [16]. Applications and services made possible by D2D communication can be manifold. For example, a local restaurant can push advertisements to the potential consumers in its vicinity. Emergency messages can also be broadcast to the MTs when the wireless network is out of service in the event of natural disasters [15].

1.1.2 Local Caching

In addition to the dramatic increase in the volume during recent years, the current wireless data traffic is also becoming more and more concentrated to hotspots [17] and this may lead to the congestion of the wireless network. In

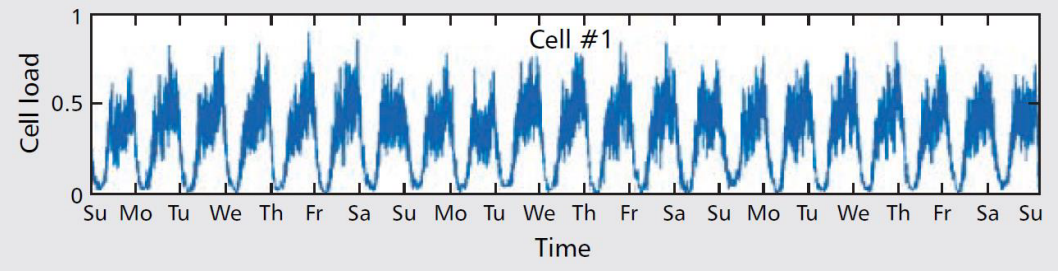


Figure 1.2: Typical fluctuations of wireless data traffic during a month [21].

order to avoid congestions, various technologies have been proposed to increase the wireless network's capacity, such as the densification of the wireless network [18], millimeter-wave communication [19], massive multiple-input multiple-output (MIMO) [20], etc. However, they do not fully exploit the characteristics of the wireless data traffic: wireless data traffic fluctuates significantly during a day, an example of which is shown in Fig. 1.2. Also, video traffic accounts for about 70% of the total wireless data traffic [1], while a large proportion of the video traffic is attributed to only a few popular files during each day. Motivated by this, cooperative local caching [22,23] has been proposed recently as an effective approach to overcome the future throughput bottleneck of the wireless network. Due to its various advantages, it has also become a promising candidate for the future 5G technologies [24].

Cooperative local caching trades off the precious bandwidth and backhaul resources with the data storage at the edge of the wireless network (e.g., evolved packet core (EPC) network, BSs or MTs), which is highly under-utilized and the price of which is currently becoming lower. Popular files can be first pro-actively cached in the EPC, BSs or MTs. Then, the cached files are shared locally to the MTs during their file requests such that they do not need to resort to the remote servers fetching for the files. As such, caching in the core network can reduce the latency for the MTs' file downloading. Caching in the BSs and MTs can also alleviate the traffic congestion in the core network and the backhaul during the peak hours [22]. Therefore, with cooperative local caching, the data traffic spikes similar to those in

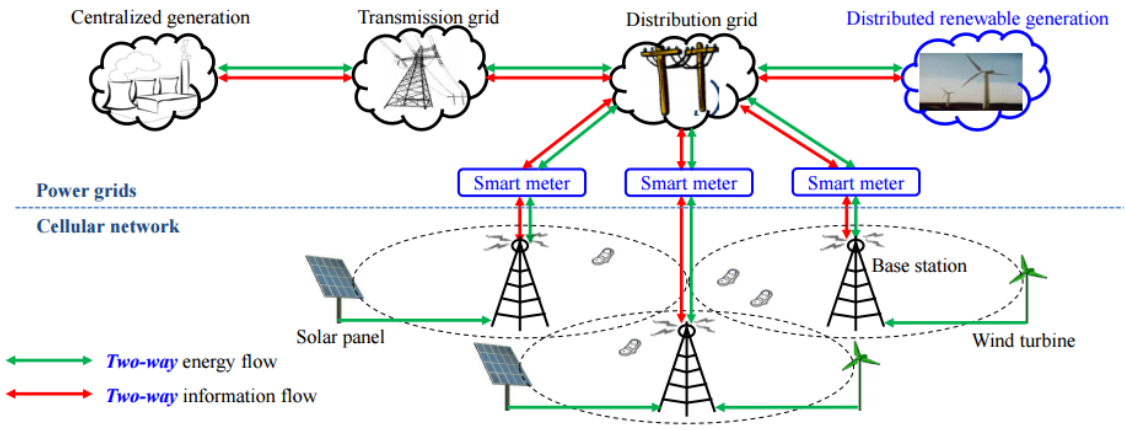


Figure 1.3: Architecture of wireless cellular network with BSs powered by smart grids and renewable energy [31].

Fig. 1.2 can be flattened and the robustness of the wireless network can be enhanced.

1.1.3 Wireless Communication Powered by Smart Grid

The electricity cost accounts for at least 40% of the total operation expenditure for the current wireless network [25]. Hence, reducing the electricity cost of the wireless network is crucially important for the wireless operators. With the growing public and governmental concerns over the global warming effects, telecom companies are paying more attention to the environmental impacts of their significant energy consumption [26]. Also, because of the governments' extensive subsidies to promote the use of renewable energy and its mass production [27], the price of renewable energy is becoming lower [28]. As a result, the integration of the renewable energy sources into the wireless network has attracted a lot of attention from both the academia [29] and the industry [30].

Recently, the fast developing smart-grid technologies, such as smart metering, renewable energy integration, two-way energy trading, etc., are revolutionizing the industry of electricity transmission [32, 33]. One of the most prominent features of the smart grid technology is the bi-directional energy flow between the power grid and its users, which enables the integration of the distributed energy sources

in the power grid. One such example of the wireless systems with renewable energy sources integrated into smart grid is shown in Fig. 1.3. Under the bi-directional energy flow (denoted by the green arrow), excess energy generated by the renewable energy sources in the wireless cellular network (e.g., solar panel or wind turbine) can be transferred back to the power grid. Conversely, the wireless cellular network can also obtain electricity from the power grid when it is short of electricity. Therefore, bi-directional energy flow offers higher flexibility to the energy management of the power grid, which leaves room for the optimization of the energy resources. Furthermore, in addition to the bi-directional energy flow, the bi-directional information flow (denoted by the red arrow) between the smart grid and the cellular systems also enables communications between different users of the power grid (e.g., different BSs in the cellular network). In this way, the buying and selling of energy with grids or between the users can be properly metered and managed [34], which makes the two-way energy trading possible.

1.2 Motivations

In this section, we present the new challenges in the design of cooperation techniques for wireless communication systems integrated with the above-mentioned new technologies. Specifically, we will discuss three new issues under the setups of energy-aware cooperative relaying, cooperative local caching with heterogeneous file preferences, and cooperation among cellular BSs powered by the smart grid, respectively.

1.2.1 Heterogeneous Battery Levels and Incentive Issue in Uplink Cooperative Relaying

First, consider cooperative relaying for the MT's energy saving in the wireless cellular network. With the fast development of the smartphones and the multimedia applications, today's MTs consume a lot more energy than before. MTs need to be

charged more frequently and this has become the biggest customer complaint about smartphones [35]. Furthermore, the communication modules account for a large proportion of the MTs' energy consumption, for either the MTs from the earlier 2G and 3G era [36] or the modern 4G mobile phones [37]. Hence, in order to reduce the MT's energy consumption, cooperative relaying on the MT side has been proposed as an effective solution. However, the heterogeneous battery levels across different MTs have been largely ignored. For MTs within a wireless cellular network, some MTs are low in battery level, while others are higher. Clearly, it would be desirable for the MTs with high battery levels to help those in low battery levels in the cooperative relaying. However, in spite of the advantages of cooperative relaying for energy saving, the incentive issues in such cooperation are still less understood. In most of the works in the literature, it is generally assumed that different MTs cooperate with each other without self-interests. However, for the MTs in the wireless cellular network, this can hardly be true since different MTs belong to different users, each with his own interest. Hence, in order to practically implement cooperative relaying between the MTs, incentive design must be developed.

1.2.2 Heterogeneous File Preferences in Cooperative Local Caching

Next, consider cooperative local caching in the wireless cellular network. Cooperative local caching has been proposed as an effective approach to exploit the under-utilized memory in the MTs, by which files are first pre-cached in the local MTs, and then cooperatively shared between the MTs upon their file requests. However, prior works on cooperative local caching largely assume that different MTs have the same preference over the files, and they cooperate with each other without self-interests. In this thesis, we revise this assumption and study the cooperative local caching under heterogeneous file preferences. Given such heterogeneity, conflict of interests between different MTs arises in the cooperative caching. For example, consider a certain file that is not popular in one group of MTs but popular in the

other groups: if the former group caches this file, the performances of the other groups will increase because of the file's popularity. However, this will reduce the memory for the former group to cache popular files and thus deteriorate its performance. It thus further motivates us to study the performance trade-offs between different MTs and the impacts of their selfish behaviours in cooperative local caching.

1.2.3 Incentive Issue in Cooperative Wireless Networks Exploiting Renewable Energy

Last but not least, consider the wireless cellular network with BSs powered by both conventional and renewable energy in the smart grid. With the fast development of the wireless technologies, there has also been a tremendous increase in the energy consumption and the resulted energy cost constitutes a significant part of the wireless systems' total operational costs [25]. In order to reduce the energy costs, more and more cellular operators are considering to power their wireless systems with renewable energy sources. However, renewable energy is intermittent in nature and brings uncertainties to the operation of the wireless network. To help mitigate such fluctuations and uncertainties from renewable energy, *energy cooperation* by sharing some systems' excessive energy to the others with insufficient energy has been proposed [38, 39]. However, one key problem that remains unaddressed for this technique is how to motivate different systems to cooperate. For the practical implementation of energy cooperation between the wireless systems with self-interests, the incentive issues must also be addressed.

1.3 Overview of the Thesis and Major Contributions

1.3.1 Overview of the Thesis

Motivated by the above discussions, in this thesis, we investigate three problems related to the new cooperation techniques in the wireless network. The rest of the thesis is organized as follows.

In Chapter 2, we consider the cooperative relaying for the energy saving in the uplink transmission of wireless cellular network. We propose a pricing-based mechanism design to address the incentive issues in the cooperative relaying. We consider two scenarios where the MTs in the network belong to the same entity or different entities, respectively. For the former scenario with full cooperation, the energy-saving problem is formulated as a deterministic relay selection problem and the optimal solution is presented. For the latter case with partial cooperation, the pricing and load sharing problem is formulated as an optimization problem under the uncertainties of the channel conditions and battery levels. An efficient algorithm based on alternating optimization is proposed to solve the problem. Also, the energy-saving performance of the cooperative relaying is demonstrated under a realistic simulation setup with the evolution of the system.

In Chapter 3, we study cooperative local caching under heterogeneous file preferences. We define the utility of each group of MTs as the probability of successfully discovering the file within the range of SRC. In order to characterize the utility trade-offs between different MTs, under the case of full cooperation, we formulate the cooperation problem as a weighted-sum utility maximization problem, in which the caching decisions of all the MTs are jointly optimized. We also consider the case of partial cooperation with limited inter-group file sharing to demonstrate the effects of selfish behaviours on the cooperation, and the case of no cooperation to show the gain brought by inter-group file-sharing. The optimal caching distributions for these two cases are also derived.

In Chapter 4, we investigate the cooperation between wireless systems powered by both renewable and conventional energy. We propose the model of joint energy and spectrum cooperation to incentivize the cooperation between the wireless systems and reduce their energy costs. A practical modelling is provided for the renewable energy availability, spectrum cooperation, and the conventional as well as the renewable energy costs. Under the scenario with full cooperation, the problem is formulated as a convex optimization problem, which minimizes the weighted-sum cost of the two systems. The closed-form optimal solution is derived, and an efficient algorithm based on ellipsoid method is proposed to obtain the optimal solution. On the other hand, in the case with partial cooperation, where cooperation is only possible when it brings mutual benefits to all the systems, we propose a distributed algorithm for different systems to achieve the Pareto optimum autonomously and simultaneously from the non-cooperation benchmark.

Last, Chapter 5 concludes this thesis and discusses the possible future work.

1.3.2 Major Contributions

The major contributions of this thesis are summarized as follows.

1. Characterization of the Complete Performance Trade-offs among MTs or BSs under Full Cooperation

Under full cooperation, all the MTs or BSs belong to the same entity and they can fully cooperate to achieve the common performance objective for all the MTs or BSs. Centralized decisions can be made by jointly optimizing the decisions for all the MTs or BSs. However, even under full cooperation, there is still conflict of interests for the performance improvement among different MTs or BSs. In this thesis, we aim to investigate the fundamental performance trade-offs between MTs or BSs under the following setups.

- *Utility Trade-offs between MTs in Cooperative Local Caching with Heterogeneous File Preference*: In Chapter 3, for cooperative local caching,

we first define the MTs' utility as the probability of successfully discovering the file within the range of SRC. Under heterogeneous file preference, different MTs have different interests over the files, which results in conflict of interests in the file caching. In order to characterize such utility trade-offs, we formulate and solve a weighted-sum utility maximization problem. Efficient algorithms based on coordinate descent method are proposed to solve this problem.

- *Cost Trade-offs between BSs in Joint Energy and Spectrum Cooperation*: In Chapter 4, under the model of joint energy and spectrum cooperation, different BSs have conflict of interests in the sharing of energy and spectrum. In order to characterize the trade-offs in the energy costs, we formulate the cooperation design as a convex optimization problem, which minimizes the weighted-sum energy cost of the two systems. We obtain the optimal solution to this problem in closed-form based on the Karush-Kuhn-Tucker (KKT) conditions. An efficient algorithm based on the ellipsoid method is proposed for characterizing the Pareto boundary of different BSs' energy costs.

2. Novel Design of Incentive Mechanism and Cooperation Model Addressing the Incentive Issues under Partial Cooperation

Under partial cooperation, different MTs or BSs belong to different entities and have their own interests. Consequently, no centralized decisions can be made due to their individual interests. But the BSs or MTs can still cooperate even without centralized coordination, provided that the cooperation can result in performance gain for them simultaneously. Therefore, incentive issues are critical for the practical design and implementation of cooperative techniques for partially cooperative MTs or BSs.

- *Pricing-based Mechanism Design for Incentivizing Cooperative Relaying*: In Chapter 2, we propose a novel pricing-based mechanism for motivating the cooperative relaying between the MTs. In the incentive mechanism, a price is paid to the helping MT in exchange for the cooperative relaying of the source

MT's data. Under the uncertainties of the helping MTs' channel conditions and battery levels, the pricing and load sharing problem is formulated as an optimization problem and efficient algorithms based on dichotomous search and alternating optimization are proposed to obtain the solutions. Our pricing-based mechanism design provides an elegant solution to the energy saving of the MTs with cooperative relaying under a realistic setup and also is easy to implement.

- *Joint Energy and Spectrum Cooperation for Incentivizing BS-side Cooperation*: In Chapter 4, we propose a novel joint energy and spectrum cooperation model, in which complementarity is possible to be found in the energy and spectrum resources between different systems. We specify the conditions for feasible partial cooperation between the BSs in closed-form. We then propose a distributed algorithm for the BSs to reduce the energy costs gradually and simultaneously from the non-cooperative benchmark to the Pareto optimum. Our proposed joint energy and spectrum cooperation model nicely resolves the incentive issues in the cooperation between the wireless systems exploiting renewable energy and significantly reduces their energy costs.

3. Performance Comparison between Different Scenarios of Cooperation

Under no cooperation, unlike the case of full or partial cooperation, MTs or BSs cannot cooperate. Hence, this scenario serves as the benchmark for quantifying the performance gain by the full or partial cooperation. In this thesis, we perform extensive numerical simulations for validating the performance gain of various cooperation techniques compared with no cooperation. For the cooperation on the MT side, network-level simulation under the evolution of the system validates the performance improvement more convincingly. For the cooperative relaying in Chapter 2 and the cooperative local caching in Chapter 3, we both perform network-level simulation under practical simulation setups. The simulation results show that the proposed cooperative relaying scheme in Chapter 2 can effectively

prolong the battery life. It is also shown that cooperative local caching in Chapter 3 can achieve performance improvement by increasing the probability of successful file sharing among the MTs. On the other hand, for the cooperation on the BS side in Chapter 4, we also confirm with numerical simulations that the wireless systems with joint energy and spectrum cooperation also achieve significant cost reduction compared with no cooperation.

Chapter 2

MT-side Cooperative Relaying

2.1 Introduction

In this chapter, we study the cooperative relaying in the uplink transmission of the MTs with the consideration of heterogeneous battery levels. In order to address the incentive issues, we propose a pricing-based mechanism to incentivize the uplink cooperative relaying for the energy saving of MTs. We first consider the ideal case of full cooperation under complete information, where private information of the helping MTs, such as battery levels and channel conditions are known by the source MTs. For this scenario as the benchmark, the problem is formulated as a relay selection problem. It is shown that the optimal solution for the pricing and load sharing follows a simple threshold structure and is easy to implement. Then, for the practical case of partial cooperation with incomplete information, the MTs need to cooperate under the uncertainties of the helping MTs' channel and battery conditions. We formulate the source MT's pricing and load sharing as an optimization problem. Efficient algorithms based on dichotomous search and alternating optimization are proposed to solve the problem for the cases of splittable and non-splittable data at the source MTs, respectively. Finally, extensive numerical results are provided to validate our proposed scheme. The result shows that cooperative relaying under the pricing-based mechanism can significantly decrease both the communication and battery outages for the MTs, and can also increase the average battery levels during the MTs' operations.

The rest of this chapter is organized as follows. Section 2.2 presents the literature review. Section 2.3 introduces the system model with the cooperative relaying for MTs' energy saving. Section 2.4 discusses our proposed pricing-based mechanism under complete information as the performance benchmark. Section 2.5 studies the general case of cooperative relaying under incomplete information. Section 2.6 presents numerical examples to validate the performance of our proposed scheme. Finally, Section 2.7 concludes this chapter.

2.2 Literature Review

2.2.1 Energy Saving for MTs

Due to the increasing energy consumption of the MTs and the limited advancement in the battery technologies during recent years, there has been increasing attention to the MT-side energy saving in the literature [40–44]. In particular, [40] studies the optimal modulation scheme to minimize the total energy consumption for transmitting a data packet of a given size. Both uncoded and coded systems are considered for the modulation optimization. [41] studies the optimal power control problem for the minimization of the average MT energy consumption in the multi-cell time-division-multiple-access (TDMA) system. In [42], the authors consider the energy saving of the MTs by leveraging the spare capacity at the BSs in cellular networks. The optimal design is obtained by solving the optimization problems for the scenarios of real-time data traffic and data files transmission, respectively. Recently, [43, 44] show that there is, in general, a trade-off between minimizing the energy consumption at the BSs and that at the MTs for meeting given QoS requirements of the MTs.

2.2.2 Cooperative Relaying for MTs' Energy Saving

Among the techniques for energy saving of the MTs, cooperative relaying [45] is particularly appealing and has been investigated in the literature on wireless sensor

and cellular networks [46–49]. In particular, [46] studies the optimal timer-based relay selection scheme for the minimization of the sum energy consumption and maximization of the network lifetime. [47] proposes a space-time coding scheme for the MTs to cooperatively transmit to the BS under given outage and capacity requirements such that the total transmit energy is minimized. [48] considers the minimization of energy consumption under QoS constraint with cooperative spectrum sharing in the cognitive radio network. [49] considers extending the lifetime of the machine-to-machine (M2M) communication network by cooperative multiple access control (MAC) protocols with cooperative retransmission schemes.

2.2.3 Incentive Issues in Cooperative Relaying

Although cooperative relaying techniques for the energy saving of the MTs have been proposed in the literature, the works before generally assume that all the MTs can cooperate without self-interests. However, this is difficult to hold in reality due to the lack of incentives that motivate the MTs to cooperate [50]. From this perspective, prior works [51–55] have also considered various incentive mechanisms to motivate cooperative relaying in wireless communication systems. Specifically, the idea of *virtual currency* for incentivizing the cooperation between self-organized entities is first investigated in [51] under the setup of wireless sensor network. [52] proposes a distributed game-theoretical framework over multiuser cooperative relaying networks to achieve optimal relay selection and power allocation. A two-stage Stackelberg game is formulated to consider the interests of the source and relay, where the source node is modelled as a buyer and the relay nodes are modelled as sellers for providing relay for the source. [53] studies the dynamic bargaining-based cooperative spectrum sharing between a primary user (PU) and a secondary user (SU), where the PU shares spectrum to the SU and the SU helps relay the signal of the PU in return. Different from the above approachs based on money returns, [54] proposes a so-called *reputation system* under the cooperation between the nodes in the long run. Based on auction theory, indirect incentives are

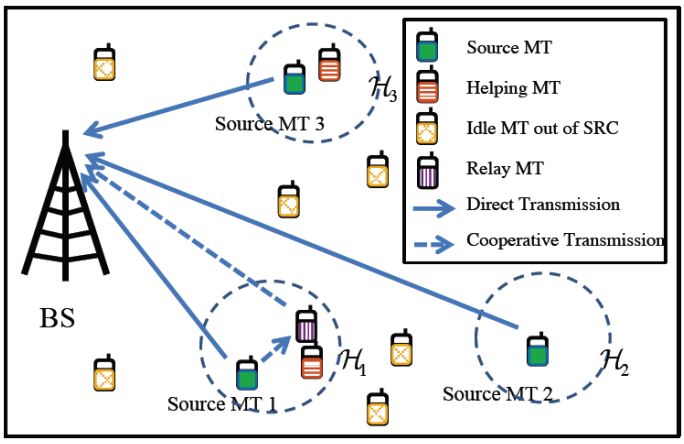


Figure 2.1: System model for direct and cooperative data transmission.

provided for stimulating node cooperation in green wireless networks. [55] proposes a business model for the power-trading between MTs in the cooperative networking problem based on auction theory.

Unlike the above works in the literature, we study the energy saving of the MTs with the new consideration of the battery levels of the MTs in cooperative relaying in the wireless cellular network. Moreover, our proposed pricing-based mechanism is also different from the existing approaches for motivating the cooperation in the literature.

2.3 System Model and Energy-saving Cooperative Relaying

As shown in Fig. 2.1, we consider the uplink data transmission within one single cell of a cellular network.[1] Different roles of the MTs will be introduced later in the chapter. Within the cell, there is one BS serving K MTs denoted by the set $\mathcal{K} = \{1, 2, \cdots, K\}$. We assume that the locations of the MTs follow

[1]Our results can be extended to the case of multiple cells by applying our results to each cell independently.

a two-dimensional independent HPPP with average spatial density $\bar{\lambda}$[2] [56]. We consider that the MTs within the cell initiate their data traffic independently with probability ρ. Then, according to the Marking Theorem [57], these *source MTs* (i.e., MTs initiating data traffic) also form an HPPP with density $\rho\bar{\lambda}$ and the remaining *idle MTs* form another HPPP with density $(1-\rho)\bar{\lambda}$. We denote these sets of source MTs and idle MTs as $\mathcal{K}_S$ and $\mathcal{K}_I$, respectively, such that $\mathcal{K}_S \cup \mathcal{K}_I = \mathcal{K}$ and $\mathcal{K}_S \cap \mathcal{K}_I = \emptyset$.

We consider the uplink data transmission of all MTs and assume the narrow-band block fading channel model. To support multiple MTs, orthogonal data transmission is assumed, e.g., by applying Single-carrier Frequency Division Multiple Access (SC-FDMA) in the uplink of the LTE system [58]. We denote the complex baseband channel coefficient from MT $k \in \mathcal{K}$ to the BS as h_k, which follows a simplified channel model incorporating the large-scale power attenuation with loss exponent $\alpha > 2$ and the small-scale Rayleigh fading. More specifically, we denote r_k as the distance between MT $k \in \mathcal{K}$ and the BS, and r_0 as a reference distance. Then, the channel coefficient h_k is expressed as

$$h_k = \begin{cases} \bar{h}_k \sqrt{G_0 \left(\frac{r_k}{r_0}\right)^{-\alpha}}, & r_k > r_0 \\ \bar{h}_k \sqrt{G_0}, & \text{otherwise} \end{cases}, \ k \in \mathcal{K}, \tag{2.1}$$

where $\bar{h}_k \sim \mathcal{CN}(0,1),\ k \in \mathcal{K}$ is an independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian (CSCG) random variable (RV) with zero mean and unit variance modelling the small-scale Rayleigh fading, and G_0 is the constant path-loss between the MT and the BS at the reference distance r_0. Therefore, the channel power gain between the MT k and the BS is

$$g_k = |h_k|^2 = \eta_k G_k,\ k \in \mathcal{K}. \tag{2.2}$$

[2]For the spatial user density $\bar{\lambda}$, it can be readily obtained by dividing the total number of MTs within the cell over the total area of the cell. The number of the MTs can be estimated by the history data of the cell or by real-time monitoring.

Here, we denote $\eta_k \sim \exp(1)$ as an exponential RV with unit mean modelling the power envelope of the Rayleigh fading and

$$G_k = \begin{cases} G_0 \left(\frac{r_k}{r_0}\right)^{-\alpha}, & r_k > r_0 \\ G_0, & \text{otherwise} \end{cases}, \ k \in \mathcal{K} \tag{2.3}$$

as the power attenuation between the BS and the MT k at the distance of r_k.

For simplicity, we consider a time-slotted system, where symbols for the message are transmitted in each time slot. For convenience, the number of symbols transmitted per time slot is normalized to unity. If MT $k \in \mathcal{K}$ initiates its data traffic, a message from the set $\{1, 2, \cdots, 2^{D_k}\}$ is sent, where D_k is the transmitted rate in bits per symbol. Without loss of generality, we also normalize the duration of one symbol time to unity such that the two terms energy and power can be used interchangeably in the chapter. Then, if the achievable data rate D_k is normalized by the available bandwidth at the MT, for given transmission energy per symbol E_k, the (normalized) achievable data rate for MT $k \in \mathcal{K}$ in bits/sec/Hz (bps/Hz) is

$$D_k = \log_2 \left(1 + \frac{g_k E_k}{\sigma^2}\right), \tag{2.4}$$

where σ^2 denotes the power of the noise at the receiver of BS.

For source MT $k \in \mathcal{K}_S$, in order to accomplish the uplink transmission at (normalized) data rate D_k, it can choose between the following two transmission modes.

Direct Transmission Mode (DT Mode)

In this mode, the source MT transmits to the BS directly with normalized data rate D_k. Hence, according to (2.4) the required energy per symbol for transmitting with data rate D_k is

$$E_k^{(D,S)} = \frac{\sigma^2}{g_k} \left(2^{D_k} - 1\right), \ k \in \mathcal{K}_S. \tag{2.5}$$

Cooperative Transmission Mode (CT Mode)

In this mode, for a certain source MT $k \in \mathcal{K}_S$, as shown in Fig. 2.1, it can *associate* with one idle MT (if any) within the distance d as its *relay* MT (illustrated by the solid circle in Fig. 2.1) that can help relay the data to the BS, where d is the range of the D2D communication such as LTE-Direct [15].[3] We denote this set of idle MTs within the distance d from the source MT $k \in \mathcal{K}_S$ as its set of *neighbouring MTs* $\mathcal{H}_k \subset \mathcal{K}_I$ (denoted by the dashed circle in Fig. 2.1), where $|\mathcal{H}_k| = N_k$ is the number of MTs within the set. Then, it follows that N_k is a Poisson RV with mean $\bar{\mu}_{N_k} = (1-\rho)\bar{\lambda}\pi d^2,\ k \in \mathcal{K}_S$ and its probability mass function (PMF) is given by

$$\mathbb{P}(N_k = n) = \frac{\bar{\mu}_{N_k}^n}{n!} e^{-\bar{\mu}_{N_k}},\ n = 0, 1, \cdots,\ k \in \mathcal{K}_S. \tag{2.6}$$

From (2.6), we observe that the PMF of N_k is approximately proportional to the range of the D2D communication d when μ_{N_k} is small, an MT's probability of remaining idle $1-\rho$ and the spatial density $\bar{\lambda}$. Note that if $N_k = 0$ or $\mathcal{H}_k = \emptyset$, source MT $k \in \mathcal{K}_S$ will operate in DT mode, i.e., transmit directly to the BS; while if $N_k \geq 1$, source MT $k \in \mathcal{K}_S$ can operate in CT mode by selecting one from its neighbouring MTs in $\mathcal{H}_k$ to relay the data.

For the CT mode, the source MT $k \in \mathcal{K}_S$ in general splits data D_k into two non-overlapping parts with $D_k = D_k^{(S)} + D_k^{(R)}$: $D_k^{(S)}$ for the source MT to transmit directly to the BS and $D_k^{(R)}$ for its relay MT to transmit. Because the MTs operate on the licensed band in D2D communication, MTs' transmission with D2D communication is orthogonal to both the other D2D communication and the wireless cellular communication. For transmitting the data $D_k^{(S)}$ (denoted by the solid line

[3] In this chapter, we assume single relay selection to keep the overhead low. Similar approach has been used in [46].

in Fig. 2.1), similar to (2.5), the required energy for the source MT $k \in \mathcal{K}_S$ is[4]

$$E_k^{(C,S)} = \frac{\sigma^2}{g_k}\left(2^{D_k^{(S)}} - 1\right), \ k \in \mathcal{K}_S. \tag{2.7}$$

Then, for the other part of data $D_k^{(R)}$, the source MT first transmits it to the selected relay MT and then the relay MT decodes and forwards the signal to the BS (denoted by the dashed line in Fig. 2.1). In practice, D2D communication offers high communication data rate with low transmit power. The energy consumption and the transmission time is also small compared to that in the wireless cellular network. Hence, we ignore them for this short range data transmission.[5] Also due to the small range, the source MT $k \in \mathcal{K}_S$ and its neighbouring MT $j \in \mathcal{H}_k$ have roughly the same distance to the BS (i.e., $r_k = r_j$) and can be assumed to have the same power attenuation G_k. Hence, the channel power gain between the neighbouring MT $j \in \mathcal{H}_k$ of the source MT $k \in \mathcal{K}_S$ and the BS is

$$g_j = \eta_j G_k, \tag{2.8}$$

where the short-term Rayleigh fading of the channel power η_j is still independently distributed among the MTs. Hence, if neighbouring MT $j \in \mathcal{H}_k$ is selected as the relay MT, the energy consumption for this data transmission is

$$E_j^{(C,R)} = \frac{\sigma^2}{g_j}\left(2^{D_k^{(R)}} - 1\right), \ j \in \mathcal{H}_k, \ k \in \mathcal{K}_S. \tag{2.9}$$

[4]In this chapter, we do not directly consider the maximum power constraint of the MT to obtain tractable problem formulation and insights. However, as will be shown later, we have already implicitly considered the issue of large transmit power. When the transmit power is very large, it incurs a substantial cost on the MT and the MT will try to get help from the other MTs. Hence, large energy consumption can be avoided. The simulation results in Section 2.6.2 will corroborate the effectiveness of our scheme.

[5]The transmit power of the LTE D2D communication between MTs is lower, while the data rate is higher compared to that of the cellular communication in the uplink and can be ignored. Furthermore, the analysis can be easily extended to the case that the energy consumption and transmission time of LTE D2D communication are constants, and there will not be major changes in the results.

2.3.1 Definition of Costs and Utilities

With different battery levels, an MT has different valuations of the remaining energy in its battery. The energy stored in the battery is generally more valuable when the battery level is low since further energy consumption may lead to an empty battery. Hence, we define the unit energy cost ζ_k for each MT $k \in \mathcal{K}$ as a function of its battery level B_k, i.e.,

$$\zeta_k = f(B_k), \tag{2.10}$$

where $B_k \in [0, B_{\max}]$ is the battery level of MT k with its range from zero to the maximum storage $B_{\max}$[6], and $f : [0, B_{\max}] \to [0, \zeta_{\max}]$ is a monotonically decreasing function of B_k whose range is from zero to the maximum energy cost $\zeta_{\max} > 0$.[7]

In order to motivate the neighbouring MT's participation in the cooperation, if a neighbouring MT $j \in \mathcal{H}_k$ is selected by the source MT $k \in \mathcal{K}_S$ as the relay MT, it will receive a price π_k for transmitting the source MT's data with data rate $D_k^{(R)}$. The payment can be in the form of currency or credits in a multimedia application. Hence, the utility of neighbouring MT $j \in \mathcal{H}_k$ by participating in the cooperation is $\pi_k - \zeta_j E_j^{(C,R)}$, where $E_j^{(C,R)}$ is the energy consumption for transmitting with data rate $D_k^{(R)}$ as defined in (2.9). Furthermore, the neighbouring MT has a reservation utility of $\epsilon \geq 0$ for accepting the request. That is, neighbouring MT $j \in \mathcal{H}_k$ will only accept the relay request from source MT $k \in \mathcal{K}_S$ if $\pi_k - \zeta_j E_j^{(C,R)} \geq \epsilon$. Therefore, the utility of the neighbouring MT $j \in \mathcal{H}_k$ for source MT $k \in \mathcal{K}_S$ is the difference between the price and the energy cost if the difference is larger than ϵ and zero

[6]For analytical tractability, in this chapter, we assume that all the MTs have the same battery capacity $B_{\max}$.

[7]This design of function f is reasonable as a user will value energy more when facing low battery, and we assume the minimum energy cost equal to zero when $B_k = B_{\max}$.

otherwise, which is defined as[8]

$$U_j = \begin{cases} \pi_k - \zeta_j E_j^{(C,R)}, & \text{if } \pi_k - \zeta_j E_j^{(C,R)} \geq \epsilon, \\ 0, & \text{otherwise.} \end{cases} \tag{2.11}$$

For the source MT $k \in \mathcal{K}_S$, if there is at least one neighbouring MT accepting the price π_k, the cost of the source MT $k \in \mathcal{K}_S$ is the sum of the price π_k and the energy cost by direct transmission $\zeta_k E_k^{(C,S)}$. Otherwise, it needs to directly transmit to the BS with rate D_k at the cost of $\zeta_k E_k^{(D,S)}$. Thus, the energy cost of source MT $k \in \mathcal{K}_S$ is

$$C_k = \begin{cases} \pi_k + \zeta_k E_k^{(C,S)}, & \text{if } \exists j \in \mathcal{H}_k, \pi_k - \zeta_j E_j^{(C,R)} \geq \epsilon, \\ \zeta_k E_k^{(D,S)}, & \text{otherwise.} \end{cases} \tag{2.12}$$

Hence, with the above cost function, the feasible price and load sharing for the source MT to choose CT mode instead of the DT mode is

$$\pi_k + \zeta_k E_k^{(C,S)} \leq \zeta_k E_k^{(D,S)}. \tag{2.13}$$

Furthermore, to ensure the benefits of the source MT $k \in \mathcal{K}_S$ in the cooperation, the price π_k should be larger than the reservation utility $\pi_k \geq \epsilon$. Hence, the following inequality should be satisfied:

$$\epsilon \leq \pi_k \leq \zeta_k E_k^{(D,S)} - \zeta_k E_k^{(C,S)}. \tag{2.14}$$

Note that $\zeta_k E_k^{(D,S)} \geq \epsilon$ must hold for the feasibility of the CT mode. That is, the value of the energy consumed by direct transmission at the source MT must be larger than the reservation utility of the neighbouring MT. If not satisfied, the source MT

[8]Here, note that the utility function is a concave function of $D_k^{(S)}$ with diminishing return and the cost function to be defined in (2.12) is a convex and monotonically increasing function with respect to $D_k^{(S)}$. Hence, these definitions conform to the classic definition of cost and utility functions in economics [60].

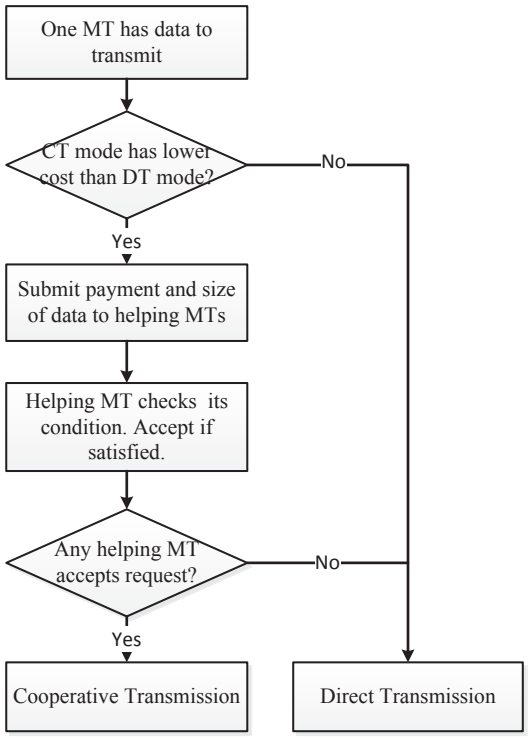


Figure 2.2: Protocol for the cooperative relaying.

should directly choose the DT mode for transmission.

2.3.2 Cooperative Transmission Protocol

Next, in order for the MTs in the cellular network to cooperate with mutual benefits, we propose the following cooperative data transmission protocol, which is also shown by the flow chart in Fig. 2.2.

1) When an MT has data to transmit, it chooses between the CT mode and DT mode according to the criterion to be specified later in (2.17) and (2.23), for the cases of complete and incomplete information, respectively.

