Analysis and Design of Communication Policies for Energy-Constrained Machine-Type Devices

Sheeraz Akhtar Alvi

June, 2020

A thesis submitted for the degree of Doctor of Philosophy
of The Australian National University

Research School of Electrical, Energy, and Materials Engineering
College of Engineering and Computer Science
The Australian National University

Dedication

To my father ...

Declaration

The contents of this thesis are the results of original research and have not been submitted for a higher degree to any other university or institution.

Much of the work in this thesis has been published or has been submitted for publication as journal papers or conference proceedings.

The research work presented in this thesis has been performed jointly with Dr. Xiangyun Zhou (The Australian National University), Dr. Salman Durrani (The Australian National University) and Dr. Duy Trong Ngo (The University of Newcastle). The substantial majority of this work was my own.

Sheeraz Akhtar Alvi
Research School of Electrical, Energy, and Materials Engineering,
College of Engineering and Computer Science,
The Australian National University,
Canberra, ACT 2601,
AUSTRALIA

Acknowledgements

The work presented in this thesis would not have been possible without the support of a number of individuals and organizations and they are gratefully acknowledged below:

- I would like to express my sincere thanks to my supervisors Drs Xiangyun Zhou, Salman Durrani and Nan Yang, for their guidance, support and encouragement throughout my PhD studies. They have motivated and supported me throughout this research work. Special thanks for always entertaining my meeting requests and showing patience and providing me instrumental assessments. Despite all the assistance provided by my supervisors and others, I am responsible alone for any shortcomings or errors which may unwittingly remain.

- I would like to thank Dr. Duy Trong Ngo from The University of Newcastle for his support and prompt responses to my queries. He helped me tackle a critical mathematical problem for one of my research work included in this thesis.

- I would also like to thank Prof. Carlo Fischione from KTH Royal Institute of Technology for kindly welcoming me to visit their research groups during spring-summer 2019. The research in his research group encouraged me to explore emerging research area of machine learning for wireless networks, which I am exploring in my post PhD research.

- It is my great pleasure to study in the communication research group at the Research School of Electrical, Energy, and Materials Engineering. I would also like to thank everyone in our research group for providing me a friendly

and relaxing research environment. Special thanks to the school's administration staff for their support and assistance.

- Thanks must go to the Australian National University for providing me with the PhD scholarship, stipend, personal development funds, and the 2018 Dean's Travel Grant Award and the 2018 Vice Chancellors Travel Grant for supporting my international research visit and conference attendances.

- Finally, I would like to express my utmost gratitude thank my parents and my siblings for their love and support throughout my life.

Abstract

This thesis focuses on the modelling, analysis and design of novel communication strategies for wireless machine-type communication (MTC) systems to realize the notion of Internet of things (IoT). We consider sensor based machine-type devices (MTDs) which acquire physical information from the environment and transmit it to a base station (BS) while satisfying application specific quality-of-service (QoS) requirements. Due to the wireless and unattended operation, these MTDs are mostly battery-operated and are severely energy-constrained. In addition, MTC systems require low-latency, perpetual operation, massive-access, etc.

Motivated by these critical requirements, this thesis proposes optimal data communication policies for four different network scenarios. In the first two scenarios, each MTD transmits data on a dedicated orthogonal channel and either (i) possess an initially fully charged battery of finite capacity, or (ii) possess the ability to harvest energy and store it in a battery of finite capacity. In the other two scenarios, all MTDs share a single channel and are either (iii) allocated individual non-overlapping transmission times, or (iv) randomly transmit data on predefined time slots. The proposed novel techniques and insights gained from this thesis aim to better utilize the limited energy resources of MTDs in order to effectively serve the future wireless networks.

Firstly, we consider a sensor based MTD communicates with a BS, and devise optimal data compression and transmission policies with an objective to prolong the device-lifetime. We formulate joint optimization problems aiming to maximize the device-lifetime whilst satisfying the delay and bit-error-rate constraints. Our results show significant improvement (that ranges from 90% to 2000%) in device-lifetime. Importantly, the gain is most profound in the low latency regime.

Secondly, we consider a sensor based MTD that is served by a hybrid BS which

wirelessly transfers power to the device and receives data transmission from the device. The MTD employs data compression in order to reduce the energy cost of data transmission. Thus, we propose to jointly optimize the harvesting-time, compression and transmission design, to minimize the energy cost of the system under given delay constraint. The proposed scheme reduces energy consumption up to 19% when data compression is employed.

Thirdly, we consider multiple MTDs transmit data to a BS following the time division multiple access (TDMA). Conventionally, the energy-efficiency performance in TDMA is optimized through multi-user scheduling, i.e., changing the transmission time allocated to different devices. In such a system, the sequence of devices for transmission, i.e., who transmits first and who transmits second, etc., does not have any impact on the energy-efficiency. We consider that data compression is performed before transmission. We jointly optimize both multi-user sequencing and scheduling along with the compression and transmission rate. Our results show that multi-user sequence optimization achieves up to 45% improvement in the energy efficiency at MTDs.

Lastly, we consider contention resolution diversity slotted ALOHA (CRDSA) with transmit power diversity where each packet copy from a device is transmitted at a randomly selected power level. It results in inter-slot received power diversity, which is exploited by employing a signal-to-interference-plus-noise ratio (SINR) based successive interference cancellation (SIC) receiver. We propose a message passing algorithm to model the SIC decoding and formulate an optimization problem to determine the optimal transmit power distribution subject to energy constraints. We show that the proposed strategy provides up to 88% system load performance improvement for massive-MTC systems.

List of Publications

The work in this thesis has been published or has been submitted for publication as journal papers or conference proceedings. These papers are:

Journal articles:

J1. **S. A. Alvi**, X. Zhou, S. Durrani, and D. T. Ngo, "Sequencing and Scheduling for Multi-User Machine-Type Communication," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2459-2473, Apr. 2020.

J2. **S. A. Alvi**, X. Zhou, and S. Durrani, "Optimal Compression and Transmission Rate Control for Node-Lifetime Maximization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7774-7788, Nov. 2018.

J3. **S. A. Alvi**, S. Durrani, and X. Zhou, "Enhancing CRDSA with Transmit Power Diversity for Machine-Type Communication," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7790-7794, Aug. 2018.

Conference papers:

C1. **S. A. Alvi**, X. Zhou, S. Durrani, and D. T. Ngo, "Proportionally-Fair Sequencing and Scheduling for Machine-Type Communication," in *Proc. IEEE ICC*, Dublin, Ireland, Jun. 2020.

C2. **S. A. Alvi**, X. Zhou, and S. Durrani, "Wireless Powered Machine-Type Communication: Energy Minimization via Compressed Transmission," in *Proc. IEEE PIMRC*, Istanbul, Turkey, Sep. 2019.

C3. **S. A. Alvi**, X. Zhou, and S. Durrani, "A Lifetime Maximization Scheme for a Sensor Based MTC Device ," in *Proc. IEEE GLOBECOM*, Abu Dhabi, UAE, Dec. 2018.

List of Acronyms

HTC	Human-Type Communication
MTC	Machine-Type Communication
3GPP	Third Generation Partnership Project
ETSI	European Telecommunications Standards Institute
UMTS	Universal Mobile Telecommunications System
LTE	Long Term Evolution
V2X	Vehicle-to-X
IoT	Internet of Things
MTD	Machine-Type Device
BS	Base Station
QoS	Quality-of-Service
TDMA	Time Division Multiple Access
CRDSA	Contention Resolution Diversity Slotted ALOHA
SINR	Signal-to-Interference-plus-Noise Ratio
SIC	Successive Interference Cancellation
WSN	Wireless Sensor Network
IP	Internet Protocol
RF	Radio-Frequency

RF-EH	Radio Frequency based Energy Harvesting
HAP	Harvesting Access Point
WPT	Wireless Power Transfer
WIT	Wireless Information Transfer
uMTC	Ultra-Reliable Machine Type Communication
mMTC	Massive Machine Type Communication
DSA	Diversity Slotted ALOHA
WMSN	Wireless Multimedia Sensor Networks
IoMT	Internet of Multimedia Things
IRSA	Irregular Repetition Slotted ALOHA
IC	Interference Cancellation
BER	Bit Error Rate
RDC	Radio Duty Cycling
MCU	Micro-Controller Unit
PAR	Peak-to-Average Ratio
CGI	Channel Gain Information
SNR	Signal-to-Noise Ratio
VPA	Variable-Precision floating-point Arithmetic
EH	Energy Harvesting
SEN	Sensing
CP	Compression
TX	Transmission
CCDF	Complimentary Cumulative Distribution Function
AWGN	Additive White Gaussian Noise

GP	Geometric Programming
GPS	Global Positioning System
KKT	Karush-Kuhn-Tucker

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Radio communication has greatly influenced the way humans remotely communicate with each other, specifically using the wide-spread mobile networks. The first and second generation of wireless cellular networks were designed for human-to-human voice over radio communication using the hand-held devices [1]. However, the advancement of hardware capabilities of the mobile phones and the enlarged capacity of wireless networks allowed data communication over the Internet. The ever-increasing services and applications offered by the Internet have explosively widened the span of the global inter-network. Currently, more than 4 billion people around the globe are connected to the Internet [2] for communication (emails, social networks, chatrooms, blogs, forums, etc.), leisure and entertainment (games, books, music, videos, shopping, etc.), sharing knowledge (education, geographical information, encyclopedias, etc.), among others services. Consequently, the human-type communication (HTC) is now extremely diverse and dominated by non-audio data communication.

In parallel to the evolution of HTC, the advancements in low-cost and tiny network devices manufacturing has enabled machine-type communications (MTC), wherein the devices are programmed to autonomously exchange certain type of data with one another and/or to a centralized entity over a communication medium [3, 4]. For ease of deployment, especially when the number of devices is large,

the devices are equipped with radios to achieve wireless communication [3, 5, 6]. The MTC has a great potential in realizing an immense range of applications and services. The potential use-cases and applications of MTC systems are widespread across various industries including logistics, process automation, healthcare, manufacturing, energy, utilities, etc.

The integration of sensing capability into machine-type devices (MTDs) gave birth to wireless sensor network (WSN) wherein the multiple tiny devices autonomously acquire some physical information from the environment and transmit it over a wireless channel to a central data fusion station [7]. The MTC systems were further revolutionised when the MTDs were equipped with Internet protocol (IP) communication stack allowing them to communicate over the Internet. MTDs equipped with the capability to observe and/or interact with physical environment and the ability to communicate with other things, are extending the Internet towards the so called Internet of things (IoT) [8–10].

IoT has the potential to significantly influence our lives and the way we interact with the devices such as sensors, actuators, mobile phones, home automation devices, smart grid devices, etc. [11, 12]. It has promoted concepts of flexible designs, visions and enormous applications, some of them are depicted in [13–15]. Currently, more than 10 billions IoT devices are connected to the Internet through the cellular network connections [16]. For individual users, IoT brings useful applications in home automation, security, automated devices monitoring and management of daily tasks. For professionals, automated applications provide useful contextual information frequently to help on their works and decision making. For industrialists, Internet enabled sensors and actuators operations can be rapid, efficient and more economical. Managers who need to keep eye on many things can automate tasks by connecting digital and physical objects together.

The HTC can be characterized by the high rate downlink data communication in active periods. However, MTC applications have diverse traffic patterns mostly composed of small and infrequent packet transmissions from a large number of devices in the uplink. The MTC devices are also different from HTC devices as the former are low-cost devices that possess limited computational and/or power resources. Thus, MTC raise different technical challenges and demand application specific service requirements in terms of delay, per-link and total bit rate, reliabil-

ity, energy consumption, and security/privacy [17, 18]. Various organizations have launched the standardization activities on MTC supported by cellular mobile networks, particularly the third generation partnership project (3GPP) [17] and the european telecommunications standards institute (ETSI) [19]. The 3GPP standard is more focused on the MTC support by the wireless cellular networks, i.e., UMTS (universal mobile telecommunications system) and LTE (long term evolution) core networks, and identification of the use-cases and the associated research challenges. On the other hand, the ETSI addresses the MTC service architecture that comprises of three main parts: devices domain, network domain and application domain.

The MTDs can employ different existing wireless technologies to communicate among each other depending upon the application and environmental context. For short range or non-proprietary systems the wireless technologies such as IEEE 802.11x (WiFi) [20], IEEE 802.15.4 (ZigBee) [21], IEEE 802.15.1 (Bluetooth) [22], etc., can effectively provide point-to-point and multi-hop wireless communication in ad hoc, sensor, and mesh networks based vehicular, local, personal, body area networks. These technologies operate in the unlicensed frequency band usually at 2.4 GHz and 5 GHz. Therein, multiple MTDs share a single or multiple physical channels to transmit sensed data randomly or periodically to a specific receiver device which is usually is the same for all MTDs. However, these solutions are used for scenarios where reliability and interference is not crucial.

For commercialization and standardized operation of proprietary IoT services and applications, the cellular mobile networks are proposed for MTC [3, 4]. According to the 3GPP standard specifications, the MTDs can directly communicate to a base station (e.g., eNodeBs) or its data traffic can be relayed through gateways. For direct communication the MTDs need to have a cellular radio interface. Otherwise, the MTDs can form a capillary network and the gateway device can exchange data traffic between the capillary and cellular networks [23]. The data transmission is performed using the radio resources which are divided into multiple uplink and downlink channels. The cellular networks provide high level of security, reliability, mobility, and enhanced coverage for wide-area and massive MTC systems (that are composed of over 10 times more MTDs than the current number of cellular devices [3]).

1.2 Challenges and Requirements for MTC

The sensor based MTDs within an IoT system, are supposed to acquire information from the environment and transmit it to the intended destination while satisfying stringent technical quality of service (QoS) requirements in terms of maximal energy efficiency, ultra-low latency, and application specific rigorous data reliability. This calls for novel solutions to realize wireless connectivity across heterogeneous and autonomous wireless MTDs such as sensors, actuators, etc., [3, 5].

The use-cases of MTC systems have highly diverse properties and QoS requirements which depend on the context of the application and environmental dynamics. Different use-cases of the MTC systems can be classified based on the following major challenges and requirements:

- *Deployment diversity*: Zero to high mobility, fixed or arbitrary device locations, sparse to highly dense network.
- *Traffic heterogeneity*: Periodic, event-driven, real-time, continuous, on demand, critical uplink and/or downlink communication, heterogeneous data packet size, delay tolerant/intolerant, diverse reliability conditions.
- *Energy constraint*: High energy efficiency, low consumption rate, small and limited battery capacity, efficient energy-harvesting mechanisms for perpetual operation.
- *Operational efficiency*: Low protocol overhead, duty cycling pattern, deep sleep mode scheduling, timely resource allocation for delay intolerant transmission, different level coordination among MTDs.

The MTC systems can be divided into two classes: ultra-reliable MTC (uMTC) and massive MTC (mMTC) [24, 25]. The uMTC systems are designed for the industrial control use-cases, vehicle-to-X (V2X) communications, self-car driving, argument reality, remote surgery, etc. The network service requirements of uMTC systems are stringent latency and reliability. In contrast, the mMTC systems provide wireless access to a large number of devices. Mostly, the MTDs in a mMTC system are low-complexity and low-cost wireless devices covering a wide

area with/without indoor communication. Typical mMTC systems have their use-cases in the smart cities (parking waste management, street lighting), building (smoke detectors, alarms, automation), environment (flood, water, noise, air), and utilities (smart metering, grid management), etc.

One the most crucial requirements and challenges listed above for MTC systems is the energy constraint. Due to their wireless and unattended operation, MTDs are mostly battery operated and thus severely energy-constrained. Since the number of MTDs in the network is typically very large which makes the replacement of batteries extremely hard. The energy-efficient operation of these devices is therefore of pivotal importance and is a critical requirement [26]. Specifically, wireless communication is one of the most energy-intensive operations run by the MTDs and this calls for effective wireless solutions to prolong the operational lifespan of these constrained devices [26]. Also, the data transmission in MTC systems, unlike HTC systems, typically takes place in the uplink, i.e., from multiple MTDs towards the core network.

In this thesis, we consider various network scenarios of the MTC systems which depict different MTC use-cases and thus differ in terms of the major requirements and challenges. In particular, our overall objective is to minimize the energy cost of the MTDs. The considered network scenarios constitute a fairly generic MTC system setting. In this regard, the design policies proposed and the insights and observations given in this thesis provide novel wireless solutions for a wide range of MTC use-cases and applications.

1.3 Network Scenarios and Design Challenges

In the following, we consider four different network scenarios and identify the associated research challenges. A simple representation of these scenarios is illustrated in the Fig. 1.1. As mentioned before, our objective is to consider generic network settings which cover various use-cases and applications scenarios with requirements and challenges identified in previous sub-section.

In the first two scenarios, each MTD transmits data on a dedicated orthogonal channel and either (i) possess an initially fully charged battery of finite capacity, or (ii) possess the ability to harvest energy and store it in a battery of finite capacity.

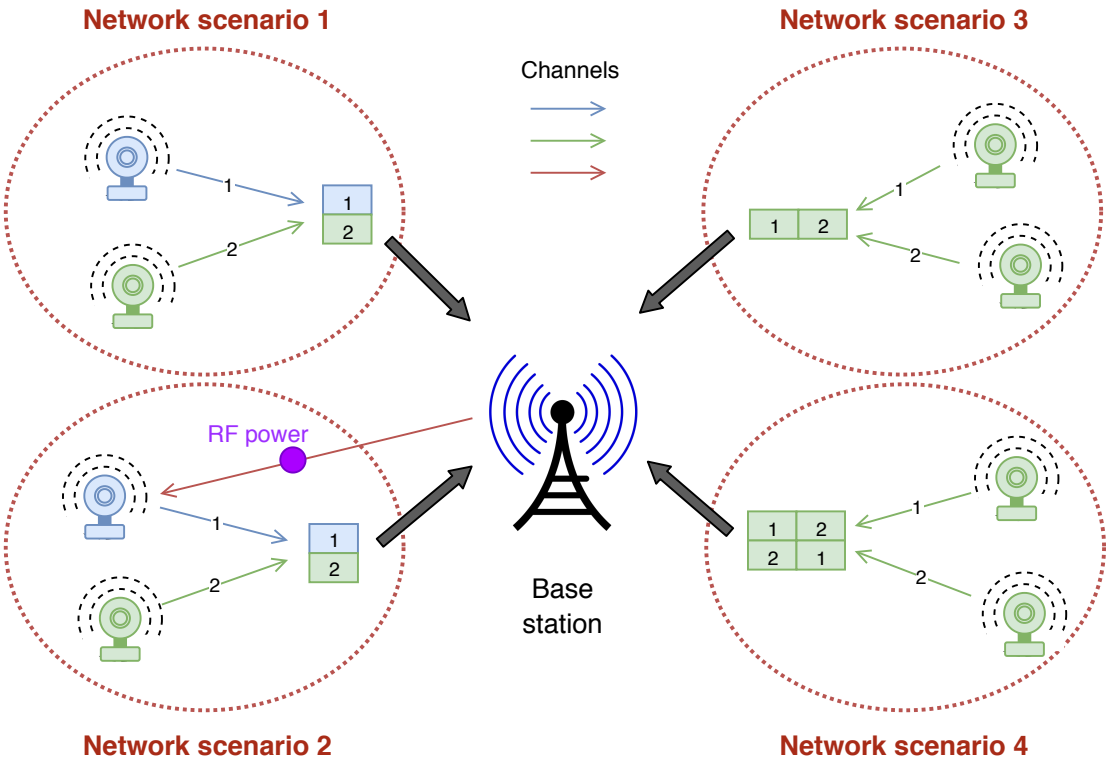


Figure 1.1: Considered network scenarios: (i) dedicated channels for each user, (ii) users have radio-frequency (RF) energy harvesting capability, (iii) single channel shared sequentially by users, and (iv) single channel shared randomly by users.

In the other two scenarios, all MTDs share a single channel and are either (iii) allocated individual non-overlapping transmission times, or (iv) randomly transmit data on predefined time slots.

The first network scenario is applicable to a multi-user multi-channel sparse MTC system. In such a system, each MTD is assigned a channel and transmits data on the dedicated channel. The data may be generated or sensed by the MTD periodically, real-time, on demand (i.e., only when requested by the receiver), or due to the occurrence of an event of interest, that is to be delivered subject to the given reliability and delay conditions. Moreover, the MTDs are battery powered and perform duty cycling to switch off radio and may also go to deep sleep by occasionally powering off the micro controller unit (MCU). This network scenario covers high reliability and low-latency MTC systems such as low-power wireless multimedia sensor networks (WMSNs) for audio/video surveillance [27], WSNs-

based monitory of heterogeneous data-traffic [7], etc. For this network scenario, designing a generic single-user communication policy would suffice and the same policy can be employed by all other users for the individual system parameters.

The second network scenario is evolved from the first scenario and considers an important aspect of energy-harvesting for a multi-user multi-channel sparse MTC system. Each MTD is battery powered, which is recharged by an energy-harvesting source for perpetual operation. This network scenario covers loss tolerant and high reliability demanding MTC systems such as low-power environmental monitoring WSNs [7], telemetry, personal and body area network MTC applications [28], etc. As before, designing a generic single-user communication policy would suffice for this network scenario.

The third network scenario is applicable to a multi-user single-channel sparse or massive MTC system. In such a system, all MTDs share a single channel and sequentially transmit data on this channel for interference free transmission. The data may be generated or sensed by the MTD periodically or due to the occurrence of an event of interest that is to be delivered under given reliability and delay conditions. Moreover, the MTDs are battery powered and perform duty cycling and deep sleep operation. This network scenario covers high reliability and low-latency MTC systems such as automotive networks FlexRay [29], mmWave based application of 5G networks [30], low-power WSNs and WMSNs, etc. For this network scenario, the communication policies for all MTDs need to be jointly devised for given system parameters.

The fourth network scenario is evolved from the third scenario. However, due to a lack of coordination among devices, the MTDs randomly transmit short data-packets on a shared channel. This network scenario covers highly reliability and dense MTC systems such as massive WSNs [31], smart cities lighting system, smart parking systems, etc. For this network scenario, the communication policies for all MTDs need to be jointly devised for given system parameters.

1.3.1 Single-User With Limited Battery Capacity

The lifetime of MTC device is defined as the time taken by the device to deplete all of its energy stored in the battery of limited capacity. For battery powered MTDs,

which are severely energy constrained, prolonging the lifetime is of paramount importance [26]. In the existing literature, the lifetime maximization problem has been approached from different perspectives such as green channel access, sleep-wake scheduling, coverage, efficient routing, network coding, data aggregation, see [26] and the references therein.

We consider a monitoring system, in which an energy constrained sensor MTD acquires some physical information from its vicinity, applies data compression on the sensed data, and transmits the compressed data to the sink node. The compression is employed to alleviate the transmission energy cost. However, applying too much compression leads to substantial energy cost. Therefore, the compression and transmission needs to be jointly optimized. Prior studies only considered the optimization of transmission policy without compression. We consider that the sink node may be able to feedback perfect or imperfect channel information to the sensor MTD. The objective is to maximize MTD's lifetime subject to the given QoS requirements in terms of data reliability.

Design Challenge: Jointly optimize the compression and transmission policy to maximize the MTD's lifetime subject to given QoS requirements when channel knowledge may or may not be available at the MTD.

1.3.2 Single-User With Energy Harvesting Capability

The limited operational time of battery operated MTDs is a major hurdle in realizing networks demanding long-term operation [3]. Radio frequency based energy harvesting (RF-EH) has recently emerged as a promising solution to provide perpetual-lifetime for MTC nodes [32–34], i.e., it may stay operational indefinitely long. Therein, the MTD is dependent on a harvesting access point (HAP) which transfers the energy to the MTD through a powerful RF signal. The MTD receives this signal and harvests the energy and stores it in a battery. This energy can later be used for data sensing and transmission.

We consider a unified wireless power transfer (WPT) and wireless information transfer (WIT) system. The objective is to minimize the energy cost of the system. The MTD solely relies on the harvested energy transferred by the HAP. Thus, the aforementioned objective can be achieved by minimizing the HAP's transmitted

energy. We seek to intelligently design the operation of the MTD and the HAP.

Design Challenge: Jointly optimize the RF energy harvesting, compression and transmission policies to minimize the power transferred by the HAP subject to given QoS requirements.

1.3.3 Multi-User Sequential Channel Access

The wireless MTDs within an IoT system usually share a single uplink channel. Therein, each device first contends for channel resources and then transmits data to a central station, while satisfying stringent QoS requirements in terms of application specific reliability and delay. In addition, different levels of fairness among the devices are considered while minimizing the energy cost of all MTDs (i.e., system energy cost).

The time division multiple access (TDMA) channel access mechanism allows deterministic scheduling for data transmission and other device operations. In previous works, TDMA is preferred for sensor networks [35], [36], wireless powered communication networks [37], and many other network technologies. In this chapter, the TDMA protocol is employed for the uplink MTC and we jointly determine the schedule for data transmission and compression operations for MTDs.