2) If the DT mode is selected, the source MT transmits directly to the BS. If the CT mode is selected, it broadcasts the proposed payment and the relay data rate to all its neighbouring MTs.[9]

[9]This can be enabled by the peer discovery function specified in LTE-Direct [15].

3) The neighbouring MT (if any) accepts the request and sends an acceptance notification to the source MT if the condition for cooperation is satisfied or rejects the request otherwise.

4) If one or more neighbouring MTs accept the relay request, the source MT randomly chooses one MT as the relay MT and transmits the data with the CT mode.[10] Otherwise, the source MT transmits with the DT mode.

In the above proposed cooperative transmission protocol, the key challenge is the mechanism design for incentivizing the cooperation of the MTs such that the MTs can mutually benefit. In the following sections, we investigate the proposed pricing-based incentive mechanism design for the cooperation under different information sharing scenarios.

2.4 Benchmark Case: Full Cooperation under Complete Information

In this section, we consider the ideal case of full cooperation under complete information, where the private information of the neighbouring MTs $\mathcal{H}_k$, including the number of neighbouring MTs N_k, their battery levels B_j's and channel conditions g_j's, is known by each source MT $k \in \mathcal{K}_S$. This case can happen when the MTs belong to a fully cooperative group (e.g., friends) that they are willing to help each other and share their private information truthfully. This case will also provide the performance benchmark (upper bound) for the partial cooperation under incomplete information in the next section.

Due to the fully cooperative nature among the MTs, the reservation utility of the neighbouring MT ϵ reduces to zero. Hence, source MT $k \in \mathcal{K}_S$ only needs to give a payment to its neighbouring MT $j \in \mathcal{H}_k$ that is just enough to cover the cost

[10] Because the relay data rate and price are already determined by the source MT and each relay candidate provides the same help to the source MT, the source MT does not care about which neighbouring MT among those accepting the request is chosen.

$\zeta_j E_j^{(C,R)}$ for transmitting $D_k^{(R)}$ such that the neighbouring MT's utility in (2.11) is non-negative. Hence, the required amount of payment to neighbouring MT $j \in \mathcal{H}_k$ from source MT $k \in \mathcal{K}_S$ is $\zeta_j E_j^{(C,R)}$. Then, source MT $k \in \mathcal{K}_S$ needs to optimize the relay data rate $D_k^{(R)}$ for each neighbouring MT $j \in \mathcal{H}_k$ to minimize the sum energy cost, i.e.,

$$\begin{aligned} C_{k,j} = \underset{D_k^{(R)} \geq 0}{\texttt{min.}} \ & \zeta_j E_j^{(C,R)} + \zeta_k E_k^{(C,S)} \\ \text{s.t.} \ & D_k^{(R)} + D_k^{(S)} = D_k. \end{aligned} \tag{2.15}$$

Problem (2.15) can be considered as a weighted-sum energy minimization problem for the source and neighbouring MTs, where the weights are the unit energy costs of the source and neighbouring MTs. It is evident that when the weights (i.e., unit energy cost) of the source and neighbouring MTs are equal, this problem reduces to the sum energy minimization problem. When the weight of one MT is larger than that of the other, the problem is more favourable for the MT with lower energy and the optimization is more similar to the max-min optimization of the battery levels.

After obtaining the minimum sum cost $C_{k,j}$ of associating with each neighbouring MT $j \in \mathcal{H}_k$, source MT $k \in \mathcal{K}_S$ chooses the best neighbouring MT with the following relay selection problem:

$$\text{(P1)}: \ \hat{C}_k = \underset{j \in \mathcal{H}_k}{\texttt{min.}} \ C_{k,j}, \tag{2.16}$$

where $C_{k,j}$ is obtained in Problem (2.15).

Next, we discuss the criterion for the mode selection of source MT $k \in \mathcal{K}_S$. For the source MT to choose the CT mode, its cost reduction from the DT mode must be larger than zero. Hence, if source MT $k \in \mathcal{K}_S$ chooses the CT mode, the following condition has to be satisfied:

$$\hat{C}_k < \zeta_k E_k^{(D,S)}. \tag{2.17}$$

In the following, we discuss the solution for the minimum cost $\hat{C}_k$ of full cooperation under complete information in two cases: non-splittable data (i.e., $D_k^{(R)} = D_k$) and splittable data (i.e., $0 \leq D_k^{(R)} \leq D_k$).

2.4.1 Cooperation with Non-splittable Data

First, we discuss the case where the data is not splittable at the source MT due to reasons such as lack of necessary processing functionalities. In this case, all the data of the cooperative relaying is transmitted by the relay MT (i.e., $D_k^{(S)} = 0$ and $D_k^{(R)} = D_k$). Hence, according to (2.15), the cost of the source MT $k \in \mathcal{K}_S$ by associating with neighbouring MT $\mathcal{H}_k$ reduces to $C_{k,j} = \zeta_j \frac{\sigma^2}{g_j} \left(2^{D_k} - 1\right)$ for Problem (P1). Then, the minimum cost of full cooperation with non-splittable data can be obtained by solving the simplified relay selection problem in (P1).

Therefore, the optimal transmission of the source MT in the case of non-splittable data follows a two-step procedure: First, the source MT computes and finds the neighbouring MT (if any) with the least energy cost. Then, it checks the condition in (2.17) and chooses between the DT mode and CT mode.

2.4.2 Cooperation with Splittable Data

It can be proved that Problem (2.15) is a convex optimization problem and the optimal solution is given by the following proposition.

Proposition 2.4.1 *The optimal data rate transmitted by the relay MT in Problem (2.15) is given by*

$$\hat{D}_k^{(R)} = \begin{cases} 0, & \text{if } \log_2 \frac{\theta_k}{\theta_j} < -D_k \\ \frac{1}{2}(D_k + \log_2 \frac{\theta_k}{\theta_j}), & \text{if } -D_k \leq \log_2 \frac{\theta_k}{\theta_j} < D_k \\ D_k, & \text{if } \log_2 \frac{\theta_k}{\theta_j} \geq D_k \end{cases}, \tag{2.18}$$

where $\theta_k = \frac{\zeta_k}{\eta_k}$ *and* $\theta_j = \frac{\zeta_j}{\eta_j}$ *can be interpreted as the effective energy cost of the*

source MT $k \in \mathcal{K}_S$ and neighbouring MT $j \in \mathcal{H}_k$, respectively.

Proof See Appendix A.

It can be observed from Proposition 2.4.1 that the optimal relay data rate follows a threshold structure with respect to the log-ratio between the effective energy costs of the source and neighbouring MTs. When the effective energy cost θ_k of the source MT $k \in \mathcal{K}_S$ is much lower than that of the relay MT $j \in \mathcal{H}_k$ to the extent that $\log_2 \frac{\theta_k}{\theta_j} < -D_k$ is satisfied, the source MT will not ask for help from this neighbouring MT and transmit all by itself. If the effective energy costs of the source and neighbouring MTs are comparable, then the source and neighbouring MT will split the data packet D_k for transmission. Finally, if the effective energy cost of the source MT is much higher than that of the neighbouring MT so that $\log_2 \frac{\theta_k}{\theta_j} \geq D_k$ is satisfied, then the neighbouring MT will transmit the whole data packet.

2.5 General Case: Partial Cooperation under Incomplete Information

In the previous section, we have considered the full cooperation under complete information, which is the optimal scenario for the source MT and can serve as the benchmark scheme. However, this scenario is not applicable if the MTs belong to different entities that are not fully cooperative and are unwilling to share private information to each other. In this section, we consider the general scenario where the MTs do not know the other MTs' channel conditions or battery levels and discuss how these MTs still can cooperate with mutual benefits under this scenario.

2.5.1 Problem Formulation

For cooperation under incomplete information between the source and neighbouring MTs, we formulate the problem of decision making under uncertainties

with the expected utility theory [61]. We denote $\mathbb{P}(\pi_k - \zeta_j E_j^{(C,R)} \leq \epsilon)$ as the probability that neighbouring MT $j \in \mathcal{H}_k$ rejects the request from source MT $k \in \mathcal{K}_S$. We assume that all the channel gains g_j's and battery states B_j's of the neighbouring MTs $j \in \mathcal{H}_k$ are independent. Hence, given the set of neighbouring MTs $\mathcal{H}_k$, the conditional expected cost of the source MT $k \in \mathcal{K}_S$ for transmitting at data rate D_k is

$$\begin{aligned}\mathbb{E}[C_k|\mathcal{H}_k] &= \left(1 - \prod_{j\in\mathcal{H}_k} \mathbb{P}(\pi_k - \zeta_j E_j^{(C,R)} \leq \epsilon)\right)(\pi_k + \zeta_k E_k^{(C,S)}) \\ &+ \left(\prod_{j\in\mathcal{H}_k} \mathbb{P}(\pi_k - \zeta_j E_j^{(C,R)} \leq \epsilon)\right)\zeta_k E_k^{(D,S)} \\ &= \left(1 - \prod_{j\in\mathcal{H}_k} \mathbb{P}(\pi_k - \zeta_j E_j^{(C,R)} \leq \epsilon)\right)(\pi_k + \zeta_k E_k^{(C,S)} - \zeta_k E_k^{(D,S)}) + \zeta_k E_k^{(D,S)}, \quad (2.19)\end{aligned}$$

where the expectation is taken over the two possible outcomes of successful and unsuccessful relay association in (2.12), respectively. By further considering all possibilities of neighbouring MT set $\mathcal{H}_k$ for source MT $k \in \mathcal{K}_S$ in (2.6), the expected cost of the source MT $k \in \mathcal{K}_S$ can be obtained by applying the law of iterated expectation, i.e.,

$$\mathbb{E}[C_k] = \mathbb{E}[\mathbb{E}[C_k|\mathcal{H}_k]] = \sum_{n=0}^{\infty} \mathbb{P}(N_k = n)\mathbb{E}[C_k|\mathcal{H}_k], \ k \in \mathcal{K}_S. \quad (2.20)$$

Here, it is worthwhile to discuss the role of reservation utility ϵ in the expected energy cost $\mathbb{E}[C_k]$. As ϵ denotes the level of minimum benefit for the relay MT in the cooperation, it can be observed that the expected energy cost $\mathbb{E}[C_k]$ should be monotonically increasing with ϵ. That is, the expected energy cost of the source MT is monotonically increasing with respect to the reservation utility of the relay MTs.

Then, we formulate the pricing and load sharing of the source MTs as the following optimization problem to minimize the expected cost of the source MT

$k \in \mathcal{K}_S$ over the price π_k and relay data $D_k^{(R)}$:

$$\begin{aligned}(\text{P2}): \quad \underset{\pi_k, D_k^{(R)} \geq 0}{\texttt{min.}} \quad & \mathbb{E}[C_k] \\ \texttt{s.t.} \quad & \epsilon \leq \pi_k \leq \zeta_k E_k^{(D,S)} - \zeta_k E_k^{(C,S)}, \qquad (2.21) \\ & D_k^{(S)} + D_k^{(R)} = D_k. \qquad (2.22)\end{aligned}$$

Next, we discuss the criterion for the mode selection between the DT mode and CT mode. Similar to the condition for the full cooperation case in (2.17), the condition for the source MT to choose the CT mode is that the cost of the CT mode should be lower than that of the DT mode:

$$\zeta_k E_k^{(D,S)} \geq \max\{\mathbb{E}[C_k^*], \epsilon\}, \qquad (2.23)$$

where $\mathbb{E}[C_k^*]$ is the minimum expected cost obtained in Problem (P2). It should be noted that the convexity of Problem (P2) is hard to prove due to its complex objective function. Thus, it is difficult to obtain its global optimal solution in general. In the following two subsections, similar to Section 2.4, we discuss the minimum expected cost $\mathbb{E}[C_k^*]$ of the source MT $k \in \mathcal{K}_S$ in details under the two cases of splittable and non-splittable data.

2.5.2 Proposed Solution

In this subsection, we first simplify Problem (P2) under some further assumptions. Then, similar to Section 2.4, we discuss the solution of the problem under the cases of non-splittable and splittable data, respectively. With the energy consumption $E_j^{(C,R)}$ for the neighbouring MT $j \in \mathcal{H}_k$ defined in (2.9), the probability of successful association between source MT $k \in \mathcal{K}_S$ and its neighbouring MT $j \in \mathcal{H}_k$

in (2.19) is

$$\mathbb{P}(\pi_k - \zeta_j E_j^{(C,R)} \geq \epsilon) = \mathbb{P}\left(\frac{\zeta_j}{\eta_j} \leq \frac{G_k(\pi_k - \epsilon)}{\sigma^2(2^{D_k^{(R)}} - 1)}\right) = \mathbb{P}\left(\frac{\zeta_j}{\eta_j} \leq \phi_k\right), \tag{2.24}$$

where ϕ_k is denoted as

$$\phi_k = \frac{G_k(\pi_k - \epsilon)}{\sigma^2(2^{D_k^{(R)}} - 1)}. \tag{2.25}$$

For simplicity, we further assume that the relation between an MT's unit energy cost ζ_k and its battery level B_k in (2.10) follows a linear function[11]

$$\zeta_k = \zeta_{\max}\left(1 - \frac{B_k}{B_{\max}}\right). \tag{2.26}$$

We also assume that the battery level B_j of the neighbouring MT $j \in \mathcal{H}_k$ is known to the source MT $k \in \mathcal{K}_S$ as uniform distribution, i.e., $B_j \sim u(0, B_{\max})$.[12] Then, due to the linear function in (2.26), the energy cost ζ_j is also uniformly distributed as $\zeta_j \sim u(0, \zeta_{\max})$. Hence, the probability of successful association between the source MT $k \in \mathcal{K}_S$ and neighbouring MT $j \in \mathcal{H}_k$ is

$$\begin{aligned}\mathbb{P}(\pi_k - \zeta_j E_j^{(C,R)} \geq \epsilon) &= \mathbb{P}\left(\eta_j \geq \frac{\zeta_j}{\phi_k}\right) \\ &= \frac{1}{\zeta_{\max}} \int_0^{\zeta_{\max}} \int_{\frac{\zeta_j}{\phi_k}}^{\infty} e^{-\eta_j} d\eta_j d\zeta_j = \frac{\phi_k}{\zeta_{\max}}(1 - e^{-\frac{\zeta_{\max}}{\phi_k}}).\end{aligned} \tag{2.27}$$

With the results in (2.19) and (2.27), the objective of Problem (P2) in (2.20) can

[11] It should be noted that our analysis can be extended to the other monotonically non-increasing functions, whose analysis will be technically more involved but offers essentially similar engineering insights. It should also be noted that the choice of the function f reflects the sensitivity of the MTs towards the usage of the energy in the battery. By adopting a function that is *in-different* to the battery level, the design objective is more similar to minimizing the total energy consumption. Instead, by adopting a function that is *sensitive* to the battery level, this design is more favourable for the MTs with low battery level.

[12] Our proposed scheme can still be applicable to the case of heterogeneous battery capacity. One heuristic is that, based on the statistics of the battery capacities of the MTs, the source MT $k \in \mathcal{K}_S$ can obtain the average battery capacities of the MTs as $\bar{B}_{\max}$ and predict the battery level of the neighbouring MT $j \in \mathcal{H}_k$ as $B_j \sim u(0, \bar{B}_{\max})$. Then, the proposed cooperative relaying protocol with pricing under uncertainty still applies.

be simplified as

$$\mathbb{E}[C_k] = \sum_{n=0}^{\infty} \mathbb{P}(N_k = n) \Bigg\{ \left[1 - \left(1 - \frac{\phi_k}{\zeta_{\max}} \left(1 - e^{-\frac{\zeta_{\max}}{\phi_k}} \right) \right)^n \right] \times \left(\pi_k + \zeta_k E_k^{(C,S)} - \zeta_k E_k^{(D,S)} \right) + \zeta_k E_k^{(D,S)} \Bigg\}. \tag{2.28}$$

Next, we discuss the convexity of Problem (P2) by the following proposition.

Proposition 2.5.1 *Problem* (P2) *is marginally convex with respect to* π_k *and* $D_k^{(R)}$.

Proof See Appendix B.

It should be noted that the objective of the Problem (P2) is not jointly convex with respect to π_k and $D_k^{(R)}$. In the following, similar to Section 2.4, we discuss the optimal solution for Problem (P2) to obtain $\mathbb{E}[C_k^*]$ under two cases: non-splittable (i.e., $D_k^{(R)} = D_k$) and splittable data (i.e., $0 \leq D_k^{(R)} \leq D_k$).

Optimal Pricing for Non-splittable Data

First, we discuss the case where the data is not splittable. In this case, all the data of the source MTs is transmitted by the relay MT (i.e., $D_k^{(S)} = 0$ and $D_k^{(R)} = D_k$). As ϕ_k in (2.25) is now reduced to $\phi_k = \frac{G_k(\pi_k - \epsilon)}{\sigma^2(2^{D_k}-1)}$, Problem (P2) is simplified to the following problem without load sharing:

$$(\text{P2}'): \quad \min_{\pi_k} \sum_{n=0}^{\infty} \mathbb{P}(N_k = n) \Bigg\{ \left[1 - \left(1 - \frac{\phi_k}{\zeta_{\max}} \left(1 - e^{-\frac{\zeta_{\max}}{\phi_k}} \right) \right)^n \right] \left(\pi_k - \zeta_k E_k^{(D,S)} \right) + \zeta_k E_k^{(D,S)} \Bigg\}$$
$$\text{s.t. } \epsilon \leq \pi_k \leq \zeta_k E_k^{(D,S)}.$$

Because the data transmitted by the relay MT is fixed at $D_k^{(R)} = D_k$, according to Proposition 2.5.1, the problem is convex with respect to π_k. Therefore, for this uni-variable convex optimization problem, the optimal solution can be obtained by checking the first-order condition of optimality. However, the objective function

Table 2.1: Algorithm I: One-dimensional dichotomous search algorithm for solving Problem (P2′) with precision δ_{π_k} and $\tau \ll 1$.

1. **Initialize:** $\pi_k^{(l)} \coloneqq \epsilon$, $\pi_k^{(h)} \coloneqq \zeta_k E_k^{(D,S)}$, $\Delta_{\pi_k} \coloneqq |\pi_k^{(l)} - \pi_k^{(h)}|$;
2. **Repeat:**
   1) Set temporary parameters: $\tilde{\pi}_k^{(l)} \coloneqq \frac{1}{2}(\pi_k^{(l)} + \pi_k^{(h)}) - \tau\Delta_{\pi_k}$, $\tilde{\pi}_k^{(h)} \coloneqq \frac{1}{2}(\pi_k^{(l)} + \pi_k^{(h)}) + \tau\Delta_{\pi_k}$;
   2) If $\mathbb{E}[C_k(\pi_k^{(l)})] < \mathbb{E}[C_k(\pi_k^{(h)})]$, set the price as $\pi_k^{(h)} \coloneqq \tilde{\pi}_k^{(l)}$;
   3) If $\mathbb{E}[C_k(\pi_k^{(l)})] > \mathbb{E}[C_k(\pi_k^{(h)})]$, set the price as $\pi_k^{(l)} \coloneqq \tilde{\pi}_k^{(h)}$;
   4) Otherwise, set $\pi_k^{(h)} \coloneqq \tilde{\pi}_k^{(h)}$ and $\pi_k^{(l)} \coloneqq \tilde{\pi}_k^{(l)}$;
   5) $\Delta_{\pi_k} \coloneqq |\pi_k^{(l)} - \pi_k^{(h)}|$;
3. **Until:** the condition $\Delta_{\pi_k} > \delta_{\pi_k}$ is violated;
4. $\pi_k^* \coloneqq (\pi_k^{(h)} + \pi_k^{(l)})/2, \quad \mathbb{E}[C_k^*] \coloneqq \mathbb{E}[C_k(\pi_k^*)]$.

of Problem (P2′) is still complicated, for which the derivative is hard to obtain. Hence, we propose Algorithm I in Table 2.1 based on the derivative-free dichotomous search [62] to obtain the optimal solution numerically for Problem (P2′).

Joint Pricing and Load Sharing for Splittable Data

Next, we discuss the general case where the data is splittable at the source MT in Problem (P2). According to Proposition 2.5.1, the objective function of Problem (P2) is convex with respect to π_k given a fixed $D_k^{(R)}$ and to $D_k^{(R)}$ given a fixed π_k. Hence, based on the dichotomous search algorithm in Algorithm I, we propose Algorithm II that minimizes the expected cost of the source MT $k \in \mathcal{K}_S$ with alternating optimization.

For Algorithm II, it starts with the optimal solution obtained in Algorithm I with $D_k^{(R)} = D_k$. The algorithm then proceeds by iteratively optimizing and updating π_k and $D_k^{(R)}$ with the other fixed until the stopping condition is satisfied. It should be noted that the algorithm always converges to a certain value within the

Table 2.2: Algorithm II: Alternating optimization algorithm for solving Problem (P2) with precision δ_{C_k}.

1. **Initialize:** $n \coloneqq 0$, $D_k^{(R)} \coloneqq D_k$, $\mathbb{E}[C_k^{(0)}] \coloneqq \zeta_k E_k^{(D,S)}$;
2. **Repeat:**
   1) Optimize the objective of Problem (P2) with respect to π_k by dichotomous search with $D_k^{(R)}$ fixed ;
   2) Optimize the objective of Problem (P2) with respect to $D_k^{(R)}$ by dichotomous search with π_k fixed ;
   3) $n \coloneqq n + 1$.
3. **Until:** the condition $|\mathbb{E}[C_k^{(n)}] - \mathbb{E}[C_k^{(n-1)}]| > \delta_{C_k}$ is violated.

range of δ_{C_k} from at least a locally optimal solution. This is because each iteration of the algorithm reduces the objective value and the optimal value of Problem (P2) is lower bounded.

Finally, we consider the complexities of the above two algorithms. For the complexity of Algorithm I, the maximum number of iterations required for the searching of the optimal pricing π_k with precision δ_{π_k} is $O(\log_2 \frac{\zeta_k E_k^{(D,S)}}{\delta_{\pi_k}})$. Next, for the complexity of Algorithm II, the upper bound of each line search for $D_k^{(R)}$ and π_k are $M_{D_k^{(R)}} = \log_2(\frac{D_k}{\delta_{D_k^{(R)}}})$ and $M_{\pi_k} = \log_2 \frac{\zeta_k E_k^{(D,S)}}{\delta_{\pi_k}}$, respectively, where $\delta_{D_k^{(R)}}$ is the precision requirement for the line search of $D_k^{(R)}$. The upper bound for the total number of iterations in the above alternating optimization is $M = \frac{\zeta_k E_k^{(D,S)}}{\delta_{C_k}}$. Hence, the upper bound for the complexity of Algorithm II is $O(M(M_{\pi_k} + M_{D_k^{(R)}}))$. Moreover, given the data rate of the source MT, user density, energy cost and channel condition, the optimal solution can be computed off-line and stored in a look-up table for practical implementation.

2.6 Numerical Results

In this section, we first show the convergence of the algorithm for the single

Table 2.3: General simulation parameters for cooperative relaying.

Simulation Parameters	Values
Noise power	$\sigma^2 = -110$ dBm
Path-loss exponent	$\alpha = 3.6$
Reference distance	$r_0 = 10$ m
Path-loss at r_0	$G_0 = -70$ dB
Relay MT reservation utility	$\epsilon = 0.2$
Maximum battery level	$B_{\max} = 100$ J
Maximum unit energy cost	$\zeta_{\max} = 1$

Table 2.4: Simulation parameters for single source MT.

Simulation Parameters	Values
Distance from the source MT $k \in \mathcal{K}_S$ to BS[13]	$r_k = 50$ m
Short-term fading of this single source MT k	$\eta_k = 0.5$
Initial battery level of the source MT k	$B_k = 10$ J

source MT and examine its performance under different transmission schemes. Then, the simulation of multiple source MTs is given to demonstrate their real-time operation under our proposed protocol in a single-cell system. The general simulation parameters are given in Table 2.3 and specific simulation setup and parameters for the cases of single source MT and multiple source MTs will be elaborated later in each subsection.

2.6.1 Single Source MT

In this subsection, we consider the simulation for single source MT. We first show the convergence of Algorithm II for the partial cooperation with splittable data rate and compare the convergent cost to that with non-splittable data rate by Algorithm I. Then, we show the simulation results for the expected cost of the single source MT versus battery levels under different schemes. The specific simulation parameters for this case of single source MT are given in Table 2.4.

[13]The typical range of LTE in the urban environment is 1-5 km [63].

Convergence of Algorithm II for partial cooperation

First, we show the convergence of Algorithm II for the partial cooperation with splittable data compared with that with non-splittable data by Algorithm I in Section 2.5.2 for the data transmission of a single source MT $k \in \mathcal{K}_S$ under different data rates D_k and average number of neighbouring MTs $\bar{\mu}_{N_k}$. First, exhaustive search on π_k and $D_k^{(R)}$ with quantization of 0.2 and 0.1 in the feasible regions is conducted for three cases with different pairs of $\bar{\mu}_{N_k}$ and D_k and the minimum expected costs are 11.51, 4.46 and 1.55, respectively. Then, the result of the joint optimization of π_k and D_k by Algorithm II with splittable data (*SD*) is shown in Fig. 2.3 with the solid line. The expected cost at iteration $\{0\}$ denotes the cost by the direct transmission, which are 19.96, 19.96 and 2.22, respectively. The procedure in Algorithm II has executed four iterations, and each of the univariable dichotomous search sub-routines in Algorithm II is executed eight times, with sub-routines $\{1, 3, 5, 7\}$ for minimization with respect to π_k and sub-routines $\{2, 4, 6, 8\}$ for that with respect to $D_k^{(R)}$ in Algorithm II. For comparison, the result by Algorithm I with non-splittable data (*NSD*) is shown by the three dash lines.

The simulation result is shown in Fig. 2.3 and it can be observed that the convergence is fast. The converged expected energy costs for the three cases are 11.51, 4.46 and 1.55, respectively, which are the same as the results by exhaustive search. Hence, the global optimal solution is obtained in this case. Furthermore, the expected cost reduction from direct transmission for the three cases are 8.45, 15.50 and 0.67, respectively. Hence, according to the condition in (2.23), the transmission mode selected by the source MT for the three cases will be CT, CT, and DT, respectively. By comparing the two cases with $D_k = 6$ bps/Hz, it can be observed that a higher density of neighbouring MTs can further reduce the expected energy cost. By comparing the two cases with splittable and non-splittable data, it can also be observed that, in addition to the optimal pricing, load sharing can indeed further reduce the expected energy cost. Furthermore, the cost reductions with load sharing are 1.57, 1.25 and 1.12 times of those without load sharing, respectively.

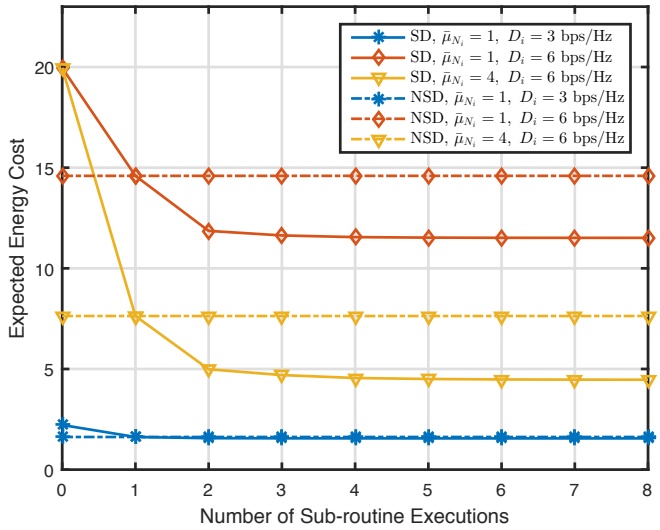


Figure 2.3: Expected energy cost of partial cooperation under incomplete information with splittable and non-splittable data versus the number of sub-routine executions under different $\bar{\mu}_{N_k}$'s and D_k's.

Therefore, load sharing is more cost-effective when the size of the data is large and the average number of neighbouring MTs is small. Finally, it should be noted that the case of splittable data leads to a lower energy cost compared with that of non-splittable data. This is because the energy consumption is exponentially increasing with respect to the transmission data rate and splitting the data and further optimizing the relay data rate result in a smaller total energy consumption.

In summary, for the cooperative transmission under complete information with splittable data, the optimal solution is obtained by the following three-step procedure: First, the source MT computes the optimal data rate for each neighbouring MT according to Proposition 2.4.1. Then, it searches for the one with the lowest energy cost from all the neighbouring MTs by Problem (P1). Last, it checks the condition in (2.17) and chooses between the DT and CT mode.

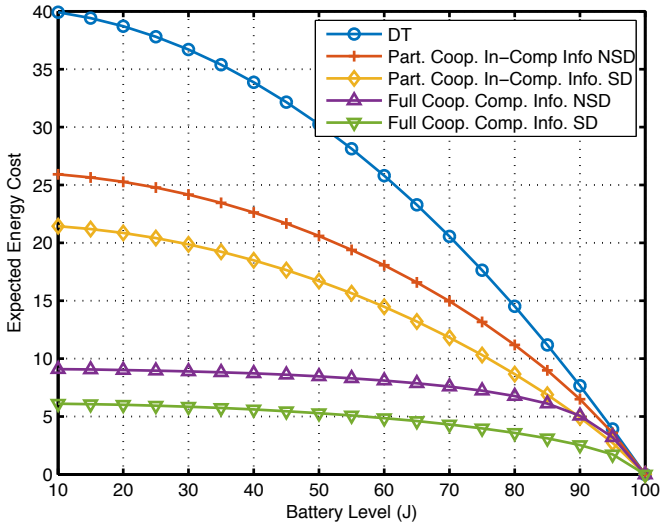


Figure 2.4: Expected cost of the single source MT versus its battery level under different schemes.

Expected energy cost under different battery levels and transmission schemes

Next, we show the expected cost of different schemes under different battery levels. The simulation setup is shown as follows. We consider the simulation under the following five schemes:

- **Direct Transmission** *(DT)* in (2.5).
- **Full Cooperation under Complete Information with Non-Splittable Data** *(Full Coop. Comp. Info. NSD)* in Section 2.4.1.
- **Full Cooperation under Complete Information with Splittable Data** *(Full Coop. Comp. Info. SD)* in Section 2.4.2.
- **Partial Cooperation under Incomplete Information with Non-Splittable Data** *(Part. Coop. In-Comp. Info. NSD)* in Section 2.5.2.

- **Partial Cooperation under Incomplete Information with Splittable Data** *(Part. Coop. In-Comp. Info. SD)* in Section 2.5.2.

Specifically, for the schemes of partial cooperation under incomplete information with splittable and non-splittable data, the minimum expected costs are obtained by Algorithms I and II, respectively. For the schemes of full cooperation under complete information with splittable and non-splittable data, the number of neighbouring MTs N_k for source MT $k \in \mathcal{K}_S$ is generated according to the Poisson distribution with $\bar{\mu}_{N_k} = 2$ and the source MT transmits the data at the rate of $D_k = 4$ bps/Hz. Their battery levels $B_j,\ j \in \mathcal{H}_k$ are uniformly generated on $[0, B_{\max}]$ and short term Rayleigh fading $\eta_j,\ j \in \mathcal{H}_k$ is generated according to exp(1). The minimum energy costs can be obtained by the results in Sections 2.4.1 and 2.4.2, respectively, and the results are averaged over 1000 independent realizations for accurately obtaining the expected energy costs for comparison. The expected cost of the transmission with DT mode is also obtained by averaging over 1000 independent realizations.

The simulation result is shown in Fig. 2.4. It can be observed that our proposed cooperative relaying scheme performs significantly better than the direct transmission benchmark. Moreover, cooperative relaying is more effective when the battery level of the source MT is low. This is because when the battery level of the source MT is high and cost for direct transmission is low, it is less likely to seek help from the other MTs. Furthermore, it can also be observed that there are gaps in expected energy costs between the schemes with complete information in Section 2.4 and those with incomplete information in Section 2.5, which are explained as follows:

1) In the case of complete information with non-splittable data, the source MT can observe the set of neighbouring MTs as well as their channel conditions and battery levels, and choose the most cost-efficient one as the relay. While for the incomplete information case, the source MT can only randomly choose one from the possible neighbouring MTs that accept the offer with the risk of ending up with direct transmission.

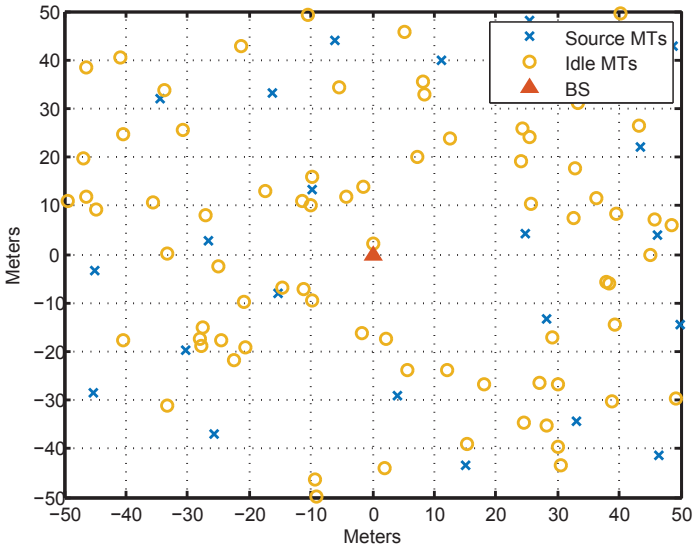


Figure 2.5: Setup for the simulation of multiple source MTs with $|\mathcal{K}| = 100$ MTs.

2) In the case of complete information with splittable data, in addition to the reason in 1), the source MT can jointly optimize the relay data rate and the payment and choose the neighbouring MT that leads to the minimum sum energy cost (as in Proposition 2.4.1). While in the case of incomplete information with splittable data, the source can only optimize the payment and relay data rate with respect to the expected energy cost, which has the possibility that the source MT ends up with direct transmission due to the lack of neighbouring MTs' information.

3) In both the cases with and without splittable data under complete information, the reservation utility margin ϵ for the neighbouring MT is zero, which further reduces the cost of the source MT from that under incomplete information and fully motivates the cooperation.

Finally, the figure shows that, when battery levels equal 100 J, the expected energy costs for all cases also become zero. This is due to our assumption that the energy cost at full battery capacity (i.e., battery level equals 100 J) is zero and in this case, there is no cooperation between the source and relay MTs.

Table 2.5: Simulation parameters for multiple source MTs.

Simulation Parameters	Values
Total number of MTs	$\|\mathcal{K}\| = 100$
Probability at which MTs initiate data transmission	$\rho = 0.2$
Normalized data rate[14]	$D_k = 6$ bps/Hz
Range of the D2D communication for a source MT	$d = 10$ m

2.6.2 Multiple Source MTs

In this subsection, we conduct a simulation with multiple source MTs and show the real-time operation of our proposed cooperative relaying protocol within a single cell. We examine the five schemes considered in the previous subsection and show the performance improvement in terms of battery and communication outage, average battery level and battery level distribution, under a single-cell setup. The specific simulation parameters for multiple source MTs are given in Table 2.5 and the simulation setup is described as follows.

We consider our simulation within a 100×100 m^2 square area as shown in Fig. 2.5. The operation of the system begins with the battery levels of the MTs uniformly generated on $[0, B_{\max}]$. For the purpose of investigation, at the beginning of each time slot, the positions of the MTs are uniformly re-generated within the above-mentioned area. In this setup, it is possible that there is overlap between the set of neighbouring MTs for different source MTs, where one neighbouring MT can possibly be associated with two source MTs. In order to avoid this situation, we re-generate the positions of the source MTs if there is overlap between the neighbouring MTs. According to the function $\bar{\mu}_{N_k} = (1-\rho)\bar{\lambda}\pi d^2,\ k \in \mathcal{K}_S$, the average number of neighbouring MTs in this setup is $\bar{\mu}_{N_k} = 1.2$. Due to the physical constraint of the MT, we set the maximum transmit energy of the MT as $E_{\max} = 3$ J for any time slot. If the transmit energy of the MT exceeds $E_{\max}$, a *communication outage* will be declared by the MT and the data packet is discarded. During the operation of the system, if the battery of a certain MT is drained out, this MT

[14]The uplink spectrum efficiency of the LTE system is $3.75 \sim 15$ bps/Hz [59].

Table 2.6: Number of communication and battery outages for the five cases after 300 time slots.

		Part. Coop. In-Comp. Info.		Full Coop. Comp. Info.	
	DT	NSD	SD	NSD	SD
Commun. Outage	289	209	153	47	30
Battery Outage	44	32	26	9	6

declares a *battery outage* and ceases any operation from that time on, including data transmission as a source MT or cooperative relay for the other source MTs as relay MT.

We show the total number of communication and battery outages for the 100 MTs after 300 time slots for the same five schemes as in Section 2.6.1. The simulation results are shown in Table 2.6. It can be observed that, compared with the benchmark case of DT, all of the four proposed schemes with cooperative relaying perform better in terms of communication and battery outage. The reduction of the battery outages reflects the effectiveness of our protocol design for the energy saving of the MT, especially for those MTs that are low in battery level. In addition to the reduction of the battery outage, our proposed scheme also shows a significant reduction in the number of communication outages. This is because, in the case of direct transmission, if the channel condition of the source MT is poor, the transmission power will exceed the peak power constraint $E_{\max}$ and communication outage will occur. While, under the same circumstances with cooperative relaying, the source MT can seek help from the other neighbouring MTs, whose transmit power is possibly lower than the peak power constraint and the transmission can be successful. Hence, our proposed schemes can improve the uplink data transmission of the MTs in terms of both the communication reliability and battery sustainability.

Next, we show the average battery level $\sum_k B_k/|\mathcal{K}|$ of the MTs during the 300 time slots in Fig. 2.6. It can be observed that the average battery levels of different schemes drop with different rates. Compared with the benchmark case of DT, our

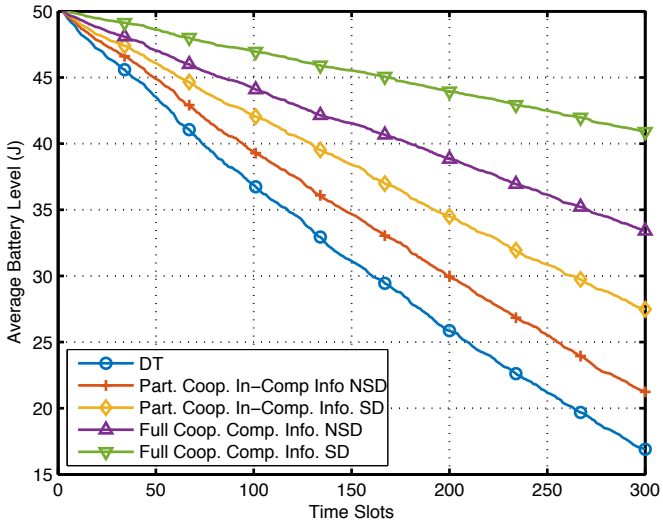


Figure 2.6: Average battery level $\sum_k B_k/|\mathcal{K}|$ of the MTs over time.

proposed protocol can effectively increase the average battery levels of the MTs over time. Even though the MTs under cooperative relaying successfully deliver more data packets as shown in Table 2.6, these schemes still perform better in terms of average battery level.

Finally, we show the distribution of the battery levels of the 100 MTs at the end of the 300 time slots in Fig. 2.7. It can be observed that, for the benchmark case of DT, a large proportion of the MTs have drained out their batteries. While for the other cases with cooperation, their battery levels remain on the relatively higher level than the direct transmission case by the distribution. It should also be noted that although a lot of MTs under cooperative relaying stay in the low battery region (i.e., $0-20$ J), their batteries are not empty according to Table 2.6. This is because, when their battery levels are low, these MTs can possibly receive help from the other MTs such that their battery levels can be sustained. While, for the direct transmission case, the batteries of a lot of MTs in this region are empty due to the lack of help from the other MTs.

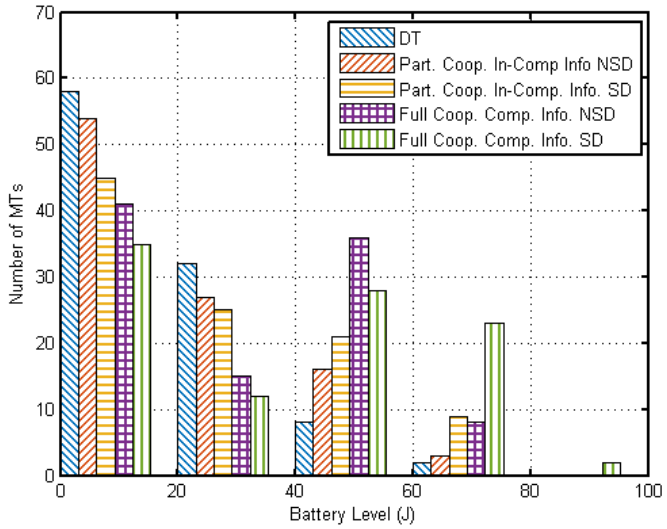


Figure 2.7: Distribution of the battery levels of 100 MTs after 300 time slots.

2.7 Chapter Summary

In this chapter, we studied the cooperative relaying for the energy saving of MTs under heterogeneous battery levels in the wireless cellular network. The benchmark case of full cooperation under complete information was first considered for the cases of splittable and non-splittable data, respectively. The problem was formulated as a relay selection problem and a simple threshold structure was shown for the optimal solution. Then, the general case of partial cooperation under incomplete information was considered. We formulated the MTs' decision-making under uncertainties as an optimization problem for minimizing the expected energy cost of source MT. An efficient algorithm based on dichotomous search and alternating optimization was proposed for the solution. Finally, simulations with single source MT and multiple source MTs were provided, which showed that our proposed cooperative relaying protocol could significantly decrease the number of communication and battery outages for the MTs and improve the average battery level during their operations. Overall, we proposed a new pricing-based mechanism for cooperative relaying and our results revealed new insights to the practical design of cooperative relaying for energy saving.

Chapter 3

MT-side Cooperative Local Caching

3.1 Introduction

In Chapter 2, we have studied the cooperation techniques in the uplink transmission of the MTs. In this chapter, we consider the downlink transmission of the MTs and the cooperative local caching for the file sharing between the MTs with heterogeneous file preferences. We first practically categorize the MTs into different interest groups according to their preferences. The MTs in each group belong to the same entity and aim to increase the probability of successful file discovery from the neighbouring MTs (MTs from the same or different groups). Hence, we define the group's utility as the probability for the MTs in the group to successfully discover the file in the neighbouring MTs, which can be maximized by optimizing the caching strategies. By modelling MTs' mobilities as independent HPPPs, we analytically characterize different groups' utilities in closed-form. We first consider the fully cooperative case where a centralizer helps all groups to make the caching decisions. We formulate the problem as a weighted-sum utility maximization problem, through which the performance trade-offs of different groups are characterized. Next, we study two benchmark cases under selfish caching, namely, partial cooperation and no cooperation, with and without inter-group file sharing, respectively. The former case of partial cooperation is investigated for demonstrating the impacts of selfish behaviours on both the group itself and the other groups. The latter case of no cooperation serves as the benchmark for assessing the performance gain due to

full/partial cooperation. The optimal caching distributions for these two cases are derived based on the KKT conditions. Finally, numerical examples are presented to validate the utility gain of different scenarios of cooperation compared with the case of no cooperation and show the impacts of selfish behaviours to cooperative caching.

The rest of this chapter is organized as follows. Section 3.2 presents the literature review. Section 3.3 introduces the system model for the cooperative local caching and the definition of the utility functions for different interest groups. Section 3.4 discusses the case of full cooperation among different groups of MTs. Section 3.5 discusses two benchmark cases of partial cooperation and no cooperation. Section 3.6 presents numerical examples to compare the performances of different caching schemes. Finally, Section 3.7 concludes this chapter.

3.2 Literature Review

Although traditional caching has long been proposed for reducing the download delay in the wired network [64], the performance of local devices only depends on the caching in the local server and there is no inter-dependence between them. In contrast, by cooperatively caching and sharing the files between the local devices, cooperative local caching utilizes the aggregate cache of the local devices in the whole network [65]. Depending on which part of network edge to cache the popular files, the literature of cooperative local caching can be reviewed on two lines of works, namely caching on the BS side and the MT side, respectively. Furthermore, the similarities, as well as the differences between cooperative local caching and wireless social network are also discussed.

3.2.1 Local Caching on the BS side

First, for local caching on the BS side [66–68], popular files are selected and cached in the BSs and the BSs cooperatively serve the MTs in the wireless cellular

network. [66] considers that BSs cooperatively serve MTs by deterministically caching files under the coded and uncoded scenarios. The drawback of this approach is that the location information of the MTs is assumed and the cache needs to be updated accordingly. In [67], an efficient caching protocol is proposed, which allows the MTs to dynamically select BSs that can serve the MTs and adaptively adjust the quality of transmission. [68] considers a broadcasting scheme where multiple BSs cooperatively serve MTs with multicast beamforming for reducing the backhaul data transmission in the CRAN network under given cached files in the BSs.

3.2.2 Local Caching on the MT side

In contrast, for local caching on the MT side [69–71, 73], files are pre-cached in the MTs during the off-peak hours and MTs share files with each other in the file delivery during the peak hours. [69] derives the power scaling law of the wireless network's capacity under cooperative local caching with respect to the range of the communications. [70] considers the probability of failing to accomplish the file sharing and the fairness issue among MTs by characterizing the capacity-outage trade-off. [71] extends the work on the BS side caching in [65] to the MT side by proposing both deterministic and random strategies for file placement under coded multicast. [73] further studies the effectiveness of cooperative local caching under different file request distributions, network sizes and cache sizes with multi-hop communications.

3.2.3 Local Caching and Wireless Social Network

Finally, there are also some works [74–76] in the literature discussing the file or information sharing in wireless social networks [77]. They exploit the fact that people who are socially connected are more likely to share similar interests or have similar mobility patterns, which can be exploited for the file and information sharing in the wireless network. [74] discusses the off-loading of mobile data traffic through opportunistic communications between neighbouring MTs. With the social

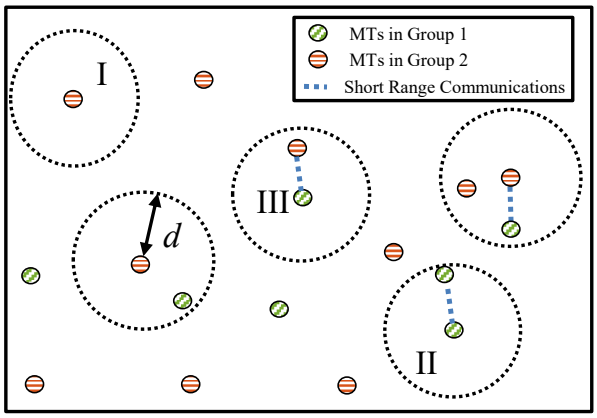


Figure 3.1: Cooperative local caching enabled file sharing among MTs in two groups. Different cases of file sharing: I. no file sharing, II. intra-group file sharing, and III. inter-group file sharing.

information between the mobile phones, a minimum target set is selected for the maximization of data off-loading. [75] exploits the social information to devise a social-network-assisted cooperation for the device-to-device communications. Also, with the social structure information, [76] develops a cooperative relaying framework in the wireless social network that reveals the trade-offs between cooperation gain and relay probing cost.

Our work and those on the wireless social network are similar in terms of the direct file or information sharing between MTs in the network. The difference is that, in the wireless social network, the social information between the MTs is required such that information or file sharing can achieve a better performance. While in cooperative local caching, we utilize the personal file preference information of the MTs. Moreover, we also investigate the heterogeneity in the file preference and the impacts of selfish behaviours on cooperation, which are not addressed in wireless social networking.

3.3 System Model

In this section, we introduce the system model of cooperative local caching. As illustrated in Fig. 3.1, we consider a large number of MTs served by the wireless

cellular network and these MTs can also leverage SRC (e.g., Bluetooth, LTE D2D, etc.) to share cached files (e.g., videos) with each other upon file requests during the peak hours. According to these MTs' file preferences, we practically categorize them into G groups, denoted by the set $\mathcal{G} = \{1, 2, \cdots, G\}$. Fig. 3.1 shows an example of two groups (i.e., $G = 2$) in the network and also illustrates different cases of file sharing between the MTs. We denote the MTs in each group $i \in \mathcal{G}$ by $\mathcal{K}_i$ and we assume that their locations follow a two-dimensional independent HPPP $\Phi_i = \{X_{i,k}\}$, $k \in \mathcal{K}_i$ [57] with average spatial density $\bar{\lambda}_i$, $i \in \mathcal{G}$ and the MTs' coordinates $X_{i,k} \in \mathbb{R}^2$, which are independent of the other groups.[15] By the superposition theorem of Poisson process [57], we also denote the *social density* of the MTs as the sum of all the groups' spatial densities, i.e., $\bar{\lambda}_0 = \sum_{i\in\mathcal{G}} \bar{\lambda}_i$. For a certain MT in each group, it can potentially acquire its requested files from its *neighbouring MTs* via SRC (in the same or different groups) within the distance d, as illustrated by the dotted circle in Fig. 3.1. Specifically, let $B(T, x) \in \mathbb{R}^2$ denote a disk of radius x centred at T. Then, considering a typical MT $k \in \mathcal{K}_i$ in group $i \in \mathcal{G}$, $B(X_{i,k}, d)$ denotes the area where file sharing is possible between the MT at location $X_{i,k} \in \Phi_k$ and all the other MTs both inside and outside the group i. Therefore, under the HPPP model, the average number of neighbouring MTs from group i for MT k in the area of $B(X_{i,k}, d)$ is $\bar{\mu}_i = \pi d^2 \bar{\lambda}_i$, $i \in \mathcal{G}$ and the average number of neighbouring MTs from all the groups in $B(X_{i,k}, d)$ is $\bar{\mu}_0 = \sum_{i\in\mathcal{G}} \bar{\mu}_i = \pi d^2 \bar{\lambda}_0$.

Cooperative local caching allows an MT to share the requested file locally with another MT (if any). However, this local file sharing highly depends on the caching strategies of all the G groups of MTs. We introduce the system model of the cooperative local caching in the following.

[15] It should be noted that the HPPP is applicable in a general setting and can serve as the spatial models for different scenarios. For example, HPPP has been used in modelling the spatial locations of the MTs in the wireless network [78] and the contact process between entities in the social network [79].

3.3.1 System Model of Cooperative Local Caching

During the off-peak hours, MTs cache files into their local memories. There are mainly two approaches for the file caching in the literature, namely *deterministic caching* (e.g., [66]) and *random caching* (e.g., [70]). Deterministic caching is generally the approach used in femto-caching [66], where the locations of the femto-BSs are fixed. However, this approach may not be suitable for cooperative local caching in our case due to the mobility of the MTs. In contrast to deterministic caching, random caching can easily address the mobility issues [70]. Therefore, we consider random caching as our caching scheme. We define the set of popular files as $\mathcal{F} = \{1, 2, \cdots, F\}$ and assume that each MT caches one file from the library $\mathcal{F}$.[16] It caches file $f \in \mathcal{F}$ based on independent random sampling from the *group caching distribution* $\boldsymbol{c}_i = [c_{i,1}, c_{i,2}, \cdots, c_{i,F}]$, which is defined as the PMF over the files $\mathcal{F}$ with

$$\sum_{f \in \mathcal{F}} c_{i,f} = 1, f \in \mathcal{F}, \tag{3.1}$$

$$0 \leq c_{i,f} \leq 1, \ i \in \mathcal{G}, \ f \in \mathcal{F}. \tag{3.2}$$

We also integrate the group caching distributions of the other groups $\mathcal{G} \setminus \{i\}$ as $\boldsymbol{c}_{-i} = [\boldsymbol{c}_1, \cdots, \boldsymbol{c}_{i-1}, \boldsymbol{c}_{i+1}, \cdots, \boldsymbol{c}_G]$. Then, given all the group caching distributions $(c_{i,f}, \ \forall i \in \mathcal{G}, \ f \in \mathcal{F})$ and the densities of the groups $(\bar{\lambda}_i, \ \forall i \in \mathcal{G})$, we denote the *social caching distribution* as the average group caching distributions with the weights equal to their normalized densities as

$$c_f = \frac{1}{\bar{\lambda}_0} \sum_{i \in \mathcal{G}} \bar{\lambda}_i c_{i,f}, \ f \in \mathcal{F}. \tag{3.3}$$

c_f has its physical meaning in our cooperative local caching scheme as it denotes the average availability of a certain file $f \in \mathcal{F}$ for a particular MT within unit area.

[16] We assume each MT only caches a single file for the tractability of analysis. However, our results can be extended to multi-file caching, which offers essentially the same insights.

During the peak hours, MTs make requests to the files based on their preferences over the files.[17] As we divide the MTs into different groups according to their interests, the MTs within the same group $\mathcal{K}_i$ have a common *group request distribution* $\boldsymbol{r}_i = [r_{i,1}, r_{i,2}, \cdots, r_{i,F}]$ with $\sum_{f\in\mathcal{F}} r_{i,f} = 1$. We assume all the files have positive request probabilities (i.e., $r_{i,f} > 0,\ \forall i \in \mathcal{G}$). In a certain realization, each MT in the group $\mathcal{K}_i,\ i \in \mathcal{G}$ makes an independent request to a certain file $f \in \mathcal{F}$ based on a random sampling from $\boldsymbol{r}_i$. We also denote $\boldsymbol{r}_{-i} = [\boldsymbol{r}_1, \cdots, \boldsymbol{r}_{i-1}, \boldsymbol{r}_{i+1}, \cdots, \boldsymbol{r}_G]$ as the aggregate group request distributions of the other groups $\mathcal{G}\setminus\{i\}$. Finally, for the MTs in all the groups, we denote the *social request distribution* as the average group request distributions with the weights equal to their normalized densities, that is

$$r_f = \frac{1}{\bar{\lambda}_0}\sum_{i\in\mathcal{G}} \bar{\lambda}_i r_{i,f},\ \ f \in \mathcal{F}. \tag{3.4}$$

Under the request of an MT in $\mathcal{K}_i,\ i \in \mathcal{G}$ for a particular file $f \in \mathcal{F}$, we are ready to specify the methods for file download.

1) *Obtaining the file from on-board cache:* If the file $f \in \mathcal{F}$ is already cached in its memory, the MT can obtain the file directly. If the file is not found, the MT requests the file from the neighbouring MTs;

2) *Obtaining the file with cooperative local caching:* The MT requests the file from its neighbouring MTs within the range d. The neighbouring MTs can be from its own group or the other groups. If multiple neighbouring MTs have the file, the MT chooses one randomly.

Based on the above-specified system model of the cooperative local caching, in the next subsection, we introduce the definition of utility function for measuring the performance of each group.

[17] In practice, this can be accomplished by defining the measure of similarity between preferences and then grouping the MTs, whose similarities of preferences are within a certain threshold, such as in [72].

3.3.2 Definition of Utility for Cooperative Local Caching

With the above-specified transmission protocol, we are ready to define the utility function for characterizing the QoS for all the groups. For quantifying their performances, there may be various methods for the definitions of the utility functions. In cooperative local caching, we observe that each MT should try to increase the probability of *successful file discovery* on its own cache or its neighbouring MTs, which can benefit the MT or the network in various aspects. Some high-level examples of such benefits are given as follows:

1) *Delay reduction of the file download:* The delay of downloading from an MT in the range of local file sharing is usually much lower than that from the remote servers [66]. Hence, higher probability of file discovery results in a lower delay in the file download, which is beneficial for the QoS of the MTs.

2) *Bandwidth usage and backhaul capacity reduction:* Due to short range of the file sharing, cooperative local caching can effectively reduce the bandwidth usage [22]. Moreover, cooperatively sharing the file locally reduces the need for high-capacity backhaul in the wireless network [68]. Hence, higher file discovery probability reduces the bandwidth and backhaul usage, which is beneficial for the QoS of the whole network.

Based on the above discussions, we give two definitions of the utility function in the following.

Group Utility

First, we define the utility of each group as the probability of successful file discovery for the MTs in the group. We define the event that, when requesting a certain file, the file is found in its own cache or that of more than one neighbouring MT as E. Then, the probability of successful file discovery for the MTs in group i under the caching distributions $\boldsymbol{c}_i$ and those of all the other groups $\boldsymbol{c}_{-i}$ is defined as $\mathbb{P}(E; \boldsymbol{c}_i, \boldsymbol{c}_{-i})$. Now, we are ready to specify the *group utility* of the MTs in group

$i \in \mathcal{G}$ as the probability of successful file discovery in its own cache as well as any of the other MTs in the range of d as

$$U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i}) = \mathbb{P}(E; \boldsymbol{c}_i, \boldsymbol{c}_{-i}). \tag{3.5}$$

With the following theorem, we provide the closed-form expression for the group utility function.

Theorem 3.3.1 *Given the group caching distribution $\boldsymbol{c}_i$ and those of the other groups $\boldsymbol{c}_{-i}$, the utility for the MTs in group $i \in \mathcal{G}$ is*

$$U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i}) = \sum_{f \in \mathcal{F}} r_{i,f} \left(1 - c'_{i,f} e^{-\bar{\mu}_0 c_f}\right), \tag{3.6}$$

where $e^{-\bar{\mu}_0 c_f}$ with c_f given in (3.3), denotes the probability that file $f \in \mathcal{F}$ is not found in any neighbouring MT within the range d, $r_{i,f}$ is defined in (3.4), $c'_{i,f} = 1 - c_{i,f}$ denotes the complement of the group caching distribution and $\bar{\mu}_0$ is defined as $\bar{\mu}_0 = \pi d^2 \bar{\lambda}_0$.

Proof See Appendix C.

Remark 3.3.1 *From (3.6) in Theorem 3.3.1, we can observe that, for a certain group $i \in \mathcal{G}$, its utility relies on its own group request distribution $r_{i,f}$ and caching distributions $c_{i,f}$, as well as the social density $\bar{\lambda}_0$ ($\bar{\mu}_0 = \pi d^2 \bar{\lambda}_0$) and the social caching distribution c_f, but regardless of the specific caching distribution of the other groups $\boldsymbol{c}_{-i}$ or their spatial densities. We can also observe that, for a certain file $f \in \mathcal{F}$, the group utility $U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i})$ for the MTs in group $\mathcal{K}_i$ monotonically increases with respect to the group caching distribution $c_{i,f}$ and the social caching distribution c_f. This is because if the MT or its neighbouring MTs cache the file with higher probability, the requested file is more available to the neighbouring MT and the utility increases. Moreover, from (3.6), the effect of the file not cached in its own MT (i.e., $c'_{i,f}$) is exponentially vanishing with respect to the social density $\bar{\lambda}_0$ and caching*

distribution c_f. This shows how the neighbouring MTs' caches and the MT's own cache compliment with each other under our mobility model.

Social Utility

Apart from the above definition of group utility, another definition of the utility function is based on the performance of the whole network, which is defined as the weighted-sum of each group's utility defined in (3.5) with respect to the normalized density of each group in the network as

$$U(\{\boldsymbol{c}_i\}) = \frac{1}{\bar{\lambda}_0} \sum_{i \in \mathcal{G}} \bar{\lambda}_i U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i}). \tag{3.7}$$

The utility function reflects the relative importance of the groups in the whole society. Hence, this definition of utility can be termed as the *social utility* of the network.

Next, in order to explicitly show the utility trade-offs between different groups under cooperative local caching, we define the feasible utility region of different groups with the above-defined utility functions.

3.3.3 Feasible Utility Region

For our proposed cooperative local caching scheme, we intend to increase the utilities of all the groups such that their QoSs can be improved. However, different groups have conflict of interests under heterogeneous file preferences. For example, consider a certain file that is not popular in one group but popular in the other groups. If the former group caches this file with high probability, the utilities of the other groups will increase because of the file's popularity. However, this will reduce the probabilities for the former group to cache popular files and thus decrease its utility. In order to explicitly characterize such conflict of interests, we define the *feasible utility region* of all the groups as follows.

Definition 3.3.1 *The feasible utility region of different groups is the set of utility*

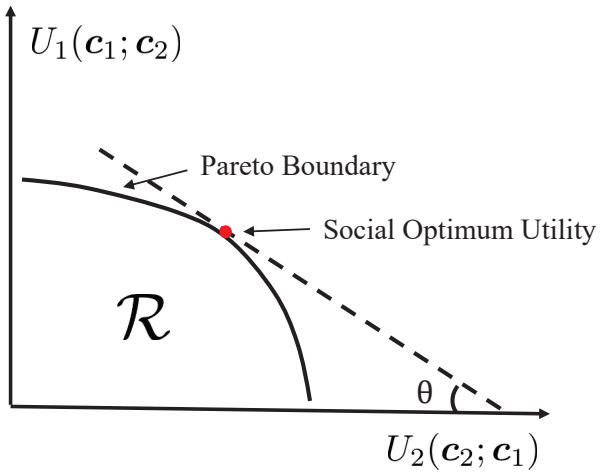


Figure 3.2: Feasible utility region, Pareto boundary and social optimum with slope $\theta = \arctan\frac{\bar{\lambda}_1}{\bar{\lambda}_2}$ for two groups.

tuples that different groups can achieve simultaneously, which is given as

$$\mathcal{R} \triangleq \bigcup \left\{ (u_1, u_2, \cdots, u_G) : \ u_i \leq U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i}), \ i \in \mathcal{G} \right\}. \tag{3.8}$$

This region denotes the *Pareto region* of all the possible utilities under different allocations of group caching distributions $\boldsymbol{c}_i$, $i \in \mathcal{G}$. An example of the feasible utility region is given in Fig. 3.2, which shows the trade-off between the utilities of two groups ($G = 2$). Within the feasible utility region, the *Pareto optimality* and *Pareto boundary* are defined as follows.

Definition 3.3.2 *A utility tuple $(u_1, u_2, \cdots, u_G) \in \mathcal{R}$ is Pareto optimal if there is no other utility tuple $(u'_1, u'_2, \cdots, u'_G) \in \mathcal{R}$ with $(u'_1, u'_2, \cdots, u'_G) \geq (u_1, u_2, \cdots, u_G)$, where the inequality is component-wise. The set of tuples in the utility region that are Pareto optimal are called the Pareto boundary of the feasible utility region $\mathcal{R}$.*

In order to characterize the complete trade-offs between different groups of MTs, in the following section, we investigate the scenario of full cooperation under centralized coordination in our cooperative local caching scheme.

3.4 Fully Cooperative Local Caching

In this section, we consider the case where the MTs in different groups $\mathcal{K}_i,\ i \in \mathcal{G}$ are *fully cooperative* and follow a centralizer's instructions to cache the files for the benefits of the whole society. It should be noted that the preferences of the MTs in each group can be readily available on the service provider's platform (such as YouTube, Netflix, etc.). The optimization of the group caching distribution can also be easily performed in the centralizer where the computation resources are abundant. In order to completely characterize the feasible utility region $\mathcal{R}$ defined in (3.8), we first define the weighted-sum utility of different groups as

$$U(\{\boldsymbol{c}_i\}, \boldsymbol{\gamma}) = \sum_{i \in \mathcal{G}} \gamma_i U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i}), \tag{3.9}$$

where $U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i})$ is the utility for group i defined in (3.6) and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \cdots, \gamma_G]$, with the weights $\gamma_i \geq 0,\ i \in \mathcal{G}$ subjected to $\sum_{i \in \mathcal{G}} \gamma_i = 1$, represents the relative importance of different groups with respect to their densities. More specifically, given group i's utility, the social utility defined in (3.7) can be obtained by equating weight γ_i to group i's normalized spatial density (i.e., $\gamma_i = \bar{\lambda}_i / \bar{\lambda}_0$). We then propose the following weighted-sum utility maximization problem for deciding the $\boldsymbol{c}_i$'s for all the groups.

$$\begin{aligned} (\text{P3}): \ \underset{\{\boldsymbol{c}_i\}}{\texttt{max.}} \ & U(\{\boldsymbol{c}_i\}, \boldsymbol{\gamma}) \\ \texttt{s.t.} \ & \sum_{f \in \mathcal{F}} c_{i,f} = 1,\ i \in \mathcal{G}, && \text{(3.10a)} \\ & 0 \leq c_{i,f} \leq 1,\ f \in \mathcal{F},\ i \in \mathcal{G}. && \text{(3.10b)} \end{aligned}$$

Among all the Pareto optimal utility tuples, the *social optimal utility* is defined as the optimal solution to Problem (P3) with $\gamma_i = \bar{\lambda}_i / \bar{\lambda}_0,\ i \in \mathcal{G}$ in the objective function $U(\{\boldsymbol{c}_i\}, \boldsymbol{\gamma})$, which corresponds to the social utility in (3.7). An example is shown in Fig. 3.2 for the case of $G = 2$, where the social optimum is the intersection

between the line with slope arctan $\frac{\bar{\lambda}_1}{\bar{\lambda}_2}$ and the Pareto boundary.

However, it can be verified that the objective function in Problem (P3) is non-concave due to the coupling terms in $c_f = \frac{1}{\bar{\lambda}_0} \sum_{i \in \mathcal{G}} \bar{\lambda}_i c_{i,f}$ and thus the global optimum solution is hard to obtain. Although the problem is not jointly concave with respect to all the $c_{i,f},\ i \in \mathcal{G},\ f \in \mathcal{F}$, it can be proved that the objective function $U(\{\boldsymbol{c}_i\}, \boldsymbol{\gamma})$ is marginally concave with respect to the caching distribution $\boldsymbol{c}_i$ for a certain group $i \in \mathcal{G}$. This is shown in the following lemma.

Lemma 3.4.1 *The weighted-sum utility $U(\{\boldsymbol{c}_i\}, \boldsymbol{\gamma})$ in Problem* (P3) *is marginally concave with respect to the group caching distribution $\boldsymbol{c}_i$ given the group caching distributions of the other groups $\boldsymbol{c}_{-i}$.*

Proof See Appendix D.

The above result on the Hessian matrix of the objective function not only shows its concavity, but also reveals the decomposition structure of the problem with respect to $c_{i,f},\ i \in \mathcal{G},\ f \in \mathcal{F}$. Based on this, we propose a coordinate descent algorithm [80] for Problem (P3), which sequentially optimizes the weighted-sum utility function with respect to different groups.

For the coordinate descent algorithm, special attentions need to be given to those groups with zero weights. We thus divide the whole set of groups $\mathcal{G}$ into two sets $\mathcal{G}_+$ and $\mathcal{G}_0$ with $\mathcal{G}_+ \cup \mathcal{G}_0 = \mathcal{G}$ and $\mathcal{G}_+ \cap \mathcal{G}_0 = \emptyset$. $\mathcal{G}_+ \subseteq \mathcal{G}$ denotes those groups with positive weights such that $\gamma_i > 0,\ \forall i \in \mathcal{G}_+$ and $\sum_{i \in \mathcal{G}_+} \gamma_i = 1$; while $\mathcal{G}_0 \subset \mathcal{G}$ denotes those groups with zero weights such that $\gamma_i = 0, \forall\ i \in \mathcal{G}_0$.

First, we consider the groups with positive weights $i \in \mathcal{G}_+$. Given the caching distributions of the other groups $\boldsymbol{c}_{-i}$, for a certain group $i \in \mathcal{G}_+$, we aim to solve for

the following problem.

$$(\mathrm{P}4-1): \max_{\boldsymbol{c}_i} \sum_{j\in\mathcal{G}_+}\gamma_j\sum_{f\in\mathcal{F}} r_{j,f}\left(1-c'_{j,f}e^{-\bar{\mu}_0 c_f}\right)$$

$$\text{s.t.} \sum_{f\in\mathcal{F}} c_{i,f}=1, \tag{3.11a}$$

$$0\le c_{i,f}\le 1,\ f\in\mathcal{F}. \tag{3.11b}$$

Thanks to the marginal concavity shown in Lemma 3.4.1, we derive the optimal solution in the following theorem by leveraging its KKT conditions [80].

Theorem 3.4.1 *The optimal group caching distribution of group $i\in\mathcal{G}$ in Problem $(\mathrm{P}4-1)$ is*

$$c^*_{i,f}=\left[1-\left(\frac{1}{\bar{\mu}_i}W\left(A_{i,f}e^{\bar{\mu}_i B_{i,f}}\right)-B_{i,f}\right)\right]_0^1, \tag{3.12}$$

with

$$A_{i,f}=\frac{\nu^*_i}{\gamma_i e^{-\bar{\mu}_0}r_{i,f}e^{\sum_{j\in\mathcal{G}\setminus\{i\}}\bar{\mu}_j c'_{j,f}}}, \tag{3.13}$$

$$B_{i,f}=\frac{1}{\bar{\mu}_i}+\sum_{j\in\mathcal{G}_+\setminus\{i\}}\frac{\gamma_j r_{j,f}}{\gamma_i r_{i,f}}c'_{j,f}, \tag{3.14}$$

where $[\cdot]_0^1=\min\{1,\max\{0,\cdot\}\}$, $W(\cdot)$ is the Lambert-W function [81] and ν^*_i, which is the optimal dual variable of constraint (3.11a), denotes a constant that satisfies $\sum_{f\in\mathcal{F}}c^*_{i,f}=1$.

Proof See Appendix E.

Next, for the groups $i\in\mathcal{G}_0$ with zero weights, they cache completely un-selfishly for the groups with positive weights in $\mathcal{G}_+$. In this case, the weighted-sum utility given the caching distributions $\boldsymbol{c}_{-i}$ of the other groups reduces to

$$\sum_{i\in\mathcal{G}}\gamma_i U_i(\boldsymbol{c}_i;\boldsymbol{c}_{-i})=1-\sum_{f\in\mathcal{F}}e^{\bar{\mu}_i c'_{i,f}+D_{i,f}}, \tag{3.15}$$

where $D_{i,f}$ is given by

$$D_{i,f} = \ln\left[\left(\sum_{j\in\mathcal{G}_+}\gamma_j r_{j,f} c'_{j,f}\right) e^{-\bar{\mu}_0} e^{\sum_{l\in\mathcal{G}\setminus\{i\}}\bar{\mu}_l c'_{l,f}}\right]. \tag{3.16}$$

Then, we need to solve the following problem.

$$\begin{aligned}(\mathrm{P4}-2):\ \max_{\boldsymbol{c}_i}\ & 1-\sum_{f\in\mathcal{F}} e^{\bar{\mu}_i c'_{i,f}+D_{i,f}} \\ \text{s.t.}\ & \sum_{f\in\mathcal{F}} c_{i,f}=1, && (3.17a)\\ & 0\le c_{i,f}\le 1,\ f\in\mathcal{F}. && (3.17b)\end{aligned}$$

Because the function e^x is convex and linear combination preserves convexity [80], the objective in Problem $(\mathrm{P4}-2)$ is a concave function. The constraints are also affine. Therefore, Problem $(\mathrm{P4}-2)$ is a convex optimization problem and its optimum can be obtained in the following theorem.

Theorem 3.4.2 *The optimal solution to Problem* $(\mathrm{P4}-2)$ *is given by*

$$c^*_{i,f} = \left[1-\frac{1}{\bar{\mu}_i}\left(\ln\left(\frac{\nu^*_i}{\bar{\mu}_i}\right)-D_{i,f}\right)\right]_0^1, \tag{3.18}$$

where ν^*_i *is the optimal dual variable for constraint (3.17a), which can be obtained by substituting the result in (3.18) into the equation* $\sum_{f\in\mathcal{F}} c^*_{i,f}=1$.

Proof The proof is similar to that of Theorem 3.4.1 and is thus omitted here.

Now, we have obtained the optimal solutions to problems $(\mathrm{P4}-1)$ and $(\mathrm{P4}-2)$ for the groups with positive and zero weights, respectively. Based on the above results, we propose Algorithm III in Table 3.1 for Problem (P3) based on the coordinate descent method [80]. Because the feasible set of the problem is a compact set and the objective function is lower-bounded, the algorithm is guaranteed to converge to at least a local maximum of Problem (P3).

Table 3.1: Algorithm III: Coordinate descent algorithm for full cooperation.

1. **Repeat for** group $i = 1, 2, \cdots, G$ in iterations $m = 1, 2, \cdots$
   1) **Initialize:** $\nu_i^{(l)} := 0$, $\nu_i^{(h)} := \infty$;
   2) **Repeat:**
      i. $\nu_i := \frac{1}{2}(\nu_i^{(l)} + \nu_i^{(h)})$;
      ii. **If** $i \in \mathcal{G}_+$, update $c_{i,f}^*$ according to (3.12) in Theorem 3.4.1;
      If $i \in \mathcal{G}_0$, update $c_{i,f}^*$ according to (3.18) in Theorem 3.4.2;
      iii. **If** $\sum_{f \in \mathcal{F}} c_{i,f}^* < 1$, set $\nu_i^{(h)} := \nu_i$; **Else**, set $\nu_i^{(l)} := \nu_i$.
   3) **Until:** the condition $|\nu_i^{(l)} - \nu_i^{(h)}| > \delta_\nu$ is violated.
2. **Until:** $|U^m(\{\boldsymbol{c}_i^*\}) - U^{m-1}(\{\boldsymbol{c}_i^*\})| \leq \delta_U$.

3.5 Partially Cooperative and Non-cooperative Local Caching

In the previous section, we have proposed a centralized algorithm under full cooperation to achieve the approximate Pareto boundary of the feasible utility region. While different groups may be fully cooperative and their caching decisions can be jointly optimized, this scheme does not apply to the scenario where different groups have selfish behaviours. In this section, we discuss two possible benchmark schemes by investigating some groups' selfish behaviours under *partial cooperation* or *no cooperation*. We first consider the partially cooperative case to access the impacts of the selfish behaviours in cooperative local caching. In this case, *inter-group file sharing* (illustrated in Fig. 3.1) is still allowed, while different groups are not willing to achieve a common social utility as in the case of full cooperation and the caching decisions are made independently by each group. Next, we consider the non-cooperative caching as a performance benchmark for the cases of full/partial cooperation. In this case, there is only *intra-group file sharing* (illustrated in Fig. 3.1) but no longer inter-group file sharing.