The performance of the TDMA protocol is enhanced by exploiting the multi-user diversity which occurs due to the difference in the signal power attenuation conditions of different devices. This performance enhancement is achieved by optimizing the multi-user scheduling, i.e., by changing the transmission time allocated to different devices within a frame, whilst maximizing a certain system objective such as energy efficiency, throughput. For example, a device allocated with a relatively longer transmission time can adapt the transmission rate for given channel gain, to achieve better energy efficiency and vice versa. Accordingly, in a TDMA-based system, the time allocated to different MTDs is adapted to achieve a given system objective and the devices transmit data successively in a fixed order. In such a system, the sequence of devices for transmission, i.e., who transmits first and who transmits second, etc., has not been considered as it does not have any impact on the energy efficiency. We consider that data compression is performed before transmission and show that the multi-user sequencing is indeed important.

Design Challenge: Jointly optimize the multi-user sequencing and scheduling along with compression and transmission policies to minimize the energy cost of MTDs subject to given QoS requirements.

1.3.4 Multi-User Random Channel Access

A vital requirement in enabling mMTC systems in next generation wireless networks is efficient access mechanisms [38]. In this regard, random access schemes are preferred over demand based schemes due to their low latency and lesser signalling to enable communication between massive uncoordinated mMTC devices [38]. One such random access mechanism is contention resolution diversity slotted ALOHA (CRDSA) [39] which is a combination of diversity slotted ALOHA (DSA) and successive interference cancellation (SIC). In CRDSA mechanism, a device transmits multiple copies (bursts) of the packet on randomly selected slots. On the receiver side, first the interference free bursts are iteratively recovered and then their copies are removed from replica slots using SIC to allow recovery of new packets. CRDSA is based on the clean packet model in which only interference-free packets are recoverable.

Recently, efforts have been made to improve the performance of CRDSA by enhancing and exploiting the capture effect [40], i.e., a packet is recoverable if its signal-to-interference-plus-noise ratio (SINR) is above a predefined threshold. The randomness in the channel gain can provide gain to the capture effect. This can be further enhanced by employing the transmit power diversity. We consider that multiple MTDs employed DSA scheme to transmit data packets. At the transmitter side, multiple bursts are transmitted on randomly selected slots. The number of the burst repetition is selected from a probability distribution. Moreover, each of these bursts are transmitted on randomly selected transmit power levels selected from a given probability distribution. At the receiver side, the bursts are iteratively resolved using SINR metric and SIC approach. The objective is to recover as many unique data packets as possible.

Design Challenge: Optimize the transmit power level distribution for the MTDs to maximize the system load subject to degree of repetition and energy constraint.

1.4 Literature Review

In the following subsections, we discuss the prior works related to our considered network scenarios and identified design challenges.

1.4.1 Single-User With Limited Battery Capacity

The low-cost and miniature sized cameras and microphones have made it possible to acquire multimedia information, i.e., image, audio, and video, from the environment enabling the notion of WMSN [41] and Internet of multimedia things (IoMT) [27]. In most of the applications of WMSN and IoMT, the amount of sensed data (raw data) can be very large, resulting in high transmission cost. In this regard, data compression schemes have been proposed [42–45], which decrease the amount of data to be transmitted and thus alleviate the transmission energy cost. Typically, the energy cost of compression and transmission is around 15% and 80% of the total energy consumed by the sensor node, respectively [46, 47]. The transmission cost depends upon the required transmission rate and the signal strength. Unlike the transmission energy cost which linearly increases with the size of data to be transmitted, the compression energy cost has a non-linear relationship with the compression ratio [48]. Owing to this non-linearity, blindly applying too much compression may even exceed the cost of transmitting raw data, thereby losing its purpose [49].

From the compressive sensing perspective, [50] considered a wireless powered cognitive radio network and optimized the time slot allocation for energy harvesting, sensing, and transmission processes. By using compressive sensing, less number of samples are collected which reduces the sensing cost as well as the data transmission cost. The reader is referred to [51] for an overview of related works on compression and transmission frame design using compressive sensing in wireless sensor networks.

Prior works on power-rate adaptation [52–56], which do not employ data compression, have considered transmission energy as a monotonically increasing function of the transmission rate. Therefore, these schemes propose to transmit data at lower transmission rates under given delay constraint to achieve energy efficiency.

These schemes assume the distance between communicating devices is large, thus the transmit power dominates the circuit power [57,58]. However, in many practical sensor networks, e.g., body area networks, the distance is fairly small and the circuit power cost cannot be ignored. In this regard, considering the power amplifier cost as a function of the constellation size and transmit power, the transmission energy is not anymore a monotonically increasing function of the transmission rate, as shown in [57]. This is the case particularly for smaller constellation sizes, which are more common in sensor networks. Therefore, simply decreasing the transmission rate may not necessarily improve energy efficiency.

To the best of our knowledge, none of the prior works has jointly considered and optimized data compression and transmission rate.

1.4.2 Single-User With Energy Harvesting Capability

The prior studies have considered the unified WPT and WIT systems from the perspective of RF-EH efficiency [34, 59–61]. The RF-EH is mainly dependent on aspects such as the transmitted power, wireless medium, rectification efficiency, etc. Firstly, the HAP's transmitted power undergoes attenuation due to the fading and pathloss, afterwards the RF-EH circuit's sensitivity threshold (-10 to -30 dBm [34]) further cuts down the received power, and lastly the RF energy transducer harvests energy with a limited efficiency ($< 50\%$ [62]). Consequently, even for a short distance, the HAP spends a substantial amount of energy on WPT process. Therefore, the energy cost of the HAP is an important component to consider for efficient operation of a RF-EH based communication system.

To achieve energy minimization at the HAP, it is critical to minimize the WIT energy cost/requirement to perform its operation subject to the system constraints. The existing literature proposes the power-rate adaptation [59–61, 63–65] to minimize the transmission energy cost under the given delay constraint. This approach is only valid when the distance is large and thus the transmit power dominates the circuit power [60,61]. For practical RF-EH based system, the distance is very short and thus the circuit power cost cannot be ignored. Thus, decreasing the transmission rate may prolong the transmission time but may not necessarily decrease the transmission energy cost.

In chapter 2, we propose using data compression to minimize the transmission energy cost by reducing the amount of data to transmit. It is shown that jointly optimizing transmission and compression in a traditional non-EH communication system minimized energy consumption significantly (typically over 90%) as compared to optimizing transmission only without compression. Motivated by this potential of data compression to reduce the energy cost, we consider this technique for a wireless powered communication system. In such systems, the time spent on compression (which itself consumes energy [42,45,49]) must be carefully controlled.

To the best of our knowledge, none of the prior works has considered the impact of data compression while devising the WPT and WIT policies for a wireless powered MTC system.

1.4.3 Multi-User Sequential Channel Access

The TDMA protocol is employed for the uplink MTC in various prior works and its performance is only optimized by controlling the time allocated to different MTC devices, i.e., the length of the allocated transmission blocks. For a wireless power transfer scenario, the multi-user scheduling is optimized for sum and max-min energy minimization objectives for TDMA and NOMA in [37] and [66], respectively. Similarly, the proportional-fairness objective is considered in [67] and [68] to balance between multi-user fairness and energy minimization for NOMA and TDMA, respectively. An optimal strategy is devised for TDMA in [35] to balance between sum throughput and energy through multi-user scheduling. In that, the data transmission activity is controlled based on the available energy at the individual devices. For a TDMA system, in [69] the multi-user scheduling is optimized and the sum throughput is maximized for energy harvesting devices subject to a total time constraint. The system energy efficiency is maximized in [70] for TDMA systems by jointly optimizing the multi-user scheduling and transmit power subject to individual QoS requirements. To the best of our knowledge, none of the above papers investigates multi-user sequencing.

In chapter 2, we consider a system model that is composed of a single user only where the compression and transmission rates for this individual user are jointly optimized to maximize its lifetime. For a multi-user TDMA scenario where

all the users share a single channel and transmit one after the other. In such a system, the compression time allowed for each user is different (unlike the system model considered in [71]). As a result, the multi-user sequencing and scheduling are tightly coupled with the design of compression and transmission for each user. The policy design in [71], which did not account for such an important property in a multiuser setting, is not applicable to the TDMA-based system model considered in this chapter.

To the best of our knowledge, none of the prior works has considered the impact of sequence of devices for transmission, i.e., who transmits first and who transmits second, etc., as it has no significance when compression is not employed, given the channel statistics do not change from one transmission block to the other.

1.4.4 Multi-User Random Channel Access

Motivated by the critical need to better support MTC in future wireless networks [3], the SIC based DSA schemes such as CRDSA [39], irregular repetition slotted ALOHA (IRSA) [72], and other variants of CRDSA [40, 73–75], have significantly improved the throughput of slotted ALOHA. The normalized throughput, defined as the number of useful packets received per slot, is around 0.37 for basic DSA, which is improved to 0.52 and 0.965 by CRDSA and IRSA, respectively [39] [72]. It is shown in [72] that iterative burst recovery approach is analogous to bipartite graph theory, which can be used to represent iterative decoding of error correcting codes on graphs and analyze message passing based decoding. This approach is used in [72] to analyze the iterative interference cancellation (IC) convergence performance of CRDSA scheme. Therein, the bipartite graph can be optimized, using the tools from error correcting codes like the AND-OR tree analysis [76], to find the optimal degree distribution of burst repetition which gives the maximum achievable load.

Most DSA schemes employ clean packet model in which only interference free packets are recoverable and the collisions are considered destructive. However, the receiver can also exploit the power unbalance between the colliding packets and possibly recover one or more packets, if the SINR is above a predefined threshold. This technique is called the capture effect. Previously, the SINR models were

applied to single packet transmission models their physical interference is minimized by controlling the access probabilities of the participating nodes [77–79], among others. A SINR metric for DSA considering same received power for each device at the receiver is proposed in [80]. Similarly, the approaches employing multiple transmit power levels and the associated probability distribution are also focused on single packet model for pure DSA only [76, 81].

While the idea of transmit power diversity is common in wireless communications [82], to the best of our knowledge, no analytical framework has been proposed in the literature to date to analyse the impact of transmit power diversity on CRDSA.

1.5 Thesis Overview and Contributions

This thesis proposes optimal data communication policies for four different network scenarios. Figure 1.2 shows an overview of the thesis.

In this regard, the device-lifetime maximization problem is considered in Chapter 2. The system energy minimization problem for a wireless power transfer communication system is considered in Chapter 3. Chapter 4 considers the multi-user sequencing and scheduling optimization problem. Chapter 5 considers channel access mechanism that is suited for massive-channel access networks.

The chapter-wise summary of the contributions by this thesis is given as follows:

Chapter 2 −− Single-User Compression and Transmission Rate Control for Device-Lifetime Maximization

This chapter presents solution to the design challenge posed in Section 1.3.1.

Existing adaptive transmission rate control schemes attempt to prolong node-lifetime without considering data compression. In this chapter, our focus is to devise joint data compression-transmission policies to optimally utilize the energy resources to maximize the lifetime of an energy constrained sensor based MTC device. We consider a monitoring system, in which an energy constrained sensor node acquires some physical information from its vicinity, applies data compression on the sensed data, and transmits the compressed data to the sink node. The sink node may be able to feedback perfect or imperfect channel information to the sensor node, depending upon the considered scenario. Accordingly, the sensor

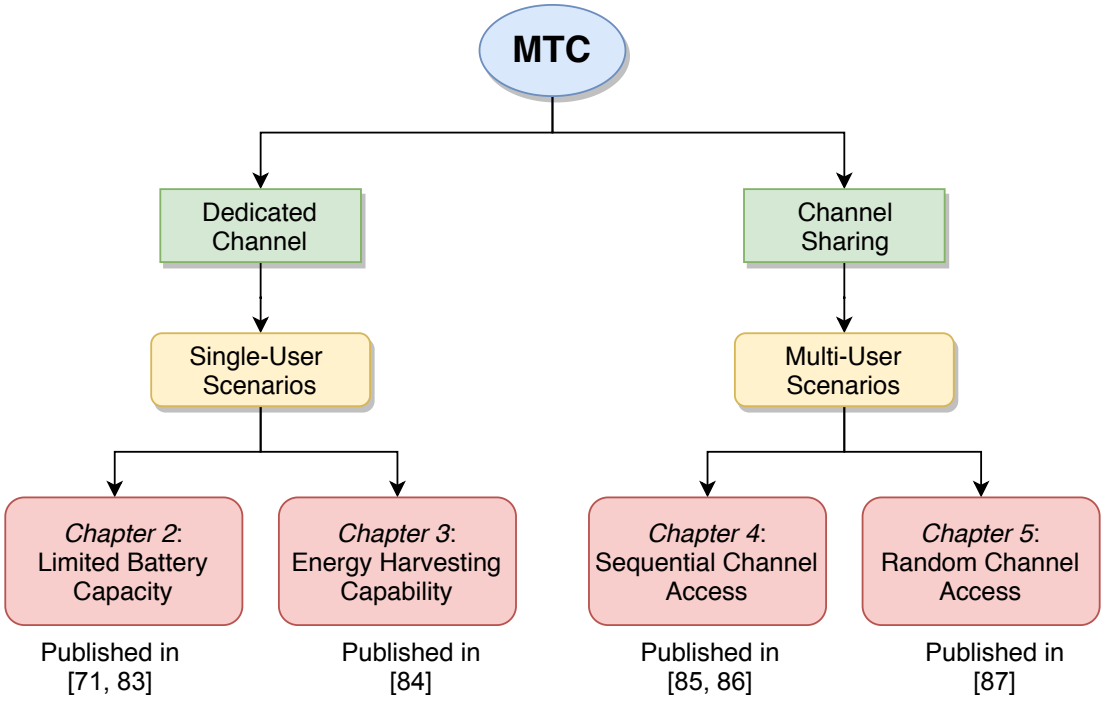


Figure 1.2: Overview of the thesis.

node devises an optimal compression and transmission policy based on the available channel information with an objective to maximize its lifetime. To the best of our knowledge, this is the first work in which the data compression and transmission are jointly designed and optimized. In contrast, prior studies only considered the optimization of transmission policy without compression. As shown in this chapter, the joint optimization of compression and transmission results in a substantial improvement in the lifetime of the sensor node. To this end, we consider three scenarios which differ in terms of the available channel information at the sensor node. The first scenario assumes the availability of the perfect instantaneous channel gain information at the sensor node and provides the benchmark theoretical performance. The remaining two scenarios assume the availability of the quantized channel gain and the statistical channel gain information, respectively, at the sensor node and provide the performance of practical sensor based wireless MTC systems.

Our investigation leads to the following observations and design guidelines:

- Our results show that a jointly optimized compression-transmission policy performs much better than optimizing transmission only without compression under any given BER and delay constraints. The performance gain observed ranges from 90% to 2000% and is most profound when the delay constraint is stringent.

- The optimal level of compression is insensitive to the change in the BER requirement. However, the optimal transmission rate increases as BER constraint gets less stringent.

- The best strategy is to reduce compression and increase the transmission rate when the delay constraint gets more stringent and vice versa. The optimal level of compression and transmission rate are more sensitive to the delay constraint when the system requires low latency, while they remain roughly unchanged when the delay constraint is relaxed beyond a certain point.

The results of this chapter have been presented in the following publications [71, 83]:

J2. **S. A. Alvi**, X. Zhou, and S. Durrani, "Optimal Compression and Transmission Rate Control for Node-Lifetime Maximization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7774-7788, Nov. 2018.

C3. **S. A. Alvi**, X. Zhou, and S. Durrani, "A Lifetime Maximization Scheme for a Sensor Based MTC Device ," in *Proc. IEEE GLOBECOM*, Abu Dhabi, UAE, Dec. 2018.

Chapter 3 −− Single-User Compression and Transmission Rate Control for Wireless Powered Machine-Type Communication

This chapter presents solution to the design challenge posed in Section 1.3.2.

In this chapter, the problem that we study is to minimize the energy cost of the system. The target is to optimally utilize the power injected into the system. The MTC node solely relies on the power transferred by the HAP. Thus, the aforementioned objective can be achieved by minimizing the HAP's transmitted energy while satisfying the given system constraints. Therefore, we design policies for the

RF energy harvesting, compression and transmission processes whilst guaranteeing the desired QoS requirements.

In this chapter, we consider a wireless powered MTC system comprising of a MTC node and a HAP. We employ a *harvest-and-use* strategy, i.e., the HAP first transfers RF power to the MTC node, which then transmits its data to the HAP.

Our investigation leads to the following contributions and observations:

- We propose to jointly optimize energy harvesting, compression and transmission times to minimize the energy cost of the system, which is injected by the HAP, when only statistical gain is known at the MTC node.
- Our results show that, for practical parameter values, employing and optimizing data compression in a wireless powered MTC system reduces the energy consumption by a further 19%, compared to only optimizing harvesting and transmission times. Specifically, the energy saving is significant when the delay constraint is less stringent.
- Our results also provide additional insights into the feasibility and actual operation of a wireless powered MTC node. For practical parameter values, the optimized time spent on sensing, compression and transmission is small and the majority of the optimized time is spent on power transfer operation.

The results of this chapter have been presented in the following publication [84]:

C2. **S. A. Alvi**, X. Zhou, and S. Durrani, "Wireless Powered Machine-Type Communication: Energy Minimization via Compressed Transmission," in *Proc. IEEE PIMRC*, Istanbul, Turkey, Sep. 2019.

Chapter 4 −− Multi-User Sequencing and Scheduling Control for Time-Division Multiple Access Systems

This chapter presents solution to the design challenge posed in Section 1.3.3.

We consider a single-channel multi-user uplink MTC communication system, in which multiple energy-constrained MTC devices transmit data to a base station (BS) within a fixed period of time, referred to as a frame, following the TDMA protocol. The BS allocates non-overlapping frame segments, referred to as transmission blocks, to individual MTC devices. Each MTC device transmits data to

the BS within its allocated transmission block. We consider that the devices apply data compression before the start of their scheduled transmission block and transmit the compressed data in the allocated transmission block. The main novelty of this work lies in the proposed multi-user sequencing, i.e., the order in which the devices are scheduled for transmission in the TDMA protocol.

In prior works, the TDMA performance is only optimized by controlling the time allocated to different MTC devices, i.e., the length of the allocated transmission blocks. In particular, the sequence of devices for transmission, i.e., who transmits first and who transmits second, etc., has no significance, given the channel statistics do not change from one transmission block to the other. However, in our proposed system the sequence of allocating the devices to the transmission blocks affects the amount of time allowed for applying data compression. As such, the energy-minimization objective can be achieved by allocating MTC devices with an optimized sequence and schedule of the transmission blocks.

To this end, we propose an optimal multi-user sequencing and scheduling scheme, and a sub-optimal multi-user scheduling scheme which does not employ multi-user sequencing. A comparative performance analysis of the two proposed schemes is carried out. Our investigation leads to the following observations and design insights:

- Our results show that the proposed optimal scheme outperforms the schemes without multi-user sequencing. The improvement due to multi-user sequence optimization is up to $35\% - 45\%$ depending on whether the length of the transmission blocks can be optimized or not.

- The energy efficiency gain of multi-user sequencing is most significant when the delay bound is stringent. In addition, multi-user sequence optimization makes the TDMA-based multi-user transmissions more likely to be feasible in the lower latency regime subject to the given power constraints.

- To solve the challenging mixed-integer nonlinear program for the proposed optimal scheme, we propose transformations to arrive at an approximate convex program which can be solved with significantly lower complexity. To solve this approximate program, we develop an algorithm that iteratively converges to a Fritz John solution.

The results of this chapter have been accepted/submitted in the following venues [85, 86]:

J1. **S. A. Alvi**, X. Zhou, S. Durrani, and D. T. Ngo, "Sequencing and Scheduling for Multi-User Machine-Type Communication," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2459-2473, Apr. 2020.

C1. **S. A. Alvi**, X. Zhou, S. Durrani, and D. T. Ngo, "Proportionally-Fair Sequencing and Scheduling for Machine-Type Communication," in *Proc. IEEE ICC*, Dublin, Ireland, Jun. 2020.

Chapter 5 −− Multi-User Transmit Power Diversity in Contention Resolution Diversity Slotted ALOHA Systems

This chapter presents solution to the design challenge posed in Section 1.3.4.

In this chapter, we propose a transmit power diversity based DSA scheme that exploits capture effect technique along with SIC. At the transmitter side, multiple bursts are transmitted on randomly selected slots. The number of the burst repetition is selected from a probability distribution. Moreover, each of these bursts are transmitted on randomly selected transmit power levels selected from a given probability distribution. At the receiver side, the bursts are iteratively resolved using SINR metric and SIC approach.

Our investigation leads to the following contributions and observations:

- We describe the proposed system by an edge-weighted bipartite graph and develop a novel graph-based message passing algorithm to perform the iterative SIC based decoding process.

- We employ the AND-OR tree analysis to derive an expression to characterize the system performance. Simulation results confirm the accuracy of the derived expression.

- We formulate and solve an optimization problem to determine the optimal transmit power probability distribution. The results show that by maximizing the probability of the gap between the power levels of two copies transmitted in a given slot, the optimal transmit power distribution leads to considerable

performance improvement. For instance, for only 2 copies per device per frame with only 5 power levels to choose from, the proposed scheme is shown to achieve a system load of 1.68 packets/slot, which shows its superiority to existing SIC based diversity slotted ALOHA methods.

The results of this chapter have been presented in the following publication [87]:

J3. **S. A. Alvi**, S. Durrani, and X. Zhou, "Enhancing CRDSA with Transmit Power Diversity for Machine-Type Communication," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7790-7794, Aug. 2018.

Finally, Chapter 6 provides a summary of the thesis results and makes suggestions for future research work.

Chapter 2

Single-User Compression and Transmission Rate Control for Device-Lifetime Maximization

In this chapter, we consider a network scenario that is applicable to a multi-user multi-channel sparse machine-type communication (MTC) system. In such a system, each machine-type device (MTD) is assigned a channel and transmit data on dedicated. The data may be generated or sensed by the MTD periodically, real-time, on demand, or due to the occurrence of an event of interest that is to be delivered under given reliability and delay conditions. Moreover, the MTDs are battery powered and perform duty cycling as well as deep sleep operation. This network scenario cover high reliability and low-latency MTC systems. For this network scenario, designing a generic single-user communication policy would suffice and the same policy can be employed by all other users for the individual system parameters.

Due to their wireless and unattended operation, MTC devices are mostly battery operated and are severely energy constrained. Thus, prolonging the lifetime of these sensor based MTC devices, which is defined as the time taken by the MTC device to deplete all of its energy, is of paramount importance [26]. In the existing literature, the lifetime maximization problem has been approached from different perspectives such as green channel access, sleep-wake scheduling, coverage, efficient

routing, network coding, data aggregation, see [26] and the references therein.

We consider a system that is composed of an energy constrained sensor node and a sink node, and devise optimal data compression and transmission policies with an objective to prolong the lifetime of the sensor node. While applying compression before transmission reduces the energy consumption of transmitting the sensed data, blindly applying too much compression may even exceed the cost of transmitting raw data, thereby losing its purpose. Hence, it is important to investigate the trade-off between data compression and transmission energy costs. In this chapter, we study the joint optimal compression-transmission design in three scenarios which differ in terms of the available channel information at the sensor node, and cover a wide range of practical situations. We formulate and solve joint optimization problems aiming to maximize the lifetime of the sensor node whilst satisfying specific delay and bit error rate (BER) constraints.

This chapter is organized as follows. The system model is presented in Section 2.1. The communication scenarios, which differ in the level of the channel knowledge at the sensor node, are discussed in Section 2.2. The lifetime maximization problems and their solutions are provided in Section 2.3. Numerical results are presented in Section 2.4. Finally, Section 2.5 concludes the chapter.

2.1 System Model

We consider a system consisting of a sensor based MTC device (sensor node) which is periodically transmitting its sensed data to a sink node, as illustrated in Fig. 2.1. Both nodes are equipped with a single omnidirectional antenna. The sensor node is battery operated and energy constrained, whereas the sink node has no energy constraint. The system follows a block-wise operation with a block of duration T, as shown in Fig. 2.2. Within each time block the sensor node performs three main functions, i.e., (i) sensing, (ii) compression, and (iii) transmission, each having individual completion time and energy cost. The block-wise operation and the duration of the block length depends on the sensor application. For example, in wireless multimedia sensor networks (WMSN) and Internet of multimedia things (IoMT) applications, the fresh data is periodically available which needs to be transmitted within a given deadline.

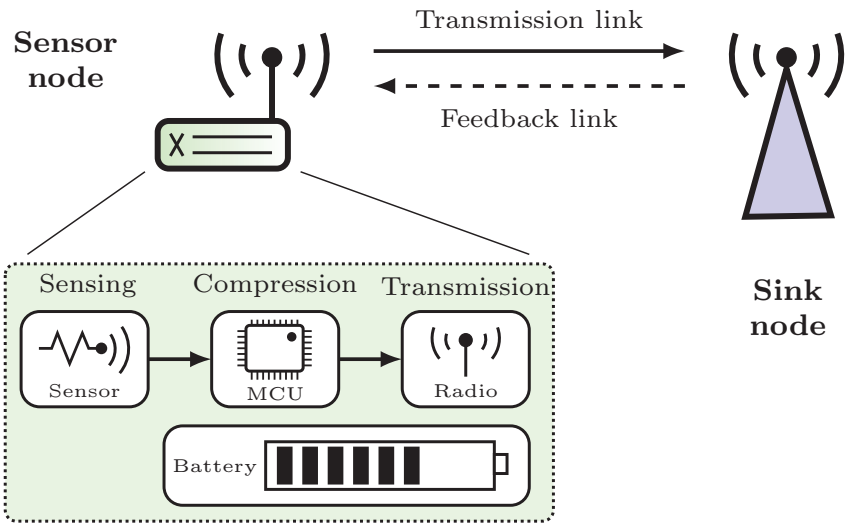


Figure 2.1: Illustration of the considered system model, comprising a sensor and a sink node.