3.5.1 Partially Cooperative Caching with Inter-group File Sharing

We assume that each group only knows its own group's file request distribution $\boldsymbol{r}_i$ and the social file request distribution $r_f, f \in \mathcal{F}$, which can be publicly accessed by all the groups when making caching decisions.[18] Based on the above information, each group makes its own decision in caching. Different groups of MTs are partially cooperative such that they are able to share files with each other. However, they are also selfish in their own caching strategies due to the heterogeneous file preferences and aim to increase their own utilities. In this case, not knowing the exact caching distribution of the other groups, it is reasonable for each group to assume that the other groups are faithful to their own preferences and they cache files according to their preferences, (i.e., $\boldsymbol{c}_{-i} = \boldsymbol{r}_{-i}$). Then, the group utility $U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i})$ in (3.6) reduces to

$$U_i(\boldsymbol{c}_i; \boldsymbol{r}_{-i}) = \sum_{f \in \mathcal{F}} r_{i,f} \left(1 - c'_{i,f} e^{-\bar{\mu}_0 r_f + \bar{\mu}_i r_{i,f} - \bar{\mu}_i c_{i,f}}\right). \tag{3.19}$$

From the above result, we can see that, under the assumption of $\boldsymbol{c}_{-i} = \boldsymbol{r}_{-i}$ for a certain group i, it does not need to know the actual file request distributions of the other groups $\boldsymbol{r}_j,\ j \neq i$ to obtain its own utility $U_i(\boldsymbol{c}_i; \boldsymbol{r}_{-i})$, but only needs to know the aggregate social request distribution $r_f,\ f \in \mathcal{F}$ defined in (3.4).

Given the above derivation of the utility of group i, we aim to answer the following question: how should a group optimally cache files to maximize its own utility by exploiting the social file preference information r_f? We formulate the above decision making for group i as the following optimization problem that maximizes

[18] This is possible when the MTs are under the same platform (such as YouTube, Netflix, etc.), where the statistics of the preference of the whole society can be publicly available. However, due to the issues of privacy, the exact preference of a certain group may not be available to the other groups.

its utility given the caching distributions of the other groups as $\boldsymbol{c}_{-i} = \boldsymbol{r}_{-i}$.

$$\begin{aligned} (\text{P5}): \quad \max_{\boldsymbol{c}_i} \quad & U_i(\boldsymbol{c}_i; \boldsymbol{r}_{-i}) \\ \text{s.t.} \quad & \sum_{f\in\mathcal{F}} c_{i,f} = 1, && (3.20a) \\ & 0 \le c_{i,f} \le 1, \ f \in \mathcal{F}. && (3.20b) \end{aligned}$$

The objective function is concave and this can be verified by checking Lemma 3.4.1 with $\gamma_j = 0, \ \forall j \in \mathcal{G} \setminus \{i\}$ and $\boldsymbol{c}_{-i} = \boldsymbol{r}_{-i}$. With all the constraints of this problem being affine, Problem (P5) is a convex optimization problem and it is always feasible. In the following theorem, we provide the closed-form solution for Problem (P5) by checking the KKT conditions.

Theorem 3.5.1 *Group i's optimal group caching distribution for file $f \in \mathcal{F}$ under the partial cooperation is*

$$c_{i,f}^* = \left[1 - \frac{1}{\bar{\mu}_i}\left\{W\left(\frac{\nu_i^* e^{1+\bar{\mu}_0}}{r_{i,f} e^{\bar{\mu}_0 r_f' - \bar{\mu}_i r_{i,f}'}}\right) - 1\right\}\right]_0^1. \tag{3.21}$$

where $r_f' = 1 - r_f$, $r_{i,f}' = 1 - r_{i,f}$ and ν_i^*, which is the optimal dual variable of constraint (3.20a), denotes a constant that satisfies $\sum_{f\in\mathcal{F}} c_{i,f}^* = 1$.

Proof The proof is similar to that of Theorem 3.4.1 and is thus omitted here.

In Theorem 3.5.1, the allocation of caching distribution can be interpreted as water-filling over different files with ν_i^* being the optimal water-level satisfying (3.20a). Also, please be noted that in the above partial cooperation scheme, group $i \in \mathcal{G}$ only needs to know about the social file request distribution $r_f, \ f \in \mathcal{F}$ instead of the specific request distributions of all the other groups $\boldsymbol{r}_j, \ j \in \mathcal{G} \setminus \{i\}$. Hence, the amount of information needed in partial cooperation is manageable in the individual decision of each group. Based on the above optimal solution, we propose the algorithm based on the bisection method [80] for Problem (P5) in the case of partial cooperation in Table 3.2.

Table 3.2: Algorithm IV: Optimal algorithm for partial cooperation.

1. **Initialize:** $\nu_i^{(l)} \coloneqq 0$, $\nu_i^{(h)} \coloneqq \infty$;
2. **Repeat:**
   1) $\nu_i \coloneqq \frac{1}{2}(\nu_i^{(l)} + \nu_i^{(h)})$;
   2) $c_{i,f} = \left[1 - \frac{1}{\bar{\mu}_i}\left\{W\left(\frac{\nu_i}{r_{i,f} e^{\bar{\mu}_0 r'_f - \bar{\mu}_i r'_{i,f}}}\right) - 1\right\}\right]_0^1$,
   3) **If** $\sum_{f\in\mathcal{F}} c_{i,f} < 1$, set $\nu_i^{(h)} \coloneqq \nu$; **Else**, set $\nu_i^{(l)} \coloneqq \nu$.
3. **Until:** the condition $|\nu_i^{(l)} - \nu_i^{(h)}| > \delta_\nu$ is violated.

In the following, we give illustrative numerical examples to show the impacts of different system parameters on the optimal group caching distribution $c_{i,f}^*$ in Theorem 3.5.1 for the case of partial cooperation.

Example 3.5.1 *The settings for this example are given as follows: there are $G = 2$ groups and $F = 5$ files in total. The densities of MTs in groups 1 and 2 are $\bar{\lambda}_1 = 0.05$ and $\bar{\lambda}_2 = 0.05$ MTs per unit area, respectively. The range of communications is $d =$ 10 m[19]. We separately examine the effects of one of the system parameters $r_{1,f}$ and $r_{2,f}$ on the optimal caching distribution $c_{1,f}^*$ of group 1 in Theorem 3.5.1 while keeping the others equal across the 5 files. First, for the equal values of these parameters across the 5 files: The group request distributions are uniform distributions with $r_{1,f} = 1/5,\ \forall f \in \mathcal{F}$ and $r_{2,f} = 1/5,\ \forall f \in \mathcal{F}$. Next, for the system parameters whose effects we want to examine: The PMFs $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are randomly generated with $\sum_{f\in\mathcal{F}} r_{1,f} = 1$ and $\sum_{f\in\mathcal{F}} r_{2,f} = 1$ for all the files $f \in \mathcal{F}$. The result is shown in Fig. 3.3. It can be observed that the optimal caching distribution $c_{1,f}^*$ has water-filling structure [82] with respect to the request distributions of group 1 and group 2: $c_{1,f}^*$'s are monotonically increasing and decreasing in $r_{1,f}$ and $r_{2,f}$, respectively. This is because if the group request probability $r_{1,f}$ is high, it is desirable to match the caching*

[19]For example, the typical range of Class 2 Bluetooth is $5 - 10$ m [8].

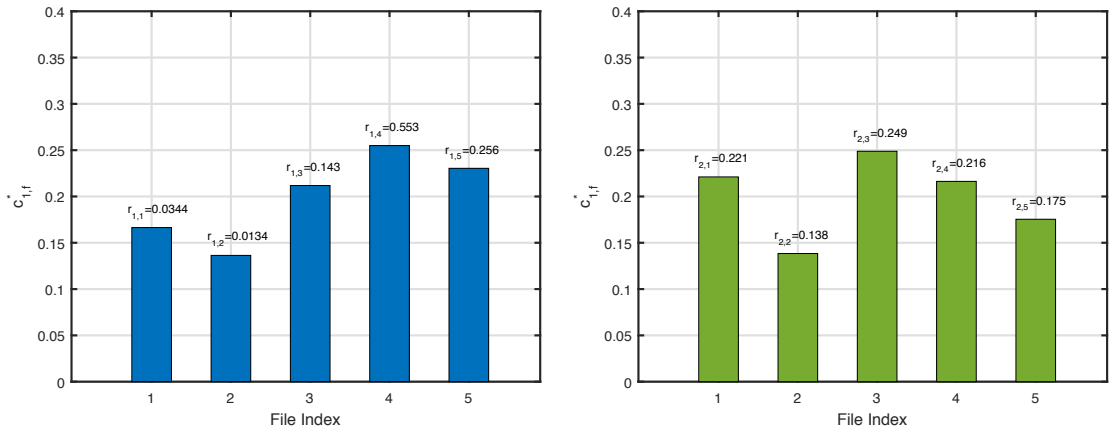


Figure 3.3: Group 1's optimal caching distribution $c^*_{1,f}$ under (a) group 1's request distribution $r_{1,f}$, and (b) group 2's request distribution $r_{2,f}$.

distribution to the request distribution to increase the utility. On the other hand, if group 2 caches the file with high probability due to high $r_{2,f}$, there is less need for group 1 to cache the file and group 1 can exploit group 2 by caching the file with a lower probability.

3.5.2 Non-cooperative Caching without Inter-group File Sharing

Next, we consider the benchmark case of no cooperation, where different groups cannot share files with the other groups, but file sharing is still possible inside the group. This situation can be possible when different groups of MTs belong to different self-enclosed social communities, such as families, schools, etc., and can only share files within the group. With this benchmark case, we can examine the gain due to inter-group cooperation, either in the case of full cooperation or partial cooperation. The optimal caching distribution of a certain group $i \in \mathcal{G}$ can be seen as a special case of Theorem 3.5.1 with $\bar{\lambda}_i > 0$ and $\bar{\lambda}_j = 0$, $j \in \mathcal{G} \setminus \{i\}$, which is equivalent to the result when there is only one group in the society in the

case of full/partial cooperation. The optimal caching distribution is specified in the following corollary.

Corollary 3.5.1 *The optimal group caching distribution for group $i \in \mathcal{G}$ in the non-cooperative local caching scheme is*

$$c_{i,f}^{*} = \left[1 - \frac{1}{\bar{\mu}_i}\left\{W\left(\frac{\nu_i^{*} e^{1+\bar{\mu}_i}}{r_{i,f}}\right) - 1\right\}\right]_0^1, \tag{3.22}$$

where ν_i^{*} denotes a constant that satisfies $\sum_{f\in\mathcal{F}} c_{i,f}^{*} = 1$.

Please be noted that the above result is also optimal for the special case of homogeneous file preference (i.e., $\boldsymbol{r}_i = \boldsymbol{r}_j,\ \forall j, i \in \mathcal{G}$) where all the MTs belong to one group.

3.6 Numerical Results

In this section, we first validate the performance of the system with various parameters under two groups. We then proceed to the network simulation under the mobility of the MTs to show the evolution of the system. The general simulation setup is given in the following. For the group request distributions, we assume that they follow Zipf distribution as in [83]

$$r_{i,f} = \frac{f^{-\gamma_r}}{\sum_{i=1}^{F} i^{-\gamma_r}},\ f \in \mathcal{F},\ i \in \mathcal{G}, \tag{3.23}$$

where γ_r is denoted as the *Zipf exponent*, measuring the concentration of the file popularity. The typical Zipf exponent with variations in different sources is around $\gamma_r = 0.9$ according to the measurements in [83]. Hence, we set $\gamma_r = 0.9$ for all the following simulations unless specified otherwise. The range of communications is $d = 10$ m and the total number of files is $F = 100$.

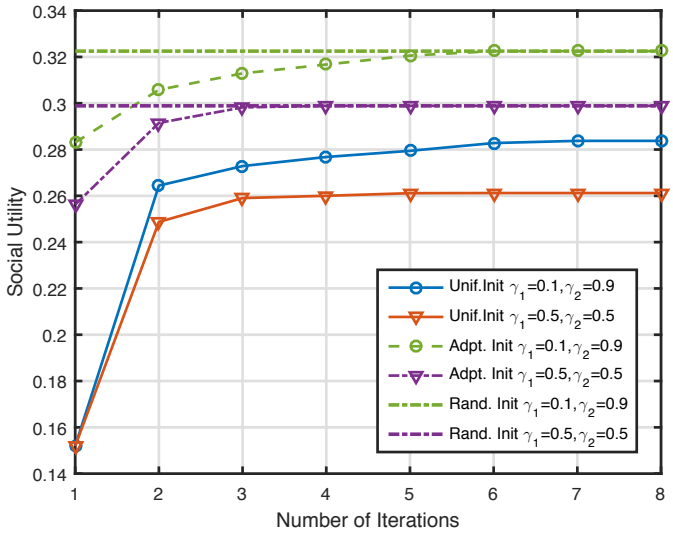


Figure 3.4: Convergence of the coordinate descent algorithm under various initialization methods.

3.6.1 Simulation under Two Groups

In this sub-section, we provide numerical results for evaluating the performance of cooperative local caching under full, partial and no cooperation with two groups. The group request distribution $\boldsymbol{r}_1$ of group 1 follows the Zipf distribution in (3.23) and $\boldsymbol{r}_2$ of group 2 follows the same distribution, except that it is a random permutation of $\boldsymbol{r}_1$ across the files.

Convergence of the Coordinate Descent Algorithm under Various Initializations

First, we show the convergence of the Algorithm III in Table 3.1. For the approximation algorithm based on coordinate descent, although the convergence can be guaranteed, different initializations may lead to different local maximums due to the non-convexity of the problem. Hence, we investigate the impacts of different initialization methods, which are specified in the following.

1. **Uniform Initialization** *(Unif. Init.)*: In uniform initialization, the initial caching distribution is uniform across the files regardless of their own or the

other groups' request distributions: $\boldsymbol{c}_1 = \frac{1}{F}\mathbf{1}_F^T,\ \boldsymbol{c}_2 = \frac{1}{F}\mathbf{1}_F^T$, where $\mathbf{1}_F$ denotes a vector of all ones with length F.

2. **Adaptive Initialization** *(Adapt. Init.)*: In adaptive initialization, the basic idea is to give the groups with higher weights more favourable initial distributions: $\boldsymbol{c}_1 = \boldsymbol{r}_1,\ \boldsymbol{c}_2 = \boldsymbol{r}_1$, if $0.66 \leq \gamma_1 \leq 1$, $\boldsymbol{c}_1 = \boldsymbol{r}_1,\ \boldsymbol{c}_2 = \boldsymbol{r}_2$, if $0.33 \leq \gamma_1 < 0.66$ and $\boldsymbol{c}_1 = \boldsymbol{r}_2,\ \boldsymbol{c}_2 = \boldsymbol{r}_2$, if $0 \leq \gamma_1 < 0.33$.

3. **Randomized Initialization** *(Rand. Init.)*: In randomized initialization, the initial distribution is randomly generated with $\sum_{f\in\mathcal{F}} c_{1,f} = 1$ and $\sum_{f\in\mathcal{F}} c_{2,f} = 1$. Algorithm under randomized initialization is performed 20,000 times and the highest converged utility is shown.

The simulation results are shown in Fig. 3.4 with two different sets of weights $\gamma_1 = 0.1$, $\gamma_2 = 0.9$ and $\gamma_1 = 0.5$, $\gamma_2 = 0.5$. It is observed that the algorithm converges under all cases. While, the algorithm with adaptive initialization offers a higher social utility compared with uniform initialization and it has the same result as randomized initialization (with the best result shown in the figure). This is because the adaptive initialization provides initial points that are closer to the potential optimum. Therefore, for all of the simulations in the following, adaptive initialization is used in the coordinate descent algorithm for the optimization in the scheme of full cooperation.

Group Utility with Partial Cooperation

Next, we study the effects of selfish behaviours of an individual group on both itself and the other groups under partial cooperation. We examine two different kinds of behaviours of the groups in partial cooperation, namely, *selfish caching* and *un-selfish caching*. For the selfish group, its behaviour is defined in Section 3.5.1 in partial cooperation, which is to take advantage of the other groups with the knowledge of the social preference $r_f,\ f \in \mathcal{F}$. In contrast, for the unselfish group, it faithfully caches files according to their request distribution (i.e., $\boldsymbol{c}_i = \boldsymbol{r}_i$). We

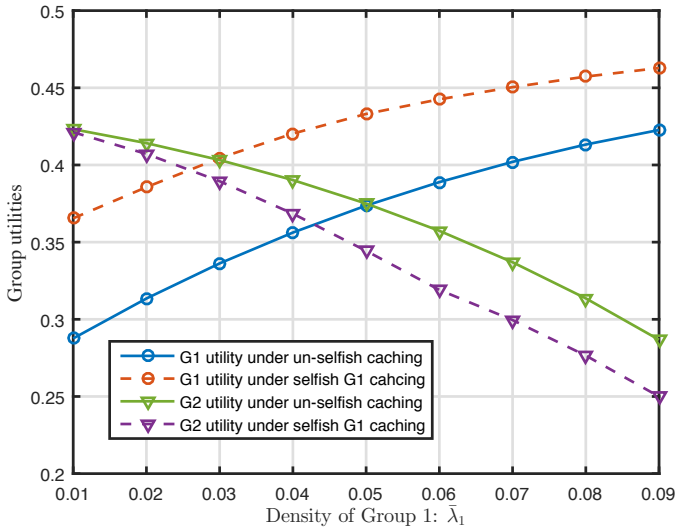


Figure 3.5: Group utilities in partial cooperation under different densities.

consider the following utilities for evaluating the performances of group 1 and 2.

1. **Group 1 utility under un-selfish caching**: $U_1(\boldsymbol{r}_1; \boldsymbol{r}_2)$;
2. **Group 2 utility under un-selfish caching**: $U_2(\boldsymbol{r}_2; \boldsymbol{r}_1)$;
3. **Group 1 utility under selfish group 1 caching**: $U_1(\boldsymbol{c}_1^*; \boldsymbol{r}_2)$;
4. **Group 2 utility under selfish group 1 caching**: $U_2(\boldsymbol{r}_2; \boldsymbol{c}_1^*)$.

The simulation result is shown in Fig. 3.5 by varying the density $\bar{\lambda}_1$ of group 1, while keeping the sum density of the two groups $\bar{\lambda}_1 + \bar{\lambda}_2 = \bar{\lambda}_0$ fixed. It can be observed that the utilities of both groups monotonically increase with their own densities. This is obvious because higher group density leads to larger probability for file sharing within the group of MTs with the same file preference. It can also be observed that, compared with the utilities of unselfish groups 1 and 2 ($U_1(\boldsymbol{r}_1; \boldsymbol{r}_2)$ and $U_2(\boldsymbol{r}_2; \boldsymbol{r}_1)$), the utility $U_1(\boldsymbol{c}_1^*; \boldsymbol{r}_2)$ of selfish group 1 increases, while the utility $U_2(\boldsymbol{r}_2; \boldsymbol{c}_1^*)$ of the un-selfish group 2 decreases. This shows that if a selfish group can exploit the social information r_f, it can indeed increase its own utility, while this may decrease the utility of the other un-selfish group. Hence, the simulation results in Fig. 3.5 verify

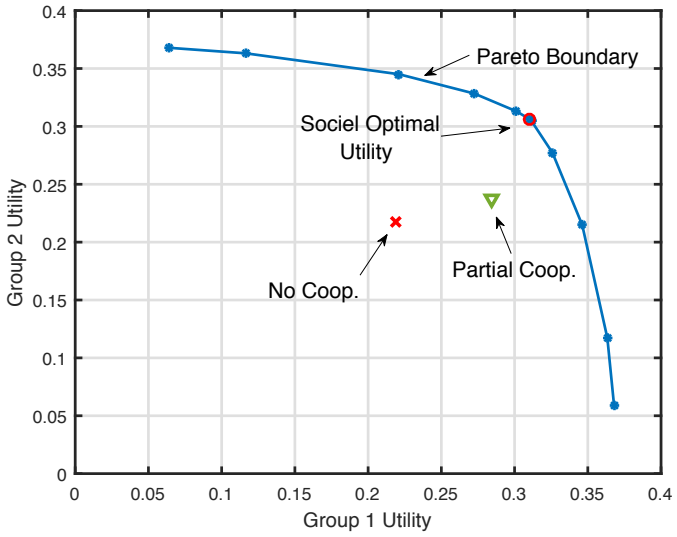


Figure 3.6: Group utility region under full cooperation versus utilities under partial and no cooperation.

that there is indeed a conflict of interests between groups under heterogeneous file preferences. This motivates us to further investigate the trade-offs between different groups, which is presented in the next simulation.

Feasible Utility Region under Full Cooperation and the Benchmark Cases

Next, we show the feasible utility region defined in (3.8). The densities for the two groups are $\bar{\lambda}_1 = \bar{\lambda}_2 = 0.05$ MT per square meter. The Pareto boundary is obtained by solving optimization problems (P4) with varying weights γ_1 and γ_2 of the two groups. Specifically, the results for weights $\gamma_1 = 1,\ \gamma_2 = 0$ and $\gamma_1 = 0,\ \gamma_2 = 1$ are obtained based on Theorem 3.4.2 and the results for the weights $\gamma_1,\ \gamma_2 > 0$ are obtained based on Theorem 3.4.1. In comparison, the results for partial cooperation and no cooperation are also shown, which are based on Theorem 3.5.1 and Corollary 3.5.1, respectively.

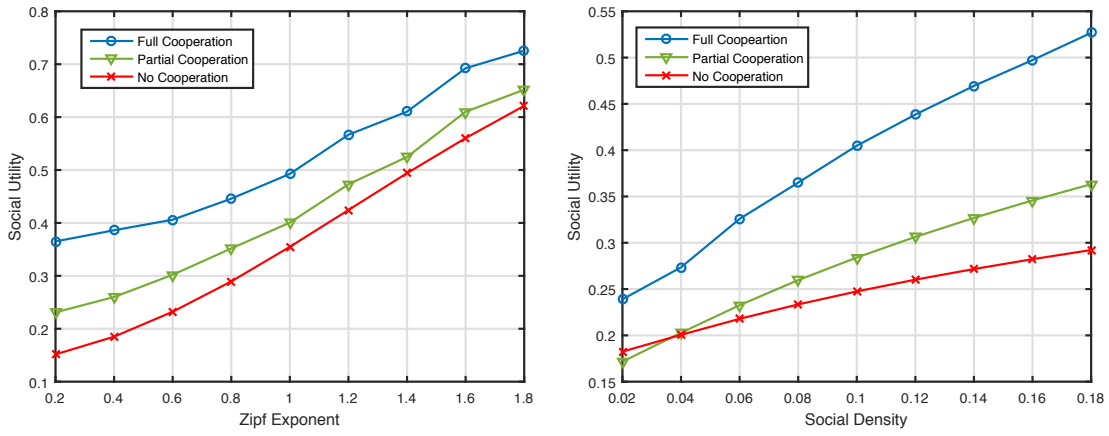
The simulation result for the feasible utility region is shown in Fig. 3.6. It can be observed that there is indeed a trade-off between the two groups in their utilities

under full cooperation in cooperative local caching. Moreover, the group utilities in partial cooperation and no cooperation both lie within the Pareto boundary achieved by full cooperation, which are at $(0.2839, 0.2371)$ and $(0.2189, 0.2173)$, respectively, both strictly lower than the social optimum utility at $(0.2937, 0.2941)$ and lies inside the Pareto boundary. This shows the benefits of full cooperation, while partial and no cooperation may lead to decreased utilities strictly inside the Pareto region. Finally, it should be noted that the utility of group 1 at $\gamma_1 = 0$ or that of group 2 at $\gamma_2 = 0$ is non-zero. This is because the preferences of the two groups are not completely different and cooperative file sharing still can bring utilities to the group with even zero weight.

Social Utility under Different Zipf Exponents

Furthermore, we evaluate the social utility under the schemes of full, partial and no cooperation with respect to (a) different Zipf exponents and (b) different social densities. In case (a), since the typical Zipf exponent is around $\gamma_r = 0.9$, we assess the influences of the Zipf exponent in the range of $0.2 \leq \gamma_r \leq 1.8$. The densities for the two groups are $\bar{\lambda}_1 = \bar{\lambda}_2 = 0.05$ MT per unit area. In case (b), the densities of the two groups are $\bar{\lambda}_1 = \bar{\lambda}_2 = \bar{\lambda}_0/2$ with varying $\bar{\lambda}_0$ in the range of $0.02 \leq \bar{\lambda}_0 \leq 0.18$.

The simulation results are shown in Fig. 3.7. First, from the result in Fig. 3.7a, it can be observed that for all of the three cases, the social utilities are monotonically increasing with respect to the Zipf exponent. This can be explained by the majorization theory [84]. With a large Zipf exponent, the popularity of the files is more concentrated on fewer files and the request distributions ($\boldsymbol{r}_1$ and $\boldsymbol{r}_2$) of the two groups will be more different. Accordingly, both groups will try to match their caching distributions $\boldsymbol{c}_i$ with the request distributions $\boldsymbol{r}_i$. Hence, if a request distribution is more concentrated (i.e., it majorizes over another distribution that is less concentrated), the utility will be lower. It can also be observed that the gap between the full cooperation and partial or no cooperation decreases with the



(a) (b)

Figure 3.7: Social utilities under (a) different Zipf exponents, and (b) different social densities.

increase of the Zipf exponent (i.e., request distributions become more different). This is because, for partial cooperation and no cooperation, as the interests of different groups become more similar, the distribution on the popular files is more concentrated and it is easier for the MTs to discover these files. However, for the cases where the Zipf component is low, the popular files are more diverse and cooperation is more beneficial; thus, the gap between full cooperation and no or partial cooperation gets larger.

Next, from the result in Fig. 3.7b, it can be observed that, for all the three schemes, the social utility increases with respect to the social density. This is obvious since higher social density means larger number of MTs in the neighbour and more opportunities for file sharing and the gain is more substantial for the full cooperation. It can also be observed that partial cooperation can be very close or even worse than no cooperation in the low-density region. This is because, in partial cooperation, groups try to exploit the distribution of the other groups, which may not be their actual caching distributions. In this case, the caching distribution may deviate far from the optimal caching distribution. As a result, the MTs' own caches are not matched to their group request distribution and they could get little help from the

other groups. While for no cooperation, each group can at least cache according to its own preferences, which promotes intra-group file sharing.

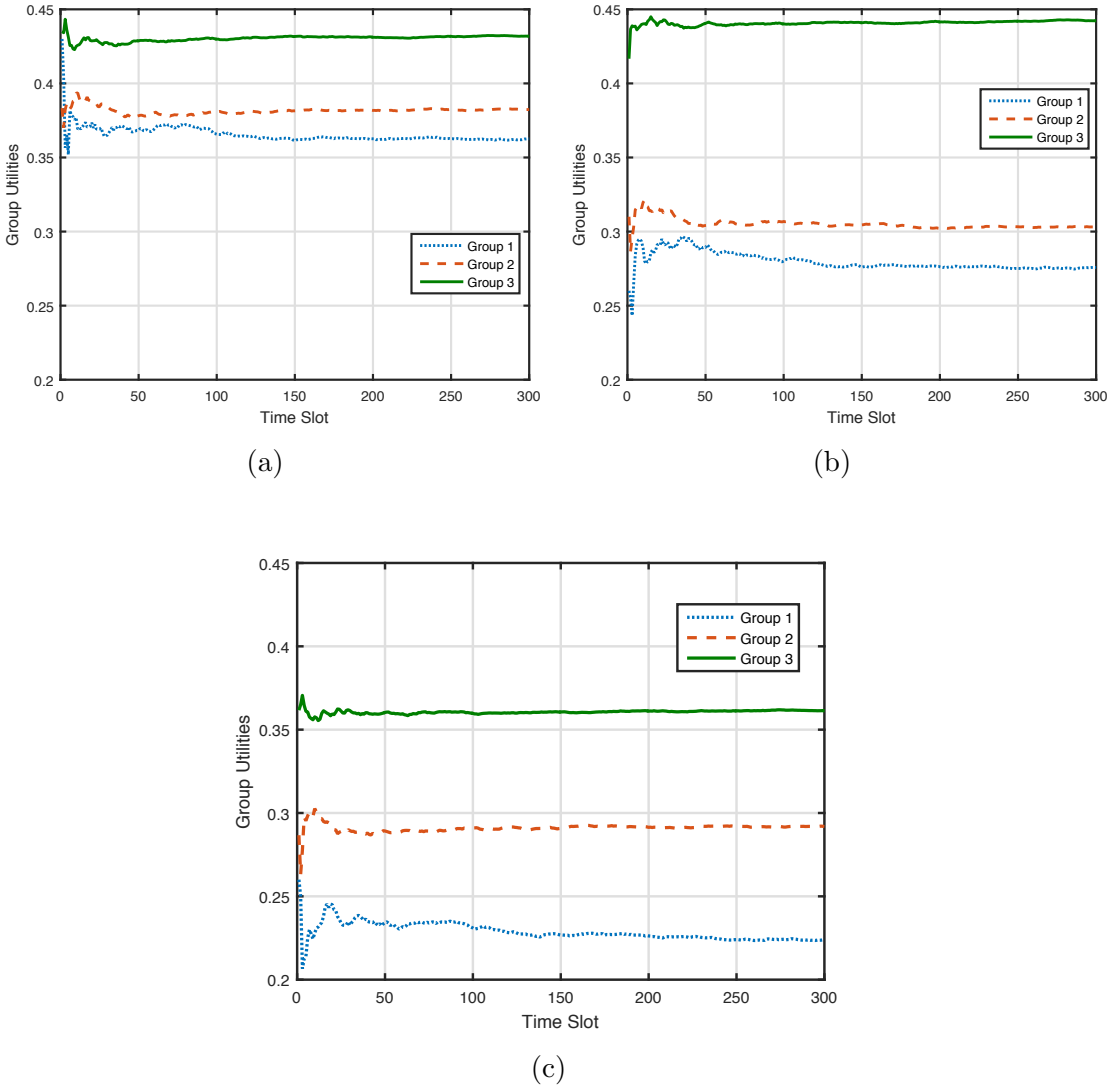


Figure 3.8: Group utilities under (a) full cooperation, (b) partial cooperation, and (c) no cooperation.

3.6.2 Network Simulation

In this sub-section, we perform a network-level simulation under the evolution of the system to validate the accuracy of the theoretical results for the cooperative

local caching under three cases. The simulation setup is specified as follows: Three groups, each having 100 MTs, 300 MTs, 600 MTs, are scattered within a square area of 100×100 m^2. Hence, the density of the three groups are $\bar{\lambda}_1 = 0.01, \bar{\lambda}_2 = 0.03, \bar{\lambda}_3 = 0.06$ MTs per square meter and their weights are $\boldsymbol{\gamma} = [0.1, 0.3, 0.6]$, respectively. Initially, the locations of all the MTs are randomly generated within this area. The simulation is operated under 300 time slots. During each time slot, each MT moves 1m to the available positions in equal probabilities. We first obtain the optimal caching distribution $c^*_{i,f},\ f \in \mathcal{F}$ for each group $i \in \mathcal{G}$ in full, partial and no cooperation with the results from Theorem 3.3.1, Theorem 3.5.1 and Corollary 3.5.1, respectively. According to these optimal distributions, which are multinomial distributions with one trial, the MTs in each group perform random samplings on the set of files for obtaining the cached file and they keep the cached file during the whole simulation. The file request during each time slot is also a random sampling from the request distribution $r_{i,f},\ f \in \mathcal{F}$ of each group $i \in \mathcal{G}$. The MTs search for the requested file within the range of $d = 5$m. During time slot n, the utility of each group $i \in \mathcal{G}$ is defined as

$$U_i(n) = \frac{\text{Total successful file discoveries of group } i \text{ up to time slot } n}{\text{Total file requests of group } i \text{ up to time slot } n}. \tag{3.24}$$

With results in Theorem 3.3.1, Theorem 3.5.1 and Corollary 3.5.1, three groups' utilities under full, partial and no cooperation are obtained as $(0.3648, 0.3821, 0.4306)$, $(0.2747, 0.3016, 0.4426)$ and $(0.2281, 0.2993, 0.3650)$, respectively. The network simulation results after 300 time slots for full cooperation, partial cooperation and no cooperation are shown in Fig. 3.8a, 3.8b and 3.8c, respectively. From the simulation results, it can be observed that the converged values of the cooperative local caching system under real-time simulation are very close to their corresponding theoretical results. This validates the effectiveness of our system under real operation and shows the accuracy of our model.

3.7 Chapter Summary

In this chapter, we studied the cooperative local caching with heterogeneous file preferences among different groups of MTs. To quantify each group's performance, we defined each group's utility as the probability for the MTs in the group to successfully discover the file in the neighbouring MTs, which can be maximized by optimizing its caching strategies. By modelling MTs' mobilities as independent HPPPs, we analytically characterized MTs' utilities in closed-form. We first considered the fully cooperative case where a centralizer helps all groups to make caching decisions. We formulated the problem as a weighted-sum utility maximization problem, through which the trade-offs of different groups were characterized. Then, we studied two benchmark cases under selfish caching, namely, partial and no cooperation, with and without inter-group file sharing, respectively. The optimal caching distributions for these two cases were derived. Our work in this chapter provides a practical modelling for cooperative local caching under heterogeneous file preferences and demonstrates essential insights on the design of wireless systems with cooperative local caching in practice.

Chapter 4

BS-side Joint Energy and Spectrum Cooperation

4.1 Introduction

In the previous two chapters, we have discussed the new cooperation techniques in the uplink and downlink transmissions of the MTs, respectively. In this chapter, we proceed to the BS side and study the cooperation among the BSs in the wireless systems exploiting renewable energy sources. In order to address the incentive issues in cooperation for reducing the wireless network's energy costs, we propose a new cooperation technique of joint energy and spectrum cooperation. A practical modelling is given for the renewable energy availability, spectrum cooperation, and the conventional as well as the renewable energy costs. When the two systems are fully cooperative (e.g., belonging to the same wireless operator), we formulate the joint cooperation as a weighted-sum cost minimization problem to characterize the trade-offs in their energy costs and the optimal solution is derived in closed-form. We also study another partially cooperative scenario where the two systems have their own interests. We show that the two systems can cooperate in partial cooperation if they could find inter-system complementarity between the energy and spectrum resources. Under the condition for partial cooperation, we propose a distributed algorithm for the two systems to gradually and simultaneously reduce their costs from the non-cooperative benchmark to the Pareto optimum. Finally, we provide numerical results to validate the convergence of the distributed algorithm and compare the centralized and distributed cost reduction approaches under various

scenarios.

The rest of this chapter is organized as follows. Section 4.2 presents the literature review. Section 4.3 introduces the system model of joint energy and spectrum cooperation. Section 4.4 presents the problem formulation under the full cooperation and also discusses the benchmark case of no cooperation. Section 4.5 discusses the centralized optimization under the full cooperation and efficient algorithm is proposed for the optimal solution. Section 4.6 discusses the benchmark case of partial cooperation and proposes a distributed algorithm for achieving the Pareto optimality. Section 4.7 presents numerical examples to compare the performances of different cooperation schemes. Finally, Section 4.8 concludes this chapter.

4.2 Literature Review

4.2.1 Energy Cooperation

In the literature, there are some recent works studying energy cooperation in the wireless systems (e.g., [38, 39, 85]). [38] first considers the energy cooperation in a two-BS cellular network to minimize the total energy drawn from conventional grid subject to practical requirements. Both off-line and on-line algorithms for the cases of unavailable and available future energy information are developed. In [39], the authors propose a joint communication and energy cooperation approach in CoMP cellular systems powered by energy harvesting. They maximize the downlink sum-rate by jointly optimizing energy sharing and zero-forcing pre-coding. Nevertheless, both works [38, 39] only consider the cooperation within one single cellular system instead of inter-system cooperation. Another work worth mentioning is [85], which studies the wireless energy cooperation in different setups of wireless systems such as the one-way and two-way relay channels. However, unlike the previously mentioned works, which implement the energy cooperation via wired power transmission, the energy sharing in [85] is enabled by wireless power transfer

with limited energy sharing efficiency in energy cooperation. [86] also discussed wireless cellular network powered with both renewable and conventional energy and the energy cost of the network is minimized. However, the incentive issues in the cooperation are not considered and the spectrum cooperation is not in the model.

4.2.2 Spectrum Cooperation

The idea of spectrum cooperation is similar to the *cooperative spectrum sharing* in the literature of cognitive radio network (e.g., [87–89]), where SUs cooperate with PUs to co-use PUs' spectrum. In order to create incentives for sharing, there are basically two approaches: *resource-exchange* [87, 88] and *money-exchange* [89]. For the resource-exchange approach, SUs relay traffics for PUs in exchange for dedicated spectrum resources for SUs' own communications [87, 88]. Specifically, in [87], the problem is formulated as a Stackelberg game, where the PU aims to optimize its QoS, while the SUs compete among themselves for transmission within the shared spectrum from the PU. In [88], the PU-SU interactions under incomplete information are modelled as a labour market using contract theory, in which the optimal contracts are designed. For the money-exchange approach, PU sells its idle spectrum to SUs for money returns. The authors in [89] model the spectrum trading process as a monopoly market and accordingly design a monopolist-dominated quality price contract, where the necessary and sufficient conditions for the optimal contract are derived.

4.2.3 Cooperation by Traffic Off-loading

It is worth noting that there has been another line of research on improving the energy efficiency in wireless networks by off-loading traffic across different transmitters and/or systems [90–92]. [90] studies a cognitive radio network, where the PU reduces its energy consumption by off-loading part of its traffic to the SU, while in return the PU shares its licensed spectrum bands to the SU. [91] and [92] consider a single cellular system, in which some BSs with light traffic load can off-load

their traffic to the neighbouring BSs and then turn off communication circuits for saving energy. Although these schemes can be viewed as another approach to realize the spectrum and energy cooperation, they are different from direct energy and spectrum sharing among BSs, as will be considered in this chapter.

Compared to the above existing works, the novelty of the results in this chapter is two-fold. First, under the new model of joint energy and spectrum cooperation, a comprehensive study on this model by considering the uncertainty of renewable energy and the relationship between the two resources is provided. Second, the conflict of interests between different systems is investigated and both centralized and distributed algorithms for the cases of fully and partially cooperative systems are proposed.

4.3 System Model

We consider two cellular systems that operate over different frequency bands. The two systems can either belong to the same entity (e.g., their associated operators are merged as a single party like Sprint and T-Mobile in some states of US [93]) or relate to different entities. For the purpose of the initial investigation, as shown in Fig. 4.1, we focus our study on the downlink transmission of two (partially) overlapping cells each belonging to one cellular system.[20] In each cell $i \in \{1, 2\}$, there is a BS serving K_i MTs. The sets of MTs associated with the two BSs are denoted by $\mathcal{K}_1$ and $\mathcal{K}_2$, respectively, with $|\mathcal{K}_1| = K_1$ and $|\mathcal{K}_2| = K_2$. We consider that the two BSs purchase energy from both conventional power grid and their dedicated local renewable utility firms. For example, as shown in Fig. 4.1, the local utility firm with solar source connects to BS 1 via a direct current (DC)/DC converter, the other one with wind source connects to BS 2 through an alternating current (AC)/DC converter, and power grid connects to both BSs by using AC/DC converters. By combining energy from the two different supplies, BS 1 and BS 2 can operate on their

[20] Our results can be extended to the multi-cell setting for each system by properly pairing the BSs in different systems.

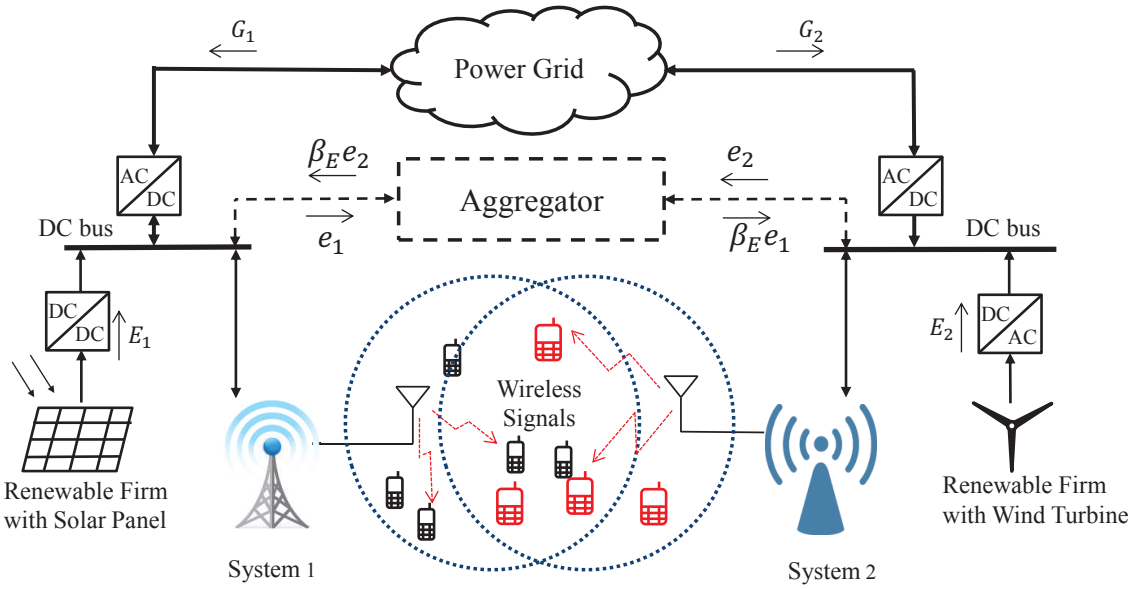


Figure 4.1: Two neighbouring cellular systems powered by power grid and renewable energy with joint energy and spectrum cooperation.

respective DC buses. Here, different power electronic circuits (AC/DC converter, DC/DC converter, etc.), which connect the BSs, renewable utility firms and the grids, are based on the types of power line connections (AC or DC buses) and the properties of different nodes (e.g., the BS often runs on a DC bus [30]).

We consider a time-slotted model, where the energy harvesting rate remains constant in each slot and may change from one slot to another. In practice, the harvested energy from solar and wind remains constant over a window of seconds and we choose our time slot of the same duration.[21] We further assume that a BS's operation in one-time slot is based on its observation of the energy availability, channel conditions, traffic loads, etc., and is independent of its operation in another time slot. This is reasonable as current energy storage devices are expensive and capacity-limited compared to the power consumption of the BSs and many cellular systems do not rely on storage for dynamic energy management. Thus, we can analyse the two systems' cooperation problem in each time slot individually. In the following, we first introduce the operation details of two systems' energy and

[21] Without loss of generality, we can further normalize the duration of each slot to be a unit of time so that the terms "power" and "energy" can be used interchangeably.

spectrum cooperation, and then propose the downlink transmission model for both systems.

4.3.1 Energy Cooperation Model

Recall that each BS can purchase energy from both conventional grid and renewable utility firms. The two different types of energy supplies are characterized as follows.

- *Conventional energy from the power grid*: Let the energy drawn by BS $i \in \{1, 2\}$ from the grid be denoted by $G_i \geq 0$. Since the practical energy demand from an individual BS is relatively small compared to the whole demand and production of the power grid network, the available energy from the grid is assumed to be infinite for BS i. We denote $\alpha_i^G > 0$ as the price per unit energy purchased from the grid by BS i. Accordingly, BS i's payment to obtain energy from grid is $\alpha_i^G G_i$.

- *Renewable energy from the renewable utility firm*: Let the energy purchased by BS $i \in \{1, 2\}$ from the dedicated local renewable utility firms be denoted as $E_i > 0$. Different from conventional energy from the grid, the local renewable energy firm is capacity-limited and subject to uncertain power supply due to environmental changes. Therefore, BS i cannot purchase more than $\bar{E}_i$, which is the electricity production of the utility firm produces in the corresponding time slot. That is,

$$E_i \leq \bar{E}_i, i \in \{1, 2\}. \tag{4.1}$$

  Furthermore, we denote $\alpha_i^E > 0$ as the price of renewable energy at BS i and BS i's payment to obtain energy E_i from renewable utility firm is $\alpha_i^E E_i$.

By combing the conventional and renewable energy costs, the total cost at BS i to obtain energy $G_i + E_i$ is thus denoted as

$$C_i = \alpha_i^E E_i + \alpha_i^G G_i, \ i \in \{1, 2\}. \tag{4.2}$$

The price to obtain a unit of renewable energy is lower than that of conventional energy (i.e., $\alpha_i^E < \alpha_i^G, \forall i \in \{1, 2\}$). This can be valid in reality thanks to governmental subsidy, potential environmental cost of conventional energy, and the high cost of delivering conventional energy to remote areas, etc.

Next, we consider the energy cooperation between the two systems. Let the transferred energy from BS 1 to BS 2 be denoted as $e_1 \geq 0$ and that from BS 2 to BS 1 as $e_2 \geq 0$. Practically, the energy cooperation between two systems can be implemented by connecting the two BSs to a common aggregator as shown in Fig. 4.1.[22] When BS i wants to share energy with BS $\bar{\imath}$, where $\bar{\imath} \in \{1, 2\} \setminus \{i\}$, BS i first notifies BS $\bar{\imath}$ the transmitted energy amount e_i. Then, at the appointed time, BS i injects e_i amount of energy to the aggregator and BS $\bar{\imath}$ draws $\beta_E e_i$ amount of energy out. Thus, energy sharing without disturbing balance in the total demand and supply can be accomplished via the aggregator. Here, $0 \leq \beta_E \leq 1$ is the energy transfer efficiency factor between the two BSs that specifies the unit energy loss through the aggregator for the transferred amount of power.[23]

As depicted in Fig. 4.2, the energy management scheme at each BS $i \in \{1, 2\}$ operates as follows. First, at the beginning of each time slot, BS i purchases the conventional energy G_i and the renewable energy E_i. Second, it performs energy cooperation by either transferring e_i amount of energy to BS $\bar{\imath}$ or collecting

[22] Aggregator is a virtual entity in the emerging smart grid that aggregates and controls the demands at distributed end users (e.g., BSs in cellular systems) [94]. In order to manage the distributed loads more efficiently, the aggregator allows the end users to either draw or inject energy from/to it under different demand/supply conditions. By utilizing the two-way energy transfer between the end users and the aggregator, the energy sharing between the BSs can be enabled.

[23] It is worth noting that there also exists another alternative approach to realize the energy sharing by direct powerline connection between the BSs [38], in which dedicated power lines may need to be newly deployed. Such approach has been implemented in the smart grid deployments, e.g., to realize the energy transfer among different micro-grids [95].

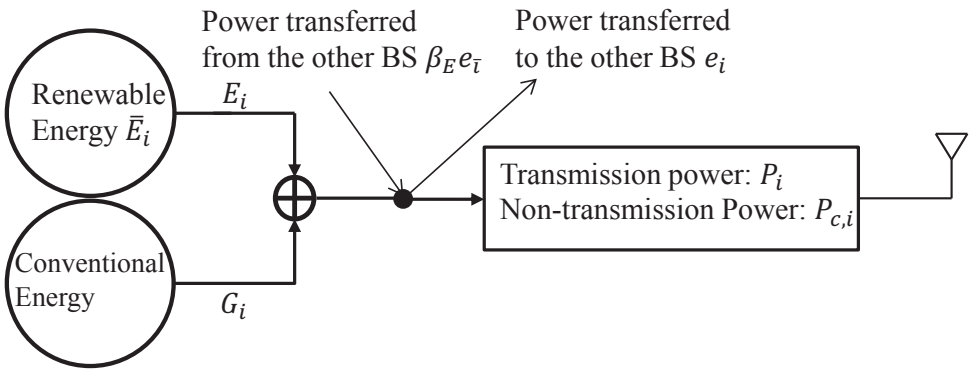


Figure 4.2: Energy management schematic of BS i.

the exchanged energy $\beta_E e_{\bar{\imath}}$ from BS $\bar{\imath}$. Finally, BS i consumes a constant non-transmission power $P_{c,i}$ to maintain its routine operation and a transmission power P_i for flexible downlink transmission. By considering transmission power, non-transmission power, and shared power between BSs, we can obtain the total power consumption at BS i, which is constrained by the total power supply:

$$\frac{1}{\eta}P_i + P_{c,i} \leq E_i + G_i + \beta_E e_{\bar{\imath}} - e_i, \ i \in \{1, 2\}, \tag{4.3}$$

where $0 < \eta \leq 1$ is the power amplifier (PA) efficiency. Since η is a constant, we normalize it as $\eta = 1$ in the sequel.

4.3.2 Spectrum Cooperation Model

We now explain the spectrum cooperation by considering two cases: adjacent and non-adjacent frequency bands. First, consider the case of adjacent frequency bands in Fig. 4.3a. As shown in Fig. 4.3a, BS 1 and BS 2 operate in the blue and red shaded frequency bands W_1 and W_2, respectively. Between W_1 and W_2 a guard band W_G is inserted to avoid interference due to out-of-band emissions. Let the shared spectrum bandwidth from BS 1 to BS 2 be denoted as $w_1 \geq 0$ and that from BS 2 to BS 1 as $w_2 \geq 0$. In this case, the shared bandwidth from BS i (i.e., w_i) can be fully used at BS $\bar{\imath}$, $i \in \{1, 2\}$. This can be implemented by carefully moving the guard band as shown in Fig. 4.3a. When BS 2 shares a bandwidth w_2 to BS

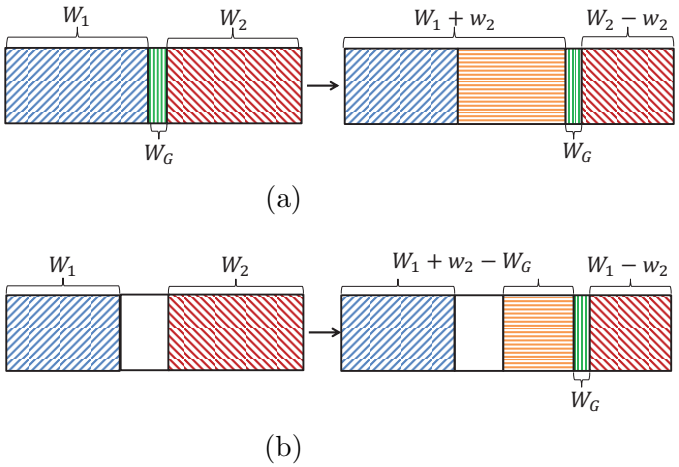


Figure 4.3: An example of spectrum cooperation between two BSs: (a) adjacent frequency bands, and (b) non-adjacent frequency bands.

1, the green shaded guard band with bandwidth W_G is moved accordingly between $W_1 + w_2$ and $W_2 - w_2$, such that the shared spectrum is fully utilized.

Next, consider the case of non-adjacent frequency bands in Fig. 4.3b. For this non-adjacent frequency band, a guard band is also needed to avoid the inter-system interference. Fig. 4.3b shows an example of the spectrum cooperation when BS 2 shares a bandwidth w_2 to BS 1. After spectrum cooperation, the total usable spectrum of BS 1 is $W_1 + w_2 - W_G$, since a green shaded guard band W_G is inserted between $W_1 + w_2 - W_G$ and $W_2 - w_2$ to avoid the inter-system interference. As a result, spectrum cooperation loss will occur.[24]

For the ease of investigation, we only focus on the former case of adjacent frequency bands.[25] We define a spectrum-cooperation factor $\beta_B \in \{0, 1\}$, for which $\beta_B = 1$ denotes that spectrum cooperation is implementable between BSs and $\beta_B = 0$ represents that spectrum cooperation is infeasible. Considering the spectrum

[24]It is technically challenging to gather non-adjacent pieces of bandwidth together at one BS. To overcome this issue, the carrier aggregation solution for the LTE-Advanced system [96] can be utilized here.

[25]It should be noticed that our result can also be extended to the non-adjacent bandwidth case by considering the bandwidth loss due to the guard band.

cooperation between the BSs, the bandwidth used by BS i can be expressed as

$$B_i \leq W_i + \beta_B w_{\bar{i}} - w_i, \ i \in \{1, 2\}. \tag{4.4}$$

Note that in our investigated spectrum cooperation, the (shared) spectrum resources can only be utilized by either BS 1 or BS 2 to avoid the interference between the two systems. If the same spectrum resources can be used by the two systems at the same time, then the spectrum utilization efficiency can be further improved while also introducing inter-system interference [97]. In this case, more sophisticated interference coordination should be implemented, which is beyond the scope of this work.

4.3.3 Downlink Transmission under Energy and Spectrum Cooperation

We now introduce the downlink transmission at each BS by incorporating the energy and spectrum cooperation. We consider a flat fading channel model for each user's downlink transmission, and denote the channel gain from BS $i \in \{1, 2\}$ to its associated MT k as g_k, $k \in \mathcal{K}_i$, which in general includes the pathloss, shadowing and antenna gains. Within each system, we assume orthogonal transmission to support multiple MTs, e.g., by applying OFDMA. Accordingly, the signal-to-noise-ratio (SNR) at each MT k is given by

$$\mathtt{SNR}_k = \frac{g_k p_k}{b_k N_0}, k \in \mathcal{K}_1 \cup \mathcal{K}_2, \tag{4.5}$$

where N_0 denotes the power spectral density (PSD) of the additive white Gaussian noise (AWGN), $p_k \geq 0$ and $b_k \geq 0$ denote the allocated power and bandwidth to MT $k \in \mathcal{K}_1 \cup \mathcal{K}_2$, respectively. We can aggregate all the transmission power and

bandwidth used by each BS as (cf. (4.3) and (4.4))

$$P_i = \sum_{k \in \mathcal{K}_i} p_k, \ B_i = \sum_{k \in \mathcal{K}_i} b_k, \ i \in \{1, 2\}.$$

Note that the bandwidth and power allocation are performed in each slot on the order of seconds, which is much longer than the coherence time of wireless channels (on the order of milliseconds). As a result, the SNR defined in (4.5) is time-averaged over the dynamics of wireless channels, and thus the fast fading is averaged out from the channel gain g_k's.

To characterize the QoS requirements of each MT, we define its performance metric as a utility function $U_k(b_k, p_k)$, and assume that it satisfies the following three properties:

1) The utility function is non-negative, i.e., $U_k(b_k, p_k) \geq 0, \ \forall p_k \geq 0, \ b_k \geq 0$, where $U_k(0, p_k) = 0$ and $U_k(b_k, 0) = 0$;

2) The utility increases as a function of allocated power and bandwidth, i.e., $U_k(b_k, p_k)$ is monotonically increasing with respect to b_k and p_k, $\forall p_k \geq 0$, $b_k \geq 0$;

3) The marginal utility decreases as the allocated power and bandwidth increase, i.e., $U_k(b_k, p_k)$ is jointly concave over b_k and p_k, $\forall p_k \geq 0$, $b_k \geq 0$.

For example, the achievable data rate at MT k defined as follows is a feasible utility function that satisfies the above three properties [98]

$$U_k(b_k, p_k) = b_k \log_2(1 + \mathtt{SNR}_k) = b_k \log_2 \left(1 + \frac{g_k p_k}{b_k N_0}\right). \tag{4.6}$$

In the rest of this chapter, we employ the utility function in (4.6) for all MTs in the two systems and averting to another function will not change our main engineering insights. Due to the fact that most cellular network services are QoS guaranteed (e.g., minimum data rate in video call), we ensure the QoS requirement at each MT i by setting a minimum utility threshold $r_k \geq 0$, $k \in \mathcal{K}_1 \cup \mathcal{K}_2$. The value of r_k is

chosen according to the type of service at MT k. Accordingly, the resultant QoS constraint is given by

$$b_k \log_2 \left(1 + \frac{g_k p_k}{b_k N_0}\right) \geq r_k, \ \forall k \in \mathcal{K}_1 \cup \mathcal{K}_2. \tag{4.7}$$

4.4 Problem Formulation

We aim to reduce the costs C_1 and C_2 in (4.2) at both BSs while guaranteeing the QoS requirements of all MTs. We denote the intra-system decision vector for BS $i \in \{1, 2\}$ as $\boldsymbol{x}_i^{\text{in}}$, which consists of its energy drawn from renewable energy E_i, energy drawn from the grid G_i, power allocation p_k's, and bandwidth allocation b_k's with $k \in \mathcal{K}_i$. We also denote $\boldsymbol{x}^{\text{ex}} = [e_1, e_2, w_1, w_2]^T$ as the inter-system energy and spectrum cooperation vector. For convenience, we aggregate all the decision variables of the two BSs as $\boldsymbol{x} = [\boldsymbol{x}_1^{\text{in}T}, \boldsymbol{x}_2^{\text{in}T}, \boldsymbol{x}^{\text{ex}T}]^T$.

It can be shown that the two systems (if not belonging to the same entity) have conflict of interests in cost reduction under joint energy and spectrum cooperation. For example, if BS 1 shares both energy and spectrum to BS 2, then the cost of BS 2 is reduced while the cost of BS 1 increases. To characterize such conflicts, we define the achievable cost region under the joint energy and spectrum cooperation as the cost tuples that the two BSs can achieve simultaneously, which is explicitly characterized by

$$\mathcal{C} \triangleq \bigcup_{\boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathcal{X}} \{(c_1, c_2) : C_i(\boldsymbol{x}) \leq c_i, i \in \{1, 2\}\}, \tag{4.8}$$

where $\mathcal{X}$ is the feasible set of $\boldsymbol{x}$ specified by (4.1), (4.3), (4.4) and (4.7), and $C_i(\boldsymbol{x})$ is the achieved cost of BS $i \in \{1, 2\}$ in (4.2) under given $\boldsymbol{x}$. The boundary of this region is then called the *Pareto boundary*, which consists of the Pareto optimal cost tuples at which it is impossible to decrease one's cost without increasing the other's. Since the feasible region $\mathcal{X}$ can be shown to be convex and the cost in (4.2) is affine, the cost region in (4.8) is convex. Also, because the Pareto optimal points

of any convex region can be found by solving a series of weighted sum minimization problems with different weights [80], we can achieve different Pareto optimal cost tuples by solving the following weighted sum cost minimization problems

$$(\text{P6}): \min_{\boldsymbol{x}\geq \mathbf{0}} \quad \sum_{i=1}^{2} \gamma_i(\alpha_i^E E_i + \alpha_i^G G_i) \tag{4.9a}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{K}_i} p_k + P_{c,i} \leq E_i + G_i + \beta_E e_{\bar{i}} - e_i, \ \forall i \in \{1,2\}, \tag{4.9b}$$

$$\sum_{k\in\mathcal{K}_i} b_k \leq W_i + \beta_B w_{\bar{i}} - w_i, \forall i \in \{1,2\}, \tag{4.9c}$$

$$E_i \leq \bar{E}_i, \ \forall i \in \{1,2\}, \tag{4.9d}$$

$$b_k \log_2\left(1 + \frac{g_k p_k}{b_k N_0}\right) \geq r_k, \ \forall k \in \mathcal{K}_1 \cup \mathcal{K}_2, \tag{4.9e}$$

where $\gamma_i \geq 0, i \in \{1,2\}$ is the cost weight for BS i, which specifies the trade-offs between the two BSs' costs. By solving (P6) given different γ_i's, we can characterize the entire Pareto boundary of the cost region. For the solution of the optimization problem, standard convex optimization techniques such as the interior point method can be employed to solve (P6) [80]. However, in order to gain more engineering insights, we propose an efficient algorithm for (P6) by applying the Lagrange duality method in Section 4.5. However, before we present the solution for (P6), we first consider a special case where there is neither energy nor spectrum cooperation between the two systems (i.e., $\beta_E = \beta_B = 0$). This serves as a performance benchmark for comparison with fully or partially cooperative systems in Sections 4.5 and 4.6.

4.4.1 Benchmark Case: Non-cooperative Systems

With $\beta_E = \beta_B = 0$, the two systems will not cooperate and the optimal solution of (P6) given any γ_i's is attained with zero inter-system exchange, i.e., $w_1 = w_2 = e_1 = e_2 = 0$. In this case, the constraints in (4.9b) and (4.9c) reduce to $\sum_{k\in\mathcal{K}_i} p_k + P_{c,i} \leq E_i + G_i$ and $\sum_{k\in\mathcal{K}_i} b_k \leq W_i, \ \forall i \in \{1,2\}$, respectively. It thus follows that the

intra-system energy and bandwidth allocation vectors $\boldsymbol{x}_1^{\text{in}}$ and $\boldsymbol{x}_2^{\text{in}}$ are decoupled in both the objective and the constraints of Problem (P6). As a result, (P6) degenerates to two cost-minimization problems as follows (one for each BS $i \in \{1, 2\}$):

$$
\begin{aligned}
(\text{P7}): \quad & \underset{\boldsymbol{x}_i^{\text{in}} \geq 0}{\texttt{min.}} \quad \alpha_i^E E_i + \alpha_i^G G_i \\
& \texttt{s.t.} \quad \sum_{k \in \mathcal{K}_i} p_k + P_{c,i} \leq E_i + G_i, && \text{(4.10a)} \\
& \qquad E_i \leq \bar{E}_i, && \text{(4.10b)} \\
& \qquad \sum_{k \in \mathcal{K}_i} b_k \leq W_i, && \text{(4.10c)} \\
& \qquad b_k \log_2 \left(1 + \frac{g_k p_k}{b_k N_0}\right) \geq r_k, \ \forall k \in \mathcal{K}_i. && \text{(4.10d)}
\end{aligned}
$$

Note that Problem (P7) is always feasible due to the fact that the BS can purchase energy from the grid without limit. Therefore, we can always find one feasible solution to satisfy all the constraints in (4.10a)-(4.10d). It is easy to show that at the optimality of Problem (P7), the constraints (4.10c) and (4.10d) are both tight, otherwise, one can reduce the cost by reducing the allocated power p_k (and/or increasing the allocated bandwidth b_k) to MT k. Then, the power allocation for each user can be expressed as

$$
p_k = \frac{b_k N_0}{g_k} \left(2^{\frac{r_k}{b_k}} - 1\right), \ \forall k \in \mathcal{K}_i. \tag{4.11}
$$

By substituting (4.11) into (P7) and applying the KKT condition, we have the closed-form optimal solution to (P7) in the following proposition. Note that the optimal solution is unique, since the constraints in (4.10d) are strictly convex over b_k's and p_k's, $\forall k \in \mathcal{K}$.

Proposition 4.4.1 *The optimal bandwidth allocation for* (P7) *is given by*

$$
b_k^\star = \frac{\ln 2 \cdot r_k}{W(\frac{1}{e}(\frac{\nu_i^\star g_k}{N_0} - 1)) + 1}, \ \forall k \in \mathcal{K}_i, \tag{4.12}
$$

where $W(\cdot)$ *is Lambert-W function and* $\nu_i^\star \geq 0$ *denotes the water-level that satisfies* $\sum_{k\in\mathcal{K}_i} b_k^\star = W_i$. *Furthermore, the optimal power allocation and energy management in* (P7) *are given by*

$$p_k^\star = \frac{b_k^\star N_0}{g_k}\left(2^{\frac{r_k}{b_k^\star}} - 1\right),\ \forall k \in \mathcal{K}_i,$$
$$E_i^\star = \max\left(\sum_{k\in\mathcal{K}_i} p_k^\star + P_{c,i}, \bar{E}_i\right),$$
$$G_i^\star = \max\left(\sum_{k\in\mathcal{K}_i} p_k^\star + P_{c,i} - \bar{E}_i, 0\right).$$

Proof See Appendix F.

In Proposition 4.4.1, the bandwidth allocation $b_k^\star$ can be interpreted as water-filling over different MTs with $\nu_i^\star$ being the water level, and the power allocation $p_k^\star$ follows from (4.11). Furthermore, the optimal solution of $E_i^\star$ and $G_i^\star$ indicate that BS i first purchases energy from the renewable energy firm, and (if not enough) then from the grid. This is intuitive due to the fact that the renewable energy is cheaper ($\alpha_i^E < \alpha_i^G$).

4.5 Centralized Energy and Spectrum Cooperation for Fully Cooperative Systems

In this section, we consider Problem (P6) with given weights γ_1 and γ_2 for the general case of $\beta_B \in \{0, 1\}$ and $0 \leq \beta_E \leq 1$. This corresponds to the scenario where the two BSs belong to the same entity and thus can fully cooperate to solve (P6) to minimize the weighted sum cost. Similar to (4.10d) in (P7), we can show that the QoS constraints in (4.9e) should always be tight for the optimal solution of (P6). As a result, the power allocation for each MT in (P6) can also be expressed as (4.11) for all $i \in \{1, 2\}$. By substituting (4.11) into the power constraint (4.9b) in (P6) and then applying the Lagrange duality method, we obtain the closed-form solution to (P6) in the following proposition.

Proposition 4.5.1 *The optimal bandwidth and power allocation solutions to Problem* (P6) *are given by*

$$b_k^\star = \frac{\ln 2 \cdot r_k}{W\left(\frac{1}{e}\left(\frac{\lambda_i^\star g_k}{\mu_i^\star N_0} - 1\right)\right) + 1}, \ \forall k \in \mathcal{K}_1 \cup \mathcal{K}_2, \tag{4.13}$$

$$p_k^\star = \frac{b_k^\star N_0}{g_k}\left(2^{\frac{r_k}{b_k^\star}} - 1\right), \ \forall k \in \mathcal{K}_1 \cup \mathcal{K}_2, \tag{4.14}$$

where $\lambda_i^\star$ *and* $\mu_i^\star$ *are non-negative constants (dual variables) corresponding to the power constraint (4.9b) and the bandwidth constraint (4.9c) for BS* $i \in \{1, 2\}$, *respectively.*[26] *Moreover, the optimal spectrum sharing between the two BSs are*

$$w_i^\star = \max\left(-\sum_{k \in \mathcal{K}_i} b_k^\star + W_i, 0\right), \forall i \in \{1, 2\}. \tag{4.15}$$

Finally, the optimal energy decisions at two BSs $\{E_i^\star\}, \{G_i^\star\}$ *and* $\{e_i^\star\}$ *are the solutions to the following problem.*

$$\begin{aligned}
(\text{P8}): \min_{\{E_i, G_i, e_i\}} \quad & \sum_{i=1}^{2} \gamma_i(\alpha_i^E E_i + \alpha_i^G G_i) \\
\text{s.t.} \quad & \sum_{k \in \mathcal{K}_i} p_k^\star + P_{c,i} = E_i + G_i + \beta_E e_{\bar{i}} - e_i, \ \forall i \in \{1, 2\}, \\
& 0 \le E_i \le \bar{E}_i, \ G_i \ge 0, \ e_i \ge 0, \ \forall i \in \{1, 2\}.
\end{aligned}$$

Proof See Appendix G.

Note that Problem (P8) is a simple linear program (LP) and thus can be solved by existing software such as CVX [99]. Also note that there always exists an optimal solution of $\{e_i^\star\}$ in (P8) with $e_1^\star \cdot e_2^\star = 0$.[27] It should also be noted that the optimal solution in Proposition 4.5.1 can only be obtained in a centralized manner. Specifically, to perform the joint energy and spectrum cooperation, the information

[26] The optimal dual variables $\{\lambda_i^\star\}_{i=1}^2$ and $\{\mu_i^\star\}_{i=1}^2$ can be obtained by solving the dual problem of (P6).

[27] If $e_1^\star \cdot e_2^\star = 0$ does not hold, we can find another feasible energy cooperation solution $e_i^{\star'} = e_i^\star - \min(e_1^\star, e_2^\star), \ \forall i \in \{1, 2\}$, with $e_1^{\star'} \cdot e_2^{\star'} = 0$, to achieve no larger weighted sum cost.

Table 4.1: Algorithm V: Algorithm based on ellipsoid method and linear programming for solving Problem (P6).

1. Initialize $\mu_1 > 0$, $\mu_2 > 0$, $\lambda_1 > 0$ and $\lambda_2 > 0$ that satisfying the conditions in Lemma G.0.1 in Appendix G;
2. **Repeat:**
   1) Obtain $b_k^{(\{\mu_i\},\{\lambda_i\})}$, $p_k^{(\{\mu_i\},\{\lambda_i\})}$, $w_i^{(\{\mu_i\},\{\lambda_i\})}$ using (4.13)-(4.15) with the given μ_1, μ_2, λ_1 and λ_2;
   2) Compute the sub-gradient of the dual function $g(\{\mu_i\},\{\lambda_i\})$ as $-\sum_{k\in\mathcal{K}_i} p_k^{(\mu_i,\lambda_i)} - P_{c,i} + E_i^{(\mu_i)} + G_i^{(\mu_i)} + \beta_E e_{\bar{i}}^{(\mu_{\bar{i}})} + e_i^{(\mu_i)}$, $i = 1, 2$, where $p_k^{(\mu_i,\lambda_i)}$ is obtained via (4.14) and $-\sum_{k\in\mathcal{K}_i} b_k^{(\mu_i,\lambda_i)} + W_i + \beta_B w_{\bar{i}}^{(\lambda_{\bar{i}})} - w_i^{(\lambda_i)}$, $i = 1, 2$, respectively. Update μ_1, μ_2, λ_1 and λ_2 accordingly based on the ellipsoid method.
3. **Until:** $\{\mu_i\}$ and $\{\lambda_i\}$ converge to a prescribed accuracy;
4. Set b_k^* from (4.13) and p_k^* from (4.14). Determine the optimal value of the other optimization variables by solving the LP (P8).

at both systems (i.e., the energy price α_i^E and α_i^G, the available renewable energy $\bar{E}_i$, the circuit power consumption $P_{c,i}$, the channel gain g_k and their QoS requirement $\bar{r}_k$, $\forall k \in \mathcal{K}_i, i \in \{1, 2\}$) should be gathered at a central unit, which can be one of the two BSs or a third-party controller. Since the limited information is exchanged over the time scale of power and bandwidth allocation, which is on the order of seconds, while the communication block usually has a length of several milliseconds, the information exchange can be efficiently implemented. To summarize, the distributed algorithm for partial cooperation is presented in Table 4.1.

It is interesting to make a comparison between the optimal solution of Problem (P6) in Proposition 4.5.1 and that of Problem (P7) in Proposition 4.4.1. First, it follows from the solution of $w_i^\star$ in (4.15) that if $\beta_B = 1$, then the bandwidth can be allocated in the two systems more flexibly, and thus resulting in a spectrum-cooperation gain in terms of cost reduction as compared to the non-cooperative benchmark in Proposition 4.4.1. Next, from the LP in (P7) with

$0 < \beta_E \leq 1$, it is evident that the BSs will purchase energy by comparing the weighted energy prices given as $\gamma_i \alpha_i^E$ and $\gamma_i \alpha_i^G, i \in \{1, 2\}$. For instance, when system i's weighted renewable energy price $\gamma_i \alpha_i^E$ is higher than $\gamma_{\bar{i}} \alpha_{\bar{i}}^E$ of the other system, then this system i will try to first request the other system's renewable energy rather than drawing energy from its own dedicated renewable utility firm. In contrast, for the non-cooperative benchmark in Proposition 4.4.1, each BS always draws energy first from its own renewable energy, and then from the grid. Therefore, the energy cooperation changes the energy management behaviour at each BS, and thus results in an energy cooperation gain in terms of cost reduction. It is worth noting that to minimize the weighted sum cost in the fully cooperative system, it is possible for one system to contribute both spectrum and energy resources to the other (i.e., $w_i^\star > 0$ and $e_i^\star > 0$ for any $i \in \{1, 2\}$), or one system exchanges its energy while the other shares its spectrum in return (i.e., $w_i^\star > 0$ and $e_i^\star > 0$ for any $i \in \{1, 2\}$). These two scenarios are referred to as *uni-directional cooperation* and *bi-directional cooperation*, respectively.

4.6 Distributed Energy and Spectrum Cooperation for Partially Cooperative Systems

In the previous section, we have proposed an optimal centralized algorithm to achieve the whole Pareto boundary of the cost region. However, this requires the fully cooperative nature and does not apply to the scenario where the two systems have their own interests (e.g., belonging to different selfish entities). Regarding this, we proceed to present a partially cooperative system that implements the joint energy and spectrum cooperation ($0 < \beta_E \leq 1$, $\beta_B = 1$) to achieve a Pareto optimum with limited information exchange in coordination. Different from the fully cooperative system that can perform both uni-directional and bi-directional cooperation, the partially cooperative systems seek mutual benefits to decrease

both systems' costs simultaneously, in which only bi-directional cooperation is feasible.[28] In the following, we first analytically characterize the conditions for partial cooperation. Then, we propose a distributed algorithm with limited information exchange that can achieve the Pareto optimality.

4.6.1 Necessary Conditions for Feasibility of Partial Cooperation

We define a function $\bar{C}_i(\boldsymbol{x}^{\text{ex}})$ to represent the minimum cost at BS i under any given energy and spectrum cooperation scheme $\boldsymbol{x}^{\text{ex}}$, which is given as:

$$\bar{C}_i(\boldsymbol{x}^{\text{ex}}) = \underset{\boldsymbol{x}_i^{\text{in}} \succeq \mathbf{0}}{\texttt{min.}} \ \alpha_i^E E_i + \alpha_i^G G_i$$
$$\texttt{s.t.} \ (4.9\text{b}), (4.9\text{c}), (4.9\text{d}) \text{ and } (4.9\text{e}). \tag{4.16}$$

Note that based on Proposition 4.5.1, we only need to consider $\boldsymbol{x}^{\text{ex}}$ with $e_1 \cdot e_2 = 0$ and $w_1 \cdot w_2 = 0$ without loss of optimality.[29] The problem in (4.16) has a similar structure as Problem (P7), which is a special case of (4.16) with $\boldsymbol{x}^{\text{ex}} = \mathbf{0}$. Thus, we can obtain its optimal solution similarly as in Proposition 4.4.1 and the details are omitted here. We denote the optimal solution to Problem (4.16) by $E_i^{(\boldsymbol{x}^{\text{ex}})}, G_i^{(\boldsymbol{x}^{\text{ex}})}, \{b_k^{(\boldsymbol{x}^{\text{ex}})}\}$, and $\{p_k^{(\boldsymbol{x}^{\text{ex}})}\}$ and the bandwidth water-level $\nu_i^{(\boldsymbol{x}^{\text{ex}})}$. Furthermore, let the optimal dual solution associated with (4.9b) and (4.9d) be denoted by $\lambda_i^{(\boldsymbol{x}^{\text{ex}})}$ and $\mu_i^{(\boldsymbol{x}^{\text{ex}})}$,

[28] Due to the mutual benefits, we believe that both systems have incentives for partial cooperation. Moreover, such incentives can be further strengthened in the future wireless systems envisioned to have more expensive energy and spectrum.