For energy efficient operation, we employ radio duty cycling (RDC), i.e., radio is kept in the inactive state except during the transmission process. Moreover, the micro-controller unit (MCU) is kept in the inactive state, when it is neither compressing nor transmitting data, referred to as deep sleep. The transition periods from active to inactive states and vice versa are fast enough to be negligible for both radio and MCU. We assume the power consumed by the radio and MCU in inactive states is negligible [46, 47].

Sensing: The sensing operation is as follows. Firstly, the sensor node acquires the required information from the physical environment and encodes it to a data of size D bits. The sensed data during a given time block, is available for transmission at the start of next time block. In this chapter, we make the following assumptions regarding data sensing:

- The data sensing can be done in parallel while the compression and transmission processes are being executed.
- The sensed data is always periodically available, which is in line with prior works in this area [33, 52, 56, 64, 65, 88].
- The periodic sensing to acquire a fixed amount of physical information from

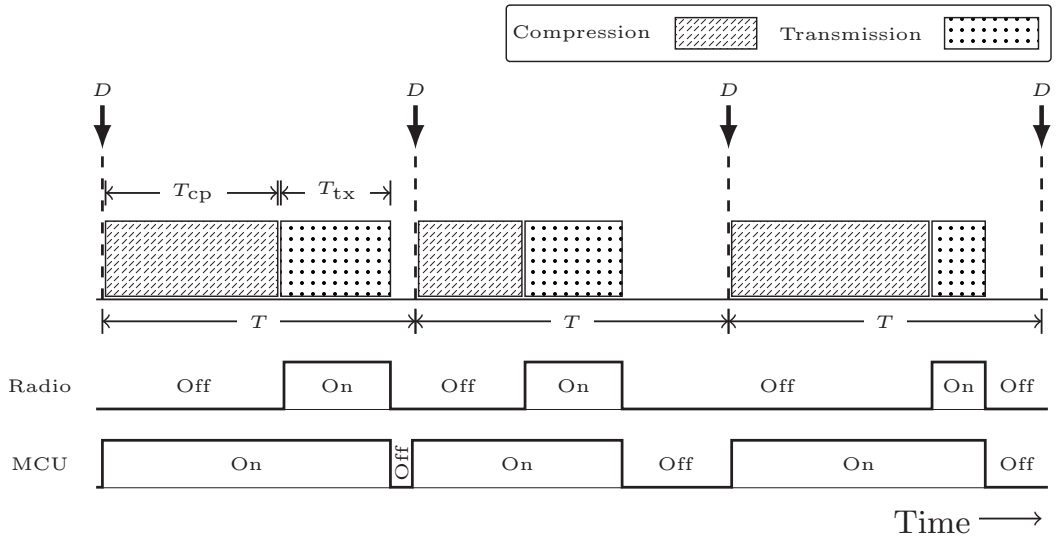


Figure 2.2: Timing diagram for compression and transmission processes and corresponding radio and MCU activity cycles.

the environment typically consumes a constant time and energy [89].

- The amount of data to be sensed, and the associated cost of sensing, depend only on the application and is independent of the compression and transmission processes.

Let the time and power spent by the sensor node to sense data of size D bits be denoted by T_{sen} and P_{sen}, respectively.

Compression: Before transmission, the sensed data of size D bits is compressed into D_{cp} bits as per the given compression ratio $\frac{D_{\text{cp}}}{D}$. The *compression time*, denoted by T_{cp}, is defined as the time required to compress raw data, D, into compressed data, D_{cp}. A non-linear compression cost model is proposed in [48] to evaluate the compression cost. This model is validated for JPEG and JPEG2000 compression algorithms in [48]. In this chapter, we adopt this non-linear model to compute the compression time as a function of the compression ratio, $\frac{D_{\text{cp}}}{D}$, as

$$T_{\text{cp}} = \tau D\left(\left(\frac{D}{D_{\text{cp}}}\right)^{\beta} - 1\right), \tag{2.1}$$

where β is the compression algorithm dependent parameter and τ is the per bit processing time. In general, β is proportional to the compression algorithm complexity and it determines the time cost for achieving a given compression ratio for given hardware resources. β can be calculated off-line for any specified compression algorithm and given hardware resources. τ depends upon the MCU processing resources and the number of program instructions executed to process 1 bit of data. Note that τ does not represent the compression time per bit. It can be given as

$$\tau = \underbrace{\frac{\text{instructions}}{\text{program}}}_{(i)} \times \underbrace{\frac{\text{clocks}}{\text{instruction}}}_{(ii)} \times \underbrace{\frac{\text{seconds}}{\text{clock}}}_{(iii)} \times \underbrace{\frac{1}{\text{reg}}}_{(iv)}, \tag{2.2}$$

The explanation for the terms in (2.2) is as follows:

(i) We assume a single-instruction program that is able to process 1 bit of information.

(ii) Most instructions in a typical sensor mote MCU are executed in 1 clock cycle. We assume a single instruction is executed in 1 clock cycle.

(iii) Seconds per clock represents the clock speed, i.e., the inverse of the MCU operational frequency which typically is between few MHz to hundreds of MHz.

(iv) reg represents MCU register size and its value for typical sensor motes is 8-bit. For 8-bit processor, the execution time to process 1 bit or up to 8 bits is the same. We assume D is large (thousands of bits) and it will be processed in chunks of 8 bits.

Let P_{cp} denote the power consumed by the sensor node during data compression process. P_{cp} is the same as the power consumed while sensor mote's MCU is processing information. Its value is predefined for a given sensor mote with a given hardware processing capability.

Transmission: Once the compression process is complete, the sensor node needs to transmit the compressed data, D_{cp}, within the next $T - T_{\text{cp}}$ seconds. The sensed data needs to be compressed and transmitted within each time block, hence

the delay constraint is T seconds. The *transmission time*, denoted by T_{tx}, depends upon the compressed data size, D_{cp}, and the link transmission rate, r. We consider the sensor node uses M-QAM modulation scheme with constellation size equal to $M = 2^l$, where $l = 1, 2, 3, .., L$. Thereby, T_{tx} is given as

$$T_{\text{tx}} = \frac{D_{\text{cp}}}{r}, \tag{2.3}$$

where

$$r = \frac{\log_2\left(M\right)}{T_{\text{s}}}, \tag{2.4}$$

where T_{s} is the symbol period and M is the modulation constellation size.

To compute the data transmission power cost, denoted by P_{tx}, we adopt a practical model as given in [57]. In this model, the total power cost of the sensor node while transmitting data, denoted by P_{tx}, is divided into three main components, the transmitted power, RF power amplifier, and the communication module circuitry power, denoted by P_{t}, P_{amp}, and P_{o}, respectively. Accordingly, P_{tx} is given as follows

$$P_{\text{tx}} = P_{\text{t}} + P_{\text{amp}} + P_{\text{o}}. \tag{2.5}$$

P_{o} is further divided into different communication circuitry modules [57]

$$P_{\text{o}} = P_{\text{fil}} + P_{\text{mix}} + P_{\text{syn}}, \tag{2.6}$$

where $P_{\text{fil}}, P_{\text{mix}}$, and P_{syn} is the power consumed by filter, mixer, and frequency synthesizer, respectively. The P_{amp} is a function of transmitted power and it can be given as [57]

$$P_{\text{amp}} = \Big(\frac{\varepsilon}{\mu} - 1\Big)P_{\text{t}}, \tag{2.7}$$

where μ represents the drain efficiency of the power amplifier [90] and ε represents the peak-to-average ratio (PAR) which depends upon the modulation scheme and the associated constellation size. Since, we consider M-QAM modulation, ε is given as [57, 91]

$$\varepsilon = 3\frac{M^{\frac{1}{2}} - 1}{M^{\frac{1}{2}} + 1}. \tag{2.8}$$

Substituting the value of P_{amp} in (2.5) and re-arranging, P_{tx} can be rewritten as

$$P_{\text{tx}} = \frac{\varepsilon}{\mu} P_{\text{t}} + P_{\text{o}}. \tag{2.9}$$

We assume that the battery used for sensor node possesses a limited charge storage capacity as well as a maximum current withdrawal limit. Therefore, instantaneous power demand of any process at any state should not exceed the maximum allowable limit. Specifically, the transmission power cost, which is a function of transmit power, needs to meet this power bound in order to ensure the feasibility of the system for practical sensor networks.

Channel model: The sensor node is located at a distance d from the sink node. The channel between the two nodes is composed of a large-scale path loss, with path loss exponent α, and small-scale quasi-static flat Rayleigh fading channel, i.e., the fading channel coefficient h remains constant over a time block and is independently and identically distributed from one time block to the next. The additive noise is assumed to be AWGN with zero mean and variance σ^2.

The *pdf* of the instantaneous channel gain, $|h|^2$, is exponentially distributed and is given as

$$f\big(|h|^2\big) \triangleq \frac{1}{\varsigma} \exp\left(- \frac{|h|^2}{\varsigma} \right), \quad |h|^2 \geqslant 0, \tag{2.10}$$

where ς represents the scale parameter of the probability distribution.

Node-Lifetime: We assume that the sensor node's battery is initially fully charged. Based on the battery capacity, operating voltage, and rate of energy consumption, we can calculate the node-lifetime, denoted by T_{NL}, which is defined as the time taken by the node to deplete all of its battery energy. The node-lifetime, T_{NL}, can be given as

$$T_{\text{NL}} = \frac{B_{\text{cap}} V_{\text{op}}}{P_{\text{avg}}}, \tag{2.11}$$

where B_{cap} represents the battery capacity that is a measure of the charge stored by the battery, V_{op} is the operating voltage, and P_{avg} represents the average power consumption by the sensing, compression, and transmission processes and is given as

$$P_{\text{avg}} = \frac{T_{\text{sen}} P_{\text{sen}} + \mathbb{E}\big[\Psi\big]}{T}, \tag{2.12}$$

where $\mathbb{E}[\cdot]$ is the expectation operator and Ψ is the energy consumed by the compression and transmission processes in a given time block which is given as

$$\Psi = T_{\text{cp}} P_{\text{cp}} + T_{\text{tx}} P_{\text{tx}}. \tag{2.13}$$

Note that the compression and transmission energy costs may change from one time block to the next. However, the sensing energy cost, $T_{\text{sen}} P_{\text{sen}}$, is the same for each time block.

2.2 Communication Scenarios and Channel Knowledge

The level of channel information available at the sensor node changes the compression and transmission policy design, since it imposes different constraints on the system which needs to comply with the channel knowledge. We consider three scenarios serving different important purposes. The perfect instantaneous channel gain information (CGI) (Scenario 1) is commonly used for theoretical performance analysis as a benchmark in wireless sensor networks. In reality when a feedback link with limited bandwidth is available then a quantized (imperfect) CGI knowledge is shared (Scenario 2). Finally, when the feedback link is not available, then the sensor node relies on the statistical channel information (Scenario 3). The considered scenarios are summarized as follows:

Scenario 1: Perfect CGI is available at the sensor node, when perfect feedback is available.

Scenario 2: Imperfect CGI is available at the sensor node, when limited feedback is available.

Scenario 3: Statistical CGI is available at the sensor node, when there is no feedback.

Note that the instantaneous CGI is assumed to be known at the sink node, which is a reasonable assumption when the sink node has no constraint on energy and data processing capability [92].

2.2.1 Instantaneous CGI Available at the Sensor Device

The sensor node is able to adaptively control the compression and transmission rate if the instantaneous CGI is available. Therein, we consider the following two scenarios.

Scenario 1: In this scenario, perfect instantaneous CGI is available at the sensor node. Therein, at the start of each time block, the sink node perfectly estimates the instantaneous fading channel coefficient and computes the CGI, $|h|^2$. The sink node then feeds back this CGI to the sensor node perfectly. Hence, perfect knowledge of instantaneous CGI is available at the sensor node.

Scenario 2: In this scenario, imperfect instantaneous CGI is available at the sensor node. Therein, at the start of each time block, the sink node perfectly estimates the instantaneous fading channel coefficient and computes the CGI, $|h|^2$. Moreover, the sink node is required to quantize the CGI to bound the feedback overhead, since in practical systems only limited feedback is available. Note, the CGI is a real and positive value which allows efficient quantization using a small number of bits [93]. The quantization of the CGI at the sink node, due to limited feedback, results in channel uncertainty at the sensor node side. We assume a perfect feedback link, hence CGI available at the sensor node is only subject to imperfection due to the quantization process.

The actual instantaneous CGI, $|h|^2$, can fall anywhere in the range $[0, \infty)$. We divide this range into 2^B quantization intervals, where B represents the number of feedback bits. The range of each of these quantization intervals is selected such that the probability of instantaneous CGI, $|h|^2$, falling in any of the given intervals is the same. Corresponding to these intervals, let the set of quantization levels be denoted by $\mathcal{C} = \{c_1, c_2, ..., c_{2^B}, c_{2^B+1}\}$, where c_j represents the jth quantization level and $c_1 = 0$ and $c_{2^B+1} = \infty$. For a given CGI, $|h|^2$, the sink node determines the interval $[c_i, c_{i+1})$, for $i \in \{1, 2, ..., 2^B\}$ such that $c_i \leqslant |h|^2 < c_{i+1}$, and feeds back the index i corresponding to the quantization interval $[c_i, c_{i+1})$ to the sensor node. Note that the CGI, $|h|^2$, falls in each of these intervals with equal probability, i.e.,

$$\mathbb{P}\big\{c_i \leqslant |h|^2 < c_{i+1}\big\} = \frac{1}{2^B}, \quad \forall\, i \in \{1, 2, ..., 2^B\}, \tag{2.14}$$

where $\mathbb{P}\{\cdot\}$ represents the probability. We assume the time spent in estimation,

quantization, and feedback is negligible, thus the sensor node can effectively exploit the provided CGI.

2.2.2 Statistical CGI Available at the Sensor Device

Scenario 3: In this scenario, the sink node has perfect estimate of the channel but no instantaneous feedback is available. Only the statistical CGI is available at the sensor node. Therefore, the sensor node cannot adapt compression and transmission policies to varying channel conditions in different time blocks. Instead a constant compression ratio and transmission rate is determined, and subsequently used in each time block, which maximizes the node-lifetime.

2.2.3 BER Expression

We consider the sensor node is able to control the transmission rate, wherein the best M value needs to be determined which will allow the sensor node to achieve the required system performance. Note that many different BER expressions exist in the literature for M-QAM. Here, we use the following BER bound defined for M-QAM modulation scheme [94], since it is easy to invert in order to obtain M as a function of the required BER

$$\text{BER} \leqslant \omega_2 \exp\bigg(- \frac{\omega_1}{(M-1)}\gamma \bigg), \tag{2.15}$$

where ω_1, ω_2 are constants and γ represents the received signal-to-noise ratio (SNR) which is defined as follows [82]

$$\gamma = \kappa \frac{P_{\text{t}}|h|^2}{\sigma^2 d^{\alpha}}, \tag{2.16}$$

where $\kappa = \left(\frac{\lambda}{4\pi}\right)^2$ is the attenuation factor, λ is the wavelength and P_{t} is the transmit power. For $M \geqslant 4$ and $0 \leqslant \gamma \leqslant 20$ dB, the bound in (2.15) with $\omega_1 = 1.5$ and $\omega_2 = 0.2$, is tight to within 1 dB of the exact result in [82].

In Section 2.3, we will use (2.15) to determine the optimal design parameters for each of the scenarios defined above.

2.3 Device-Lifetime Maximization Problem

The main problem we address is to determine the optimal compression and transmission policies which will maximize the node-lifetime under given delay constraint and BER performance. From (2.11), we can see that the node-lifetime is calculated using a predefined initial energy level, $B_{\text{cap}}V_{\text{op}}$, and the controllable rate of energy consumption, P_{avg}, which is a function of $\mathbb{E}[\Psi]$. Node-lifetime is inversely proportional to $\mathbb{E}[\Psi]$, which implies that maximizing node-lifetime is equivalent to minimizing $\mathbb{E}[\Psi]$. Based on the compression and transmission energy cost models, defined in Section II, Ψ inherits the tradeoff between data compression and transmission.

In this section, we study three different problems based on the scenarios defined in Section 2.2.1. In Scenario 1 and Scenario 2, the knowledge of the instantaneous CGI is used by the sensor node to optimally choose the design parameters in order to adapt to each realization of the channel, with an objective to minimize Ψ in each time block. On the other hand, in Scenario 3, when no knowledge of the instantaneous CGI is available, the optimal design parameters are set to be fixed for each time block, wherein, the objective is still to minimize Ψ. It is because, in Scenario 3, the value of Ψ is the same for all time blocks.

2.3.1 Instantaneous CGI Available at the Sensor Device

In this subsection, we consider an adaptive compression and transmission rate control system in which for a given time block, the design target is to minimize the energy cost of compression and transmission under the given delay and BER constraints. Based on the availability of the instantaneous CGI at the sensor node we study the following two problems.

2.3.1.1 Perfect CGI Availability

The first problem we study, considering the availability of perfect CGI at the sensor node, can be summarized as follows:

Problem 1: What is the optimal compression and transmission policy that minimizes the compression and transmission energy cost under specific delay and

BER constraints, when perfect channel gain is known at the sensor node?

The energy consumption cost of compression and transmission, Ψ, is defined in (2.13). Now given the value of instantaneous CGI, $|h|^2$, Problem 1 can be expressed as follows

$$\begin{aligned} \underset{M,P_{\mathrm{t}},D_{\mathrm{cp}}}{\text{minimize}} \quad & \Psi(M, P_{\mathrm{t}}, D_{\mathrm{cp}}) \\ \text{subject to} \quad & T_{\mathrm{cp}} + T_{\mathrm{tx}} \leqslant T, \\ & \mathrm{BER}(M, P_{\mathrm{t}}) \leqslant \phi, \\ & P_{\mathrm{t}} \geqslant 0, \\ & M \geqslant 2, \\ & M \leqslant M_{\max}, \\ & D_{\mathrm{cp}} \geqslant D_{\min}, \\ & D_{\mathrm{cp}} \leqslant D. \end{aligned} \tag{2.17}$$

where the first constraint defines the delay constraint for the data delivery, thus both compression and transmission processes should be completed within the deadline and the second constraint mandates that the BER should be below or equal to a specific threshold value denoted by ϕ. The remaining constraints reflect practical range of values for M, P_{t}, and D_{cp}. In the fifth constraint, $M_{\max}$ is the maximum value of the constellation size which can be used by the sensor node. In the sixth constraint, $D_{\min}$ represents the maximum achievable compression that can be applied using a given compression algorithm and the compressed data is completely transmitted within the same time block at the highest allowed transmission rate. Thus, $D_{\min}$ is the lower bound on D_{cp} and a function of $M_{\max}$, T, D and β. $D_{\min}$ is computed by numerically solving the following equation using the variable-precision floating-point arithmetic (VPA) method

$$\tau D^{\beta+1} D_{\min}^{-\beta} - \tau D + \frac{D_{\min} T_{\mathrm{s}} \ln(2)}{\ln(M_{\max})} = T. \tag{2.18}$$

In order to solve (2.17), we first present Proposition 1, which allows the solution to (2.17) to be given by Theorem 1.

Proposition 1. *The optimal P_{t} to minimize $\Psi(M, P_{\mathrm{t}}, D_{\mathrm{cp}})$ for given values of M*

and D_{cp} *while satisfying the constraints in* (2.17) *is given by*

$$P_{\text{t}} = (1 - M)\frac{\Omega}{|h|^2}, \tag{2.19}$$

where

$$\Omega = \frac{\sigma^2 d^{\alpha} \ln(\phi/\omega_2)}{\omega_1 \kappa}. \tag{2.20}$$

Proof. The proof is provided in Appendix D.1. ■

Using the result in Proposition 1, substituting T_{cp}, T_{tx}, r, P_{tx}, ε and P_{t} from (2.1), (2.3), (2.4), (2.9), (2.8) and (2.19), respectively, in (2.13) yields Ψ as a function of M and D_{cp} for given instantaneous CGI, $|h|^2$, which can be expressed as follows

$$\Psi(M, D_{\text{cp}}) \;=\; \tau D^{\beta+1} D_{\text{cp}}^{-\beta} P_{\text{cp}} \;-\; \tau D P_{\text{cp}} \;-\; \frac{D_{\text{cp}} T_{\text{s}} \ln(2)}{\ln(M)} \left(\frac{3\Omega (M^{\frac{1}{2}} - 1)^2}{\mu |h|^2} \;-\; P_{\text{o}} \right). \tag{2.21}$$

Now a simpler equivalent optimization problem with only two design parameters, i.e., M, D_{cp}, needs to be solved and the third parameter P_{t} can be obtained using the result in Proposition 1. Accordingly, the solution to the optimization problem in (2.17) is given by the following theorem.

Theorem 1. *In solving the optimization problem in* (2.17)*, the optimal constellation size is given by the following conditional expression*

$$M^* = \begin{cases} \widetilde{M}, & \text{if } \mathcal{Q}(\widetilde{M}, \widetilde{D}_{\text{cp}}) < T. \\ \min\left(\widehat{M}, M_{\max}\right), & \text{otherwise.} \end{cases} \tag{2.22}$$

where $\widetilde{M}$ *and* $\widehat{M}$ *are given by the solution of the following equations which can be solved numerically using the* VPA *method*

$$\frac{3\Omega}{\mu |h|^2} \left(\widetilde{M}^{\frac{1}{2}} - 1\right) \left((\ln(\widetilde{M}) - 1)\widetilde{M}^{\frac{1}{2}} + 1 \right) + P_{\text{o}} = 0, \tag{2.23}$$

$$\frac{T}{D} + \tau - \tau \xi^{\frac{-\beta}{\beta+1}} = \frac{T_{\text{s}} \ln(2)}{\ln\left(\widehat{M}\right)} \xi^{\frac{1}{\beta+1}}, \tag{2.24}$$

where

$$\xi = \frac{P_{\rm cp} - P_{\rm o} - \frac{3\Omega}{\mu|h|^2}\left(\widehat{M}^{\frac{1}{2}}-1\right)\left((\ln(\widehat{M})-1)\widehat{M}^{\frac{1}{2}}+1\right)}{\frac{3\Omega}{\tau\beta\mu|h|^2}\left(\widehat{M}^{\frac{1}{2}}-\widehat{M}\right)}, \tag{2.25}$$

respectively,

$$\mathcal{Q}(\widetilde{M},\widetilde{D}_{\rm cp}) \triangleq \tau D^{\beta+1}\widetilde{D}_{\rm cp}^{-\beta} - \tau D + \frac{\widetilde{D}_{\rm cp}T_{\rm s}}{\log_2\left(\widetilde{M}\right)}, \tag{2.26}$$

and

$$\frac{\widetilde{D}_{\rm cp}}{D} = \left(\frac{\tau\beta P_{\rm cp}\ln(\widetilde{M})}{\frac{3\Omega T_{\rm s}\ln(2)}{\mu|h|^2}\left(\widetilde{M}^{\frac{1}{2}}-1\right)^2 + P_{\rm o}T_{\rm s}\ln(2)}\right)^{\frac{1}{\beta+1}}, \tag{2.27}$$

and the optimal transmit power is given by

$$P_{\rm t}^* = \left(1-M^*\right)\frac{\Omega}{|h|^2}, \tag{2.28}$$

and the optimal compression ratio is given by

$$\frac{D_{\rm cp}^*}{D} = \begin{cases} \frac{\widetilde{D}_{\rm cp}}{D}, & \text{if } \mathcal{Q}(\widetilde{M},\widetilde{D}_{\rm cp}) < T. \\ \max\left(\frac{D_{\min}}{D}, \frac{\widehat{D}_{\rm cp}}{D}\right), & \text{otherwise.} \end{cases} \tag{2.29}$$

where $\frac{\widehat{D}_{\rm cp}}{D} = \xi^{\frac{1}{\beta+1}}$ *and* ξ *is defined in* (2.25).

Proof. The proof is provided in Appendix C.2. ■

The insights from Theorem 1 are discussed in the following four remarks.

Remark 1. *$\widetilde{M}$ and $\widetilde{D}_{\rm cp}$ provide a lower bound on the optimization problem in* (2.17) *for given instantaneous CGI, $|h|^2$. $\widetilde{M}$ and $\widetilde{D}_{\rm cp}$ are optimal design parameters when the first constraint in* (2.17) *is slack, i.e., $\mathcal{Q}(\widetilde{M},\widetilde{D}_{\rm cp}) < T$, and other constraints are also slack. On the other hand, $\widehat{M}$ and $\widehat{D}_{\rm cp}$ are optimal design parameters for optimization problem in* (2.17) *for given instantaneous CGI, $|h|^2$, when all constraints in* (2.17) *are slack except for the first constraint.*

Remark 2. *In prior power-rate adaptation schemes* [52–56, 95]*, which do not consider data compression, the transmission policy is adapted with channel variations considering delay constraint under a fixed* BER*. In these schemes, for a fixed*

amount of data to be transmitted, the transmission policy is adapted for the channel realization in each time block. However, in our case the transmitted compressed data size changes from one time block to the next, as a result of joint optimization of the transmission and compression policies adapting the channel realization in each time block.