[29] For any given energy and spectrum cooperation scheme $\boldsymbol{x}^{\text{ex}}$ with $e_1 \cdot e_2 \neq 0$ or $w_1 \cdot w_2 \neq 0$, we can always trivially find an alternative scheme $\boldsymbol{x}^{\text{ex}\prime} = [e_1', e_2', w_1', w_2']^T$ with $e_i' = e_i - \min(e_1, e_2)$ and $w_i' = w_i - \min(w_1, w_2)$ to achieve the same or smaller cost at both systems as compared to $\boldsymbol{x}^{\text{ex}}$, i.e., $\bar{C}_i(\boldsymbol{x}^{\text{ex}'}) \leq C_i(\boldsymbol{x}^{\text{ex}}), i = 1, 2$. Since $e_1' \cdot e_2' = 0$ and $w_1' \cdot w_2' = 0$ always hold, it suffices to only consider $\boldsymbol{x}^{\text{ex}}$ with $e_1 \cdot e_2 = 0$ and $w_1 \cdot w_2 = 0$.

respectively. Then, it follows that[30]

$$\mu_i^{(\boldsymbol{x}^{\mathrm{ex}})} = \begin{cases} \alpha_i^E, & \sum_{k\in\mathcal{K}_i} p_k^{(\boldsymbol{x}^{\mathrm{ex}})} + P_{c,i} - \beta_E e_{\bar{\imath}} + e_i \le \bar{E}_i \\ \alpha_i^G, & \sum_{k\in\mathcal{K}_i} p_k^{(\boldsymbol{x}^{\mathrm{ex}})} + P_{c,i} - \beta_E e_{\bar{\imath}} + e_i > \bar{E}_i \end{cases}, \tag{4.17}$$

$$\lambda_i^{(\boldsymbol{x}^{\mathrm{ex}})} = \nu_i^{(\boldsymbol{x}^{\mathrm{ex}})} \cdot \mu_i^{(\boldsymbol{x}^{\mathrm{ex}})}. \tag{4.18}$$

It is easy to verify that $\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}}), i \in \{1,2\}$, is a convex function of $\boldsymbol{x}^{\mathrm{ex}}$. Therefore, under any given $\boldsymbol{x}^{\mathrm{ex}}$, two BSs can reduce their individual cost simultaneously if and only if there exists $\boldsymbol{x}^{\mathrm{ex}\prime} = \boldsymbol{x}^{\mathrm{ex}} + \Delta\boldsymbol{x}^{\mathrm{ex}} \neq \boldsymbol{x}^{\mathrm{ex}}$ with $\Delta\boldsymbol{x}^{\mathrm{ex}} = [\Delta e_1, \Delta e_2, \Delta w_1, \Delta w_2]^T$ sufficiently small and $\boldsymbol{x}^{\mathrm{ex}\prime} \ge \mathbf{0}$ and $\boldsymbol{x}^{\mathrm{ex}\prime} \neq \mathbf{0}$ such that $\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}\prime}) < \bar{C}_i(\boldsymbol{x}^{\mathrm{ex}}), \forall i \in \{1,2\}$. In particular, by considering the non-cooperative benchmark system with $\boldsymbol{x}^{\mathrm{ex}} = \mathbf{0}$, it is inferred that partial cooperation is feasible if and only if there exists $\boldsymbol{x}^{\mathrm{ex}\prime} \ge \mathbf{0}$ and $\boldsymbol{x}^{\mathrm{ex}\prime} \neq \mathbf{0}$ such that $\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}\prime}) < \bar{C}_i(\mathbf{0})$. Based on these observations, we are ready to investigate the conditions for partial cooperation by checking the existence of such $\boldsymbol{x}^{\mathrm{ex}\prime}$. First, we derive BS i's cost change $\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}\prime}) - \bar{C}_i(\boldsymbol{x}^{\mathrm{ex}})$ analytically when the energy and spectrum cooperation decision changes from any given $\boldsymbol{x}^{\mathrm{ex}}$ to $\boldsymbol{x}^{\mathrm{ex}\prime} = \boldsymbol{x}^{\mathrm{ex}} + \Delta\boldsymbol{x}^{\mathrm{ex}}$ with sufficiently small $\Delta\boldsymbol{x}^{\mathrm{ex}}$. We have the following proposition.

Lemma 4.6.1 *Under any given $\boldsymbol{x}^{\mathrm{ex}}$, BS i's cost change by adjusting the energy and spectrum cooperation decisions is expressed as*

$$\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}} + \Delta\boldsymbol{x}^{\mathrm{ex}}) - \bar{C}_i(\boldsymbol{x}^{\mathrm{ex}}) = \nabla\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}})^T \Delta\boldsymbol{x}^{\mathrm{ex}}, \tag{4.19}$$

where $\Delta\boldsymbol{x}^{\mathrm{ex}}$ is sufficiently small, $\boldsymbol{x}^{\mathrm{ex}} + \Delta\boldsymbol{x}^{\mathrm{ex}} \ge \mathbf{0}$, and

$$\nabla\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}}) = \left[\frac{\partial\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}})}{\partial e_1}, \frac{\partial\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}})}{\partial e_2}, \frac{\partial\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}})}{\partial w_1}, \frac{\partial\bar{C}_i(\boldsymbol{x}^{\mathrm{ex}})}{\partial w_2}\right]^T. \tag{4.20}$$

[30] As will be shown later, $\mu_i^{(\boldsymbol{x}^{\mathrm{ex}})}$ and $\lambda_i^{(\boldsymbol{x}^{\mathrm{ex}})}$ can be interpreted as the marginal costs with respect to the shared energy and bandwidth between two BSs, respectively. Therefore, the result in (4.17) is intuitive, since the marginal cost should be the energy price of α_i^E if the renewable energy is excessive to support the energy consumption and energy exchange, while the marginal cost should be α_i^G if the renewable energy is insufficient.

$$\nabla \bar{C}_i(\boldsymbol{x}^{\text{ex}}) = \left[\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial e_1}, \frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial e_2}, \frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial w_1}, \frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial w_2}\right]^T = [\mu_i^{(\boldsymbol{x}^{\text{ex}})}, -\beta_E \mu_i^{(\boldsymbol{x}^{\text{ex}})}, \lambda_i^{(\boldsymbol{x}^{\text{ex}})}, -\lambda_i^{(\boldsymbol{x}^{\text{ex}})}] \tag{4.21}$$

Here, $\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial e_i} = \mu_i^{(\boldsymbol{x}^{\text{ex}})}$, $\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial e_{\bar{i}}} = -\beta_E \mu_i^{(\boldsymbol{x}^{\text{ex}})}$, $\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial w_i} = \lambda_i^{(\boldsymbol{x}^{\text{ex}})}$ *and* $\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial w_{\bar{i}}} = -\lambda_i^{(\boldsymbol{x}^{\text{ex}})}$ *can be interpreted as the marginal costs at BS* i *with respect to the energy and spectrum cooperation decisions* $e_i, e_{\bar{i}}, w_i$ *and* $w_{\bar{i}}$*, respectively.*

Proof See Appendix H.

Next, based on Lemma 4.6.1, we obtain the conditions for which the two BSs' costs can be decreased at the same time under any given $\boldsymbol{x}^{\text{ex}}$, by examining whether there exists sufficiently small $\Delta\boldsymbol{x}^{\text{ex}} \neq \mathbf{0}$ with $\boldsymbol{x}^{\text{ex}} + \Delta\boldsymbol{x}^{\text{ex}} \geq \mathbf{0}$ such that $\nabla \bar{C}_i(\boldsymbol{x}^{\text{ex}})^T \Delta\boldsymbol{x}^{\text{ex}} < 0$ for both $i = 1, 2$.

Proposition 4.6.1 *For any given* $\boldsymbol{x}^{\text{ex}}$*, the necessary and sufficient conditions that the two BSs' costs can be decreased at the same time are given as follows:*

$$\begin{cases} \frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}} > \frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}\beta_E}, \frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}} > \frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}\beta_E} & \text{, if } e_1 = e_2 = 0 \\ \frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}} \neq \frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}\beta_E} & \text{, if } e_1 > 0 \\ \frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}} \neq \frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}\beta_E} & \text{, if } e_2 > 0 \end{cases}. \tag{4.22}$$

Proof See Appendix I.

Remark 4.6.1 *Proposition 4.6.1 can be intuitively explained as follows by taking* $\frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}} > \frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}\beta_E}$ *when* $e_1 = e_2 = 0$ *as an example and the other cases can be understood by similar observations: When* $e_1 = e_2 = 0$*, the condition of* $\frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}} > \frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}\beta_E}$ *implies that we can always find* $\Delta\boldsymbol{x}^{\text{ex}} = [\Delta e_1, \Delta e_2, \Delta w_1, \Delta w_2]^T$ *sufficiently small with* $\Delta e_1 > 0, \Delta e_2 = 0, \Delta w_1 = 0$ *and* $\Delta w_2 > 0$ *such that*

$$\frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}} > \frac{\Delta e_1}{\Delta w_2} > \frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}\beta_E}.$$

In other words, there exists a new joint energy and spectrum cooperation scheme for the costs of both systems to be reduced at the same time, i.e.,

$$\nabla \bar{C}_1(\boldsymbol{x}^{\text{ex}})^T \Delta \boldsymbol{x}^{\text{ex}} = \mu_1^{(\boldsymbol{x}^{\text{ex}})} \Delta e_1 - \lambda_1^{(\boldsymbol{x}^{\text{ex}})} \Delta w_2 < 0$$

and

$$\nabla \bar{C}_2(\boldsymbol{x}^{\text{ex}})^T \Delta \boldsymbol{x}^{\text{ex}} = \lambda_2^{(\boldsymbol{x}^{\text{ex}})} \Delta w_2 - \mu_2^{(\boldsymbol{x}^{\text{ex}})} \beta_E \Delta e_1 < 0.$$

By using the marginal cost interpretation in Proposition 4.6.1, the costs at both BSs can be further reduced by transferring Δe_1 amount of energy from BS 1 to BS 2 and sharing Δw_2 amount of spectrum from BS 2 to BS 1.

Finally, we can characterize the conditions for partial cooperation by examining $\boldsymbol{x}^{\text{ex}} = \mathbf{0}$ in Proposition 4.6.1. We explicitly give the conditions as follows.

Corollary 4.6.1 *Partial cooperation is feasible if and only if*

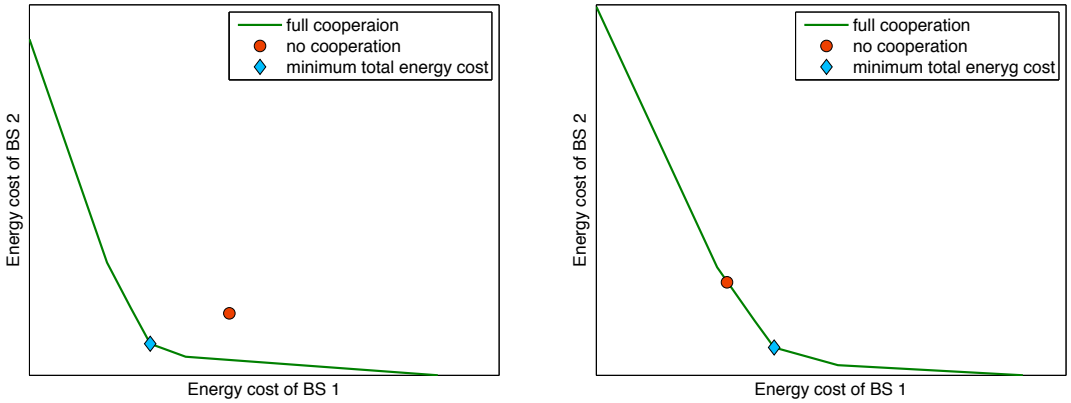
$$\frac{\lambda_1^{(\mathbf{0})}}{\mu_1^{(\mathbf{0})}} > \frac{\lambda_2^{(\mathbf{0})}}{\mu_2^{(\mathbf{0})} \beta_E}$$

and

$$\frac{\lambda_2^{(\mathbf{0})}}{\mu_2^{(\mathbf{0})}} > \frac{\lambda_1^{(\mathbf{0})}}{\mu_1^{(\mathbf{0})} \beta_E}.$$

Corollary 4.6.1 is implied by Proposition 4.6.1. More intuitively, under the condition of $\frac{\lambda_1^{(\mathbf{0})}}{\mu_1^{(\mathbf{0})}} > \frac{\lambda_2^{(\mathbf{0})}}{\mu_2^{(\mathbf{0})} \beta_E}$, it follows from Remark 4.6.1 that BS 1 is more spectrum-hungry than BS 2, while BS 2 is more insufficient of energy than BS 1. Hence, the costs at both BS can be reduced at the same time by BS 1 transferring spectrum to BS 2 and BS 2 transferring energy to BS 1. Similarly, if $\frac{\lambda_2^{(\mathbf{0})}}{\mu_2^{(\mathbf{0})}} > \frac{\lambda_1^{(\mathbf{0})}}{\mu_1^{(\mathbf{0})} \beta_E}$, the opposite is true. This shows that partial cooperation is only feasible when two systems find inter-system complementarity in energy and spectrum resources.

Example 4.6.1 *We provide an example in Fig. 4.4 to illustrate partial cooperation conditions in Corollary 4.6.1. We plot the Pareto boundary achieved by full*



(a) (b)

Figure 4.4: Two different scenarios with joint energy and spectrum cooperation: (a) partial cooperation feasible scenario, and (b) partial cooperation infeasible scenario.

cooperation, non-cooperation benchmark and the point corresponding to the minimum total cost (full cooperation with $\gamma_1 = \gamma_2 = 1$*). The joint energy and spectrum cooperation results in two scenarios as shown in Figs. 4.4a and 4.4b, which correspond to cases where partial cooperation is feasible and infeasible, respectively.*

- *Fig. 4.4a shows the feasible partial cooperation scenario, in which the partial cooperation conditions are satisfied. In this scenario, the non-cooperative benchmark is observed to lie within the Pareto boundary of cost region. As a result, from the non-cooperative benchmark, the costs of both BSs can be reduced at the same time until reaching the Pareto boundary.*

- *Fig. 4.4b shows the scenario when the partial cooperation conditions are not satisfied, where the non-cooperation benchmark is observed to lie on the Pareto boundary. From this result, it is evident that the two BSs' costs cannot be reduced at the same time. That is, the partial cooperation is infeasible.*

- *In both scenarios of Figs. 4.4a and 4.4b, it is observed that the minimum total cost point differs from the non-cooperation benchmark. This shows that full cooperation can decrease the total cost at two BSs from the non-cooperative*

benchmark even when partial cooperation is infeasible, which can be realized by uni-directional cooperation (e.g., in Fig. 4.4b).

The results in this example motivate us to propose distributed algorithms for the partial cooperation scenario to reduce two BSs' costs from non-cooperative benchmark to Pareto optimality, as will be discussed next.

4.6.2 Distributed Algorithm

In this subsection, we design a distributed algorithm to implement the energy and spectrum cooperation for two partially cooperative systems (satisfying Corollary 4.6.1).[31] Since the two systems are selfish, we need to ensure that they can improve their performance fairly. We design our algorithm based on the proportionally fair cost reduction, which is defined as follows.

Definition 4.6.1 *Proportional fair cost reduction is achieved by both systems if, for the resultant cost tuple* $(\tilde{C}_1, \tilde{C}_2)$*, the cost reduction ratio between two BSs equals the ratio of their costs in the non-cooperative scenario, i.e.,*

$$\frac{\bar{C}_1(\mathbf{0}) - \tilde{C}_1}{\bar{C}_2(\mathbf{0}) - \tilde{C}_2} = \frac{\bar{C}_1(\mathbf{0})}{\bar{C}_2(\mathbf{0})}. \tag{4.23}$$

Next, we proceed to elaborate on the key issue of the update of the energy and spectrum cooperation decision vector $\boldsymbol{x}^{\text{ex}}$ to have proportionally fair cost reductions. Our algorithm begins with the non-cooperative benchmark (i.e., $\boldsymbol{x}^{\text{ex}} = \mathbf{0}$). Then, the inter-system energy and spectrum cooperation is adjusted to decrease the costs at both BSs in each iteration. Specifically, under any given $\boldsymbol{x}^{\text{ex}}$, if the conditions in Proposition 4.6.1 are satisfied, then the two BSs cooperate by updating their energy and spectrum cooperation decision vector according to

$$\boldsymbol{x}^{\text{ex}\prime} = \boldsymbol{x}^{\text{ex}} + \delta \boldsymbol{d}, \tag{4.24}$$

[31] As long as each system agrees to install the algorithm to benefit from its efficiency and fairness, the system will not make any deviation in its decisions as the algorithm runs automatically.

where $\delta > 0$ is a sufficiently small step size and $\boldsymbol{d} \in \mathbb{R}^4$ is the direction of the update that satisfies $\nabla \bar{C}_i(\boldsymbol{x}^{\text{ex}})^T \boldsymbol{d} < 0$ (cf. (4.19)). It can be observed that there are multiple solutions satisfying this condition. Here, we choose $\boldsymbol{d}$ in each iteration as follows:

- If $\lambda_1^{(\mathbf{0})} \mu_2^{(\mathbf{0})} \beta_E > \lambda_2^{(\mathbf{0})} \mu_1^{(\mathbf{0})}$ holds, which means that costs of both systems can be reduced by system 1 sharing energy to system 2 and system 2 sharing spectrum to system 1 (cf. Corollary 4.6.1), then we choose

$$\begin{aligned} \boldsymbol{d} =& \text{sign}\left(\lambda_1^{(\boldsymbol{x}^{\text{ex}})} \mu_2^{(\boldsymbol{x}^{\text{ex}})} \beta_E - \lambda_2^{(\boldsymbol{x}^{\text{ex}})} \mu_1^{(\boldsymbol{x}^{\text{ex}})}\right) \\ & \times \left[\rho \lambda_2^{(\boldsymbol{x}^{\text{ex}})} + \lambda_1^{(\boldsymbol{x}^{\text{ex}})}, 0, 0, \mu_1^{(\boldsymbol{x}^{\text{ex}})} + \rho \beta_E \mu_2^{(\boldsymbol{x}^{\text{ex}})}\right]^T. \end{aligned} \tag{4.25}$$

- If $\lambda_2^{(\mathbf{0})} \mu_1^{(\mathbf{0})} \beta_E > \lambda_1^{(\mathbf{0})} \mu_2^{(\mathbf{0})}$ holds, which means that costs of both systems can be reduced by system 1 sharing spectrum to system 2 and system 2 sharing energy to system 1 (cf. Corollary 4.6.1), then we choose

$$\begin{aligned} \boldsymbol{d} =& \text{sign}\left(\lambda_2^{(\boldsymbol{x}^{\text{ex}})} \mu_1^{(\boldsymbol{x}^{\text{ex}})} \beta_E - \lambda_1^{(\boldsymbol{x}^{\text{ex}})} \mu_2^{(\boldsymbol{x}^{\text{ex}})}\right) \\ & \times \left[0, \rho \lambda_2^{(\boldsymbol{x}^{\text{ex}})} + \lambda_1^{(\boldsymbol{x}^{\text{ex}})}, \mu_1^{(\boldsymbol{x}^{\text{ex}})} + \rho \beta_E \mu_2^{(\boldsymbol{x}^{\text{ex}})}, 0\right]^T. \end{aligned} \tag{4.26}$$

Here, $\text{sign}(x) = 1$ if $x \geq 0$ and $\text{sign}(x) = -1$ if $x < 0$, and ρ is a factor controlling the ratio of cost reduction at both BSs in each update. With the choice of $\boldsymbol{d}$ as shown above, the decrease of cost for each BS in each update is

$$\begin{bmatrix} \bar{C}_1(\boldsymbol{x}^{\text{ex}\prime}) - \bar{C}_1(\boldsymbol{x}^{\text{ex}}) \\ \bar{C}_2(\boldsymbol{x}^{\text{ex}\prime}) - \bar{C}_2(\boldsymbol{x}^{\text{ex}}) \end{bmatrix} = \begin{bmatrix} \nabla \bar{C}_1(\boldsymbol{x}^{\text{ex}})^T \boldsymbol{d} \\ \nabla \bar{C}_2(\boldsymbol{x}^{\text{ex}})^T \boldsymbol{d} \end{bmatrix} = \sigma \begin{bmatrix} \rho \\ 1 \end{bmatrix}, \tag{4.27}$$

where $\sigma \leq 0$ is obtained by substituting (4.24) into (4.19), given by

$$\sigma = \begin{cases} -(\lambda_1^{(\boldsymbol{x}^{\text{ex}})} \mu_2^{(\boldsymbol{x}^{\text{ex}})} \beta_E - \lambda_2^{(\boldsymbol{x}^{\text{ex}})} \mu_1^{(\boldsymbol{x}^{\text{ex}})}), & \lambda_1^{(\boldsymbol{x}^{\text{ex}})} \mu_2^{(\boldsymbol{x}^{\text{ex}})} \beta_E \geq \lambda_2^{(\boldsymbol{x}^{\text{ex}})} \mu_1^{(\boldsymbol{x}^{\text{ex}})} \\ -(\lambda_2^{(\boldsymbol{x}^{\text{ex}})} \mu_1^{(\boldsymbol{x}^{\text{ex}})} \beta_E - \lambda_1^{(\boldsymbol{x}^{\text{ex}})} \mu_2^{(\boldsymbol{x}^{\text{ex}})}), & \lambda_2^{(\boldsymbol{x}^{\text{ex}})} \mu_1^{(\boldsymbol{x}^{\text{ex}})} \beta_E \geq \lambda_1^{(\boldsymbol{x}^{\text{ex}})} \mu_2^{(\boldsymbol{x}^{\text{ex}})} \\ 0, & \text{otherwise} \end{cases}.$$

Table 4.2: Algorithm VI: Distributed algorithm for partial cooperation.

1. Each BS $i \in \{1,2\}$ initializes from the non-cooperative benchmark by setting $e_i = w_i = 0$ (i.e., $\boldsymbol{x}^{\text{ex}} = \mathbf{0}$). Each BS i solves the problem in (4.16) for obtaining $\lambda_i^{(\mathbf{0})}$ and $\mu_i^{(\mathbf{0})}$, and sends them to the other BS $\bar{\imath}$;

2. Each BS $i \in \{1,2\}$ tests the conditions in Corollary 4.6.1. If $\lambda_1^{(\mathbf{0})}\mu_2^{(\mathbf{0})}\beta_E > \lambda_2^{(\mathbf{0})}\mu_1^{(\mathbf{0})}$, then choose $\boldsymbol{d}$ in (4.25) as the update vector in the following iterations. If $\lambda_2^{(\mathbf{0})}\mu_1^{(\mathbf{0})}\beta_E > \lambda_1^{(\mathbf{0})}\mu_2^{(\mathbf{0})}$, then choose $\boldsymbol{d}$ in (4.26). Otherwise, the algorithm ends. Sets ρ as in (4.28);

3. **Repeat:**

   1) Each BS $i \in \{1,2\}$ computes the dual variables $\lambda_i^{(\boldsymbol{x}^{\text{ex}})}$ and $\mu_i^{(\boldsymbol{x}^{\text{ex}})}$ by solving the problem in (4.16), and sends them to the other BS $\bar{\imath}$;

   2) BS $i \in \{1,2\}$ updates the energy and spectrum cooperation vector as $\boldsymbol{x}^{ex\prime} = \boldsymbol{x}^{ex} + \delta\boldsymbol{d}$;

   3) $\boldsymbol{x}^{ex} \leftarrow \boldsymbol{x}^{ex\prime}$.

4. **Until:** The conditions in Proposition 4.6.1 are satisfied.

From (4.27), it follows that the cost reduction in each iteration satisfies $\rho = \frac{\bar{C}_1(\boldsymbol{x}^{\text{ex}})-\bar{C}_1(\boldsymbol{x}^{\text{ex}\prime})}{\bar{C}_2(\boldsymbol{x}^{\text{ex}})-\bar{C}_2(\boldsymbol{x}^{\text{ex}\prime})}$. Using this fact together with the proportional fairness criterion in Definition 5.1, ρ is determined as

$$\rho = \frac{\bar{C}_1(\mathbf{0})}{\bar{C}_2(\mathbf{0})}. \tag{4.28}$$

Remark 4.6.2 *Generally, it follows from (4.27) that ρ controls the ratio of cost reduction at the two BSs. Besides the proportionally fair choice of ρ in (4.27), we can set other values of $\rho > 1$ (or $\rho < 1$) to ensure that a larger (or smaller) cost decrease is achieved for BS 1 compared to BS 2 (provided that the step size δ is sufficiently small). By exhausting ρ from zero to infinity, we can achieve all points on the Pareto boundary that have lower costs at both BSs than the non-cooperative benchmark.*

To summarize, the distributed algorithm for partial cooperation is presented in Table 4.2 and is described as follows. Initially, each BS $i \in \{1, 2\}$ begins from the non-cooperation benchmark case with $\boldsymbol{x}^{\text{ex}} = \mathbf{0}$ and determines the update vector $\boldsymbol{d}$ that will be used in each iterations. Specifically, each BS computes $\lambda_i^{(\mathbf{0})}$ and $\mu_i^{(\mathbf{0})}$, and shares them with each other. If $\lambda_1^{(\mathbf{0})}\mu_2^{(\mathbf{0})}\beta_E > \lambda_2^{(\mathbf{0})}\mu_1^{(\mathbf{0})}$, then choose $\boldsymbol{d}$ in (4.25) as the update vector; while if $\lambda_2^{(\mathbf{0})}\mu_1^{(\mathbf{0})}\beta_E > \lambda_1^{(\mathbf{0})}\mu_2^{(\mathbf{0})}$, then choose $\boldsymbol{d}$ in (4.26). We set the cost reduction ratio as in (4.28). Then, the following procedures are implemented iteratively. In each iteration, according to the current energy and spectrum cooperation vector $\boldsymbol{x}^{\text{ex}}$, each BS computes the dual variables $\lambda_i^{(\boldsymbol{x}^{\text{ex}})}$ and $\mu_i^{(\boldsymbol{x}^{\text{ex}})}$ by solving the problem in (4.16) and sends them to the other BS. After exchanging the dual variables, the two BSs examine the conditions in Proposition 4.6.1 individually. If the conditions are satisfied, then each BS updates the cooperation scheme $\boldsymbol{x}^{\text{ex}}$ according to (4.24). The procedure shall proceed until the two BSs cannot decrease their costs at the same time, i.e., conditions in Proposition 4.6.1 are not satisfied. Due to the fact that the algorithm can guarantee the costs to decrease proportionally fair at each iteration and the Pareto optimal costs are bounded, the algorithm can always converge to a Pareto optimal point with proportional fairness provided that the step size δ is sufficiently small. Note that Algorithm VI in Table 4.2 minimizes both systems' costs simultaneously based on the gradients of two convex cost functions in (4.16), which differs from the conventional gradient descent method in convex optimization which minimizes a single convex objective [80].

Compared to the centralized joint energy and spectrum cooperation scheme, which requires a central unit to gather all channel and energy information at two systems, the distributed algorithm only needs the exchange of four scalers (i.e., the marginal spectrum and energy prices $\lambda_i^{(\boldsymbol{x}^{\text{ex}})}$ and $\mu_i^{(\boldsymbol{x}^{\text{ex}})}, \forall i \in \{1, 2\}$) between two BSs in each iteration. As a result, such distributed algorithm can preserve the two systems' privacy and greatly reduce the cooperation complexity (e.g., signalling overhead).

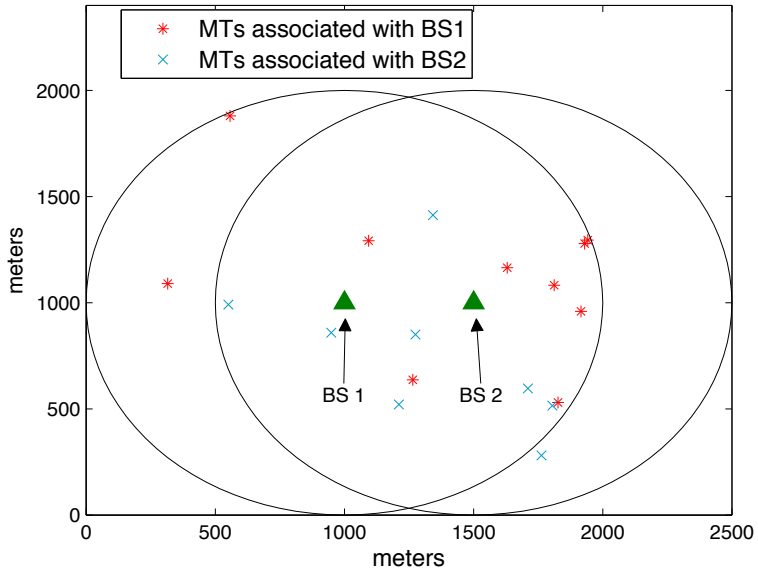


Figure 4.5: Simulation setup for joint energy and spectrum cooperation.

4.7 Numerical Results

In this section, we provide numerical results for evaluating the performance of our proposed joint energy and spectrum cooperation. For the simulation setup, we assume that BS 1 and BS 2 each covers a circular area with a radius of 500 meters (m) as shown in Fig. 4.5. $K_1 = 10$ and $K_2 = 8$ MTs are randomly generated in the two cells. We consider a simplified path loss model for the wireless channel with the channel gain set as $g_k = c_0(\frac{d_k}{d_0})^{-\zeta}$, where $c_0 = -60$dB is a constant path loss at the reference distance $d_0 = 10$ m, d_k is the distance between MT k and its associated BS in meter and $\zeta = 3$ is the path loss exponent. The noise PSD at each MT is set as $N_0 = -150$ dBm/Hz. Furthermore, we set the non-transmission power consumption for the BSs as $P_{c,1} = P_{c,2} = 100$ Watts(W). The maximum usable renewable energy at the two BSs are $\bar{E}_1 = 190$ W and $\bar{E}_2 = 130$ W, respectively. We set the energy price from renewable utility firm and power grid as $\alpha_i^E = 0.2$/W and $\alpha_i^G = 1$/W, respectively, where the price unit is normalized for simplicity. The bandwidth for the two BSs are $W_1 = 15$ MHz and $W_2 = 20$ MHz, respectively.

Fig. 4.6 shows the BSs' optimized costs by the proposed joint energy and

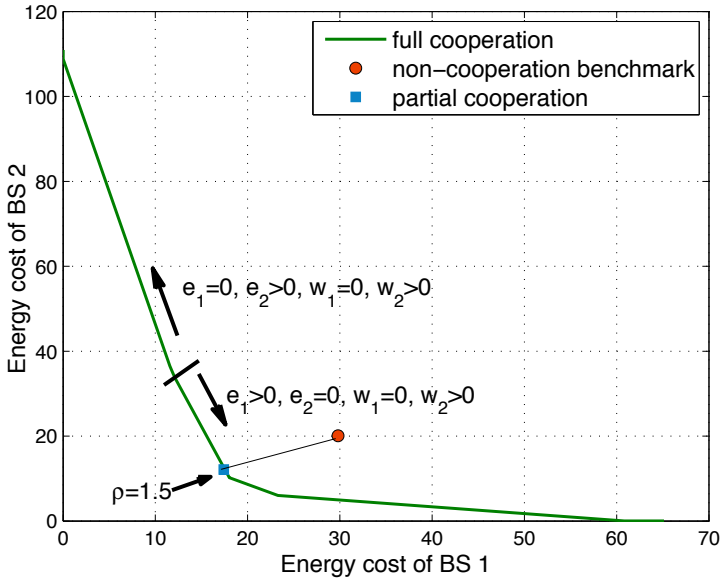


Figure 4.6: Energy cost region for the case of joint energy and spectrum cooperation versus the case without energy or spectrum cooperation.

spectrum cooperation in full cooperation with $\beta_B = 1$ and $\beta_E = 0.8$ compared with the non-cooperation benchmark. Notice that Fig. 4.6 only shows a cooperation case at a time slot that BS 2 has relatively more bandwidth (considering its realized traffic load) than BS 1 and the spectrum cooperation is from BS 2 to BS 1. Yet, in other time slots, two different BSs' traffic loads and channel realizations can change and their spectrum cooperation may follow a different direction. It is observed that the non-cooperation benchmark lies within the Pareto boundary achieved by full cooperation, while the partial cooperation lies on that Pareto boundary. This indicates the benefit of joint energy and spectrum cooperation in minimizing the two systems' costs. It is also observed that the Pareto boundary in full cooperation is achieved by either uni-directional cooperation with BS 2 transferring both energy and spectrum to BS 1 (i.e., $e_1 = 0, e_2 > 0$ and $w_1 = 0, w_2 > 0$) or bi-directional cooperation with BS 1 transferring energy to BS 2 and BS 2 transferring spectrum to BS 1 (i.e., $e_1 > 0, e_2 = 0$ and $w_1 = 0, w_2 > 0$). Specifically, the energy costs at both BSs are decreased simultaneously compared to the non-cooperation benchmark only

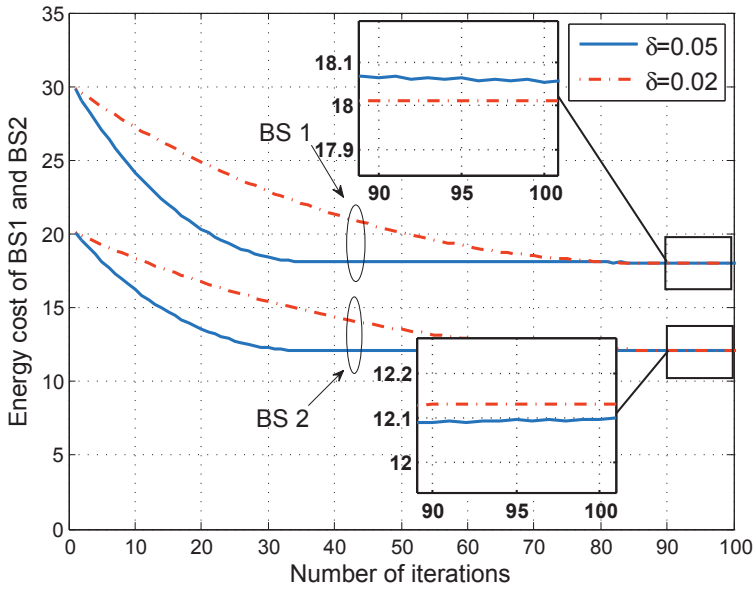
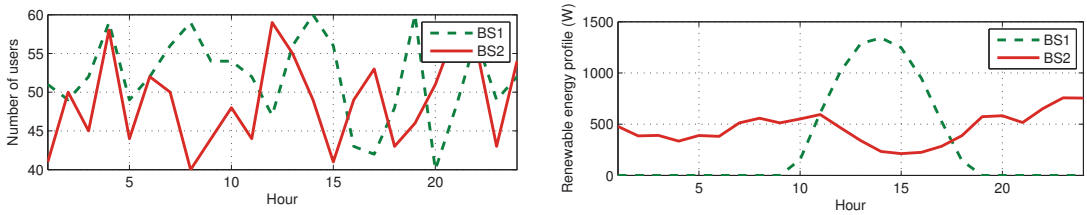


Figure 4.7: Convergence of the distributed algorithm under different step-size δ's.

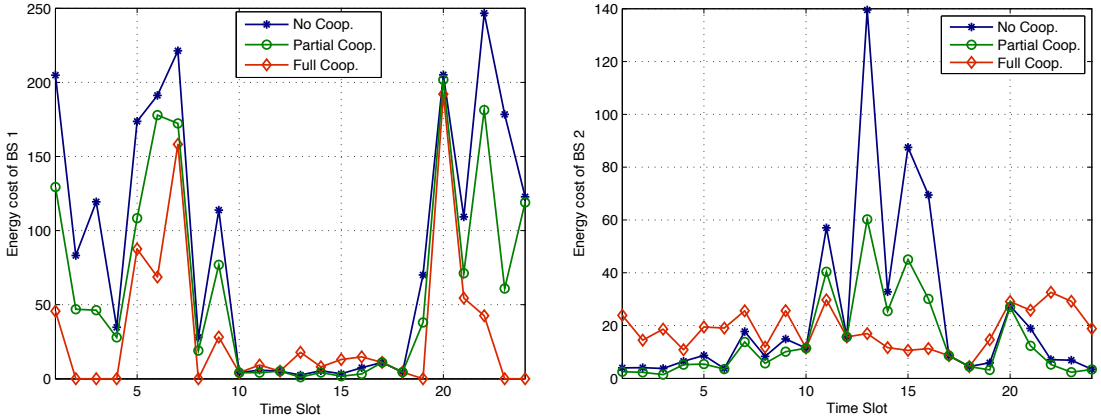
in the case of bidirectional cooperation. This is intuitive, since otherwise the cost of the BS that shares both resources will increase. Furthermore, for our proposed distributed algorithm, it is observed that it converges to the proportional fair result with $\rho = 1.5$, which lies on the Pareto boundary in this region. This is also expected, since partial cooperation is only feasible when both systems find complementarity between energy and spectrum resources.

In Fig. 4.7, we show the convergence of the partially cooperative distributed algorithm under step-sizes $\delta = 0.05$ and $\delta = 0.02$, and $\rho = 1.5$ is chosen to achieve the proportional fairness. It is observed that under different step-sizes, the costs at two BSs converge to different points on the Pareto boundary. Specifically, given $\delta = 0.05$, the cost reductions at BS 1 and BS 2 are observed to be 11.7606 (from 29.8092 to 18.0423) and 7.9825 (from 20.0860 to 12.1035), respectively, with the cost reduction ratio being 11.7606/7.9825=1.4733; while given $\delta = 0.02$, the cost reductions at BS 1 and BS 2 are observed to be 11.7968(from 29.8092 to 18.0124) and 7.9625 (from 20.0860 to 12.1235), respectively, with the cost reduction ratio being 11.7968/7.9625=1.4815. By comparing the cost reduction ratios in two cases with $\rho = 1.5$, it is inferred that the proportional fairness can be better guaranteed



(a) (b)

Figure 4.8: Profiles of (a) number of users K_1 and K_2, and (b) renewable energy $\bar{E}_1$ and $\bar{E}_2$ at the two systems for simulation.



(a) (b)

Figure 4.9: Energy costs of (a) BS1, and (b) BS2 under full, partial and no cooperation.

with smaller δ. This also validates that δ should be sufficiently small to ensure the proportionally cost reduction in each iteration step (see Section 4.6). It is also observed that the algorithm converges after about 40 iterations for $\delta = 0.05$. This indicates that under a proper choice of δ, the convergence speed is very fast provided that certain proportional fairness inaccuracies are admitted.

Finally, we show the optimized costs at both BSs over time by considering stochastically varying traffic load and harvested renewable energy. We assume that BS 1 and BS 2 are powered by solar and wind energy, respectively, and their energy harvesting rates are based on the real-world solar and wind data from Elia, a Belgium

electricity transmission system operator.[32] For demonstration, we use the average harvested energy over one hour as $\bar{E}_i$ at each slot, as shown in Fig. 4.8a, and thus our studied 24 slots correspond to the energy harvesting profile over one day. Furthermore, we consider the number of MTs served by each BS, K_i's, over slots as shown in Fig. 4.8b, which are randomly generated based on a discrete uniform distribution over the interval [40, 60] for simplification. Under this setup, Fig. 4.9 shows the optimized costs of the two BSs by different schemes. It is observed that over the 24 slots the full and partial cooperation achieve 55.68% and 33.75% total cost reduction for the two BSs as compared to the no cooperation benchmark, respectively. It is also observed that for partial cooperation, the energy costs of both BSs are reduced at the same time, while for full cooperation, at certain slots, i.e., slots 22, the cost of one BS is reduced significantly at the expense of the cost increase of the other BSs. The reason is as follows. In partial cooperation, the two systems seek for mutual benefits and thus only bidirectional cooperation is feasible. For instance, at time slot 6, BS 1 shares bandwidth to BS 2 and BS 2 transfers energy to BS 1. In contrast, under full cooperation, it is not required that the costs of both BSs be reduced at the same time and the common goal is to reduce the sum energy costs. Hence, in this case, it is possible that uni-directional cooperation happens where one BS sacrifices its interest to another one in order to reduce the sum energy cost.

4.8 Chapter Summary

In this chapter, we proposed a new cooperation technique of joint energy and spectrum cooperation to reduce the energy costs and address the incentive issues in wireless cellular systems powered by both conventional and renewable energy. We aimed to minimize the costs of both systems by considering two scenarios where different systems belong to the same entity and different entities, respectively. In the former case with full cooperation, we proposed an optimal centralized algorithm for

[32] See http://www.elia.be/en/grid-data/power-generation/

achieving the minimum weighted-sum energy cost at two BSs to characterize their trade-offs. In the latter case with partial cooperation, we developed a distributed algorithm to achieve the Pareto optimal energy costs with proportional fairness. Our results in this chapter provide key insights to the design of cooperative cellular systems powered by emerging smart grid utilizing both conventional and renewable energy.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

This thesis made an in-depth investigation of advanced cooperation techniques in the wireless cellular network under the new setups of energy-aware cooperative relaying, cooperative local caching with heterogeneous file preferences, and joint energy and spectrum cooperation among cellular BSs powered by the smart grid. We summarize the thesis as follows.

In Chapter 2, with the consideration of the MTs' heterogeneous battery levels, we investigated the cooperative relaying for the energy saving in the wireless cellular network. We proposed a novel pricing-based mechanism for motivating the cooperative relaying between MTs in practice. For the benchmark case of full cooperation, the problem was formulated as a relay selection problem and a simple threshold structure was shown for the optimal solution. Then, for the general case of partial cooperation, under the uncertainties of the MTs' channel conditions and battery levels, we formulated the pricing and load sharing problem as an optimization problem. Efficient algorithms based on dichotomous search and alternating optimization were proposed for the optimal solutions. Finally, extensive simulations were provided, which showed that our cooperative relaying scheme could significantly decrease the number of communication and battery outages for the MTs and increase the average battery level during their operations. Our proposed pricing-based mechanism provides a practical design for the MTs' energy saving by cooperative relaying in the wireless cellular network.

In Chapter 3, under the heterogeneous file preferences among the MTs, we

investigated the fundamental performance trade-offs among the MTs and discussed the impacts of their selfish behaviours to the cooperative local caching. We defined the utilities of different groups as the MTs' probability of successfully obtaining the file with the SRC. The closed-form result of the group utility was derived under the HPPP mobility model. For the case of full cooperation where all the MTs belong to the same entity, an efficient algorithm based on coordinate descent was proposed for characterizing the complete performance trade-offs between different groups. In addition, for the two benchmark cases under selfish caching, namely, partial and no cooperation, closed-form optimal caching distributions were obtained by utilizing the KKT conditions. Finally, performance comparison was made among these three schemes under various practical setups. Our study on cooperative local caching provides new insights on the design of wireless systems with local caching under heterogeneous file preferences.

In Chapter 4, with the integration of the renewable energy supply in the wireless network, we proposed the model of joint energy and spectrum cooperation for reducing the energy cost of the BSs. We minimized the costs of two cooperating systems by considering two scenarios where the wireless systems belong to the same operator and different operators, respectively. In the former case with full cooperation, we proposed an optimal centralized algorithm for achieving the minimum weighted-sum energy cost at two BSs and fully characterized the trade-offs between the two systems. In the latter case with partial cooperation, we specified the condition for partial cooperation and developed a distributed algorithm to achieve the Pareto optimal energy costs gradually from the non-cooperation benchmark. Our proposed cooperation technique of joint energy and spectrum cooperation provides an attractive new design for the practical implementation of the wireless network powered by utilizing renewable energy.

In summary, the results of this thesis shed essential and new insights on the optimal design and practical implementation of advanced cooperation techniques for the future wireless communication systems.

5.2 Future Work

Finally, we highlight several directions that we think are important and worthy of further investigation beyond the topics presented in this thesis.

For the energy-aware cooperative relaying considered in Chapter 2, we assumed single–relay selection for the tractability of analysis and the demonstration of the insights. Clearly, a more general scenario is to consider multiple relays selection, which will lead to higher flexibility for resource allocation but also higher complexity for the cooperation. Furthermore, we considered the static optimization of the pricing and load sharing over a single time slot under our proposed pricing-based mechanism. Future work can investigate the dynamic optimization of transmit energy and pricing under multiple time slots considering the dynamics of the MTs' battery levels. This leads to a problem that is more complicated to solve and the new mathematical tool of dynamic programming is required for the solution.

For the cooperative local caching problem addressed in Chapter 3, we considered the cooperative local caching between MTs with only single-hop SRC. It would be interesting to generalize the results to multiple hops, which will offer more opportunities for file sharing among MTs and the heterogeneity in the file preferences of different MTs can be further mitigated due to higher probability of file sharing. Moreover, in this thesis, we considered unicast of the file sharing. It will be interesting yet challenging to extend to the case of coded multicast [71] for the cooperative local caching with the advanced techniques of network coding. Last but not least, since different groups have conflict of interests under heterogeneous file preference, it is also pertinent to consider the behaviours of the MTs in competitive environment from the perspective of *game theory*.

For joint energy and spectrum cooperation proposed in Chapter 4, we considered energy cooperation without storage at BSs. When energy storage is implemented in the systems, the energy management of the BS can have more flexibility such that the energy supply variations in time can be mitigated and energy cooperation between the two systems can more efficiently exploit the geographical

energy diversity. Furthermore, we considered the case where two systems operate over orthogonal frequency bands. In general, allowing MTs to be associated with different BSs to share the same frequency band may further improve the spectrum efficiency. Finally, we discussed the joint energy and spectrum sharing for energy saving under the setup of hybrid energy supply. In addition, another interesting direction could be the optimization of QoS performance with energy cooperation subject to the resource (i.e., spectrum and power) constraints. Depending on different application scenarios, the QoS metrics can be delay [100], throughput [101], etc.

Appendix A

Proof of Proposition 2.4.1

First, we take the first-order derivative of the sum energy cost $\zeta_j E_j^{(C,R)} + \zeta_k E_k^{(C,S)}$ with respect to $D_k^{(R)}$ and obtain

$$-\ln 2 \frac{\zeta_k \sigma^2}{G_k \eta_k} 2^{D_k - D_k^{(R)}} + \ln 2 \frac{\zeta_j \sigma^2}{G_k \eta_j} 2^{D_k^{(R)}}. \tag{A.1}$$

Then, by setting it to zero, we obtain the stationary point of the objective function as

$$\tilde{D}_k^{(R)} = \frac{1}{2}\left(D_k + \log_2 \frac{\theta_k}{\theta_j}\right). \tag{A.2}$$

Finally, we discuss the optimal solution in the region of $[0, D_k]$ for the following 3 cases.

- $\tilde{D}_k^{(R)} \in (-\infty, 0]$: In this case, the sum energy cost is monotonically increasing within the region of $(0, D_k]$. Hence, the optimal solution is $\hat{D}_k^{(R)} = 0$.
- $\tilde{D}_k^{(R)} \in (0, D_k]$: The stationary point $\tilde{D}_k^{(R)}$ is contained in this region. Hence, the optimal solution is $\hat{D}_k^{(R)} = \frac{1}{2}\left(D_k + \log_2 \frac{\theta_k}{\theta_j}\right)$.
- $\tilde{D}_k^{(R)} \in (D_k, \infty)$: In this case, the sum energy cost is monotonically decreasing within the region $(0, D_k]$. Hence, the optimal solution is $\hat{D}_k^{(R)} = D_k$.

The proof of Proposition 3.1 is thus completed.

Appendix B

Proof of Proposition 2.5.1

We need to prove the objective of the Problem (P2) is marginally convex with respect to π_k and $D_k^{(R)}$. Denote

$$\psi_I = 1 - \left(1 - \frac{\phi_k}{\zeta_{\max}}\left(1 - e^{-\frac{\zeta_{\max}}{\phi_k}}\right)\right)^n,$$

where $\phi_k = \frac{G_i(\pi_k - \epsilon)}{\sigma^2(2^{D_k^{(R)}} - 1)}$ and

$$\psi_{II} = \pi_k + \zeta_k E_k^{(C,S)} - \zeta_k E_k^{(D,S)}.$$

We first prove that $\psi_I \psi_{II}$ is marginally convex with respect to π_k and $D_k^{(R)}$. Then, the convexity of Problem (P2)'s objecive function

$$\sum_{n=0}^{\infty} \mathbb{P}(N_k = n)\left\{\psi_I \psi_{II} + \zeta_k E_k^{(D,S)}\right\}$$

follows naturally.

- First, we prove that the $\psi_I \psi_{II}$ is marginally convex with respect to π_k. In the first place, we denote $\psi_I(\pi_k) = 1 - \psi_j(\pi_k)^{N_k},\ j \in \mathcal{H}_k$ and prove that $\psi_j(\pi_k) = 1 - \frac{\phi_k}{\zeta_{\max}}\left(1 - e^{-\frac{\zeta_{\max}}{\phi_k}}\right)$ is a convex function with respect to π_k. By taking the second-order derivative of $\psi_j(\pi_k)$, we can have

$$\psi_j''(\pi_k) = \frac{\zeta_{\max} e^{-\frac{\zeta_{\max}}{\phi_k}}}{\phi_k(\pi_k - \epsilon)^2},\ j \in \mathcal{H}_k,\ k \in \mathcal{K}_S. \tag{B.1}$$

Since $\psi_j''(\pi_k) > 0$, $\psi_j(\pi_k)$ is a convex function. It can be verified that $\psi_j(\pi_k)^{N_k}$ is also convex because $\psi_j(\pi_k)$ is monotonically decreasing and convex. Hence, $\psi_I(\pi_k)$ is concave and monotonically increasing.

Then, it can be verified that $\psi_{II}''(\pi_k) = 0$ since $\psi_{II}(\pi_k)$ is a linear function. Because $\psi_I(\pi_k)$ and $\psi_{II}(\pi_k)$ are both monotonically increasing with respect to π_k, it can be obtained that $\psi_I'(\pi_k)\psi_{II}'(\pi_k) > 0$. Also, due to the fact that $\psi_I''(\pi_k) < 0$ and $\psi_{II}(\pi_k) \leq 0$, $\psi_I''(\pi_k)\psi_{II}(\pi_k)$ is also positive. Then, the second-order derivative of $\psi_I(\pi_k)\psi_{II}(\pi_k)$ can be expressed as

$$\underbrace{\psi_I''(\pi_k)}_{\leq 0}\underbrace{\psi_{II}(\pi_k)}_{\leq 0} + 2\underbrace{\psi_I'(\pi_k)}_{\geq 0}\underbrace{\psi_{II}'(\pi_k)}_{\geq 0} + \psi_I(\pi_k)\underbrace{\psi_{II}''(\pi_k)}_{=0}. \tag{B.2}$$

Therefore, it can be verified that (B.2) is positive and thus $\psi_I(\pi_k)\psi_{II}(\pi_k)$ is a convex function with respect to π_k with $D_k^{(R)}$ fixed.

- Second, by a similar approach as above, we can prove that $\psi_I(D_k^{(R)})$ is monotaonically decreasing and concave with respect to $D_k^{(R)}$ and $\psi_{II}(D_k^{(R)})$ is montonically decreasing and convex with respect to $D_k^{(R)}$. Then, by taking the second-order derivative of $\psi_I(D_k^{(R)})\psi_{II}(D_k^{(R)})$ with respect to $D_i^{(R)}$, we have

$$\underbrace{\psi_I''(D_k^{(R)})}_{\leq 0}\underbrace{\psi_{II}(D_k^{(R)})}_{\leq 0} + 2\underbrace{\psi_I'(D_k^{(R)})}_{\leq 0}\underbrace{\psi_{II}'(D_k^{(R)})}_{\leq 0} + \underbrace{\psi_I(D_k^{(R)})}_{\geq 0}\underbrace{\psi_{II}''(D_k^{(R)})}_{\geq 0}. \tag{B.3}$$

Hence, we have proved that the function $\psi_I(D_k^{(R)})\psi_{II}(D_k^{(R)})$ is also marginally convex with respect to $D_k^{(R)}$ with π_k fixed.

Proposition 2.5.1 thus follows.

Appendix C

Proof of Theorem 3.3.1

Consider a typical MT $k \in \mathcal{K}_i,\ i \in \mathcal{G}$ in group i. We denote the MTs from group $j \in \mathcal{G}$ that falls into the range of MT k in group i as $\mathcal{N}_j = \{l \in \mathcal{K}_j | X_{j,l} \in B(X_{i,k}, d)\} \subseteq \mathcal{K}_j$ and denote the number of MTs in this set as $N_j = |\mathcal{N}_j|$. Please be noted that because of the Slyvnyak-Mecke Theorem [102], the distribution of the homogeneous Poisson point process Φ_i of group $i \in \mathcal{G}$ conditioning on any particular point, is the same as the distribution of the point process Φ_i. Hence, although we are conditioning on the typical MT $k \in \mathcal{K}_i$ for file request in group $i \in \mathcal{G}$, the density of the point process Φ_i is still $\bar{\lambda}_i$. According to the model of HPPP, it follows that N_j is a Poisson random variable with mean $\bar{\mu}_j = \pi d^2 \bar{\lambda}_j,\ j \in \mathcal{G}$ and its PMF is given by

$$\mathbb{P}[N_j] = \frac{\bar{\mu}_j^{N_j}}{N_j!} e^{-\bar{\mu}_j},\ N_j = 0, 1, \cdots, \infty. \tag{C.1}$$

With the theorem of iterated expectation, group i's utility is

$$U_i(\boldsymbol{c}_i; \boldsymbol{c}_{-i}) = \sum_{f \in \mathcal{F}} r_{i,f} \mathbb{P}(E|f; \boldsymbol{c}_i, \boldsymbol{c}_{-i}). \tag{C.2}$$

Then, given the requested file $f \in \mathcal{F}$ and the neighbouring MTs $\mathcal{N}_j \subseteq \mathcal{K}_j,\ j \in \mathcal{G}$, the probability for successful file discovery is

$$\begin{aligned} \mathbb{P}(E|f, \{\mathcal{N}_j\}; \boldsymbol{c}_i, \boldsymbol{c}_{-i}) &= c_{i,f} + (1 - c_{i,f}) \Big[1 - \prod_{j \in \mathcal{G}} (1 - c_{j,f})^{N_j}\Big] \\ &= 1 - c'_{i,f} \prod_{j \in \mathcal{G}} {c'_{j,f}}^{N_j}, \end{aligned} \tag{C.3}$$

where $\prod_{j\in\mathcal{G}}(1-c_{j,f})^{N_j}$ denotes the probability that file $f \in \mathcal{F}$ is not found in caches of the neighbouring MTs from all groups. Then, the probability of successful file discovery is

$$\begin{aligned}
\mathbb{P}(E;\boldsymbol{c}_i,\boldsymbol{c}_{-i}) &= \sum_{f\in\mathcal{F}} r_{i,f}\mathbb{P}(E|f;\boldsymbol{c}_i,\boldsymbol{c}_{-i}) \\
&= \sum_{f\in\mathcal{F}} r_{i,f} \sum_{N_1}\mathbb{P}[N_1]\sum_{N_2}\mathbb{P}[N_2]\cdots\sum_{N_K}\mathbb{P}[N_G]\mathbb{P}[E|f,\{\mathcal{N}_j\};\boldsymbol{c}_i,\boldsymbol{c}_{-i}] \\
&= \sum_{f\in\mathcal{F}} r_{i,f} \sum_{N_1=0}^{\infty} e^{-\bar{\mu}_1}\frac{\bar{\mu}_1^{N_1}}{N_1!}\sum_{N_2=0}^{\infty} e^{-\bar{\mu}_2}\frac{\bar{\mu}_2^{N_2}}{N_2!}\cdots\sum_{N_G=0}^{\infty} e^{-\bar{\mu}_G}\frac{\bar{\mu}_G^{N_G}}{N_G!}\Big(1-c'_{i,f}\prod_{j\in\mathcal{G}}{c'_{j,f}}^{N_j}\Big) \\
&\overset{(a)}{=} \sum_{f\in\mathcal{F}} r_{i,f}\left(1-c'_{i,f}e^{-\bar{\mu}_0 c_f}\right), \qquad \text{(C.4)}
\end{aligned}$$

where (a) is obtained by utilizing (3.3) and the Taylor expansion for exponential function $e^x = \sum_{i=0}^{\infty} x^i/i!$. Theorem 3.3.1 is thus proved.

Appendix D

Proof of Lemma 3.4.1

We obtain the Hessian matrix of objective function of Problem (P3) with respect to $\boldsymbol{c}_i$ as

$$\frac{\partial^2 U(\{\boldsymbol{c}_i\},\boldsymbol{\gamma})}{\partial {c'_{i,f}}^2} = -\left(\gamma_i r_{i,f} e^{-\bar{\mu}_0 c_f}\bar{\mu}_i(2+c'_{i,f}\bar{\mu}_i) + \sum_{j\in\mathcal{G}\setminus\{i\}} \gamma_j r_{j,f} c'_{j,f} e^{-\bar{\mu}_0 c_f}\bar{\mu}_i^2\right), \quad \text{(D.1)}$$

$$\frac{\partial^2 U(\{\boldsymbol{c}_i\},\boldsymbol{\gamma})}{\partial c'_{i,f}\partial c'_{i,f'}} = 0, \quad \text{(D.2)}$$

where $f' \in \mathcal{F}\setminus\{f\}$. Hence, by the negative diagonal dominance of the Hessian matrix, its negative semi-definiteness is proved. Hence, the objective function is marginally concave with respect to each individual group's caching distribution $\boldsymbol{c}_i$.

Appendix E

Proof of Theorem 3.4.1

First, it should be noted that maximizing the objective function in the original problem is equivalent to minimizing the objective function $\sum_{j\in\mathcal{G}_+}\gamma_j\sum_{f\in\mathcal{F}}r_{j,f}c'_{j,f}e^{-\bar{\mu}_0 c_f}$. By introducing the dual variable ν_i for the constraint in (3.11a), the partial Lagrangian for Problem $(\mathrm{P4}-1)$ is

$$\begin{aligned} L\left(\boldsymbol{c}_i,\nu_i\right) &= \sum_{j\in\mathcal{G}_+}\gamma_j\sum_{f\in\mathcal{F}}r_{j,f}c'_{j,f}e^{-\bar{\mu}_0}e^{\sum_{i\in\mathcal{G}}\bar{\mu}_i c'_{i,f}} - \nu_i\Big(\sum_{f\in\mathcal{F}}c'_{i,f}-F+1\Big) \\ &= \sum_{f\in\mathcal{F}}\left(\sum_{j\in\mathcal{G}_+}\gamma_j r_{j,f}c'_{j,f}e^{-\bar{\mu}_0}e^{\sum_{i\in\mathcal{G}}\bar{\mu}_i c'_{i,f}} - \nu_i c'_{i,f}\right) + \nu_i(F-1). \end{aligned} \tag{E.1}$$

The dual function is then given as

$$\begin{aligned} g(\nu_i) = \underset{\boldsymbol{c}'_i}{\mathtt{min.}}\ & L\left(\boldsymbol{c}_i,\nu_i\right) \\ \mathtt{s.t.}\ & 0\le c'_{i,f}\le 1,\ f\in\mathcal{F}. \end{aligned} \tag{E.2}$$

The associated dual problem is then defined as

$$\underset{\nu_i\ge 0}{\mathtt{max.}}\quad g(\nu_i). \tag{E.3}$$

Because the primal problem is convex and satisfies the Slater's condition [80], the duality gap between the primal and dual problem is zero. Hence, the problem can be optimally solved in the dual domain by first minimizing the Lagrangian $L\left(\boldsymbol{c}_i,\nu_i\right)$ for a given ν_i and then maximizing the dual function $g(\nu_i)$ with respect to ν_i. For a given ν_i, the dual function can be decomposed into parallel sub-problems, each for

a given file $f \in \mathcal{F}$ as follows.

$$(\mathrm{P4}-2): \quad \underset{0\le c'_{i,f}\le 1}{\texttt{min.}} \quad \sum_{j\in\mathcal{G}_+} \gamma_j r_{j,f} c'_{j,f} e^{-\bar{\mu}_0} e^{\sum_{i\in\mathcal{G}} \bar{\mu}_i c'_{i,f}} - \nu_i c'_{i,f} \tag{E.4}$$

First, according to the KKT conditions, the following equations should be satisfied by the optimal primal and dual solutions,

$$\nu_i^* \left(\sum_{f\in\mathcal{F}} {c'_{i,f}}^* - F + 1 \right) = 0, \tag{E.5}$$

$$0 \le {c'_{i,f}}^* \le 1, \ f \in \mathcal{F}. \tag{E.6}$$

Then, the Lagrangian for the sub-problem of file $f \in \mathcal{F}$ is

$$L_f(c'_{i,f}, \nu_{i,f,l}, \nu_{i,f,r}) = \sum_{j\in\mathcal{G}_+} \gamma_j r_{j,f} c'_{j,f} e^{-\bar{\mu}_0} e^{\sum_{i\in\mathcal{G}} \bar{\mu}_i c'_{i,f}} - \nu_i c'_{i,f} - \nu_{i,f,l} c'_{i,f} + \nu_{i,f,r}(c'_{i,f} - 1), \tag{E.7}$$

where $\nu_{i,f,l} \ge 0$ is the dual variable for $c'_{i,f} \ge 0$, and $\nu_{i,f,r} \ge 0$ is the dual variable for $c'_{i,f} \le 1$. By taking the partial derivative of the Lagrangian with respect to $c'_{i,f}$, we obtain

$$\frac{\partial L_f}{\partial c'_{i,f}} = e^{-\bar{\mu}_0} e^{\sum_{i\in\mathcal{G}} \bar{\mu}_i c'_{i,f}} \left(\gamma_i r_{i,f} + \sum_{j\in\mathcal{G}_+\setminus\{i\}} \gamma_j r_{j,f} c'_{j,f} \bar{\mu}_i + \gamma_i r_{i,f} \bar{\mu}_i c'_{i,f} \right) - \nu_i - \nu_{i,f,l} + \nu_{i,f,r}. \tag{E.8}$$

By the KKT conditions, we can obtain the following system of equations:

$$\frac{\partial L_f}{\partial c'_{i,f}} = 0, \tag{E.9}$$

$$\nu^*_{i,f,l} {c'_{i,f}}^* = 0, \ \nu^*_{i,f,r}({c'_{i,f}}^* - 1) = 0, \tag{E.10}$$

$$\nu^*_{i,f,l}, \nu^*_{i,f,r} \ge 0. \tag{E.11}$$

With (E.9), we can obtain that

$$e^{-\bar{\mu}_i c'_{i,f}} = \frac{A'_{i,f}}{C'_{i,f}} + \frac{B'_{i,f}}{C'_{i,f}} c'_{i,f}, \tag{E.12}$$

where

$$A'_{i,f} = \left(\gamma_i r_{i,f} + \sum_{j \in \mathcal{G}_+ \setminus \{i\}} \gamma_j r_{j,f} c'_{j,f} \bar{\mu}_i \right) e^{-\bar{\mu}_0} e^{\sum_{j \in \mathcal{G} \setminus \{i\}} \bar{\mu}_j c'_{j,f}} \tag{E.13}$$

$$B'_{i,f} = \gamma_i r_{i,f} \bar{\mu}_i e^{-\bar{\mu}_0} e^{\sum_{j \in \mathcal{G} \setminus \{i\}} \bar{\mu}_j c'_{j,f}}, \tag{E.14}$$

$$C'_{i,f} = \nu_i + \nu_{i,f,l} - \nu_{i,f,r}. \tag{E.15}$$

By utilizing the equation $e^{ax+b} = cx + d \to x = -\frac{1}{a} W(-\frac{a}{c} e^{b - \frac{ad}{c}}) - \frac{d}{c}$, where $W(\cdot)$ is the Lambert-W function, $a = -\bar{\mu}_i$, $b = 0$, $c = \frac{B'_{i,f}}{C'_{i,f}}$ and $d = \frac{A'_{i,f}}{C'_{i,f}}$. Hence, it follows that

$$c'_{i,f} = \frac{1}{\bar{\mu}_i} W\left(A''_{i,f} e^{\bar{\mu}_i B_{i,f}} \right) - B_{i,f}, \tag{E.16}$$

where

$$A''_{i,f} = \frac{\bar{\mu}_i C'_{i,f}}{B'_{i,f}} = \frac{\nu_i + \nu_{i,f,l} - \nu_{i,f,r}}{\gamma_i e^{-\bar{\mu}_0} r_{i,f} e^{\sum_{j \in \mathcal{G} \setminus \{i\}} \bar{\mu}_j c'_{j,f}}}, \tag{E.17}$$

$$B_{i,f} = \frac{A'_{i,f}}{B'_{i,f}} = \frac{1}{\bar{\mu}_i} + \sum_{j \in \mathcal{G}_+ \setminus \{i\}} \frac{\gamma_j r_{j,f}}{\gamma_i r_{i,f}} c'_{j,f}. \tag{E.18}$$

For notational convenience, in the following, we denote

$$\tilde{c}_{i,f}(\nu_i, \nu_{i,f,l}, \nu_{i,f,r}) = \frac{1}{\bar{\mu}_i} W\left(A''_{i,f} e^{\bar{\mu}_i B_{i,f}} \right) - B_{i,f}. \tag{E.19}$$

Because the Lambert-W function is monotonically increasing, the function $\tilde{c}_{i,f}(\nu_i, \nu_{i,f,l}, \nu_{i,f,r})$ is monotonically increasing with respect to $\nu_{i,f,l}$ and decreasing with respect to $\nu_{i,f,r}$.

Then, we discuss the following three cases of $\tilde{c}_{i,f}(\nu_i, \nu_{i,f,l}, \nu_{i,f,r})$ on the regions

of $(-\infty, 0]$, $(0, 1)$ and $[1, \infty)$, respectively.

- First, for the case of $(-\infty, 0]$, assume that ${c'_{i,f}}^* > 0$ when $\tilde{c}_{i,f}(\nu_i, 0, 0) \leq 0$. When ${c'_{i,f}}^* > 0$, $\nu^*_{i,f,l} = 0$ should be satisfied due to the complementary slackness condition in (E.10). Hence, $\tilde{c}_{i,f}(\nu_i, 0, \nu_{i,f,r}) > 0$. Moreover, because $\tilde{c}_{i,f}(\nu_i, 0, \nu^*_{i,f,r})$ is monotonically decreasing with respect to $\nu_{i,f,r}$ and $\nu_{i,f,r} \geq 0$, we obtain $\tilde{c}_{i,f}(\nu_i, 0, 0) \geq \tilde{c}_{i,f}(\nu_i, 0, \nu^*_{i,f,r}) > 0$. This contradicts the previous assumption. Hence, due to the primal feasibility ${c'_{i,f}}^* \geq 0$ in (E.6), when $\tilde{c}_{i,f}(\nu_i, 0, 0) \leq 0$, ${c'_{i,f}}^* = 0$ is the optimal solution;

- Next, for the case of $(0, 1)$, $\nu_{i,f,r} = \nu_{i,f,l} = 0,\ f \in \mathcal{F}$ due to the complementary slackness condition in (E.10). Hence, ${c'_{i,f}}^* = \tilde{c}_{i,f}(\nu_i, 0, 0)$;

- Finally, for the case of $[1, \infty)$, similar to the first case, it can be proved by contradiction that, when $\tilde{c}_{i,f}(\nu_i, 0, 0) \geq 1$, ${c'_{i,f}}^* = 1$.

By summarizing the above three cases, we can obtain that the optimal file caching probability for file $f \in \mathcal{F}$ of group $i \in \mathcal{G}$ is

$$c^*_{i,f} = [1 - \tilde{c}_{i,f}(\nu_i, 0, 0)]_0^1 = \left[1 - \left(\frac{1}{\bar{\mu}_i} W\left(A_{i,f} e^{\bar{\mu}_i B_{i,f}}\right) - B_{i,f}\right)\right]_0^1, \tag{E.20}$$

where

$$A_{i,f} = \frac{\nu_i}{\gamma_i e^{-\bar{\mu}_0} r_{i,f} e^{\sum_{j \in \mathcal{G} \setminus \{i\}} \bar{\mu}_j c'_{j,f}}} \tag{E.21}$$

and $B_{i,f}$ is defined in (E.18). With the above results, the optimal dual variable ν^*_i can be obtained by the constraint $\sum_{f \in \mathcal{F}} c^*_{i,f} = 1$.

Theorem 3.4.1 is thus proved.

Appendix F

Proof of Proposition 4.4.1

First, we obtain the optimal bandwidth allocation $\{b_k^\star\}$ and optimal power allocation $\{p_k^\star\}$. Given that the non-transmission power at BS i, $P_{c,i}$, is constant, it can be shown that the objective of (P7), i.e., the cost at BS i, is a monotonically increasing function of the sum transmission power $\sum_{k\in\mathcal{K}_i} p_k$ at BS i, no matter the power is purchased from the conventional grid or the renewable utility firm. Thus, deriving the optimal $\{b_k^\star\}$ and $\{p_k^\star\}$ to (P7) is equivalent to minimizing $\sum_{k\in\mathcal{K}_i} p_k$ at BS i subject to the bandwidth constraint in (4.10c) and the QoS constraints in (4.10d). Using this argument together with (4.11), we can obtain the optimal bandwidth allocation for (P7) by solving the following problem:

$$\begin{aligned} \underset{\{b_k\geq 0\}}{\texttt{min.}} \quad & \sum_{k\in\mathcal{K}_i} \frac{b_k N_0}{g_k}\left(2^{\frac{r_k}{b_k}}-1\right) \\ \texttt{s.t.} \quad & \sum_{k\in\mathcal{K}_i} b_k \leq W_i. \end{aligned} \tag{F.1}$$

Since Problem (F.1) is convex and satisfies the Slater's condition [80], the KKT conditions given as follows are necessary and sufficient for its optimal solution.

$$\frac{N_0}{g_k}\left(2^{\frac{r_k}{b_k}}-1\right) - \frac{N_0 r_k}{g_k b_k}\ln 2\cdot 2^{\frac{r_k}{b_k}} + \nu_i - \zeta_k = 0,\ \forall k\in\mathcal{K}_i, \tag{F.2}$$

$$\nu_i\left(\sum_{k\in\mathcal{K}_i} b_k - W_i\right) = 0,\ \nu_i \geq 0, \tag{F.3}$$

$$\zeta_k b_k = 0,\ \zeta_k \geq 0,\ b_k \geq 0,\ \forall k\in\mathcal{K}_i, \tag{F.4}$$

where $\nu_i \geq 0$ is the dual variable associated with the bandwidth constraint in (F.1)

and $\zeta_k \geq 0$ is the dual variable for $b_k \geq 0,\ k \in \mathcal{K}_i$. Note that the optimal bandwidth allocation $\{b_k^\star\}$ should satisfy that $b_k^\star > 0$, $\forall k \in \mathcal{K}_i$, otherwise the objective value in (F.1) will go to infinity. By using this together with (F.4), we thus have $\zeta_k^\star = 0, \forall k \in \mathcal{K}_i$. Accordingly, it follows from (F.2) that $b_k^\star$ can be obtained as in (4.15), where $\nu_i^\star > 0$ is determined by the equation $\sum_{k\in\mathcal{K}_i} b_k^\star = W_i$. Furthermore, by substituting the derived $\{b_k^\star\}$ into (4.11), the optimal $\{p_k^\star\}$ can be obtained.

Next, with $\{p_k^\star\}$ at hand, we proceed to obtain the optimal energy allocation $E_i^\star$ and $G_i^\star$. By using $\alpha_i^E < \alpha_i^G$ together with the fact that the optimal solution of (P7) is attained when the power constraint in (4.10a) is tight, it can be verified that the optimal solutions of $E_i^\star$ and $G_i^\star$ are obtained as in Proposition 4.4.1. Hence, the proof of Proposition 4.4.1 is complete.

Appendix G

Proof of Proposition 4.5.1

By substituting (4.12) into the power constraint (4.9b) in (P6), we can re-express (P6) as

$$\begin{aligned}(\text{P6.1}): \ \underset{\boldsymbol{x}\geq \mathbf{0}}{\texttt{min.}} \ & \sum_{i=1}^{2}\gamma_i(\alpha_i^E E_i+\alpha_i^G G_i)\\ \texttt{s.t.} \ & \sum_{k\in\mathcal{K}_i}\frac{b_k N_0}{g_k}\left(2^{\frac{r_k}{b_k}}-1\right)+P_{c,i}\leq E_i+G_i+\beta_E e_{\bar{\imath}}-e_i,\ i\in\{1,2\}, \quad \text{(G.1)}\\ & \text{(4.9c) and (4.9d)}.\end{aligned}$$

Denote the dual variables associated with the constraints in (G.1) and (4.9c) as $\mu_i\geq 0$ and $\lambda_i\geq 0,\ i\in\{1,2\}$, respectively. The partial Lagrangian of (P6.1) is then expressed as:

$$\begin{aligned}L(\boldsymbol{x},\{\mu_i\},\{\lambda_i\}) = & \sum_{i=1}^{2}E_i(\gamma_i\alpha_i^E-\mu_i)+\sum_{i=1}^{2}G_i(\gamma_i\alpha_i^G-\mu_i)+\sum_{i=1}^{2}\lambda_i\sum_{k\in\mathcal{K}_i}b_k-\sum_{i=1}^{2}\lambda_i W_i\\ & +\sum_{i=1}^{2}\mu_i\sum_{k\in\mathcal{K}_i}\frac{b_k N_0}{g_k}\left(2^{\frac{r_k}{b_k}}-1\right)+\sum_{i=1}^{2}\mu_i P_{c,i}+\sum_{i=1}^{2}w_i(\lambda_i-\beta_B\lambda_{\bar{\imath}})\\ & +\sum_{i=1}^{2}e_i(\mu_i-\beta_E\mu_{\bar{\imath}}). \qquad \text{(G.2)}\end{aligned}$$

Accordingly, the dual function can be obtained as

$$g(\{\mu_i\},\{\lambda_i\}) = \underset{\boldsymbol{x}\geq 0}{\texttt{min.}} \;\; L(\boldsymbol{x},\{\mu_i\},\{\lambda_i\}) \qquad \text{(G.3)}$$
$$\texttt{s.t.} \;\; (4.9\text{d}).$$

Thus, the dual problem is expressed as

$$(\text{P6.1}-\text{D}): \quad \underset{\{\mu_i\},\{\lambda_i\}}{\texttt{max.}} \;\; g(\{\mu_i\},\{\lambda_i\})$$
$$\texttt{s.t.} \;\; \lambda_i \geq 0, \mu_i \geq 0, \;\forall i \in \{1,2\}.$$

Since (P6.1) is convex and satisfies the Slater's condition, strong duality holds between (P6.1) and (P6.1 − D) [80]. Therefore, (P6.1) can be solved optimally by solving its dual Problem (P6.1 − D) as follows. We first solve the problem in (G.3) to obtain $g(\{\mu_i\},\{\lambda_i\})$ for given $\{\mu_i\}$ and $\{\lambda_i\}$, and then maximize $g(\{\mu_i\},\{\lambda_i\})$ over $\{\mu_i\}$ and $\{\lambda_i\}$.