Remark 3. *The classical works* [52, 53] *have designed the transmission rate control policies which are adaptive to the channel conditions. These schemes do not consider data compression. Their optimal solution suggests that the energy cost of data communication is a strictly increasing function of the transmission rate. Therefore, the transmission rate should be minimized for the given delay bound, T, in order to minimize the energy cost of data communication cost. That is, the lowest transmission rate which will meet the delay constraint with equality is optimal for any value of the delay bound, T. However, in our case the combined data compression and transmission rate strategy suggests that there exists a lower bound on the total energy cost of compression and transmission. In this regard, the corresponding optimal design parameters cost a finite delay, $\mathcal{Q}$. Therefore, if the required delay constraint is larger than this delay only then these design parameters will maximize the lifetime and are optimal. Hence, in general, it is not optimal to transmit at the lowest transmission rate.*

Remark 4. *The optimal constellation size given by* (2.22) *is real valued. Thus, for practical admissibility, the transmission policy should opt to select the closest value from the available set of modulation order values. If a lower value is closer then it can only be selected if the first constraint in* (2.17) *is slack, else a higher value should be selected which will surely satisfy the first constraint in* (2.17)*. This optimal practical value, denoted by M^*_{pr}, is subsequently used to determine P^*_{t} and $\frac{D^*_{\mathrm{cp}}}{D}$. M^*_{pr} can be obtained using the following conditional expression*

$$M^*_{\mathrm{pr}} = \begin{cases} \min\left(2^L, \nu_1\right), & \text{if } |M^* - \nu_1| \leqslant |M^* - \nu_2| \\ & \text{and } \mathcal{Q}(\nu_1, \breve{D}_{\mathrm{cp}}) < T. \\ \min\left(2^L, \nu_2\right), & \text{otherwise.} \end{cases} \tag{2.30}$$

where $\min(\cdot)$ *is the* min *operation,* $|\cdot|$ *represents the absolute value,* 2^L *is the maxi-*

mum modulation order supported by the system, $\nu_1 = 2^{\lfloor \log_2(M^*) \rfloor}$, $\nu_2 = 2^{\lceil \log_2(M^*) \rceil}$, $\mathcal{Q}$ *is defined in* (2.39), $\lfloor \cdot \rfloor$ *and* $\lceil \cdot \rceil$ *is the floor and ceil operations, respectively. Note, the compressed data size* $\breve{D}_{\mathrm{cp}}$*, is a function of the constellation size* $2^{\lfloor \log_2(M^*) \rfloor}$.

2.3.1.2 Imperfect CGI Availability

The second problem we study considering imperfect CGI availability at the sensor node can be given as:

Problem 2: What is the optimal compression and transmission policy that minimizes the compression and transmission energy cost under specific delay and BER constraints, when quantized channel gain is known at the sensor node?

In Scenario 2, for a given channel realization h, the sink node feeds back the quantization level index i to the sensor node, which implies CGI $|h|^2$, lies in the interval $[c_i, c_{i+1})$, i.e., $c_i \leqslant |h|^2 < c_{i+1}$. Hence, there lies some channel uncertainty at the sensor node. As discussed above, the sensor node can minimize Ψ in order to maximize the node-lifetime. Therefore, given $c_i \leqslant |h|^2 < c_{i+1}$, the Problem 2 can be expressed as follows

$$\begin{aligned} \underset{M, P_{\mathrm{t}}, D_{\mathrm{cp}}}{\text{minimize}} \quad & \Psi(M, P_{\mathrm{t}}, D_{\mathrm{cp}}) \\ \text{subject to} \quad & T_{\mathrm{cp}} + T_{\mathrm{tx}} \leqslant T, \\ & \mathrm{BER}(M, P_{\mathrm{t}}) \leqslant \phi, \, \forall \, c_i \leqslant |h|^2 < c_{i+1}, \\ & P_{\mathrm{t}} \geqslant 0, \\ & M \geqslant 2, \\ & M \leqslant M_{\max}, \\ & D_{\mathrm{cp}} \geqslant D_{\min}, \\ & D_{\mathrm{cp}} \leqslant D. \end{aligned} \tag{2.31}$$

where the first constraint in (2.31) defines the delay constraint for the data delivery, thus both compression and transmission processes should be completed within the deadline, and the second constraint in (2.31) mandates that the BER should be below or equal to a specific threshold value denoted by ϕ, where the actual CGI, $|h|^2$, lies within the interval $[c_i, c_{i+1})$.

Accordingly, the solution to the problem in (2.31) is given by the following corollary.

Corollary 1. *In solving the optimization problem in* (2.31)*, the optimal constellation size, transmit power, and compression ratio can be obtained using* (2.22)*,* (2.28)*, and* (2.29)*, respectively, by replacing* $|h|^2$ *with* c_i*.*

Proof. The BER, defined in (2.15), is a function of constellation size, M, and the CGI, $|h|^2$. Recall, the sink node feeds back the index i corresponding to the quantization interval, where the actual CGI lies in the interval $c_i \leqslant |h|^2 < c_{i+1}$, as mentioned earlier in Section 2.2.1. In order to meet the required BER constraint, defined in (2.31), the sensor node selects the smallest value, i.e., c_i, representing the worst-case choice for the CGI, in the quantization interval $[c_i, c_{i+1})$ indicated by the index i. This underestimation ensures the quantized CGI, c_i, is always less than or equal to the actual CGI, i.e., $c_i \leqslant |h|^2$. Hence the required BER constraint, defined in (2.31), is naturally satisfied for any given level of channel uncertainty due to CGI quantization, i.e., $\text{BER}(c_i) \leqslant \text{BER}(|h|^2) \leqslant \phi$. Therefore, by replacing $|h|^2$ with c_i in (2.15) and following similar steps provided in Appendix D.1, the optimal constellation size, transmit power, and compression ratio can be obtained, given $c_i \leqslant |h|^2 < c_{i+1}$. ■

Remark 5. *The values for the optimal design parameters for optimization problem defined in* (2.31) *can be obtained offline using Corollary 1. This is because there are only a finite set of quantization levels, each corresponding to a set of design parameters values. For example, if 5 feedback bits are used then only 32 different values for the optimal design parameters need to be computed. These values can then be used in a given time block based on the received quantized CGI feedback value, denoted by* c_i*. Hence, the proposed solution for Scenario 2 is practical and feasible for energy constrained wireless sensor based MTC devices.*

2.3.2 Statistical CGI Availability

The third problem we study considering availability of only statistical information of CGI at the sensor node can be summarized as follows:

Problem 3: What is the optimal compression and transmission policy that minimizes the compression and transmission energy cost under specific delay and BER constraints, when only statistical information about the channel gain is known at the sensor node?

Since, the instantaneous CGI is not available, thus the design parameters cannot be adapted to any given channel condition and cannot guarantee a certain BER, in a given time block. However, we can determine the optimal design parameters which are set to be the same for each time block with an objective to minimize the compression and transmission energy cost, Ψ. Note that $\mathbb{E}[\Psi] = \Psi$ when the design parameters are set to be the same for all time blocks. Given the fading power gain distribution, $f(|h|^2)$, Problem 3 can be expressed as follows

$$\begin{aligned} \underset{M,P_{\mathrm{t}},D_{\mathrm{cp}}}{\text{minimize}} \quad & \Psi(M, P_{\mathrm{t}}, D_{\mathrm{cp}}) \\ \text{subject to} \quad & T_{\mathrm{cp}} + T_{\mathrm{tx}} \leqslant T, \\ & \mathbb{P}\{\mathrm{BER} \leqslant \phi\} \geqslant \vartheta, \\ & P_{\mathrm{t}} \geqslant 0, \\ & M \geqslant 2, \\ & M \leqslant M_{\max}, \\ & D_{\mathrm{cp}} \geqslant D_{\min}, \\ & D_{\mathrm{cp}} \leqslant D. \end{aligned} \tag{2.32}$$

where the first constraint in (2.32) defines the delay constraint for the data delivery, thus both compression and transmission processes should be completed within the deadline, and the second constraint in (2.32) mandates that the probability of having as acceptable level of BER should be greater than certain percentage. Specifically, ϕ denote the maximum acceptable BER and ϑ denote the required minimum probability of achieving the acceptable BER performance. The third constraint in (2.32) guarantees a given BER performance with a certain probability in each time block. Note that an alternative way of constraining the BER performance is to put an upper bound on the average BER over all time blocks. But we do not adopt it because it gives minimal control over the BER performance in each time block.

Proposition 2. *The optimal* P_{t} *to minimize* $\Psi(M, P_{\mathrm{t}}, D_{\mathrm{cp}})$ *for given values of* M *and* D_{cp} *while satisfying the constraints in* (2.32) *is given by*

$$P_{\mathrm{t}} = (M-1)\frac{\Omega}{\varsigma \ln(\vartheta)}. \tag{2.33}$$

Proof. The proof is provided in Appendix C.3. ■

Using the result in Proposition 2, substituting T_{cp}, T_{tx}, r, P_{tx}, ε and P_{t} from (2.1), (2.3), (2.4), (2.9), (2.8) and (2.33), respectively, in (2.13) yields Ψ as a function of M and D_{cp} as follows

$$\Psi(M, D_{\mathrm{cp}}) \;=\; \tau D^{\beta+1} D_{\mathrm{cp}}^{-\beta} P_{\mathrm{cp}} \,-\, \tau D P_{\mathrm{cp}} \,+\, \frac{D_{\mathrm{cp}} T_{\mathrm{s}} \ln(2)}{\ln(M)} \left(\frac{3\Omega (M^{\frac{1}{2}} - 1)^2}{\mu \varsigma \ln(\vartheta)} \,+\, P_{\mathrm{o}} \right). \tag{2.34}$$

Now a simpler equivalent optimization problem with only two design parameters, i.e., M, D_{cp}, needs to be solved and the third parameter P_{t} can be obtained using the result in Proposition 2. Accordingly, the solution to the optimization problem in (2.32) is given by the following theorem.

Theorem 2. *In solving the optimization problem in* (2.32)*, the optimal constellation size is given by the following conditional expression*

$$M^* = \begin{cases} \widetilde{M}, & \text{if } \mathcal{Q}(\widetilde{M}, \widetilde{D}_{\mathrm{cp}}) < T. \\ \min\left(\widehat{M}, M_{\max}\right), & \text{otherwise.} \end{cases} \tag{2.35}$$

where $\widetilde{M}$ *and* $\widehat{M}$ *are given by the solution of the following equations which can be solved numerically using* VPA *method*

$$-\frac{3\Omega}{\mu \varsigma \ln(\vartheta)} \big(\widetilde{M}^{\frac{1}{2}} - 1\big)\big((\ln(\widetilde{M}) - 1)\widetilde{M}^{\frac{1}{2}} + 1\big) + P_{\mathrm{o}} = 0, \tag{2.36}$$

$$\frac{T}{D} + \tau - \tau \zeta^{\frac{-\beta}{\beta+1}} = \frac{T_{\mathrm{s}} \ln(2)}{\ln(\widehat{M})} \zeta^{\frac{1}{\beta+1}}, \tag{2.37}$$

where

$$\zeta = \frac{P_{\mathrm{cp}} - P_{\mathrm{o}} + \frac{3\Omega}{\mu\varsigma\ln(\vartheta)}\left(\widehat{M}^{\frac{1}{2}} - 1\right)\left((\ln(\widehat{M}) - 1)\widehat{M}^{\frac{1}{2}} + 1\right)}{-\frac{3\Omega}{\varsigma\tau\beta\mu\ln(\vartheta)}\left(\widehat{M}^{\frac{1}{2}} - \widehat{M}\right)}, \tag{2.38}$$

respectively,

$$\mathcal{Q}(\widetilde{M}, \widetilde{D}_{\mathrm{cp}}) \triangleq \tau D^{\beta+1}\widetilde{D}_{\mathrm{cp}}^{-\beta} - \tau D + \frac{\widetilde{D}_{\mathrm{cp}} T_{\mathrm{s}}}{\log_2\left(\widetilde{M}\right)}, \tag{2.39}$$

and

$$\frac{\widetilde{D}_{\mathrm{cp}}}{D} = \left(\frac{\tau\beta P_{\mathrm{cp}}\ln(\widetilde{M})}{-\frac{3\Omega T_{\mathrm{s}}\ln(2)}{\mu\varsigma\ln(\vartheta)}\left(\widetilde{M}^{\frac{1}{2}} - 1\right)^2 + P_{\mathrm{o}} T_{\mathrm{s}}\ln(2)}\right)^{\frac{1}{\beta+1}}, \tag{2.40}$$

and the optimal transmit power is given by

$$P_{\mathrm{t}}^{*} = \left(M^{*} - 1\right)\frac{\Omega}{\varsigma\ln(\vartheta)}, \tag{2.41}$$

and the optimal compression ratio is given by

$$\frac{D_{\mathrm{cp}}^{*}}{D} = \begin{cases} \frac{\widetilde{D}_{\mathrm{cp}}}{D}, & \text{if } \mathcal{Q}(\widetilde{M}, \widetilde{D}_{\mathrm{cp}}) < T. \\ \max\left(\frac{D_{\min}}{D}, \frac{\widehat{D}_{\mathrm{cp}}}{D}\right), & \text{otherwise.} \end{cases} \tag{2.42}$$

where $\frac{\widehat{D}_{\mathrm{cp}}}{D} = \zeta^{\frac{1}{\beta+1}}$ *and* ζ *is defined in* (2.38).

Proof. The proof follows similar steps as the proof of Theorem 1. Hence, it is omitted for brevity. ■

Remark 6. *As discussed earlier in* Remark 4*, for practical admissibility the optimal constellation size,* M^{*}*, needs to be scaled to the closest achievable practical value using* (2.30)*. Similarly,* P_{t}^{*} *needs to meet the maximum power bound of the battery.*

Remark 7. *The values for the optimal design parameters for optimization problem defined in* (2.32) *can be obtained offline using Theorem 2. It is because these values are fixed for all time blocks. Thus, the proposed solution for Scenario 3 is practical and feasible for energy constrained wireless sensor based MTC devices.*

Remark 8. *An alternative scenario to consider is to use an average BER constraint even when the instantaneous CGI is available. In this case, the lifetime can*

be further extended by dropping the sensed data in a given time block in which the channel is in deep fade. Let us replace the BER constraint function in (2.17) *as follows:* $\mathbb{P}\{\text{BER} \leqslant \phi\} \geqslant \vartheta$. *Thus, the probabilistic BER constraint can directly give a threshold on channel gain, given by* $\Theta = -\varsigma \ln\left(\vartheta\right)$, *above which data transmission should occur. Then the sensor only employs Theorem 1 to design compression and transmission when the instantaneous channel gain is above the threshold,* Θ.

2.4 Numerical Results

In this section, we present the numerical results to illustrate the performance of the proposed Scenarios 1-3 jointly optimizing data compression and transmission rate. Unless specified otherwise, the values for the system parameters shown in Table 5.1 are adopted.

To illustrate the advantage of joint optimization of data compression and transmission rate, we are interested in prior works which provide solutions for the data transmission design to maximize node-lifetime. To the best of our knowledge, the prior works [53, 54, 96] are the most relevant and most recent schemes which can be compared to our proposed scheme. In this regard, we adopt the data transmission design policies proposed by these schemes for our considered system model except that data compression is not employed. Moreover, when our considered bit error rate and delay constraints are applied, the resultant data transmission design problem can be given as in (17) by substituting $D_{\text{cp}} = D$. We refer to this adapted scheme as the baseline scheme. The strategy followed to optimize the transmission rate policy for the baseline scheme is essentially the same as in the state of the art [53,54,96]. The optimal constellation size for this scheme, denoted by M^*_{nc}, can be obtained using Proposition 3.

Proposition 3. *The optimal constellation size to maximize lifetime without performing compression while satisfying constraints given in* (2.17) *is given by the following conditional expression*

$$M^*_{\text{nc}} = \begin{cases} \widetilde{M}, & \text{if } \frac{DT_{\text{s}}}{\log_2\left(\widetilde{M}\right)} < T. \\ \exp\left(\frac{DT_{\text{s}}\ln(2)}{T}\right), & \text{otherwise.} \end{cases} \tag{2.43}$$

Table 2.1: System parameter values.

Name	**Symbol**	**Value**
Drain efficiency of power amplifier	μ	0.35
Operating voltage	V_{op}	3 V
Scale parameter for CGI's *pdf*	ς	1
Battery capacity	B_{cap}	9000 As
Power cost of compression	P_{cp}	24 mW
Symbol period	T_{s}	16 μs
Power cost of synthesizer	P_{syn}	50 mW
constant	ω_2	0.2
Power cost of filter	P_{fil}	2.5 mW
constant	ω_1	1.5
Power cost of mixer	P_{mix}	30.3 mW
distance	d	20 m
Uncompressed data	D	20kb
Noise power	σ^2	-174 dBm
Per bit processing time	τ	0.35 ns/b
BER constraint	ϕ	10^{-3}
Compression cost parameter	β	5
Delay constraint	T	50 ms

where $\widetilde{M}$ *is given by the solution of the following equation which can be solved numerically using the* VPA *method*

$$\frac{3\Omega}{\mu|h|^2}\big(\widetilde{M}^{\frac{1}{2}}-1\big)\big((\ln(\widetilde{M})-1)\widetilde{M}^{\frac{1}{2}}+1\big)+P_{\text{o}}=0, \tag{2.44}$$

and the optimal transmit power is given by

$$P_{\text{t}}^{*}=\big(1-M_{\text{nc}}^{*}\big)\frac{\Omega}{|h|^2}. \tag{2.45}$$

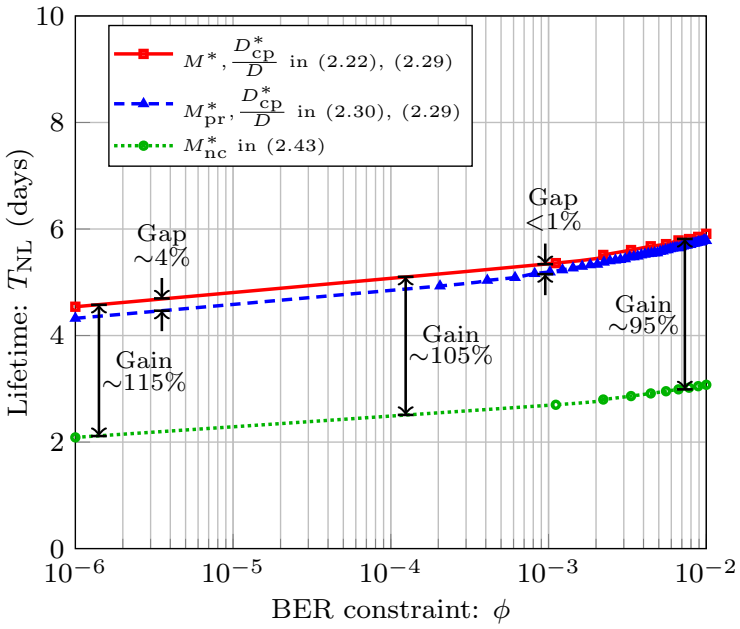
Proof. The proof follows similar steps as the proof of Theorem 1 and substituting $D_{\text{cp}}=D$. Hence, it is omitted for brevity. ■

2.4.1 Advantage of Proposed Scheme

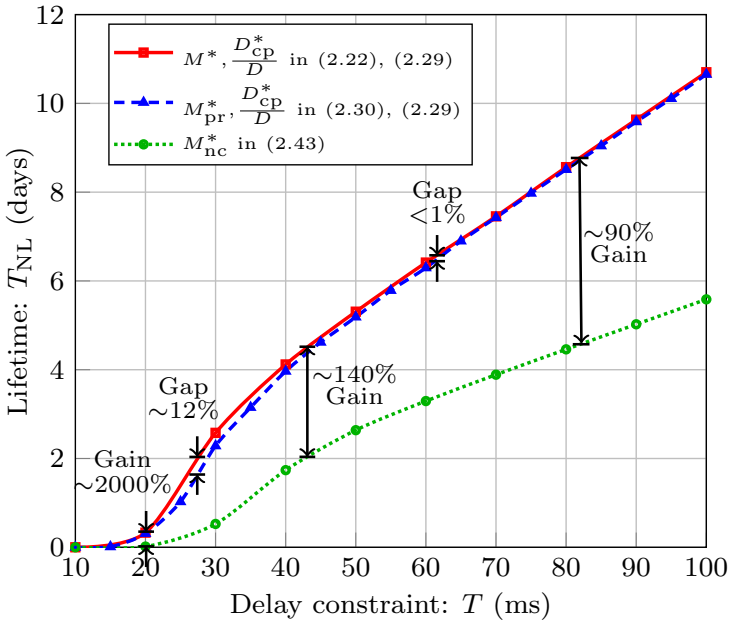
Fig. 2.3a plots the node-lifetime, T_{NL} (days), versus the BER constraint, ϕ, for Scenario 1 and system parameters in Table 5.1. The lifetime is plotted with the optimal (real valued) M^* in (2.22), the practical (quantized value) ${M_{\mathrm{pr}}}^*$ in (2.30) and the baseline scheme in ${M_{\mathrm{nc}}}^*$ in (2.43) in Fig. 2.3a. We can see that the gain compared to the baseline scheme is significant - between 95% to 115% for the considered range of BER constraint. This shows the advantage of joint optimal compression and transmission rate control. *In addition, we can see that the node-lifetime is not so significantly affected by the BER constraint.* As the BER constraint is relaxed, the lifetime slightly increases. For instance, as BER constraint is varied from stringent BER requirement, i.e., 10^{-6}, to loose BER requirement, i.e., 10^{-2}, the lifetime only changes by around 30%. Finally, the performance with practical modulation scheme is very close to the optimal performance, e.g., at $\phi = 10^{-5}$ the gap is less than 4%.

Fig. 2.3b plots the node-lifetime, T_{NL} (days), versus the delay constraint, T (ms), for Scenario 1 and system parameters in Table 5.1. The node-lifetime is plotted with the optimal (real valued) M^* in (2.22), the practical (quantized value) ${M_{\mathrm{pr}}}^*$ in (2.30) and the baseline scheme in ${M_{\mathrm{nc}}}^*$ in (2.43) in Fig. 2.3b. The compression ratio and transmission rate are optimized through M and D_{cp}. Though inherently independent, jointly optimizing M, P_{t}, and D_{cp} to maximize the node-lifetime yields their relationship as follows. As given in Theorem 1, the optimal constellation size has a direct relationship with both the optimal transmit power level and the optimal compressed data size. Note, a smaller value of compressed data size means a high level of compression is applied.

The explanation of the gains shown in Fig. 2.3b is as follows. The performance gains are comparatively lower ($\backsim$ 90%) in the case when the delay bound is loose ($\backsim$ $80ms$). This is because the data communication energy cost without employing compression almost increases exponentially with the employed transmission rate. When the delay bound is stringent then relatively lower transmission rate is good enough to transmit data whilst satisfying the BER and delay constraints. Since, the data communication energy cost without employing compression is not too significant, employing compression improves the lifetime with relatively lesser



(a)



(b)

Figure 2.3: Lifetime vs. BER (a) and delay constraints (b) for Scenario 1 with and without data compression, when $T = 50$ ms for (a) and $\phi = 10^{-3}$ for (b).

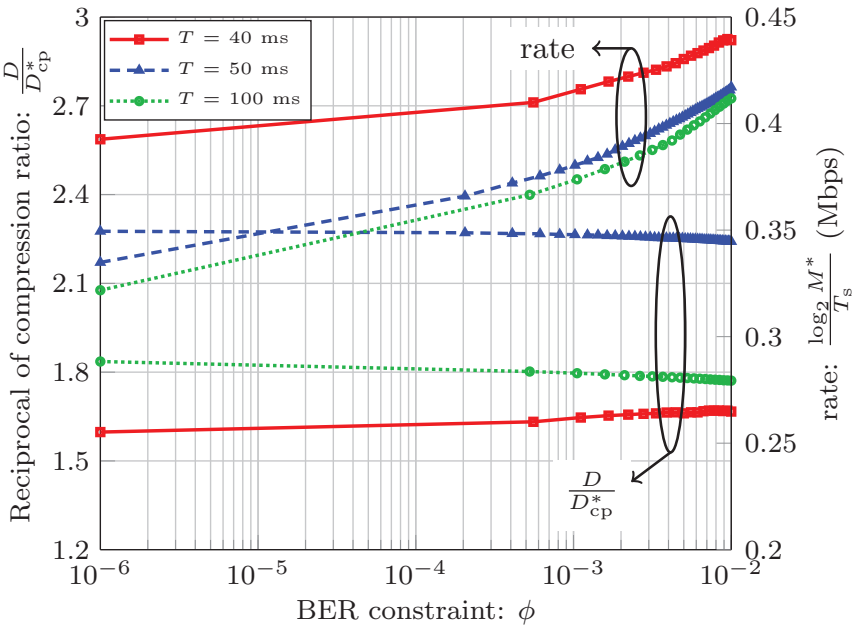
magnitude. The performance gains are huge ($\backsim$ 2000%) when the delay bound is stringent ($\backsim$ $20ms$). This is because, if the delay bound is stringent then a very high transmission rate is required to transmit the given data. Thus, the per bit transmission energy cost is very large for a stringent delay bound. However, in the proposed schemes, an optimal level of compression is applied to the data and the energy of applying (optimal level of) compression is much less than transmitting data at a very high transmission rate, i.e., per bit compression energy cost is very low. Therefore, with the help of data compression the amount of data is reduced and in addition the compressed data is transmitted at a comparatively lower transmission rate. As a result large gains are possible when optimal level of compression and transmission rate policy is employed.

From Fig. 2.3a and 2.3b, we can say that joint optimization is much better than no compression under any BER and delay constraints. In addition, the performance gain observed ranges from 90% to 2000% and is most profound when the delay constraint is stringent, which demonstrates the suitability of applying the proposed scheme in the low latency regime.

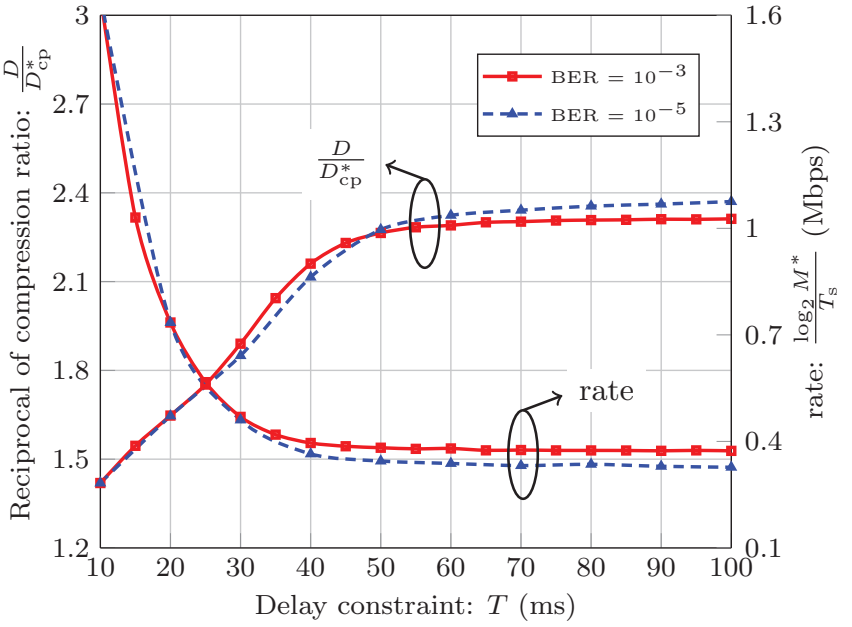
2.4.2 Impact of BER and Delay Constraints

Fig. 2.4a plots the reciprocal of the compression ratio and the transmission rate vs. the BER constraint for different delay constraint values, for Scenario 1. *Note, high compression ratio implies less compression is applied, thus we plot its reciprocal for better clarity.* The performance observed in Fig. 2.3a is explained in terms of design parameters in Fig. 2.4a. We can see that as the BER constraint becomes less stringent, the transmission rate increases, especially in the range 10^{-3} to 10^{-2}, whereas the level of compression remains almost constant. Hence, the optimal level of compression is insensitive to the change in the BER requirement. However, the optimal transmission rate increases as BER constraint is relaxed. *Thus, the best choice is to keep a constant optimal level of compression and adapt transmission rate as per the BER requirement.*

Fig. 2.4b plots the reciprocal of the compression ratio and the transmission rate vs. the delay constraint for different BER constraint values, for Scenario 1. The performance observed in Fig. 2.3b is explained in terms of design parameters



(a)



(b)

Figure 2.4: Optimal compression ratio and transmission rate vs. BER (a) and delay constraints (b) for Scenario 1, when $T = \{40, 50, 100\}$ ms for (a) and $\phi = \{10^{-3}, 10^{-5}\}$ for (b).

in Fig. 2.4b. We can see that as the delay constraint becomes less stringent, the transmission rate significantly increases and the level compression also significantly increases until the upper bound $\mathcal{Q}$ in (2.26) is reached. Beyond that, both design parameters remains almost constant.

Overall, it is best to reduce compression and increase the transmission rate when the delay constraint gets stringent and vice versa. The optimal design of compression and transmission rate is more sensitive to the delay constraint when the system requires low latency.

2.4.3 Effect of Number of Feedback Bits

Fig. 2.5 plots the lifetime versus the BER constraint with perfect feedback for Scenario 1 and quantized feedback for Scenario 2. The gap between the performance of Scenario 1 (perfect CGI availability at the sensor node) and Scenario 2 (quantized CGI availability at the sensor node) is small for practical number of feedback bits. It can be observed that a relatively small number of feedback bits are enough, e.g., 6 bits achieve within 0.6% optimal performance.

2.4.4 Impact of BER and Delay Constraints

Fig. 2.6 plots the lifetime versus the BER constraint for different level of probabilistic BER performance requirement, ϑ, for Scenario 3. The lifetime is impacted by the BER constraint in a similar manner as compared to Scenario 1 shown in Fig. 2.3a. The lifetime non-linearly decreases with increase in the BER performance requirement.

Fig. 2.7a shows that for a given BER constraint, ϕ, as ϑ increases both the transmission rate and the level of compression decrease. However, in the case of stringent BER requirement, i.e., $\vartheta = 0.99$, both the level of compression and transmission rate remains almost constant for different values of ϕ. For a given value of ϑ, the level of compression increases with ϕ, unlike the trend observed in Fig. 2.4a for Scenario 1.

Fig. 2.7b shows that for a given delay constraint, as ϑ increases, the transmission rate decreases. The level of compression displays different trend as compared to Scenario 1 in Fig. 2.3b. This is because the value of the upper bound $\mathcal{Q}$ in

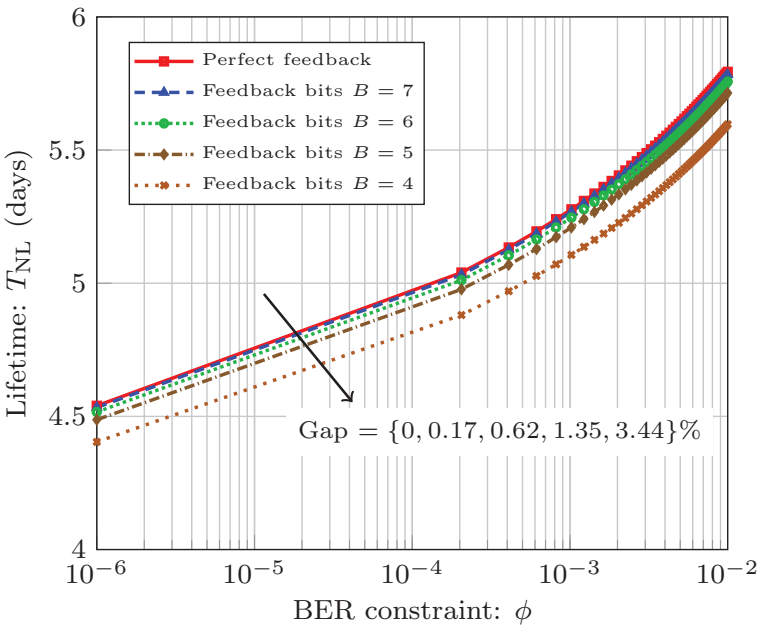


Figure 2.5: Lifetime vs. BER constraint for Scenario 2 for different level of channel uncertainty, when $T = 50$ ms.

(2.39) increases as ϑ increases. Thus, the level of compression increase until this upper bound is reached and afterwards it remains almost constant.

2.5 Summary

In this chapter, we investigated the joint optimization of compression and transmission strategy for an energy-constrained sensor node, and illustrated their tradeoff. We showed that the joint optimization performs much better than only optimizing transmission without compression under any BER and delay constraints. The performance gain observed ranges from 90% to 2000% and is most profound when the delay constraint is stringent. Overall, it is best to reduce compression and increase the transmission rate when the delay constraint gets more stringent and vice versa. The optimal level of compression is insensitive to the change in the BER requirement. However, the optimal transmission rate increases as BER constraint is relaxed. Hence, the best choice is to keep a constant optimal level of compression and adapt transmission rate as per the BER requirement. The optimal level

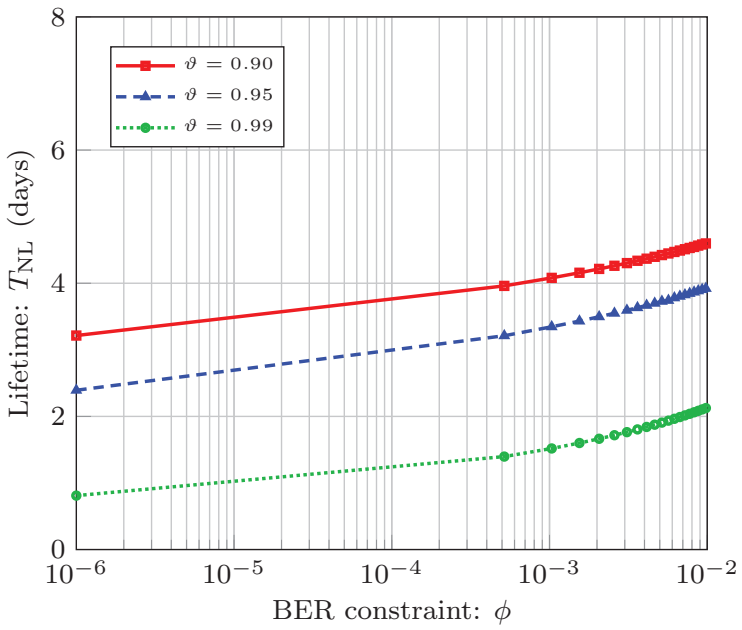
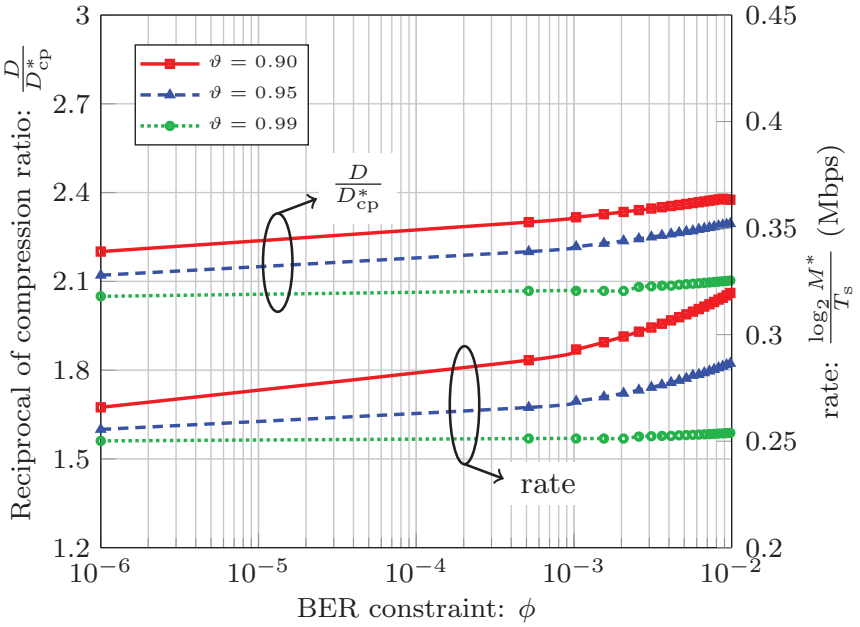
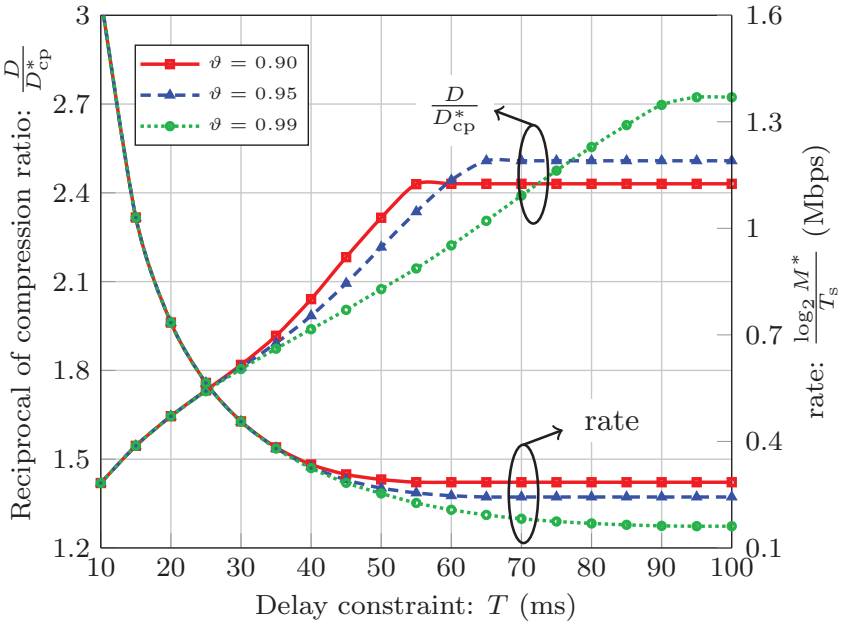


Figure 2.6: Lifetime vs. BER constraint for Scenario 3 for different probabilistic BER performance requirement, when $T = 50$ ms and $\phi = 10^{-3}$.

of compression has an inverse relationship with severity of the BER requirement when the delay constraint is stringent and vice versa. In this chapter, we have assumed that new data is generated once the previous one is processed. In some event monitoring applications, the sensed data can arrive before the previous data is completely transmitted. Future work can consider transmission and compression design for such a scenario.



(a)



(b)

Figure 2.7: Optimal compression and transmission rate vs. BER (a) and delay constraints (b) for Scenario 3 for different probabilistic BER performance requirement, when $T = 50$ ms for (a) and $\phi = 10^{-3}$ for (b).

Chapter 3

Single-User Compression and Transmission Rate Control for Wireless Powered Machine-Type Communication

In this chapter, we consider a network scenario that is evolved from the network scenario presented in chapter 2 and employs an important aspect of energy-harvesting for a multi-user multi-channel sparse machine-type communication (MTC) system. In such a system, each machine-type device (MTD) is battery powered which is recharged by an energy-harvesting source for perpetual operation. Each MTD is assigned a channel and transmit data on dedicated. The data may be generated or sensed by the MTD periodically, due to the occurrence of an event of interest, or on demand (only when requested by the receiver) that is to be delivered under given reliability and delay conditions. Moreover, the MTDs perform duty cycling as well as deep sleep operation. This network scenario cover loss tolerant and high reliability demanding MTC systems. For this network scenario, designing a generic single-user communication policy would suffice and the same policy can be employed by all other users for the individual system parameters.

The limited operational time of battery operated wireless MTC nodes is a major hurdle in fully realizing the potential of the Internet of Things (IoT) [3]. Radio

frequency based energy harvesting (RF-EH) has recently emerged as a promising solution to provide perpetual-lifetime for MTC nodes [32–34], i.e., it may stay operational indefinitely long. In this chapter, we consider a unified wireless power transfer (WPT) and wireless information transfer (WIT) system.

In this chapter, we consider a MTC node that is served by a hybrid access point (HAP) which provides RF power transfer to the node and receives data transmission from the node. Due to the lossy wireless medium and limited efficiency of RF energy transducer, the energy cost at the HAP is substantial. To minimize the energy cost while still satisfying the system requirement, the harvested energy at the MTC node must be used efficiently. To this end, we consider that the MTC node employs data compression in order to reduce the energy cost of data transmission. Data compression itself consumes time and energy, which needs to be carefully controlled. We seek to intelligently design the operation of a MTC node and a HAP with an objective to minimize the power transferred by the HAP. Thus, we propose to jointly optimize the harvesting-time, compression and transmission design, to minimize the energy cost of the system under given delay constraint.

This chapter is organized as follows. The system model is presented in Section 3.1. The system energy minimization problem composed of both transmitter and receiver side designs is formulated and solved in Section 3.2. Numerical results are presented in Section 3.3. Finally, Section 3.4 concludes the chapter.

3.1 System Model

We consider a wireless powered communication system employing a *harvest-and-use* strategy, i.e., the HAP periodically transmits a power beacon and a MTC node harvests RF energy and transmits its sensed data to the HAP, as illustrated in Fig. 3.1. For a multi-user multi-channel sparse MTC system designing a generic single-user communication policy would suffice and the same policy can be employed by all other users for the individual system parameters. The system follows a block-wise operation with a block of duration T seconds. Within each time block the MTC node performs four main functions, i.e., (i) energy harvesting, (ii) sensing, (iii) compression and (iv) transmission, as shown in Fig. 3.2. The energy spent by the MTC node to perform all its operations in a given time block is denoted by

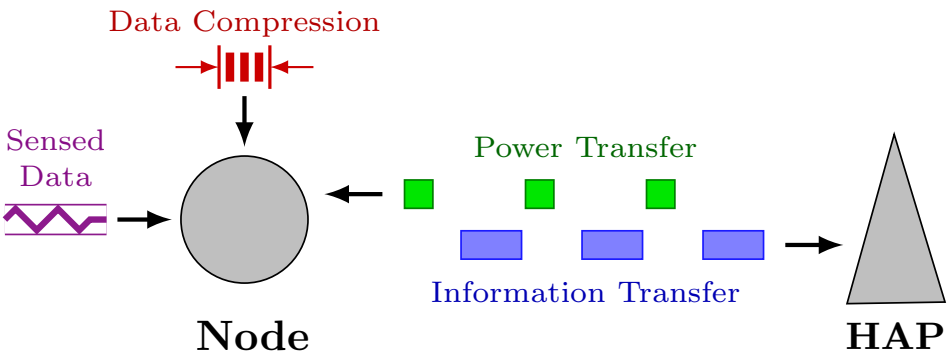


Figure 3.1: Illustration of the considered WPCN.

E_{MTC}.

Channel model: The MTC node is equipped with an omnidirectional antenna and the HAP is equipped with a directional antenna with gain G. The MTC node and the HAP are located at a distance r. We assume channel reciprocity between WPT and WIT channels [97]. The channel is composed of large scale path loss, with path loss exponent α, and small-scale quasi-static flat Rayleigh fading channel, i.e., the fading channel gain $|h|^2$ remains constant in a given time block and is independently and identically distributed from one time block to the next [97, 98]. We assume the instantaneous channel gain information (CGI) is available only at the HAP [97]. We consider the additive white Gaussian noise for the channel with zero mean and variance σ^2. The noise power density is denoted by N_0.

Energy Harvesting: The HAP is powered by an electrical grid, whereas the MTC node has simple short-term energy storage capability [97, 98]. Hence, we assume that energy accumulation is not possible at the MTC node. During the first T_{H} seconds of a time block, the HAP transmits the power beacon with a fixed transmit power level denoted by P_{PT}, and the MTC node harvests E_{H} units of energy under fading channel gain $|h|^2$, given as

$$E_{\mathrm{H}} = \eta T_{\mathrm{H}} \Big(\kappa \frac{|h|^2 P_{\mathrm{PT}}}{r^{\alpha}} - P_{\mathrm{th}} \Big)^{+}, \tag{3.1}$$

where $\kappa = \left(\frac{\nu}{4\pi}\right)^2 G$ is the attenuation factor and ν is the wavelength, $0 \leqslant \eta \leqslant 1$ is the loss in energy transducer during conversion of harvested energy into electrical

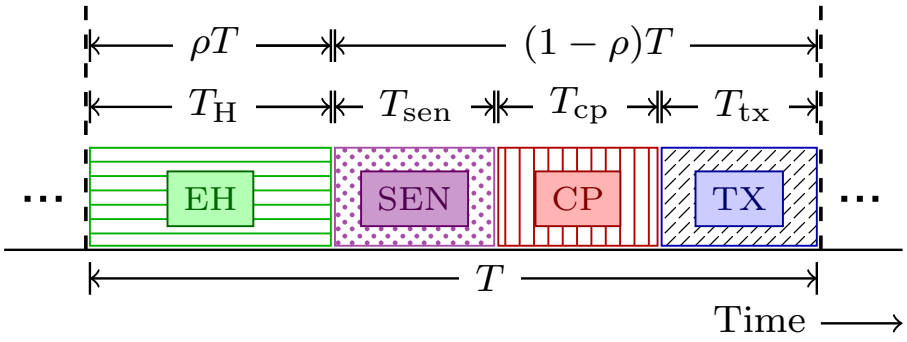


Figure 3.2: Timing diagram for energy harvesting (EH), sensing (SEN), compression (CP) and transmission (TX) processes.

energy, P_{th} is the RF-EH sensitivity threshold, and $(z)^+ = \max\{z, 0\}$. Accordingly, the RF-EH circuit can only be activated if the received signal power is greater than a RF-EH sensitivity threshold [98]. The value of the thermal noise power, σ^2, is typically very small, and thus it is not included in the energy harvesting model.

Sensing: We assume the MTC node spends a constant amount of time and energy to acquire a fixed amount of data in each time block [89]. Note that this sensing cost model is generic and may represent any other operation or can even be ignored. Let the time, energy, power spent to sense and generate a fixed data of size D bits is denoted by T_{sen}, E_{sen} and P_{sen}, respectively. E_{sen} is given as

$$E_{\mathrm{sen}} = \vartheta D P_{\mathrm{sen}}, \tag{3.2}$$

where ϑ is the per bit sensing time.

Compression: Before transmission, the sensed data of size D bits is compressed into D_{cp} bits as per the given compression ratio $\frac{D_{\mathrm{cp}}}{D}$. In this chapter, we adopt a non-linear compression cost model given in [48] to compute the *compression time*, denoted by T_{cp}, as a function of $\frac{D_{\mathrm{cp}}}{D}$, which is given as

$$T_{\mathrm{cp}} = \tau D \Big(\frac{D}{D_{\mathrm{cp}}}\Big)^{\beta} - \tau D, \tag{3.3}$$

where τ is the per bit processing time and β is the compression algorithm dependent parameter that is proportional to the compression algorithm's complexity. β deter-

mines the time cost for achieving a given compression ratio and can be calculated off-line for any specified compression algorithm and given hardware resources. τ depends upon the micro-controller unit (MCU) processing resources and the number of program instructions executed to process 1 bit of data. Note that τ does not represent the compression time per bit. Let P_{cp} denote the power consumed by the MTC node during data compression process. P_{cp} is the same as the power consumed while MCU is processing information, which is predefined.

Transmission: Once the compression process is complete, the MTC node needs to transmit the compressed data, D_{cp}, within the same time block. Let the *transmission time* is denoted by T_{tx}. Thereby, T_{tx} is given as

$$T_{\text{tx}} = \frac{D_{\text{cp}}}{R}, \tag{3.4}$$

and the transmission rate, R, is given as

$$R = B \log_2 \left(1 + \frac{\gamma}{\Gamma}\right), \tag{3.5}$$

where B is the bandwidth of the considered system, P_{IT} is the transmit power level, and Γ characterizes the gap between the achievable rate and the channel capacity due to the use of practical modulation and coding schemes [60]. γ represents the received signal-to-noise ratio (SNR) which is given by [82] as follows

$$\gamma = \frac{\kappa |h|^2 P_{\text{IT}}}{\sigma^2 d^{\alpha}}. \tag{3.6}$$

To compute the data transmission power cost, denoted by P_{tx}, we adopt a practical model as given in [60]. The transmission power cost is composed of two main components, the transmitted power, P_{IT}, and the static communication module circuitry power, denoted by P_{o}, which accounts for digital-to-analog converter, frequency synthesizer, mixer, transmit filter, and antenna circuits, etc. Accordingly, P_{tx} is given as follows

$$P_{\text{tx}} = \frac{P_{\text{IT}}}{\mu} + P_{\text{o}}, \tag{3.7}$$

where $\mu \in (0, 1]$ is the drain efficiency of the power amplifier.

The energy spent by the MTC node to perform all its operations in a given

time block is

$$E_{\mathrm{MTC}} = P_{\mathrm{sen}} T_{\mathrm{sen}} + P_{\mathrm{cp}} T_{\mathrm{cp}} + P_{\mathrm{tx}} T_{\mathrm{tx}}. \tag{3.8}$$

Note that it is possible to perform energy harvesting and sensing operation simultaneously. It allows more time for energy harvesting and thus increases the energy budget. In order to achieve this the rate of energy harvesting must be higher than the rate of energy expenditure on sensing (e.g., if a drone takes 1 hr to charge but its maximum flying time is 20 mins, then it implies that we cannot wirelessly charge and fly the drone at the same time) and if this condition is satisfied than simultaneous operation is possible. Although, in this setting, energy harvesting can be performed for longer time, however the harvesting energy will not necessarily be much higher as a fraction of energy is spent on the sensing operation. The gain in this setting depends upon the difference of the rates of energy harvesting and sensing energy expenditure, and the duration of these two operations. Nevertheless, it will complicate the device circuitry and its operation. The approach we employed wherein sensing is performed post energy harvesting is more generalized approach that is valid for any energy harvesting and sensing energy expenditure rates.

We assume that the instantaneous channel gain for MTC device is perfectly estimated by the HAP, which is inline with the relevant prior works [37, 66, 67, 99]. The details of the channel estimation process is outside the scope of this work. Note that the energy portion split by the device for pilot signal transmission does not affect the design of the optimal compression and transmission policies. Thus, we ignore this energy cost. However, the pilot signal transmission energy cost can straightforwardly be added in the final system energy cost.