We first give the following lemma.

Lemma G.0.1 *In order for $g(\{\mu_i\},\{\lambda_i\})$ to be bounded from below, it follows that*

$$\gamma_i\alpha_i^G \geq \mu_i, \;\; \beta_E\mu_{\bar{i}} \leq \mu_i, \;\; \beta_B\lambda_{\bar{i}} \leq \lambda_i, \;\; \forall i \in \{1,2\}. \qquad \text{(G.4)}$$

Proof First, suppose that $\gamma_i\alpha_i^G < \mu_i$ for any $i \in \{1,2\}$. In this case, it is easy to verify that the dual function $g(\{\mu_i\},\{\lambda_i\})$ will go to minus infinity as $G_i \to \infty$, i.e., $g(\{\mu_i\},\{\lambda_i\})$ is unbounded from below. Hence, $\gamma_i\alpha_i^G \geq \mu_i, \;\; i \in \{1,2\}$, should always hold.

Second, suppose that $\beta_E\mu_{\bar{i}} > \mu_i$ for any $i \in \{1,2\}$. In this case, it is easy to verify that the dual function $g(\{\mu_i\},\{\lambda_i\})$ will go to minus infinity as $e_i \to \infty$, i.e., $g(\{\mu_i\},\{\lambda_i\})$ is unbounded from below. Hence, $\beta_E\mu_{\bar{i}} \leq \mu_i, \;\; i \in \{1,2\}$, should always hold.

Last, suppose that $\beta_B\lambda_{\bar{i}} > \lambda_i$ for any $i \in \{1,2\}$. In this case, it is easy to

verify that the dual function $g(\{\mu_i\},\{\lambda_i\})$ will go to minus infinity as $w_i \to \infty$, i.e., $g(\{\mu_i\},\{\lambda_i\})$ is unbounded from below. Hence, $\beta_B\lambda_{\bar{i}} \le \lambda_i,\ i \in \{1,2\}$, should always hold.

By combining the above three arguments, Lemma G.0.1 is thus proved.

From Lemma G.0.1, it follows that the optimal solution of $(\mathrm{P6.1-D})$ is achieved when $\{\mu_i\}$ and $\{\lambda_i\}$ satisfies the inequalities in (G.4). As a result, we only need to solve Problem (G.3) with given $\{\mu_i\}$ and $\{\lambda_i\}$ satisfying (G.4). In this case, it can be observed that Problem (G.3) can be decomposed into the following subproblems:

$$\min_{b_k \ge 0} \quad \lambda_i b_k + \mu_i\left(\frac{b_k N_0}{g_k}\left(2^{\frac{r_k}{b_k}}-1\right)\right),\ k \in \mathcal{K}_1 \cup \mathcal{K}_2, \tag{G.5}$$

$$\min_{0 \le E_i \le \bar{E}_i} \quad E_i(\gamma_i \alpha_i^E - \mu_i),\ i \in \{1,2\}, \tag{G.6}$$

$$\min_{G_i \ge 0} \quad G_i(\gamma_i \alpha_i^G - \mu_i),\ i \in \{1,2\}, \tag{G.7}$$

$$\min_{e_i \ge 0} \quad e_i(\mu_i - \beta_E \mu_{\bar{i}}),\ i \in \{1,2\}, \tag{G.8}$$

$$\min_{w_i \ge 0} \quad w_i(\lambda_i - \beta_B \lambda_{\bar{i}}),\ i \in \{1,2\}. \tag{G.9}$$

For the $K_1 + K_2$ subproblems in (G.5), the optimal bandwidth allocation with given $\{\mu_i\}$ and $\{\lambda_i\}$ can be obtained based on the first order condition and is expressed as

$$b_k^{(\mu_i,\lambda_i)} = \begin{cases} \frac{r_k \ln 2}{W(\frac{1}{e}(\frac{\lambda_i g_k}{\mu_i N_0}-1))+1}, & \mu_i > 0 \\ 0, & \mu_i = 0 \end{cases},\ k \in \mathcal{K}_i,\ i \in \{1,2\}. \tag{G.10}$$

Furthermore, the optimal solution to the subproblems in (G.6)–(G.9) can be

obtained as follows.

$$E_i^{(\mu_i)} = \begin{cases} 0, & \gamma_i \alpha_i^E \geq \mu_i \\ \bar{E}_i, & \gamma_i \alpha_i^E < \mu_i \end{cases}, \ i \in \{1,2\}, \tag{G.11}$$

$$G_i^{(\mu_i)} = 0, \ i \in \{1,2\}, \tag{G.12}$$

$$e_i^{(\mu_i)} = 0, \ i \in \{1,2\}, \tag{G.13}$$

$$w_i^{(\lambda_i)} = 0, \ i \in \{1,2\}. \tag{G.14}$$

Note that for the subproblems in (G.6) with any $i \in \{1,2\}$, if $\gamma_i \alpha_i^E = \mu_i$, then the solution of E_i is non-unique and can be any value within its domain. For convenience, we choose $E_i^{(\mu_i)} = 0$ in this case. The similar case holds for subproblems in (G.7), (G.8) and (G.9) if $\beta_E \mu_{\bar{i}} = \mu_i$, $\gamma_i \alpha_i^G = \mu_i$ and $\beta_B \lambda_{\bar{i}} = \lambda_i$, respectively. In these cases, $G_i^{(\mu_i)} = 0$, $e_i^{(\mu_i)} = 0$ and $w_i^{(\mu_i)} = 0, i \in \{1,2\}$, are chosen as the solution for simplicity. Also note that the solutions in (G.10) – (G.14) are only for obtaining the dual function $g(\{\mu_i\},\{\lambda_i\})$ under any given μ_i and λ_i to solve the dual Problem $(\text{P6.1}-\text{D})$, while they may not be the optimal solution to the original (primal) Problem (P8) duo to their non-uniqueness.

With the results in (G.10) – (G.14), we have obtained the dual function $g(\{\mu_i\},\{\lambda_i\})$ with given $\{\mu_i\}$ and $\{\lambda_i\}$ satisfying (G.4). Next, we maximize $g(\{\mu_i\},\{\lambda_i\})$ over $\{\mu_i\}$ and $\{\lambda_i\}$ to solve $(\text{P6.1}-\text{D})$. Since $g(\{\mu_i\},\{\lambda_i\})$ is convex but in general not differentiable, subgradient based algorithms such as the ellipsoid method [103] can be applied to solve $(\text{P6.1}-\text{D})$, where the subgradients of $g(\{\mu_i\},\{\lambda_i\})$ for μ_i and λ_i are given by $-\sum_{k\in\mathcal{K}_i} \frac{b_k^{(\mu_i,\lambda_i)} N_0}{g_k} \left(2^{\frac{r_k}{b_k^{(\mu_i,\lambda_i)}}} - 1\right) - P_{c,i} + E_i^{(\mu_i)}$ and $-\sum_{k\in\mathcal{K}_1\cup\mathcal{K}_2} b_k^{(\mu_i,\lambda_i)} + W_i$, $i \in \{1,2\}$, respectively. As a result, we can obtain the optimal dual solution as $\{\mu_i^\star\}$ and $\{\lambda_i^\star\}$. Accordingly, the corresponding $\{b_k^{(\mu_i^\star,\lambda_i^\star)}\}$ becomes the optimal bandwidth allocation solution for (P6.1) and thus (P8), given by $\{b_k^\star\}$. Substituting the obtained $\{b_k^\star\}$ into (4.11), the optimal power allocation solution for (P8) is thus obtained as $\{p_k^\star\}$.

However, it is worth noting that the other optimal optimization variables

for (P8), given by $\{E_i^\star\}, \{G_i^\star\}, \{e_i^\star\}$ and $\{w_i^\star\}$, cannot be directly obtained from (G.11)−(G.14), since the solutions in (G.11)−(G.14) are in general non-unique. Nevertheless, it can be shown that the optimal solution of (P8) is always attained when the inequality constraints in (4.9b) and (4.9c) are tight. Thus, we have

$$\sum_{k\in\mathcal{K}_i} p_k^\star + P_{c,i} = E_i^\star + G_i^\star + \beta_E e_{\bar{i}}^\star - e_i^\star,\ i \in \{1,2\}, \tag{G.15}$$

$$\sum_{k\in\mathcal{K}_i} b_k^\star = W_i + \beta_B w_{\bar{i}}^\star - w_i^\star,\ i \in \{1,2\}. \tag{G.16}$$

From (G.16) and using the fact that $w_1^\star$ and $w_2^\star$ should not be positive at the same time, the optimal solution of $\{w_i^\star\}$ can be obtained. Last, the optimal optimization variables $\{E_i^\star\}, \{G_i^\star\}$ and $\{e_i^\star\}$ can be obtained by solving the LP in (P8). Therefore, Proposition 4.5.1 is proved.

Appendix H

Proof of Lemma 4.6.1

We prove Lemma 4.6.1 by considering two cases. First, consider $\boldsymbol{x}^{\text{ex}}$ with $\sum_{k\in\mathcal{K}_i} p_k^{(\boldsymbol{x}^{\text{ex}})} + P_{c,i} - \beta_E e_{\bar{\imath}} + e_i \neq \bar{E}_i$. In this case, it is evident from (4.17) and (4.18) that the optimal solutions $\mu_i^{(\boldsymbol{x}^{\text{ex}})}$ and $\lambda_i^{(\boldsymbol{x}^{\text{ex}})}$ are both unique, since the bandwidth water-level ν_i is always unique. According to Theorem 1 in [104], the left-partial derivative is equal to the right-partial derivative. Hence, $C_i(\boldsymbol{x}^{\text{ex}})$ is differentiable with the partial derivatives given in Lemma 4.6.1. Therefore, (4.19) follows directly based on the first order approximation of $\bar{C}_i(\boldsymbol{x}^{\text{ex}} + \Delta\boldsymbol{x}^{\text{ex}})$.

Next, consider $\boldsymbol{x}^{\text{ex}}$ with $\sum_{k\in\mathcal{K}_i} p_k^{(\boldsymbol{x}^{\text{ex}})} + P_{c,i} - \beta_E e_{\bar{\imath}} + e_i = \bar{E}_i$. In this case, the optimal dual solution of $\mu_i^{\boldsymbol{x}^{\text{ex}}}$ in (4.17) is not unique. More specifically, it can be shown that μ_i^{ex} can be any real number between α_i^E to α_i^G. As a result, $C_i(\boldsymbol{x}^{\text{ex}})$ is not differentiable in such point. However, it follows from [104] that the left- and right-hand derivatives of $C_i(\boldsymbol{x}^{\text{ex}})$ with respect to $e_1,\ e_2,\ w_1$ and w_2 still exist, which can be given as

$$\begin{aligned}
\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial e_i^+} &= \alpha_i^G, \quad \frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial e_i^-} = \alpha_i^E \\
\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial e_{\bar{\imath}}^+} &= -\beta_E \alpha_i^E, \quad \frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial e_{\bar{\imath}}^-} = -\beta_E \alpha_i^G \\
\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial w_i^+} &= \alpha_i^G \nu_i^{(\boldsymbol{x}^{\text{ex}})}, \quad \frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial w_i^-} = \alpha_i^E \nu_i^{(\boldsymbol{x}^{\text{ex}})} \\
\frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial w_{\bar{\imath}}^+} &= -\alpha_i^E \nu_i^{(\boldsymbol{x}^{\text{ex}})}, \quad \frac{\partial \bar{C}_i(\boldsymbol{x}^{\text{ex}})}{\partial w_{\bar{\imath}}^-} = -\alpha_i^G \nu_i^{(\boldsymbol{x}^{\text{ex}})}.
\end{aligned}$$

By replacing the partial derivatives in (4.21) as the corresponding left- or right-hand derivatives, (4.19) also follows from the first order approximation of $\bar{C}_i(\boldsymbol{x}^{\text{ex}} + \Delta\boldsymbol{x}^{\text{ex}})$.

Appendix H. Proof of Lemma 4.6.1

It is worth noting that, in Lemma 4.6.1, we do not introduce the left- and right-hand derivatives for notational convenience.

Therefore, Lemma 4.6.1 is proved.

Appendix I

Proof of Proposition 4.6.1

Since the shared energy from BS 1 to BS 2 and that from BS 2 to BS 1 cannot be zero at the same time, i.e., $e_1 \cdot e_2 = 0$, there exist three possible cases for the shared energy between the two BSs, which are (a) $e_1 = e_2 = 0$, (b) $e_1 > 0, e_2 = 0$, and (c) $e_1 = 0, e_2 > 0$. Similarly, there are three possible cases for the shared bandwidth between the two BSs, i.e., (a) $w_1 = w_2 = 0$, (b) $w_1 > 0, w_2 = 0$, and (c) $w_1 = 0, w_2 > 0$. As a result, by combining the above energy and spectrum cooperation, there are nine cases for the shared energy and the shared bandwidth. Therefore, we prove this proposition by enumerating the nine possible cases. In the following, we consider the case of $e_1 = e_2 = w_1 = w_2 = 0$ and show that in this case, $\boldsymbol{x}^{\text{ex}}$ attains the Pareto optimality if and only if $\frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}} \leq \frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}\beta_E}$ and $\frac{\lambda_2^{(\boldsymbol{x}^{\text{ex}})}}{\mu_2^{(\boldsymbol{x}^{\text{ex}})}} \leq \frac{\lambda_1^{(\boldsymbol{x}^{\text{ex}})}}{\mu_1^{(\boldsymbol{x}^{\text{ex}})}\beta_E}$. We prove the "only if" and "if" parts, respectively.

First, we show the necessary part by contradiction. Suppose that there exists an inter-system energy and bandwidth cooperation vector $\bar{\boldsymbol{x}}^{\text{ex}}$ with $\bar{e}_1 = \bar{e}_2 = \bar{w}_1 = \bar{w}_2 = 0$ attains the Pareto optimality, where $\frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}} > \frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$ or $\frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}} > \frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$. If $\frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}} > \frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$, then we can construct a new inter-system energy and bandwidth cooperation vector as

$$\tilde{\boldsymbol{x}}^{\text{ex}} = \bar{\boldsymbol{x}}^{\text{ex}} + \Delta\boldsymbol{x}^{\text{ex}}, \tag{I.1}$$

with $\Delta\boldsymbol{x}^{\text{ex}} = (\Delta e_1, \Delta e_2, \Delta w_1, \Delta w_2)^T$, where $\Delta e_2 = \Delta w_1 = 0$, while $\Delta e_1 > 0$ and

$\Delta w_2 > 0$ are sufficiently small and satisfy

$$\frac{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}} < \frac{\Delta w_2}{\Delta e_1} < \frac{\beta_E \mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}. \tag{I.2}$$

In this case, it can be shown from (4.19) that

$$\begin{bmatrix} \Delta C_1 \\ \Delta C_2 \end{bmatrix} = \begin{bmatrix} \bar{C}_1(\tilde{\boldsymbol{x}}^{\text{ex}}) - \bar{C}_1(\bar{\boldsymbol{x}}^{\text{ex}}) \\ \bar{C}_2(\tilde{\boldsymbol{x}}^{\text{ex}}) - \bar{C}_2(\bar{\boldsymbol{x}}^{\text{ex}}) \end{bmatrix} = \begin{bmatrix} \mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})} \Delta e_1 - \lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})} \Delta w_2 \\ -\beta_E \mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})} \Delta e_1 + \lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})} \Delta w_2 \end{bmatrix} < \mathbf{0}, \tag{I.3}$$

where the inequality is component-wise. In other words, we have found a new inter-system energy and bandwidth cooperation vector to achieve lower energy costs for both BSs. As a result, $\bar{\boldsymbol{x}}^{\text{ex}}$ with $\bar{e}_1 = \bar{e}_2 = \bar{w}_1 = \bar{w}_2 = 0$ does not achieve the Pareto optimality. On the other hand, if $\frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}} > \frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$, then we can also construct a new inter-system energy and bandwidth cooperation vector as in (I.1), where $\Delta e_1 = \Delta w_2 = 0$, while $\Delta e_2 > 0$ and $\Delta w_1 > 0$ are sufficiently small and satisfy

$$\frac{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}} < \frac{\Delta w_1}{\Delta e_2} < \frac{\beta_E \mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}. \tag{I.4}$$

Then, it can be shown that under this choice, the energy costs of the two BSs can be decreased at the same time. By combining the results for the two cases of $\frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}} > \frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$ and $\frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}} > \frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$, a contradiction is induced. As a result, the presumption cannot be true. Accordingly, the necessary part is proved.

Second, for the sufficient part, we can also show it by contradiction. Suppose that for an inter-system energy and bandwidth cooperation vector $\bar{\boldsymbol{x}}^{\text{ex}}$ with $\bar{e}_1 = \bar{e}_2 = \bar{w}_1 = \bar{w}_2 = 0$ satisfying $\frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}} \leq \frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$ and $\frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}} \leq \frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$, but does not achieve the Pareto optimality. This case implies that the two BSs can exchange energy and bandwidth to decrease energy cost of both at the same time. In other words, there must exist a new vector $\tilde{\boldsymbol{x}}^{\text{ex}}$ given in (I.1) satisfying either $\Delta e_2 = \Delta w_1 = 0$, $\Delta e_1 > 0, \Delta w_2 > 0$ or $\Delta e_1 = \Delta w_2 = 0$, $\Delta e_2 > 0, \Delta w_1 > 0$ such that $\bar{C}_1(\tilde{\boldsymbol{x}}^{\text{ex}}) < \bar{C}_1(\bar{\boldsymbol{x}}^{\text{ex}})$ and $\bar{C}_2(\tilde{\boldsymbol{x}}^{\text{ex}}) < \bar{C}_2(\bar{\boldsymbol{x}}^{\text{ex}})$. If $\Delta e_2 = \Delta w_1 = 0$, $\Delta e_1 > 0, \Delta w_2 > 0$, then it can be shown from

(I.3) that $\frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}} < \frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$ must hold, whereas if $\Delta e_1 = \Delta w_2 = 0$, $\Delta e_2 > 0, \Delta w_1 > 0$, then we have $\frac{\lambda_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_2^{(\bar{\boldsymbol{x}}^{\text{ex}})}} < \frac{\lambda_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}}{\mu_1^{(\bar{\boldsymbol{x}}^{\text{ex}})}\beta_E}$. As a result, we have a contradiction here and thus the presumption cannot be true. Therefore, the sufficient part is proved.

By combing the two parts, we have verified the proposition in the case of $e_1 = e_2 = w_1 = w_2 = 0$.

Next, the other eight cases remain to be proved in order to complete the proof of this proposition. Since the proof for these cases can follow the same contradiction procedure as the case of $e_1 = e_2 = w_1 = w_2 = 0$, we omit the details here for brevity. By combining the proof for the nine cases, this proposition is verified.

References

[1] Cisco. (2015). *Cisco visual networking index: global mobile data traffic forecast update 2014–2019 white paper* [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf.

[2] W. Zhuang and M. Ismail, "Cooperation in wireless communication networks," *IEEE Wireless Commun.*, vol. 19, no. 2, pp. 10–20, Apr. 2012.

[3] J. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Info. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[4] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[5] 3GPP, *Further advancements for E-UTRA physical layer aspects (release 9)* [Online]. Available: http://www.qtc.jp/3GPP/Specs/36814-900.pdf

[6] 3GPP, *Coordinated multi-point operation for LTE physical layer aspects (release 11)* [Online]. Available: http://www.qtc.jp/3GPP/Specs/36819-b10.pdf.

[7] H. Nick. *Essentials of short-range wireless.* Cambridge, UK: Cambridge University Press, 2010.

[8] Bluetooth SIG Inc. (2014). *Bluetooth low energy* [Online]. Available: https://developer.bluetooth.org/TechnologyOverview/Pages/BLE.aspx.

[9] Alliance, ZigBee. (2009). *IEEE 802.15.4, ZigBee standard* [Online]. Available: http://www.zigbee.org.

[10] L. Militano, G. Araniti, M. Condoluci, I. Farris, and A. Iera, "Device-to-Device communications for 5G Internet of things," *EAI Trans. Internet of Things*, vol. 15, no. 1, pp. 1–15, Oct. 2015.

[11] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-Device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[12] X. Lin, J. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP Device-to-Device proximity services," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 40–48, Apr. 2014.

[13] J. G. Andrews, S. Buzzi , W. Choi , S. V. Hanly , A. Lozano , A. C. K. Soong, and J. Zhang, "What will 5G be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[14] Universal Mobile Telecommunications System (UMTS). *Proximity-based services (ProSe), technical specifications (release 13)* [Online]. Available: http://www.3gpp.org/DynaReport/23303.htm.

[15] Qualcomm. (2013). *LTE Direct* [Online]. Available: https://www.qualcomm.com/invention/research/projects/lte-direct.

[16] Deutsche Telekom, Huawei, and Qualcomm. (2015). *LTE-Direct trial white paper* [Online]. Available: https://www.qualcomm.com/media/documents/files/lte-direct-trial-white-paper.pdf.

[17] Informa Telecoms & Media, *Mobile broadband access at home: the business case for Femto-cells, UMA and IMS-VCC dual mode solutions*. UK, London: Informa Telecoms & Media, 2008.

[18] V. Chandrasekhar, J. G. Andrews and A. Gatherer, "Femto-cell networks: a survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.

[19] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[20] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[21] D. Willkomm, S. Machiraju, J. Bolot and A. Wolisz, "Primary user behavior in cellular networks and implications for dynamic spectrum access," *IEEE Commun. Mag.*, vol. 47, no. 3, pp. 88–95, Mar. 2009.

[22] X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. Leung, "Cache in the air: exploiting content caching and delivery techniques for 5G systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 131–139, Feb. 2014.

[23] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: the role of proactive caching in 5G wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 82–89, Aug. 2014.

[24] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[25] Z. Hasan, H. Boostanimehr, and V. Bhargava, "Green cellular networks: a survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 524–540, 4th Quarter 2011.

[26] M. Hibberd. *Vodafones green strategies* [Online]. Available: http://telecoms.com/interview/interview-vodafones-green-strategies/.

[27] International Energy Agency. *IEA – Energy subsidies* [Online]. Available: http://www.worldenergyoutlook.org/resources/energysubsidies/.

[28] International Renewable Energy Agency. *Renewable power generation cost in 2014* [Online]. Availeble: http://www.irena.org/documentdownloads/publications/irena_re_power_costs_2014_report.pdf.

[29] J. Xu, L. Duan, and R. Zhang, "Cost-aware green cellular networks with energy and communication cooperation," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 257-263, May 2015.

[30] Huawei. *Mobile networks go green* [Online]. Available: http://www.huawei.com/en/about-huawei/publications/communicate/hw-082734.htm.

[31] R. Zhang, "Keynote talk: wireless communications in the era of energism," in *Proc. IEEE Int. Conf. Comm. (ICC), Wksp on Green Commun. and Netw. with Energy Harvesting, Smart Grids, and Renewable Energies*, 2015.

[32] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid – the new and improved power grid: a survey," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 944–980, 4th Quarter 2012.

[33] S. M. Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century," *IEEE Power Energy Mag.*, vol. 3, no. 5, pp. 34–41, Sep. 2005.

[34] S. Chen, N. B. Shroff and P. Sinha, "Energy trading in the smart grid: from end-user's perspective," in *Proc. Asilomar Conf. on Signals, Sys. and Comput.*, Pacific Grove, CA, 2013, pp. 327–331.

[35] CNN. (2005). *Battery life concerns mobile users* [Online]. Available: http://edition.cnn.com/2005/TECH/ptech/09/22/phone.study/.

[36] G. P. Perrucci, F. H. P. Fitzek, and J. Widmer, "Survey on energy consumption entities on the smartphone platform," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, May 2011, pp. 1–6.

[37] A. Pathak, Y. C. Hu, and M. Zhang, "Where is the energy spent inside my APP? fine-grained energy accounting on smartphones with Eprof," in *Proc. ACM Euro. Conf. on Computer Sys. (EuroSys)*, 2012, pp. 29–42.

[38] Y.-K. Chia, S. Sun, and R. Zhang, "Energy cooperation in cellular networks with renewable powered base stations," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*, Apr. 2013, pp. 2542–2547.

[39] J. Xu, Y. Guo, and R. Zhang, "CoMP meets energy harvesting: a new communication and energy cooperation paradigm," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 2508–2513.

[40] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

[41] L. Fu, H. Kim, J. Huang, S.-C. Liew, and M. Chiang, "Energy conservation and interference mitigation: from decoupling property to win-win strategy," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3943–3955, Nov. 2011.

[42] H. Kim and G. de Veciana, "Leveraging dynamic spare capacity in wireless systems to conserve mobile terminals' energy," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 802–815, Jun. 2010.

[43] S. Luo, R. Zhang, and T. J. Lim, "Joint transmitter and receiver energy minimization in multiuser OFDM systems," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3504–3516, Oct. 2014.

[44] S. Luo, R. Zhang, and T. J. Lim, "Downlink and uplink energy minimization through user association and beamforming in C-RAN," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 494–508, Jan. 2015.

[45] G. Kramer, I. Marić, and R. Yates, *Cooperative communications*, Foundations and Trends in Networking. Boston, MA: Now Publishers, 2007.

[46] Z. Zhou, S. Zhou, J.-H. Cui, and S. Cui, "Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3066–3078, Aug. 2008.

[47] Y. Zou, J. Zhu, and R. Zhang, "Exploiting network cooperation in green wireless communication," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 999–1010, Mar. 2013.

[48] D. Liu, W. Wang, and W. Guo, "Green cooperative spectrum sharing communication," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 459–462, Mar. 2013.

[49] G. Botter, J. Alonso-Zarate, L. Alonso, F. Granelli, and C. Verikoukis, "Extending the lifetime of M2M wireless networks through cooperation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 6003–6007.

[50] D. Yang, X. Fang, and G. Xue, "Game theory in cooperative communications," *IEEE Wireless Commun.*, vol. 19, no. 2, pp. 44–49, Apr. 2012.

[51] J.-P. Hubaux, T. Gross, J.-Y. Le Boudec, and M. Vetterli, "Toward self-organized mobile ad hoc networks: the terminodes project," *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 118–124, Jan. 2001.

[52] B. Wang, Z. Han, and K. Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using Stackelberg game," *IEEE Trans. Mobile Comput.*, vol. 8, no. 7, pp. 975–990, Jul. 2009.

[53] Y. Yan, J. Huang, and J. Wang, "Dynamic bargaining for relay-based cooperative spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1480–1493, Aug. 2013.

[54] H.-T. Lin, Y.-Y. Lin, and W.-C. Chang, "Reputation auction framework for cooperative communications in green wireless networks," in *Proc. IEEE Int. Sym. Personal Indoor and Mobile Radio Commun. (PIMRC)*, Sep. 2012, pp. 875–880.

[55] S. Kandeepan, S. Jayaweera, and R. Fedrizzi, "Power-trading in wireless communications: a cooperative networking business model," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1872–1880, May 2012.

[56] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, Sep. 2009.

[57] J. Kingman, *Poisson process*. Oxford, UK: Oxford University Press, 1993.

[58] G. Arunabha, J. Zhang, J. G. Andrews, and R. Muhamed. *Fundamentals of LTE*. Pearson Education, 2010.

[59] T. Nakamura. (2009). *Proposal for candidate radio interface technologies for IMT-Advanced based on LTE release 10 and beyond* [Online]. Available: http://www.3gpp.org/IMG/pdf/2009_10_3gpp_IMT.pdf.

[60] A. Mas-Colell, M. Whinston, and J. Green, *Microeconomic theory*. Oxford, UK: Oxford University Press, 1995.

[61] P. Fishburn, *Utility theory for decision making*. Publications in Operations Research. New York: Wiley, 1970.

[62] A. Antoniou and W.-S. Lu, *Practical optimization: algorithms and engineering applications*. Berlin, Germany: Springer, 2007.

[63] S. Yi, S. Chun, Y. Lee, S. Park, and S. Jung, *Radio protocols for LTE and LTE-advanced*. New York: John Wiley & Sons, 2012.

[64] L. W. Dowdy and D. V. Foster, "Comparative models of the file assignment problem," *ACM Comput. Survey*, vol. 14, no. 2, pp. 287–313, Jun. 1982.

[65] M. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Info. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[66] K. Shanmugam, N. Golrezaei, A. Dimakis, A. Molisch, and G. Caire, "FemtoCaching: wireless content delivery through distributed caching helpers," *IEEE Trans. Info. Theory*, vol. 59, no. 12, pp. 8402–8413, Dec. 2013.

[67] D. Bethanabhotla, G. Caire, and M. Neely, "Adaptive video streaming for wireless networks with multiple users and helpers," *IEEE Trans. Commun.*, vol. 63, no. 1, pp. 268–285, Jan. 2015.

[68] H. Zhou, M. Tao, E. Chen and W. Yu, "Content-centric multicast beamforming in cache-enabled cloud radio access networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[69] N. Golrezaei, A. Dimakis, and A. Molisch, "Scaling behavior for Device-to-Device communications with distributed caching," *IEEE Trans. Info. Theory*, vol. 60, no. 7, pp. 4286–4298, Jul. 2014.

[70] M. Ji, G. Caire, and A. F. Molisch, "The throughput-outage tradeoff of wireless one-hop caching networks," *IEEE Trans. Info. Theory*, vol. 61, no. 12, pp. 6833–6859, Dec. 2015.

[71] M. Ji, G. Caire, and A. F. Molisch, "Fundamental limits of caching in wireless D2D networks," *IEEE Trans. Info. Theory*, vol. 62, no. 2, pp. 849–869, Feb. 2016.

[72] A. Bouza, A. Bernstein, "Partial user preference similarity as classification-based model similarity," *Semantic Web*, Jan. 2014, pp. 47-64.

[73] S. Gitzenis, G. Paschos, and L. Tassiulas, "Asymptotic laws for joint content replication and delivery in wireless networks," *IEEE Trans. Info. Theory*, vol. 59, no. 5, pp. 2760–2776, May 2013.

[74] B. Han, P. Hui, V. Kumar, M. Marathe, J. Shao, and A. Srinivasan, "Mobile data offloading through opportunistic communications and social participation," *IEEE Trans. Mobile Comput.*, vol. 11, no. 5, pp. 821–834, May 2012.

[75] X. Chen, B. Proulx, X. Gong, and J. Zhang, "Exploiting social ties for cooperative D2D communications: a mobile social networking case," *IEEE/ACM Trans. Netw.*, vol. 23, no. 5, pp. 1471–1484, Oct. 2015.

[76] M. Zhang, X. Chen, and J. Zhang, "Social-aware relay selection for cooperative networking: an optimal stopping approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 2257–2262.

[77] Y. Li, T. Wu, P. Hui, D. Jin, and S. Chen, "Social-aware D2D communications: qualitative insights and quantitative analysis," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 150–158, Jun. 2014.

[78] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Aug. 2009.

[79] L. H. Wong, P. Pattison, and G. Robins, "A spatial model for social networks," *Physica A: Statistical Mechanics and its Applications*, vol. 360, no. 1, pp. 99 – 120, 2006.

[80] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge, UK: Cambridge University Press, 2004.

[81] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert-W function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

[82] A. Goldsmith, *Wireless communications*. Cambridge, UK: Cambridge University Press, 2005.

[83] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and Zipf-like distributions: evidence and implications," in *Proc. IEEE Conf. Comput. and Commun. Societies (INFOCOM)*, vol. 1, Mar. 1999, pp. 126–134.

[84] A. W. Marshall, I. Olkin, and B. Arnold, *Inequalities: theory of majorization and its applications*, Berlin, Germany: Springer Science & Business Media, 2010.

[85] B. Gurakan, O. Ozel, J. Yang, and S. Ulukus, "Energy cooperation in energy harvesting communications," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4884–4898, Dec. 2013.

[86] J. Leithon, T. J. Lim and S. Sun, "Energy exchange among base stations in a cellular network through the smart grid," *IEEE Int'l Conf. Commun. (ICC)*, Sydney, Australia, Jun. 2014, pp. 4036-4041.

[87] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad-hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 203–213, Jan. 2008.

[88] L. Duan, L. Gao, and J. Huang, "Cooperative spectrum sharing: a contract-based approach," *IEEE Trans. Mobile Compt.*, vol. 13, no. 1, pp. 174–187, Jan. 2014.

[89] G. Lin, X. Wang, Y. Xu, and Q. Zhang, "Spectrum trading in cognitive radio networks: a contract-theoretic modelling approach," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 843–855, Apr. 2011.

[90] T. Han and N. Ansari, "Auction-based energy-spectrum trading in green cognitive cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 6205–6209.

[91] E. Yaacoub, A. Imran, Z. Dawy, and A. Abu-Dayya, "A game theoretic framework for energy efficient deployment and operation of heterogeneous LTE networks," in *Proc. IEEE Int. Wksp on Comput. Aided Modelling and Design of Commun. Links and Netw. (CAMAD)*, Sep. 2013, pp. 33–37.

[92] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 74–79, Nov. 2010.

[93] E. Hardy. *Sprint and T-mobile may merge* [Online]. Available: http://www.brighthand.com/default.asp?newsID=15636&news=Sprint+T-Mobile.

[94] L. Gkatzikis, I. Koutsopoulos, and T. Salonidis, "The role of aggregators in smart grid demand response markets," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1247–1257, Jul. 2013.

[95] W. Saad, Z. Han, and H. Poor, "Coalitional game theory for cooperative micro-grid distribution networks," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Jun. 2011, pp. 1–5.

[96] Qualcomm. *LTE advanced carrier aggregation* [Online]. Available: http://www.qualcomm.com/solutions/wireless-networks/technologies/carrier-aggregation.

[97] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[98] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge, UK: Cambridge University Press, 2005.

[99] M. Grant and S. Boyd. (2011). *CVX: matlab software for disciplined convex programming* [Online]. Available: http://cvxr.com/cvx/.

[100] T. Zhang, W. Chen, Z. Han, and Z. Cao, "A cross-layer perspective on energ harvesting aided green communications over fading channels," in *Proc. IEEE Conf. Comp. Commun. Wkshops (INFOCOM WKSHPS)*, Apr. 2013, pp. 19–24.

[101] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4808–4818, Sep. 2012.

[102] F. Baccelli and B. Blaszczyszyn, *Stochastic geometry and wireless networks: Vol. I Theory*. Delft, Netherlands: Now Publishers Inc., 2009.

[103] S. Boyd. *Convex optimization II, Stanford University* [Online]. Available: http://www.stanford.edu/class/ee364b/lectures.html.

[104] R. Horst, "On the interpretation of optimal dual solutions in convex programming," *J. Oper. Res. Soc.*, vol. 35, no. 4, pp. 327–335, 1984.

List of Publications

Journal Publications

1. **Y. Guo**, J. Xu, L. Duan, and R. Zhang, "Joint energy and spectrum cooperation for cellular communication systems," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3678–3691, Oct. 2014.

2. **Y. Guo**, L. Duan, and R. Zhang, "Optimal pricing and load sharing for energy saving in communications cooperation", *IEEE Trans. Wireless Commun.*, vol. 15, no.2, pp. 951–964, Feb. 2016.

3. **Y. Guo**, L. Duan, and R. Zhang, "Cooperative local caching under heterogeneous file preferences", to appear in *IEEE Trans. Commun.*

Conference Publications

1. J. Xu, **Y. Guo**, and R. Zhang, "CoMP meets energy harvesting: A new communication and energy cooperation paradigm," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Atlanta, GA, USA, Dec. 2013, pp. 2508–2513.

2. **Y. Guo**, J. Xu, L. Duan, and R. Zhang, "Optimal energy and spectrum sharing for cooperative cellular systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, Australia, Jun. 2014, pp. 4771–4776.

3. **Y. Guo**, J. Xu, L. Duan, and R. Zhang, "Cooperative local caching and file sharing under heterogeneous file preferences," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumper, Malaysia, May 2016, pp. 1–6.