3.2 Energy Minimization Problem

In this section, we first present the design objective and then formulate the energy minimization problem. Finally, we provide the solution to the problem.

Design Objective: The main problem that we study is to minimize the energy cost of the system. The target is to optimally utilize the power injected into the system. The MTC node solely relies on the power transferred by the HAP. Thus, the aforementioned objective can be achieved by minimizing the HAP's transmitted

energy, $E_{\mathrm{HAP}} = T_{\mathrm{H}} P_{\mathrm{PT}}$, while satisfying the given system constraints. Therefore, we design policies for the RF energy harvesting, compression and transmission processes whilst guaranteeing the desired quality-of-service (QoS) requirements.

We consider a *harvest-and-use* strategy for energy harvesting and sensed data transmission. Therein, a fraction of the total completion time is spent for energy harvesting operation, which is controlled by the *harvesting-time ratio* denoted by $0 \leqslant \rho \leqslant 1$, such that $T_{\mathrm{H}} = \rho T$. Accordingly, the rest of the time is spent in sensing, compression and transmission processes $T_{\mathrm{sen}} + T_{\mathrm{cp}} + T_{\mathrm{tx}} = (1-\rho)T$, as illustrated in Fig. 3.2. Moreover, the system requires specific QoS in terms of data reliability and delivery time. This imposes a constraint in terms of the delay bound, which mandates that the cumulative completion time of the MTC node's operations, must meet the deadline, T. Moreover, the system imposes another probabilistic constraint on the successful data delivery. Specifically, the data should be successfully delivered with probability equal to a given value, denoted by δ.

The channel varies independently in different time blocks. In a given time block, if the channel is in deep fade, the required level of energy to transfer, E_{HAP}, would be too high to achieve the given system performance. This high energy cost might be infeasible for a loss tolerant system, a typical case in sensor networks. Moreover, when the channel is in deep fade, the delay bound may also hinder the HAP to transfer the required amount of energy and the MTC node may not be able to complete all of its operations within the given deadline. Also, when the channel is in deep fade, the MTC node would not be able to harvest enough energy to perform its operations within the deadline, referred to as the *energy outage*.

The probability distribution function of the instantaneous channel gain, $|h|^2$, is exponentially distributed and is given as

$$f\big(|h|^2\big) \triangleq \frac{1}{\lambda} \exp\Big(- |h|^2 \frac{1}{\lambda}\Big), \quad |h|^2 \geqslant 0, \tag{3.9}$$

where λ is the scale parameter of the probability distribution.

To minimize HAP's transmitted energy, E_{HAP}, under given data delivery constraint, the best strategy is to exploit the channel diversity in different time blocks. Therein, the system is operated only when the channel gain is larger than a specific value, denoted by Θ, which satisfies the probabilistic data delivery constraint.

Accordingly, the value of Θ can be computed using the following expression

$$\mathbb{P}\Big\{|h|^2 \geqslant \Theta\Big\} = \delta. \tag{3.10}$$

The left hand side of the above equation represents the complimentary cumulative distribution function (CCDF) for $|h|^2$. For the considered Rayleigh fading channel, the fading power gain, $|h|^2$, is exponentially distributed, which yields

$$1 - \Big[1 - \exp\Big(-\Theta\frac{1}{\lambda}\Big)\Big] = \delta. \tag{3.11}$$

Solving the above equation for Θ yields

$$\Theta = -\lambda \ln\big(\delta\big).$$

Hence, in a given time block if channel gain is lower than the threshold value, i.e., $|h|^2 < -\lambda \ln(\delta)$, then HAP simply informs the MTC not to perform any of its operations and, for simplicity, we ignore the energy consumed by this action from the HAP.

3.2.1 Optimal Design Policies for the MTC Node

When the channel condition is good enough to preclude the energy outage, i.e., $|h|^2 \geqslant -\lambda \ln(\delta)$, the HAP transmits the power beacon and the MTC node performs all of its operations. This is referred to as an active time block.

Considering the operations at the MTC node, we optimize the harvesting-time, ρ_{MTC}, compressed data size, D_{cp}, and transmitted power, P_{IT}, for a wireless powered MTC node to minimize HAP's transmitted energy, E_{HAP}, under given system constraints. We devise the optimal design considering the worst-case channel condition that is possible in an active time block, i.e., $|h|^2 = -\lambda \ln(\delta)$. It is because the instantaneous CGI is not available at the MTC node, and therefore MTC node cannot adapt design parameters to different channel conditions in different time blocks. The worst-case channel based design ensures that the required system performance would be met in all active time blocks using the fixed design parameters.

Minimizing ρ_{MTC} actually minimizes E_{HAP}, because T and P_{PT} are both fixed

parameters. Thus, solving the following problem, for given channel realization $|h|^2 = -\lambda \ln(\delta)$, yields optimal MTC node design parameters

$$\begin{aligned}
\underset{\rho_{\mathrm{MTC}}, P_{\mathrm{IT}}, D_{\mathrm{cp}}}{\text{minimize}} \quad & \rho_{\mathrm{MTC}} \\
\text{subject to} \quad & \vartheta D P_{\mathrm{sen}} + \tau D P_{\mathrm{cp}} \Big(\frac{D}{D_{\mathrm{cp}}}\Big)^{\beta} - \tau D P_{\mathrm{cp}} + \frac{D_{\mathrm{cp}}}{B \log_2\big(1 - \frac{\kappa \lambda \ln(\delta) P_{\mathrm{IT}}}{\sigma^2 d^{\alpha} \Gamma}\big)} \Big(\frac{P_{\mathrm{IT}}}{\mu} + P_{\mathrm{o}}\Big) \\
& \qquad \leqslant -\eta \rho_{\mathrm{MTC}} T \Big(\frac{\kappa \lambda \ln(\delta) P_{\mathrm{PT}}}{r^{\alpha}} + P_{\mathrm{th}}\Big), \\
& \vartheta D + \tau D \Big(\frac{D}{D_{\mathrm{cp}}}\Big)^{\beta} - \tau D + \frac{D_{\mathrm{cp}}}{B \log_2\big(1 - \frac{\kappa \lambda \ln(\delta) P_{\mathrm{IT}}}{\sigma^2 d^{\alpha} \Gamma}\big)} \leqslant \big(1 - \rho_{\mathrm{MTC}}\big) T, \\
& 0 \leqslant \rho_{\mathrm{MTC}} \leqslant 1, \\
& P_{\mathrm{IT}} \geqslant 0, \\
& D_{\min} \leqslant D_{\mathrm{cp}} \leqslant D.
\end{aligned} \tag{3.12}$$

where $D_{\min}$ is the lower bound on the compressed data size and its value depends on the nature of the data and the system application. The first constraint in (3.12) mandates that the harvested energy by MTC node must be large enough to cover the cumulative energy cost of its operations. The second constraint in (3.12) mandates that the completion time for all the MTC nodes' operations must meet the delay bound. The remaining constraints reflect practical range of values for harvesting-time ratio, ρ_{MTC}, MTC node's transmitted power, P_{IT}, and the compressed data size, D_{cp}.

The solution to (3.12) yields optimal MTC node design parameters $\rho^*_{\mathrm{MTC}}, P^*_{\mathrm{IT}}, D^*_{\mathrm{cp}}$ to be used by the MTC node in all active time blocks. Substituting the values of P^*_{IT} and D^*_{cp} in (3.8), yields the fixed optimal minimal $E^*_{\mathrm{MTC}}\big(P^*_{\mathrm{IT}}, D^*_{\mathrm{cp}}\big)$ value which is used in all active time blocks.

It can be shown that the problem defined in (3.12) is non-convex, because the first constraint is non-convex in P_{IT}. By substitution of variable $\ln\big(1 - \frac{\kappa \lambda \ln(\delta) P_{\mathrm{IT}}}{\sigma^2 r^{\alpha} \Gamma}\big) = z$ in (3.8), E_{MTC} can equivalently be defined as

$$E_{\mathrm{MTC}}\big(z, D_{\mathrm{cp}}\big) = \vartheta D P_{\mathrm{sen}} + \tau D P_{\mathrm{cp}} \Big(\frac{D}{D_{\mathrm{cp}}}\Big)^{\beta} - \tau D P_{\mathrm{cp}} + \frac{D_{\mathrm{cp}} b}{z} \big(\exp(z) + c\big). \tag{3.13}$$

where $a = -\frac{\kappa \lambda \ln(\delta)}{\sigma^2 r^{\alpha} \Gamma}$, $b = \frac{\ln(2)}{\mu a B}$, $c = \mu a P_{\mathrm{o}} - 1$.

Accordingly, the problem defined in (3.12) can equivalently be given as follows

$$\begin{aligned}
\underset{\rho_{\mathrm{MTC}},\, z,\, D_{\mathrm{cp}}}{\text{minimize}} \quad & \rho_{\mathrm{MTC}} \\
\text{subject to} \quad & \vartheta D P_{\mathrm{sen}} + \tau D P_{\mathrm{cp}} \Big(\frac{D}{D_{\mathrm{cp}}}\Big)^{\beta} - \tau D P_{\mathrm{cp}} + \frac{D_{\mathrm{cp}} b}{z}\big(\exp(z) + c\big) \\
& \qquad\qquad \leqslant -\eta \rho_{\mathrm{MTC}} T \Big(\frac{\kappa \lambda \ln(\delta) P_{\mathrm{PT}}}{r^{\alpha}} + P_{\mathrm{th}}\Big), \\
& \vartheta D + \tau D \Big(\frac{D}{D_{\mathrm{cp}}}\Big)^{\beta} - \tau D + \frac{D_{\mathrm{cp}} \ln(2)}{B z} \leqslant \big(1 - \rho_{\mathrm{MTC}}\big) T, \\
& 0 \leqslant \rho_{\mathrm{MTC}} \leqslant 1, \\
& z \geqslant \frac{2}{\ln(2)}, \\
& D_{\min} \leqslant D_{\mathrm{cp}} \leqslant D.
\end{aligned} \tag{3.14}$$

For brevity we omit the proof, however using basic calculus and with some algebraic manipulation, it can be shown that the problem in (3.14) is a convex optimization problem.

The solution to the optimization problem defined in (3.14) is given by the following theorem.

Theorem 3. *In solving the optimization problem in* (3.12)*, the optimal* MTC *node's transmitted power is given by*

$$P_{\mathrm{IT}}^{*} = \begin{cases} \widetilde{P}_{\mathrm{IT}}, & \text{if } \mathcal{Q}(\widetilde{P}_{\mathrm{IT}}, \widetilde{D}_{\mathrm{cp}}) < (1 - \widetilde{\rho}_{\mathrm{MTC}}) T. \\ \widehat{P}_{\mathrm{IT}}, & \text{otherwise.} \end{cases} \tag{3.15}$$

where $\widetilde{P}_{\mathrm{IT}}$ *is given as*

$$\widetilde{P}_{\mathrm{IT}} = \frac{1}{a} \exp\Big(\mathrm{W}_0\Big(\exp\big(\ln(c) - 1\big)\Big) + 1\Big) - \frac{1}{a}, \tag{3.16}$$

where $\mathrm{W}_0(\cdot)$ *is the principle branch of the Lambert* W *function.* $\widehat{P}_{\mathrm{IT}}$ *is given by the solution of the following equation which can be solved numerically using Bisection method*

$$\frac{T_{\mathrm{sen}}}{P_{\mathrm{sen}}^{-1}} + \frac{\tau D^{\beta+1}}{P_{\mathrm{cp}}^{-1} \widehat{D}_{\mathrm{cp}}^{\beta}} + \frac{\widehat{D}_{\mathrm{cp}} b\big(1 + a\widehat{P}_{\mathrm{IT}} + c\big)}{\ln\big(1 + a\widehat{P}_{\mathrm{IT}}\big)} + \frac{\eta T d}{\widehat{\rho}_{\mathrm{MTC}}^{-1}} = \frac{\tau D}{P_{\mathrm{cp}}^{-1}}, \tag{3.17}$$

and the optimal compression ratio is given by

$$\frac{D^*_{\rm cp}}{D} = \begin{cases} \frac{\widetilde{D}_{\rm cp}}{D}, & \text{if } \mathcal{Q}(\widetilde{P}_{\rm IT}, \widetilde{D}_{\rm cp}) < (1-\widetilde{\rho}_{\rm MTC})T. \\ \frac{\widehat{D}_{\rm cp}}{D}, & \text{otherwise.} \end{cases} \tag{3.18}$$

and the optimal MTC *node's harvesting-time ratio is given by*

$$\rho^*_{\rm MTC} = \begin{cases} \widetilde{\rho}_{\rm MTC}, & \text{if } \mathcal{Q}(\widetilde{P}_{\rm IT}, \widetilde{D}_{\rm cp}) < (1-\widetilde{\rho}_{\rm MTC})T. \\ \widehat{\rho}_{\rm MTC}, & \text{otherwise.} \end{cases} \tag{3.19}$$

where

$$\frac{\widetilde{D}_{\rm cp}}{D} = \left(\frac{\tau\beta P_{\rm cp}\ln(1+a\widetilde{P}_{\rm IT})}{b\big(1+a\widetilde{P}_{\rm IT}\big)+bc}\right)^{\frac{1}{\beta+1}}, \tag{3.20}$$

$$\frac{\widehat{D}_{\rm cp}}{D} = \left(\frac{\tau\beta\big(T^{-1}\widehat{\Lambda}_1^{-1}+\eta d-P_{\rm cp}\big)\ln(1+a\widehat{P}_{\rm IT})}{\frac{\widehat{\Lambda}_1^{-1}\ln(2)}{TB}+\frac{\eta d\ln(2)}{B}-b(1+a\widehat{P}_{\rm IT})-bc}\right)^{\frac{1}{\beta+1}}, \tag{3.21}$$

$$\widetilde{\rho}_{\rm MTC}=\frac{1}{\eta T d}\left(\frac{\tau D}{P_{\rm cp}^{-1}}-\frac{T_{\rm sen}}{P_{\rm sen}^{-1}}-\frac{\tau D^{\beta+1}}{P_{\rm cp}^{-1}\widetilde{D}_{\rm cp}^{\beta}}-\frac{\widetilde{D}_{\rm cp}b(1+a\widetilde{P}_{\rm IT}+c)}{\ln\big(1+a\widetilde{P}_{\rm IT}\big)}\right), \tag{3.22}$$

$$\widehat{\rho}_{\rm MTC} = \frac{1}{T}\left(T-T_{\rm sen}-\frac{\tau D^{\beta+1}}{\widehat{D}_{\rm cp}^{\beta}}+\tau D-\frac{\widehat{D}_{\rm cp}\ln(2)}{B\ln\big(1+a\widehat{P}_{\rm IT}\big)}\right). \tag{3.23}$$

$$\mathcal{Q}(\widetilde{P}_{\rm IT}, \widetilde{D}_{\rm cp}) \triangleq T_{\rm sen}+\frac{\tau D^{\beta+1}}{\widetilde{D}_{\rm cp}^{\beta}}-\tau D+\frac{B^{-1}\widetilde{D}_{\rm cp}\ln(2)}{\ln\big(1+a\widetilde{P}_{\rm IT}\big)}, \tag{3.24}$$

$$\widehat{\Lambda}_1=\frac{T^{-1}B^{-1}\ln(2)}{b\big(\ln\big(1+a\widehat{P}_{\rm IT}\big)-1\big)\big(1+a\widehat{P}_{\rm IT}\big)-bc-\eta dB^{-1}\ln(2)}. \tag{3.25}$$

Proof. The proof is provided in Appendix D.1. ■

Following remark discusses an insight from Theorem 3.

Remark 9. *$\widetilde{P}_{\rm IT}$, $\widetilde{D}_{\rm cp}$, $\widetilde{\rho}_{\rm MTC}$ provide an upper bound on the HAP's energy cost in any active time block, $\rho^*_{\rm MTC}TP_{\rm PT}$, and are optimal design parameters when all constraints in* (3.12) *are slack, i.e., $\mathcal{Q}(\widetilde{P}_{\rm IT}, \widetilde{D}_{\rm cp}) < (1-\widetilde{\rho}_{\rm MTC})T$, except the first constraint. On the other hand, $\widehat{P}_{\rm IT}$, $\widehat{D}_{\rm cp}$, $\widehat{\rho}_{\rm MTC}$ are optimal design parameters, when all constraints in* (3.12) *are slack except for the first and second constraint.*

3.2.2 Optimal Design Policies for the HAP

Finally, we can relate the design for the MTC node, given in Theorem 3, to determine the HAP's optimal design. The MTC node employs a fixed harvesting-time ratio ρ^*_{MTC}, defined in (3.19), in all active time blocks. ρ^*_{MTC} is determined based on the worst-case channel condition due to the unavailability of instantaneous CGI at the MTC node. This design yields the upper bound on the HAP's energy cost in any active time block, i.e., $E^{\mathrm{ub}}_{\mathrm{HAP}} = \rho^*_{\mathrm{MTC}} T P_{\mathrm{PT}}$. However, the instantaneous CGI is available at the HAP, thus it can adapt the harvesting-time for different channel conditions in different active time blocks. Accordingly, the actual harvesting-time used by the HAP to transfer power is given by $\rho^*_{\mathrm{HAP}} T$, where $\rho^*_{\mathrm{HAP}} \leqslant \rho^*_{\mathrm{MTC}}$ is defined as

$$\rho^*_{\mathrm{HAP}}\big(|h|^2\big) \triangleq \frac{E^*_{\mathrm{MTC}}\big(P^*_{\mathrm{IT}}, D^*_{\mathrm{cp}}\big)}{\eta T\big(\kappa |h|^2 r^{-\alpha} P_{\mathrm{PT}} + P_{\mathrm{th}}\big)}. \tag{3.26}$$

where $|h|^2 \geqslant -\lambda \ln\big(\delta\big)$ is the instantaneous channel gain in an active time block. Note that the actual harvesting-time, ρ^*_{HAP}, is almost always less than ρ^*_{MTC}.

Lastly, the minimal HAP's transmitted energy is given as

$$E^*_{\mathrm{HAP}} = \rho^*_{\mathrm{HAP}} T P_{\mathrm{PT}}. \tag{3.27}$$

The system model considers that the MTC node lacks the ability to store energy for later blocks which makes it more cost efficient and helps keep the size of the device small. To minimize MTC node's and HAP energy cost, the HAP only transfers the amount energy that is the minimum energy required by the MTC node to perform its operations. Accordingly, there will be no energy left at the MTC node at the end of the block to accumulate in the later blocks. For this reason, there is no need to enable the MTC node to store energy for next blocks.

3.3 Numerical Results

In this section, we present the numerical results to observe the performance of the proposed scheme. Unless specified otherwise, the values adopted for the system parameters are shown in Table 3.1. Note that we assume realistic values, consistent with prior works, for EH [63] and sensing, compression, transmission operations.

Table 3.1: System parameter values.

Name	Symbol	Value
RF-EH sensitivity threshold	P_{th}	-30 dBm
Energy transducer efficiency	η	0.5
Pathloss exponent	α	3
Sensing power cost	P_{sen}	70 mW
Power amplifier efficiency	μ	0.35
Antenna gain	G	10
Compression power cost	P_{cp}	24 mW
Compression cost parameter	β	5
Bandwidth	B	2 MHz
HAP transmitted power	P_{PT}	40 dBm
Per bit processing time	τ	3.9 ns/b
Noise power density	N_0	-174 dBm
Radio circuitry power	P_{o}	82.5 mW
CGI's *pdf* scale parameter	λ	1
Data size	D	200 bits
Data delivery probability	δ	0.80
Per bit sensing time	ϑ	37 ns/b
Distance	r	10 m
Min. compressed data size	$D_{\min}$	$0.5\times D$
Modulation gap parameter	Γ	0 dB
Wavelength	ν	0.125 m

Baseline scheme: To illustrate the advantage of joint optimization of harvesting-time ratio, compression and transmission rate, we also consider a baseline scheme which minimizes E_{HAP} by optimizing harvesting-time ratio and transmission rate only (no data compression is employed) whilst satisfying the system constraints. This problem can be given by substituting $D_{\text{cp}} = D$ in (3.12). The optimal MTC node's transmitted power, denoted by $P^*_{\text{IT,nc}}$, and harvesting-time ratio, denoted by $\rho^*_{\text{MTC,nc}}$, for the baseline scheme can be obtained using Proposition 1.

Proposition 4. *The optimal* MTC *node's transmitted power, without employing compression, is given by*

$$P^*_{\text{IT,nc}} = \begin{cases} \widetilde{P}_{\text{IT,nc}}, & \text{if } T_{\text{sen}} + \frac{B^{-1} D \ln(2)}{\ln\left(1+a\widetilde{P}_{\text{IT}}\right)} < \frac{1-\widetilde{\rho}_{\text{MTC,nc}}}{T^{-1}}. \\ \widehat{P}_{\text{IT,nc}}, & \text{otherwise.} \end{cases} \tag{3.28}$$

where

$$\widetilde{P}_{\mathrm{IT,nc}} = \frac{1}{a}\exp\Big(\mathrm{W}_0\Big(\exp\big(\ln(c)-1\big)\Big)+1\Big)-\frac{1}{a}, \tag{3.29}$$

$$\widehat{P}_{\mathrm{IT,nc}} = \frac{1}{a}\exp\left(\frac{u}{v}-\mathrm{W}_0\Big(-\frac{bD}{\eta dv}\exp\Big(\frac{u}{v}\Big)\Big)\right)-\frac{1}{a}, \tag{3.30}$$

where $u = \frac{cbD}{\eta d} - \frac{D\ln(2)}{B}$ *and* $v = T_{\mathrm{sen}} - \frac{T_{\mathrm{sen}}P_{\mathrm{sen}}}{\eta d} - T$. *The optimal harvesting-time ratio is given by*

$$\rho^*_{\mathrm{MTC,nc}} = \begin{cases} \widetilde{\rho}_{\mathrm{MTC,nc}}, & \text{if } T_{\mathrm{sen}}+\frac{B^{-1}D\ln(2)}{\ln\left(1+a\widetilde{P}_{\mathrm{IT}}\right)} < \frac{1-\widetilde{\rho}_{\mathrm{MTC,nc}}}{T^{-1}}. \\ \widehat{\rho}_{\mathrm{MTC,nc}}, & \text{otherwise.} \end{cases} \tag{3.31}$$

where

$$\widetilde{\rho}_{\mathrm{MTC,nc}} = -\frac{1}{\eta T d}\left(T_{\mathrm{sen}}P_{\mathrm{sen}} + \frac{Db(1+a\widetilde{P}_{\mathrm{IT,nc}}+c)}{\ln\left(1+a\widetilde{P}_{\mathrm{IT,nc}}\right)}\right), \tag{3.32}$$

$$\widehat{\rho}_{\mathrm{MTC,nc}} = \frac{1}{T}\left(T - T_{\mathrm{sen}} - \frac{D\ln(2)}{B\ln\left(1+a\widehat{P}_{\mathrm{IT,nc}}\right)}\right). \tag{3.33}$$

Proof. The proof follows similar steps as the proof of Theorem 3, substituting $D_{\mathrm{cp}} = D$. For brevity, we omit it. ■

Note that, E^*_{HAP} is now only proportional to E^*_{MTC}, since all other parameters in (3.26) and (3.27) are same for both the schemes. Thus, we focus on the MTC node's performance.

3.3.1 Advantage of Proposed Scheme

Fig. 3.3 plots the MTC node's energy cost, E_{MTC}, versus the delay constraint, T, for system parameters in Table 3.1. E_{MTC} is plotted using the optimal MTC node design parameters for the proposed scheme, when it employs data compression, and for the baseline scheme, when compression is not employed.

The MTC node's energy cost, E_{MTC}, is fixed for all active time-blocks, however the energy of the HAP is different for different active time block. Nevertheless, the HAP enjoys a constant performance gain (decrease in its energy cost) when the MTC node employs data compression. We can see that the gain compared to the baseline scheme is significant - up to 19% for the considered range of delay

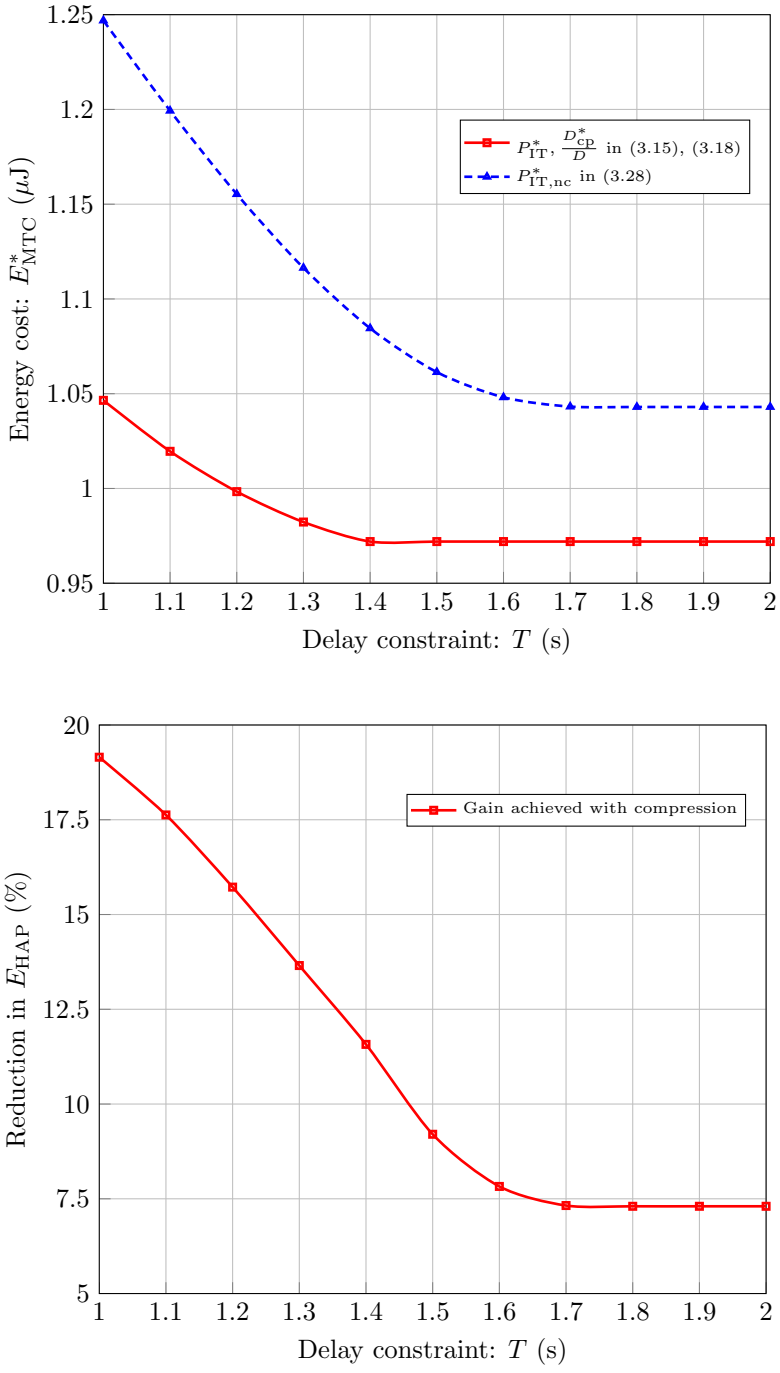


Figure 3.3: MTC node's energy cost vs. delay and performance gain achieved by HAP when MTC node employs compression.

constraint. This shows the advantage of employing compression and jointly optimizing harvesting-time ratio, compression and transmission. In both the cases, the performance gain straightens out when the upper-bound on the transmission time is achieved. This bound means that increasing the transmission time, by reducing the transmission rate further, would not decrease the energy cost of the MTC node.

3.3.2 Impact of Delay Constraint

Fig. 3.4 plots the transmission rate and compression ratio, and the MTC node's operational timing, versus the delay constraint, T, for system parameters in Table 3.1. We can see that the transmission rate requirement is significantly higher for the baseline scheme. However, in the proposed scheme the rate is almost constant and the compression ratio is adapted to the delay constant. Once the upper-bound on the transmission time is reached, both schemes do not need to change the design parameters to adapt to the delay constraint. Thanks to the data compression, this upper-bound value is relatively smaller for the proposed scheme. Thus, employing compression notably decreases the transmission rate requirement at the MTC node.

For the proposed scheme, the timing information for different operations of the MTC node in an active time block is given in Table 3.2. The percentage of time dedicated for the power transfer operation is $99\%, 93\%, 80\%$, for delay bound values $1\,\text{s}, 1.5\,\text{s}, 2\,\text{s}$, respectively. Hence, due to the inherent limited efficiency of the RF-EH process, specifically when the delay constraint is stringent, almost all the active time-block is dedicated for the power transfer operation. On the other hand, the time available for the MTC node to perform its operations is $20\,\mu\,\text{s}, 101\,\text{ms}, 391\,\text{ms}$, for delay bound values $1\,\text{s}, 1.5\,\text{s}, 2\,\text{s}$, respectively. Note that only $97.78\%, 0.014\%, 0.0037\%$, of the available time is actually spent by the MTC node, for delay bound values $1\,\text{s}, 1.5\,\text{s}, 2\,\text{s}$, respectively. Thus, the actual spent time is very small and most of the time is spent on the RF-EH process. This is because, the design considers worst-case channel due to the lack of instantaneous CGI at the MTC node. For practical parameter values, the results in Table 3.2 confirm the feasibility and provide insights into the actual operation of a RF-EH powered MTC node.

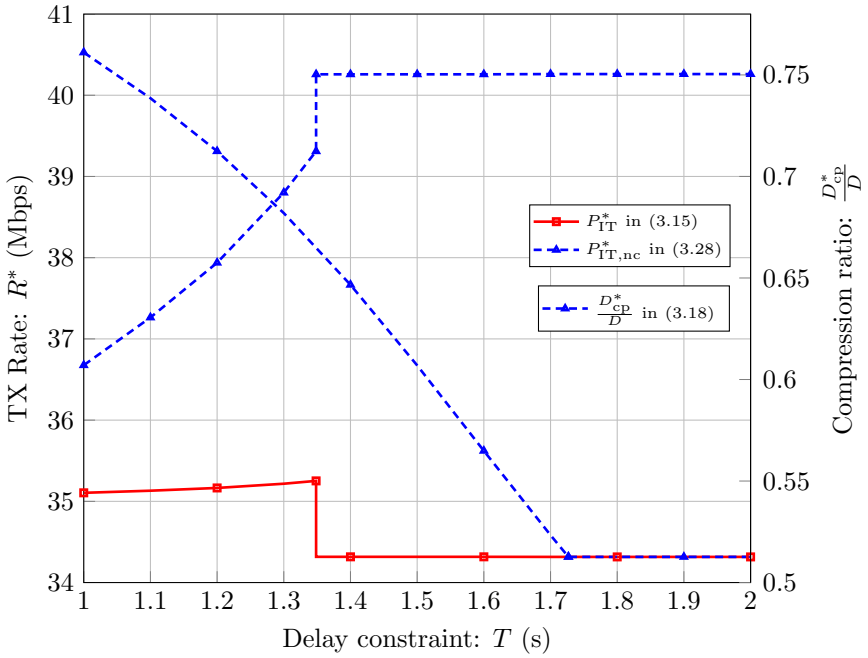


Figure 3.4: MTC node's optimal transmission rate and compression ratio, vs. delay bound.

Table 3.2: MTC node's operational timing vs. delay bound.

T (s)	$\rho^*_{\text{MTC}}T$ (s)	T_{sen} (μs)	T^*_{cp} (μs)	T^*_{tx} (μs)	$T_{\text{sen}}+T^*_{\text{cp}}+T^*_{\text{tx}}$ (μs)
1	0.999	7.400	8.696	3.458	19.555
1.5	1.399	7.400	2.504	4.373	14.277
2	1.609	7.400	2.502	4.374	14.275

3.4 Summary

We investigated the joint WPT and WIT policies, employing data compression, to minimize the energy transferred by the HAP under given system constraints. For practical parameter values, the joint optimization performs significantly better than only optimizing harvesting-time ratio and transmission rate without compression. Specifically, the gain is relatively large, up to 19%, when the delay constraint is stringent. Due to the lack of instantaneous CGI at the MTC node the design considers a worst-case channel, thus most of the total time is dedicated to RF-EH when the delay constraint is stringent.

Chapter 4

Multi-User Sequencing and Scheduling Control for Time-Division Multiple Access Systems

In this chapter, we consider a network scenario is applicable to a multi-user single-channel sparse/massive machine-type communication (MTC) system. In such a system, all machine-type devices (MTDs) share a single channel and sequentially transmit data on this channel. The data may be generated or sensed by the MTD periodically, on demand, or due to the occurrence of an event of interest. The data is to be delivered under given reliability and delay conditions. The source from which the data is acquired and the time spent on its acquisition is not relevant to the TDMA sequencing and scheduling design. Moreover, the MTDs are battery powered and perform duty cycling as well as deep sleep operation. This network scenario cover high reliability and low-latency MTC systems For this network scenario, the communication policies for all MTDs are jointly devised for given system parameters.

The Internet of Things (IoT) is largely based on the uplink communication from heterogeneous and autonomous wireless devices such as sensors, actuators which are often referred to as MTC devices [3, 5]. Due to their wireless and unattended

operation, MTC devices are mostly battery operated and thus severely energy-constrained. The energy efficient operation of these devices is therefore of pivotal importance [26]. Specifically, wireless communication is one of the most energy-intensive operations run by the MTC devices and this calls for effective wireless solutions to prolonging the device lifetime [26].

In this chapter, we propose joint sequencing and scheduling optimization for uplink MTC. We consider multiple energy-constrained MTC devices that transmit data to a base station following the time division multiple access (TDMA) protocol. Conventionally, the energy efficiency performance in TDMA is optimized through multi-user scheduling, i.e., changing the transmission block length allocated to different devices. In such a system, the sequence of devices for transmission, i.e., who transmits first and who transmits second, etc., has not been considered as it does not have any impact on the energy efficiency. In this chapter, we consider that data compression is performed before transmission and show that the multi-user sequencing is indeed important. We apply three popular energy-minimization system objectives, which differ in terms of the overall system performance and fairness among the devices. We jointly optimize both multi-user sequencing and scheduling along with the compression and transmission rate control.

This chapter is organized as follows. The system model is presented in Section 4.1. The proposed multi-user sequencing and scheduling problems for different system objectives are formulated in Section 4.2. The problem transformation and its solution strategy are presented in Section 4.3. The sub-optimal scheme is presented in Section 4.4. Numerical results are presented in Section 4.5. Finally, Section 4.6 concludes the chapter.

4.1 System Model

We consider a system consisting of multiple MTC devices transmitting data packets to a BS, as shown in Fig. 4.1. The devices are battery-operated and energy-constrained, whereas the BS has no energy constraint. Each device has a data packet of a specific length and the data packets of all devices need to be transmitted within a frame of length T_{frame} seconds. The devices employ the TDMA channel access mechanism for data transmission, as shown in Fig. 4.2. We assume perfect

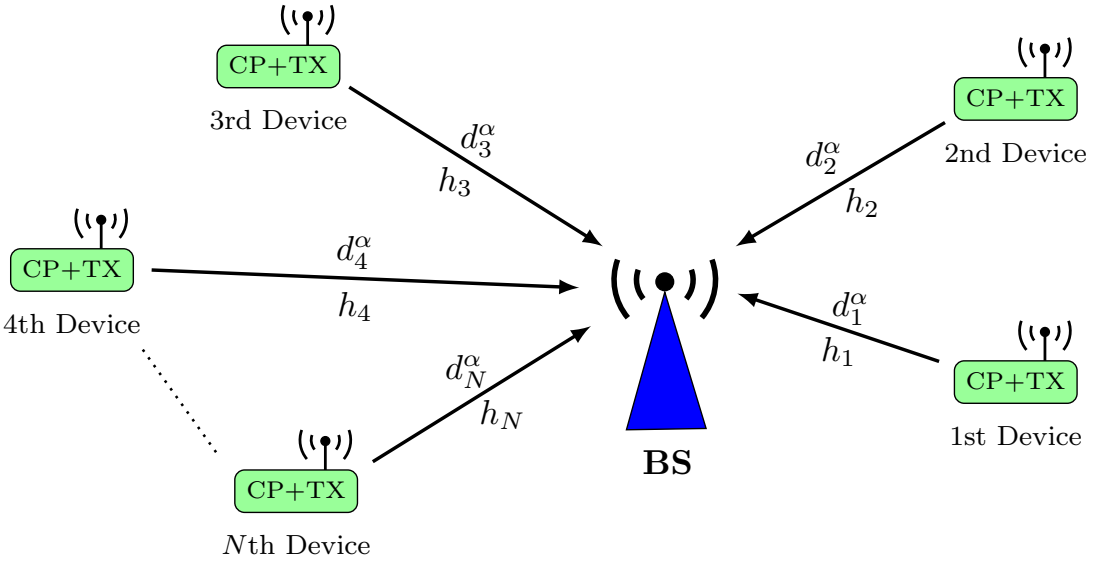


Figure 4.1: Illustration of the considered system model comprising multiple MTC devices and a base station. (CP = data compression, TX = data transmission)

synchronization among devices, which is in line with recent works [37, 66–68, 88]. The devices that have data to transmit, contend for the channel by transmitting a control packet, and the BS then grants channel access to N devices, i.e., allocates transmission blocks to N devices. As the details of channel contention mechanism are outside the scope of this paper, interested readers are referred to [100] for more information.

The BS determines the TDMA sequence and schedule and allocates the non-overlapping frame segments (referred to as the transmission blocks) to individual devices. Each device is allocated a single transmission block. Both the sequence and schedule of the transmission blocks are shared with the devices by the BS before the start of the frame. Each device applies data compression before the start of its scheduled transmission block and then transmits the compressed data in the allocated transmission block, as shown in Fig. 4.2. The device allocated with the first transmission block in the frame performs both the data compression and transmission operations within its allocated transmission block. Note that the transmission block length can be different for different devices. Moreover, a device may not necessarily use all of its allocated time for compression and/or

transmission.

For an energy-efficient operation, the devices allocated with transmission blocks are kept in power saving state when they are neither compressing nor transmitting data. We assume that the power consumed by the device in power saving state is negligible [46, 47].

Channel model: The BS and all the devices are equipped with an omnidirectional antenna. The devices are located at arbitrary distances from the BS. The distance between the ith device and the BS is d_i meters. The channel between each device and the BS is composed of a large-scale path loss, with path loss exponent α, and a small-scale quasi-static frequency-flat Rayleigh fading channel. The fading channel coefficient for the ith device is denoted as h_i. We assume that the channel response/coefficient remains unchanged over one frame. For static network devices, as in our case, the coherence time is of the orders of 100 ms for sub-GHz carrier frequencies (typically < 6 GHz) and the channel can be predicted with high accuracy [101]. The noise is assumed to be additive white Gaussian noise (AWGN) with zero mean and variance σ^2. The noise spectral density is given by N_0. The probability distribution function (*pdf*) of the instantaneous channel gain, $|h_i|^2$, is exponentially distributed as

$$f\big(|h_i|^2\big) \triangleq \frac{1}{\varsigma} \exp\Big(- \frac{|h_i|^2}{\varsigma}\Big), \quad |h_i|^2 \geqslant 0, \quad \forall\, i \in \{1, 2, \cdots, N\}, \tag{4.1}$$

where ς is the scale parameter for the *pdf*. We assume that the instantaneous channel gain for each device is perfectly estimated by the BS, which is inline with the relevant prior works [37, 66, 67, 99]. Note that the energy portion split by the devices for pilot signal transmission does not affect the design of the multi-user sequencing and scheduling. Thus, we ignore this energy cost. However, the pilot signal transmission energy cost can straightforwardly be added in the final system energy cost.

MTC device sequencing and scheduling: In response to the channel access requests, the BS broadcasts a control packet which contains the sequence and schedule of the device transmission blocks and the optimal compression and transmission parameters for each device.

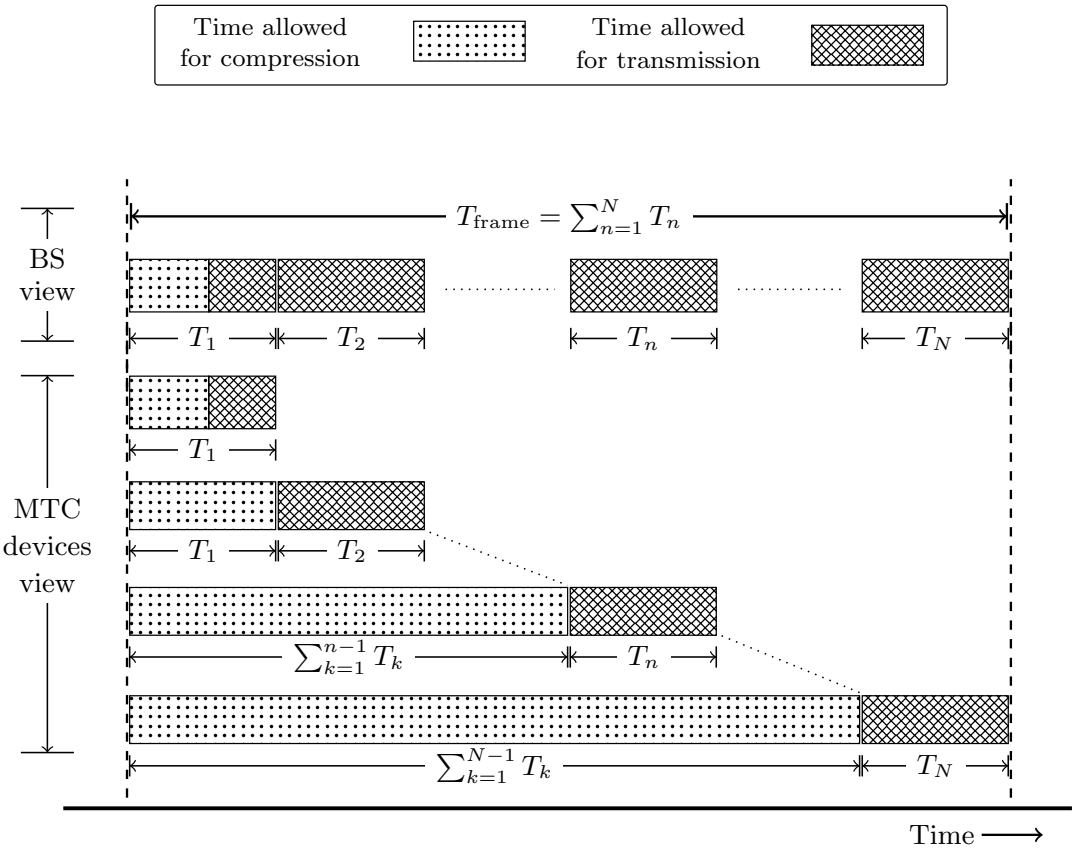


Figure 4.2: Timing diagram for the compression and transmission processes within a frame of uplink MTC. For simplicity, this figure only shows the scenario with the same block length.

The frame duration, T_{frame}, is divided into N transmission blocks as

$$T_{\text{frame}} = \sum_{n=1}^{N} T_n, \tag{4.2}$$

where T_n is the duration of nth transmission block.

In conventional settings, only the length of the transmission blocks affects the energy efficiency performance of a TDMA-based system. On the other hand, the sequence of the transmission block has no affect on the performance. In our case, however, the devices perform data compression before transmission. As a result, a device allocated with a later transmission block has more time to perform data

compression as compared to a device allocated with an earlier transmission block. Therefore, the position of the transmission block, which depends upon the multi-user sequence, influences the achievable energy efficiency performance.

Let us define $x_{n,i}$ as:

$$x_{n,i} = \begin{cases} 1, & \text{if the } n\text{th transmission block is allocated to the } i\text{th device,} \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

As each transmission block is allocated to only one device, we have

$$\sum_{i=1}^{N} x_{n,i} = 1, \quad \forall\, n. \tag{4.4}$$

Also, each device is assigned only one transmission block, implying that

$$\sum_{n=1}^{N} x_{n,i} = 1, \quad \forall\, i. \tag{4.5}$$

Compression: Before the start of its allocated transmission block, a device applies data compression on the raw data, as shown in Fig. 4.2. For the ith device, the D_i bits of raw data is compressed into $D_{\mathrm{cp},i}$ bits, resulting in a compression ratio of $\frac{D_{\mathrm{cp},i}}{D_i}$. The *compression time*, $T_{\mathrm{cp},i}$, is defined as the time required by the ith device to compress the raw data, D_i, into the compressed data, $D_{\mathrm{cp},i}$. We employ a generic non-linear compression cost model as proposed in [48]. The parameters of this compression model can be determined off-line for a given compression algorithm using data fitting. Specifically, the performance of this compression model is validated for the JPEG and JPEG2000 compression algorithms in [48]. Thereby, the compression time, $T_{\mathrm{cp},i}$, is given as a function of compression ratio, $\frac{D_{\mathrm{cp},i}}{D_i}$, as

$$T_{\mathrm{cp},i} = \tau D_i \Big(\Big(\frac{D_i}{D_{\mathrm{cp},i}} \Big)^{\beta} - 1 \Big), \tag{4.6}$$

where τ is the per-bit processing time and β is a compression algorithm dependent parameter that is proportional to the compression algorithm's complexity. τ depends upon the MCU processing resources and the number of program instructions

executed to process one bit of data. β determines the time taken to achieve a given compression ratio and is calculated off-line for a specific compression algorithm and a given hardware configuration. Let P_{cp} be the power consumed by a device during the data compression process. P_{cp} is predefined and constant for a given MTC device hardware.

Transmission: Once the compression process is complete, each device needs to transmit its compressed data within the allocated transmission block. The *transmission time* for the ith device, $T_{\text{tx},i}$, depends upon its compressed data size, $D_{\text{cp},i}$, and its link transmission rate, R_i, as

$$T_{\text{tx},i} = \frac{D_{\text{cp},i}}{R_i}. \tag{4.7}$$

Here, the transmission rate, R_i, is given as

$$R_i = B \log_2 \left(1 + \frac{\gamma_i}{\Gamma}\right), \tag{4.8}$$

where B is the bandwidth of the considered system, γ_i is the received signal-to-noise ratio (SNR) for the ith device, and Γ characterizes the gap between the achievable rate and the channel capacity due to the use of practical modulation and coding schemes [37, 60]. The received SNR for the ith device, γ_i, is defined as [82]

$$\gamma_i = \kappa \frac{P_i |h_i|^2}{\sigma^2 d_i^{\alpha}}, \tag{4.9}$$

where $\kappa = \left(\frac{\lambda}{4\pi}\right)^2$ is the path loss factor, λ is the wavelength, P_i is the transmit power for the ith device. To compute the data transmission power cost $P_{\text{tx},i}$ for the ith device, we adopt the practical model of [102]. The transmission power cost is composed of two components: (i) the transmit power P_i and (ii) the static communication module circuitry power P_{o}, which accounts for the operation of the digital-to-analog converter, frequency synthesizer, mixer, transmit filter, antenna circuits, etc. Specifically,

$$P_{\text{tx},i} = \frac{P_i}{\mu} + P_{\text{o}}, \tag{4.10}$$

where $\mu \in (0, 1]$ is the drain efficiency of the power amplifier.

4.2 Optimal Multi-User Sequencing and Scheduling Scheme

In this section, we first present the proposed multi-user sequencing and scheduling scheme and formulate the main optimization problem. Next, we present the system objectives and formulate the corresponding optimization problems.

The MTC devices perform two main operations (i) compression and (ii) transmission, each having individual completion time and energy cost. Recall that the device allocated with the first transmission block (T_1) in the frame, performs both the data compression and transmission operations within its allocated transmission block, i.e.,

$$x_{1,i}T_{\mathrm{cp},i} + x_{1,i}T_{\mathrm{tx},i} \leqslant T_1, \quad \forall\, i. \tag{4.11}$$

For all other devices that are allocated the remaining transmission blocks, they can apply data compression on the raw data during the period between the start of the frame and the start of its allocated transmission block, as illustrated in Fig. 4.2. This implies that the compression time for the ith device is upper bounded by the following constraint

$$T_{\mathrm{cp},i} \leqslant \sum_{k=1}^{n-1} T_k, \ \forall\, n \geqslant 2, \text{ if } x_{n,i} = 1,$$
$$\text{or equivalently, } \sum_{n=2}^{N} x_{n,i}T_{\mathrm{cp},i} \leqslant \sum_{n=2}^{N}\sum_{k=1}^{n-1} x_{n,i}T_k, \ \forall\, i. \tag{4.12}$$

One may ask the option of using part of the allocated transmission block for compression. This is unnecessary once the scheduling scheme is optimized, because the available time for compression for the device allocated into the $(n+1)$th transmission block is optimally controlled by optimizing the durations of all previous n transmission blocks.

After compression, each device transmits the compressed data within the transmission block allocated through multi-user scheduling. Accordingly, the transmission time for the ith device is upper bounded by the following constraint

$$T_{\mathrm{tx},i} \leqslant T_n, \ \forall\, n \geqslant 2, \text{ if } x_{n,i} = 1,$$

$$\text{or equivalently, } \sum_{n=2}^{N} x_{n,i} T_{\mathrm{tx},i} \leqslant \sum_{n=2}^{N} x_{n,i} T_n, \ \forall\, i. \quad (4.13)$$

The formulated problem allows fully flexible multi-user scheduling and thus there is no need to allow the associated devices to use some portion of the second and later transmission blocks for compression. Our design ensures that the length of the second and later transmission blocks does not go unnecessarily under utilized, if it is beneficial to have a longer compression time. This is achieved by allocating more compression time to the devices allocated the second or later transmission blocks by controlling the length of the first transmission block. Thereby, the extension in the length of first transmission block, beyond its associated device's need, reduces the length of the other transmission blocks to give more time for compression.

Each device needs to know the following parameters for its operation: (i) the starting time for its compression and transmission processes, (ii) the processing time allowed for its compression and transmission processes, (iii) the optimal compression ratio, and (iv) the optimal transmission rate. In the considered system, the multi-user scheduling mechanism ensures there is no gap or overlap between individual transmission blocks allocated to different devices. Therefore, the starting time of both compression and transmission processes for all the devices can be determined using the starting time of the frame, the length of transmission blocks and the multi-user sequence. The starting time for compression process for all devices is equal to the starting time of the frame. The starting time for the transmission process is determined using the transmission block lengths and the multi-user sequence. Note that the transmission rate of a device is controlled through its transmit power.

4.2.1 General Optimization Problem Formulation

The main problem we address is to determine the optimal length of transmission blocks allocated to devices (i.e., scheduling), the sequence of allocated transmission blocks, the compression and transmission policies for all devices. The aim is to

minimize some certain energy minimization objective (to be defined in the next subsection), under given delay and power constraints. The energy cost of the ith device is given as $E_i = P_{\text{cp}} T_{\text{cp},i} + P_{\text{tx},i} T_{\text{tx},i}$. Substituting the values for $T_{\text{cp},i}$, $P_{\text{tx},i}$ and $T_{\text{tx},i}$ from (4.6), (4.7), (4.10) here yields

$$E_i = P_{\text{cp}} \tau D_i \left(\left(\frac{D_i}{D_{\text{cp},i}} \right)^{\beta} - 1 \right) + \frac{D_{\text{cp},i}}{B \log_2 \left(1 + \kappa \frac{P_i |h_i|^2}{\Gamma \sigma^2 d_i^{\alpha}} \right)} \left(\frac{P_i}{\mu} + P_{\text{o}} \right). \tag{4.14}$$

The main optimization problem for the proposed scheme is formulated as follows.

$$\begin{aligned}
\mathbf{P^o}: \quad & \underset{\substack{P_i,\, D_{\text{cp},i},\, T_n,\\ x_{n,i},\, \forall\, n,i}}{\text{minimize}} && \mathcal{E}(E_1, E_2, \cdots, E_N) && \text{(4.15a)}\\
& \text{subject to} && \sum_{n=1}^{N} T_n = T_{\text{frame}}, && \text{(4.15b)}\\
& && x_{1,i} \tau D_i \left(\left(\frac{D_i}{D_{\text{cp},i}} \right)^{\beta} - 1 \right) && \\
& && + x_{1,i} \frac{D_{\text{cp},i}}{B \log_2 \left(1 + \kappa \frac{P_i |h_i|^2}{\Gamma \sigma^2 d_i^{\alpha}} \right)} \leqslant T_1, \quad \forall\, i, && \text{(4.15c)}\\
& && \sum_{n=2}^{N} x_{n,i} \tau D_i \left(\left(\frac{D_i}{D_{\text{cp},i}} \right)^{\beta} - 1 \right) \leqslant \sum_{n=2}^{N} \sum_{k=1}^{n-1} x_{n,i} T_k, \quad \forall\, i, && \text{(4.15d)}\\
& && \sum_{n=2}^{N} x_{n,i} \frac{D_{\text{cp},i}}{B \log_2 \left(1 + \kappa \frac{P_i |h_i|^2}{\Gamma \sigma^2 d_i^{\alpha}} \right)} \leqslant \sum_{n=2}^{N} x_{n,i} T_n, \quad \forall\, i, && \text{(4.15e)}\\
& && \sum_{i=1}^{N} x_{n,i} = 1, \qquad \forall\, n, && \text{(4.15f)}\\
& && \sum_{n=1}^{N} x_{n,i} = 1, \qquad \forall\, i, && \text{(4.15g)}\\
& && 0 \leqslant P_i \leqslant P_{\max}, \qquad \forall\, i, && \text{(4.15h)}\\
& && D_{\min,i} \leqslant D_{\text{cp},i} \leqslant D_i, \qquad \forall\, i, && \text{(4.15i)}\\
& && 0 \leqslant T_n \leqslant T_{\text{frame}}, \qquad \forall\, n, && \text{(4.15j)}\\
& && x_{n,i} \in \{0, 1\}, \qquad \forall\, n, i, && \text{(4.15k)}
\end{aligned}$$

where $\mathcal{E}(E_1, E_2, \cdots, E_N)$ is the objective function imposed by the considered energy minimization strategy to be defined in the next subsection. (4.15c), (4.15d) and (4.15e) are obtained by substituting the values of compression and transmission time from (4.6) and (4.7), in to inequalities (4.11), (4.12) and (4.13), respectively. $P_{\max}$ is the maximum transmit power constraint for each device. $D_{\min,i}$ is the lower bound on the compressed data size for the ith device. Thus, the maximum compression that can be applied is given by the minimum compression ratio defined as $\frac{D_{\min,i}}{D_i} \forall i$. The maximum compression ratio depends on the nature of the data and the system application. Note that a device may not fully utilize its allocated transmission block, depending upon its optimal compressed data size and/or the optimal transmission rate.

4.2.2 Considered System Objectives

In the literature, there are three popular system objectives for energy minimization, which differ in terms of the overall system performance and fairness among the MTC devices. These system objectives are (i) sum energy minimization, (ii) min-max energy minimization, and (iii) proportionally-fairness energy minimization. In the following, we discuss each of these system objectives and formulate the corresponding optimization problems. The system model under investigation allows the MTC devices to be located at various distances from the BS and experience different path attenuation. Moreover, the channel gain fluctuates independently for different devices in a given frame. This results in multi-user diversity due to the difference in the signal power attenuation conditions. The purpose behind different energy minimization objectives is to exploit this multi-user diversity while considering the trade-off between the system energy cost and the level of fairness in terms of energy cost of individual devices.

4.2.2.1 Sum energy minimization

The motivation behind the sum energy minimization objective is to prioritize the system performance over fairness among the devices. This objective attempts to achieve the maximum energy efficiency performance by fully exploiting the multi-user diversity. The strategy followed is to minimize the overall energy cost of all

the devices while ensuring each device completely transmits its data within the frame duration. The energy-fairness among devices is not considered. Therein, the devices with high signal power attenuation are provided with limited system resources, thus spend more energy than other devices and vice versa.

Mathematically, the objective is to minimize the total energy cost of all the devices in the given frame, i.e., minimize $\sum_{i=1}^{N} E_i$, where E_i is defined in (4.14). Accordingly, for the sum energy minimization objective, (4.15) becomes:

$$\begin{aligned} \mathbf{P^o_{SUM}}: \quad & \underset{\substack{P_i,\, D_{\mathrm{cp},i},\, T_n,\\ x_{n,i},\, \forall\, n,i}}{\text{minimize}} \quad \sum_{i=1}^{N} E_i\big(P_i, D_{\mathrm{cp},i}\big) \\ & \text{subject to} \quad (4.15\text{b}) - (4.15\text{k}). \end{aligned} \tag{4.16}$$

4.2.2.2 Min-max energy minimization

This system objective aims to prioritize fairness over system performance. It attempts to guarantee fairness in terms of individual device energy cost while maximizing the overall system energy efficiency performance. In particular, the maximum value of the energy spent by a device is minimized. Thereby, each device spends an equal amount of energy to transmit its data, irrespective of its signal power attenuation conditions.

The objective is to minimize the maximum energy cost among the devices in the given frame, i.e., minimize $\max\limits_{1 \leqslant i \leqslant N} \{E_i\}$, where E_i is defined in (4.14). This design strategy is Pareto-efficient [103], i.e., the energy cost of a device cannot be further decreased without increasing the energy cost of another device. This system objective provides strict energy fairness. For the min-max energy minimization objective, (4.15) becomes:

$$\begin{aligned} \mathbf{P^o_{MM}}: \quad & \underset{\substack{P_i,\, D_{\mathrm{cp},i},\, T_n,\\ x_{n,i},\, \forall\, n,i}}{\text{minimize}} \quad \max_{1 \leqslant i \leqslant N} \Big\{ E_i\big(P_i, D_{\mathrm{cp},i}\big) \Big\} \\ & \text{subject to} \quad (4.15\text{b}) - (4.15\text{k}). \end{aligned} \tag{4.17}$$

4.2.2.3 Proportionally-fair energy minimization

The sum energy minimization objective prioritizes the devices with better signal power attenuation performance, thereby allocating more system resources to boost their energy efficiency. As a result, the overall system energy efficiency performance increases at the cost of energy-unfairness among the devices. On the other hand, the min-max energy minimization objective targets strict energy fairness at the cost of reduced overall system energy efficiency performance. The motivation behind proportionally-fair energy minimization objective is to strike a balance between the system energy efficiency and device energy-fairness. This objective achieves some level of fairness among devices by providing each device with a performance that is proportional to its signal power attenuation conditions. This is achieved by reducing the opportunity of the devices with low signal power attenuation, getting more share of system resources to the weak devices. More system resources are allocated to the devices when their instantaneous signal power attenuation is low relative to their own signal power attenuation statistics. Thereby, proportional-fairness is achieved without compromising much energy efficiency performance. Since the signal power attenuation fluctuates independently for different devices, this strategy effectively exploits multi-user diversity. This can be achieved by minimizing the sum of logarithmic energy cost function of the individual devices [67, 99, 104], i.e., $\sum_{i=1}^{N} \log(E_i)$, where E_i is defined in (4.14).

For the proportionally-fair energy minimization objective, (4.15) becomes:

$$\mathbf{P^{o}_{PF}}: \quad \underset{\substack{P_i, D_{\mathrm{cp},i}, T_n, \\ x_{n,i}, \forall\, n,i}}{\text{minimize}} \quad \sum_{i=1}^{N} \log\Big(E_i\big(P_i, D_{\mathrm{cp},i}\big)\Big) \qquad \text{subject to} \quad (4.15\mathrm{b}) - (4.15\mathrm{k}). \tag{4.18}$$

4.3 Problem Solution and Optimality

The optimization problem defined in (4.15) is a mixed-integer nonlinear program which is non-convex in its natural form. Therefore, it is very challenging to determine the globally optimal solution or even to determine if the globally optimal solution exists [105]. In this regard, we first consider the binary variables to be

deterministic (i.e., a known sequence) and transform the non-convex optimization problem into convex sub-problems using methods that preserve equivalence. This will prove that the globally optimal solution exists. Later, we propose an alternative problem modelling approach to handle the binary constraints, which targets a more computationally feasible implementation. Therein, we perform transformations on the binary variables into real variables to get an approximate convex program which can be solved with significantly lower complexity. Moreover, we propose an iterative solution approach to push these real variables towards optimal binary values, which correspond to the binary variables in (4.15). These strategies are adopted for each of the considered objectives.

4.3.1 Existence of Globally Optimal Solution

For a fixed multi-user sequence $\big($i.e., $\{x_{n,i},\, \forall\, n, i\}$ is known$\big)$, problems (4.16), (4.17) and (4.18) can be remodelled as the following sub-problems, respectively, which are still non-convex in their natural form:

$$\widehat{\mathbf{P}}_{\mathbf{SUM}}: \quad \underset{P_i,\, D_{\mathrm{cp},i},\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \sum_{i=1}^{N} E_i(P_i, D_{\mathrm{cp},i}) \qquad \text{subject to} \quad (4.15\text{b})-(4.15\text{e}),\ (4.15\text{h})-(4.15\text{j}), \tag{4.19}$$

$$\widehat{\mathbf{P}}_{\mathbf{MM}}: \quad \underset{P_i,\, D_{\mathrm{cp},i},\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \max_{1\leqslant i\leqslant N} \Big\{ E_i(P_i, D_{\mathrm{cp},i}) \Big\} \qquad \text{subject to} \quad (4.15\text{b})-(4.15\text{e}),\ (4.15\text{h})-(4.15\text{j}), \tag{4.20}$$

$$\widehat{\mathbf{P}}_{\mathbf{PF}}: \quad \underset{P_i,\, D_{\mathrm{cp},i},\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \sum_{i=1}^{N} \log\Big(E_i(P_i, D_{\mathrm{cp},i}) \Big) \qquad \text{subject to} \quad (4.15\text{b})-(4.15\text{e}),\ (4.15\text{h})-(4.15\text{j}). \tag{4.21}$$

Lemma 1. *For a given sequence, the optimization problems* (4.19)*,* (4.20) *and* (4.21) *can be transformed into corresponding equivalent convex problems. Thus, a globally optimal solution exists for each of these problems* (4.19)*,* (4.20) *and* (4.21)*.*

Proof. The proof is provided in Appendix D.1. ■

From Lemma 1, we have the following equivalent convex problems for (4.19), (4.20) and (4.21), respectively:

$$\mathbf{P}^{\mathbf{so}}_{\mathbf{SUM}}: \quad \underset{Z_i,\, D_{\mathrm{cp},i},\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \sum_{i=1}^{N}\left(\tau D_i P_{\mathrm{cp}}\left(\left(\frac{D_i}{D_{\mathrm{cp},i}}\right)^{\beta}-1\right)+\frac{D_{\mathrm{cp},i}b_i}{Z_i}\big(\exp(Z_i)+c_i\big)\right) \tag{4.22a}$$

$$\text{subject to} \quad \text{(4.15b), (4.15i), (4.15j)},$$

$$x_{1,i}\tau D_i\left(\left(\frac{D_i}{D_{\mathrm{cp},i}}\right)^{\beta}-1\right)+x_{1,i}\frac{D_{\mathrm{cp},i}}{BZ_i}\leqslant T_1, \quad \forall\, i, \tag{4.22b}$$

$$\sum_{n=2}^{N} x_{n,i}\tau D_i\left(\left(\frac{D_i}{D_{\mathrm{cp},i}}\right)^{\beta}-1\right)\leqslant \sum_{n=2}^{N}\sum_{k=1}^{n-1} x_{n,i}T_k, \quad \forall\, i, \tag{4.22c}$$

$$\sum_{n=2}^{N} x_{n,i}\frac{D_{\mathrm{cp},i}}{BZ_i}\leqslant \sum_{n=2}^{N} x_{n,i}T_n, \quad \forall\, i, \tag{4.22d}$$

$$0\leqslant Z_i\leqslant Z_{\max}, \qquad \forall\, i, \tag{4.22e}$$

$$\mathbf{P}^{\mathbf{so}}_{\mathbf{MM}}: \quad \underset{Z_i,\, D_{\mathrm{cp},i},\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \max_{1\leqslant i\leqslant N}\left\{\tau D_i P_{\mathrm{cp}}\left(\left(\frac{D_i}{D_{\mathrm{cp},i}}\right)^{\beta}-1\right)+\frac{D_{\mathrm{cp},i}b_i}{Z_i}\big(\exp(Z_i)+c_i\big)\right\}$$

$$\text{subject to} \quad \text{(4.15b), (4.15i), (4.15j), (4.22b)} - \text{(4.22e)}, \tag{4.23}$$

$$\mathbf{P}^{\mathbf{so}}_{\mathbf{PF}}: \quad \underset{Z_i,\, V_i,\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \sum_{i=1}^{N}\log\left(\tau D_i P_{\mathrm{cp}}\left(\frac{D_i^{\beta}}{\exp\left(\beta V_i\right)}-1\right)+\frac{b_i\exp\left(V_i\right)}{Z_i}\Big(\exp\left(Z_i\right)+c_i\Big)\right) \tag{4.24a}$$

$$\text{subject to} \quad \text{(4.15b), (4.15j), (4.22e)},$$

$$x_{1,i}\tau D_i\left(\frac{D_i^{\beta}}{\exp\left(\beta V_i\right)}-1\right)+x_{1,i}\frac{\exp\left(V_i\right)\ln(2)}{BZ_i}\leqslant T_1, \quad \forall\, i, \tag{4.24b}$$

$$\sum_{n=2}^{N} x_{n,i}\tau D_i\left(\frac{D_i^{\beta}}{\exp\left(\beta V_i\right)}-1\right)\leqslant \sum_{n=2}^{N}\sum_{k=1}^{n-1} x_{n,i}T_k, \quad \forall\, i, \tag{4.24c}$$

$$\sum_{n=2}^{N} x_{n,i} \frac{\exp\big(V_i\big)\ln(2)}{BZ_i} \leqslant \sum_{n=2}^{N} x_{n,i} T_n, \quad \forall\, i, \tag{4.24d}$$

$$\ln(D_{\min,i}) \leqslant V_i \leqslant \ln(D_i), \qquad \forall\, i, \tag{4.24e}$$

where $b_i = \frac{\Gamma\sigma^2 d_i^{\alpha}\ln(2)}{\mu B\kappa|h_i|^2}$, $c_i = \frac{\mu\kappa|h_i|^2 P_{\mathrm{o}}}{\Gamma\sigma^2 d_i^{\alpha}} - 1$, $Z_i = \ln\big(1+\kappa\frac{P_i|h_i|^2}{\Gamma\sigma^2 d_i^{\alpha}}\big)$, $Z_{\max} = \ln\big(1+\kappa\frac{P_{\max}|h_i|^2}{\Gamma\sigma^2 d_i^{\alpha}}\big)$, $V_i = \ln\big(D_{\mathrm{cp},i}\big)$.

Let us consider the exhaustive search approach for the sake of proving the existence of the globally optimal solution. Note that it is not adopted to actually solve the proposed problems. In this search approach, we first find the globally optimal solution for all possible multi-user sequence permutations for problems in (4.19). Then, we determine which multi-user sequence minimizer and its associated optimal solution gives the minimum objective value. This multi-user sequence and solution is the globally optimal solution of problem (4.16). The same argument applies to problems (4.17) and (4.18).

Lemma 2. *The minimum of the globally optimal solutions of problems* (4.19)*,* (4.20) *and* (4.21) *for all possible sequences,* $\{x_{n,i},\, \forall\, n, i\}$*, is the corresponding globally optimal solution of problems* (4.16)*,* (4.17) *and* (4.18)*, respectively.*

From Lemmas 1 and 2, the globally optimal solution for each of the mixed-integer nonlinear programs in (4.16), (4.17) and (4.18), respectively, exists and can be found.

4.3.2 Handling Binary Variables

In the following, we consider problem (4.16) and first apply transformation on continuous variables as in (4.22). Next, we transform the binary variables into real variables to get an approximate convex program. Let us first deal with the binary variables $x_{n,i} \in \{0,1\},\ \forall\, n, i$. Note that for a real variable $x_{n,i} \in [0,1]$, we have $x_{n,i} \geqslant x_{n,i}^2,\ \forall\, n, i$. To this end, we can write

$$x_{n,i} \in \{0,1\} \Leftrightarrow x_{n,i} - x_{n,i}^2 = 0 \Leftrightarrow \big(x_{n,i} \in [0,1] \ \ \&\ \ x_{n,i} - x_{n,i}^2 \leqslant 0\big), \quad \forall\, n, i, \tag{4.25}$$

and adopt the approach of [106–109] to rewrite the constraint function (4.15k) as

$$\sum_{n=1}^{N}\sum_{i=1}^{N}\left(x_{n,i}-x_{n,i}^2\right)\leqslant 0, \tag{4.26}$$

$$0\leqslant x_{n,i}\leqslant 1,\quad \forall\, n,i. \tag{4.27}$$

In this way, we can relax the binary variables $x_{n,i}\in\{0,1\}$, $\forall\, n,i$, in (4.16) to real variables $x_{n,i}\in[0,1]$, $\forall\, n,i$, and introduce a cost function that penalizes the objective in (4.16) to impose $x_{n,i}=x_{n,i}^2$, $\forall\, n,i$ [107]. Therefore, the binary to real variables transformation and the continuous variables transformation as in (4.22) leads to the following equivalent problem

$$\begin{aligned}\mathbf{P}_{\mathbb{R}}:\quad &\underset{\substack{Z_i,\,D_{\mathrm{cp},i},\,T_n,\\ x_{n,i},\,\forall\, n,i,}}{\text{minimize}} && \sum_{i=1}^{N}E_i(Z_i,D_{\mathrm{cp},i})\;+\;\Lambda\sum_{n=1}^{N}\sum_{i=1}^{N}\left(x_{n,i}-x_{n,i}^2\right)\\ &\text{subject to} && (4.15\text{b}),\ (4.15\text{i}),\ (4.15\text{j}),\ (4.22\text{b})-(4.22\text{e}),\ (4.27),\end{aligned} \tag{4.28}$$

where $\Lambda\geqslant 0$ is a constant penalty factor. The term $\sum_{n=1}^{N}\sum_{i=1}^{N}\left(x_{n,i}-x_{n,i}^2\right)$ in (4.28) is the penalizing function on violation of the binary constraints over the energy minimization objective. Its magnitude quantifies the degree of violation from the binary constraints. Λ embodies the cost of this violation from the binary values $x_{n,i}$, $\forall\, n,i$. The minimizer of (4.28) will satisfy the binary constraints, $x_{n,i}\in\{0,1\}$, $\forall\, n,i$, for a finite value of Λ, i.e., the penalization is exact [105]. Thus, the optimization problems defined in (4.16) and (4.28) are equivalent, and the same optimal solution minimizes both the objective functions for a suitable value of the penalty factor [107].

The non-negative term $\sum_{n=1}^{N}\sum_{i=1}^{N}\left(x_{n,i}-x_{n,i}^2\right)$ in (4.28) decreases to 0 as $\Lambda\to+\infty$. Ideally, we need this term to be zero, and for that we would have to derive the optimal value of the penalty factor, Λ^*. For practical computational feasibility, let us introduce a numerical tolerance level such that it is acceptable to have $\sum_{n=1}^{N}\sum_{i=1}^{N}\left(x_{n,i}-x_{n,i}^2\right)<\epsilon$, where ϵ is very small and Λ is sufficiently large. Following [106] and [107], in our numerical experiments we found $\Lambda\geqslant 200$ is large enough to satisfy a tolerance level of $\epsilon=10^{-6}$ such that $\sum_{n=1}^{N}\sum_{i=1}^{N}\left(x_{n,i}-x_{n,i}^2\right)\leqslant\epsilon$.

Note that the penalty function in (4.28) is non-convex in $x_{n,i}$, $\forall\, n,i,$. Consider

a non-convex quadratic function $g(x) \triangleq x - x^2$, where $x \in [0,1]$. If we apply the first-order Taylor series expansion at a given point $x^{(j)} \in [0,1]$, we can obtain the convex lower bound on $g(x)$ as [109]

$$x\big(1-2x^{(j)}\big) + \big(x^{(j)}\big)^2 \leqslant x - x^2. \tag{4.29}$$

Similarly, the convex lower bound on the penalty function can also be given as

$$\sum_{n=1}^{N}\sum_{i=1}^{N}\Big(x_{n,i}\big(1-2x_{n,i}^{(j)}\big) + \big(x_{n,i}^{(j)}\big)^2\Big) \leqslant \sum_{n=1}^{N}\sum_{i=1}^{N}\big(x_{n,i} - x_{n,i}^2\big). \tag{4.30}$$

Accordingly, for a given point $x_{n,i}^{(j)} \in [0,1]$, the global upper bound minimization for problem (4.28) is given as

$$\begin{aligned}\mathbf{P}_{\mathbb{R}}^{\mathbf{UB}}: \quad & \underset{\substack{Z_i, D_{\mathrm{cp},i}, T_n,\\ x_{n,i}, \forall\, n,i,}}{\text{minimize}} \quad \sum_{i=1}^{N} E_i(Z_i, D_{\mathrm{cp},i}) \;+\; \Lambda \sum_{n=1}^{N}\sum_{i=1}^{N}\Big(x_{n,i}\big(1-2x_{n,i}^{(j)}\big) + \big(x_{n,i}^{(j)}\big)^2\Big) \\ & \text{subject to} \quad \text{(4.15b), (4.15i), (4.15j), (4.22b)} - \text{(4.22e), (4.27).}\end{aligned} \tag{4.31}$$

4.3.3 Solution Approach

Algorithm 1 outlines the steps to find the solution to the nonconvex problem (4.16) by iteratively solving a number of convex problem (4.31). In the first iteration, $j = 1$, problem (4.31) is solved using the initially guessed points, $x_{n,i}^{(j)}$, $\forall\, n, i$. The solution for the jth iteration $x_{n,i}^{*}$, $\forall\, n, i$ is used as an initial point for next iteration $j+1$. This process is repeated until convergence is achieved. The final solution yields the optimal parameters for multi-user sequencing and scheduling and compression and transmission rates for problem (4.16) due to its equivalence to problem (4.31).

Lemma 3. *Algorithm 1 converges to a Fritz John solution of problem* (4.16).

Proof. The proof is based on [109], [110], [111] and it is provided in Appendix C.2. ■

Algorithm 1 Iterative Approach for Multi-User Sequencing and Scheduling Optimization

1: **Initialization**: Set iteration count $j = 0$. Set initial point for $x_{n,i}^{(j)} = 0.5$, $\forall\, n, i$. Select a reasonably high penalty value $\Lambda = 200$ and low tolerance value $\epsilon = 10^{-6}$.
2: **while** $\sum_{n=1}^{N} \sum_{i=1}^{N} \left(x_{n,i}^{(j)} - \left(x_{n,i}^{(j)}\right)^2\right) \geqslant \epsilon$ **do**
3: Solve (4.31) using point $x_{n,i}^{(j)}$, $\forall\, n, i$ and get solution parameters Z_i^*, $D_{\mathrm{cp},i}^*$, T_n^*, $x_{n,i}^*$, $\forall\, n, i$.
4: Update point $x_{n,i}^{(j+1)} = x_{n,i}^*$, $\forall\, n, i$
5: Update iteration count $j = j + 1$
6: **end while**

From Lemma 3, we know that the solution of (4.31) found through Algorithm 1 converges to a Fritz John solution of nonlinear problem (4.31). The convergence to a Fritz John solution is a necessary condition for a solution to be optimal for a non-linear program. Due to the equivalence between problems (4.31) and (4.16), the solution by Algorithm 1 is an optimal solution for nonlinear problem (4.31) and correspondingly it is also an optimal solution for problem (4.16).

Problem (4.31) is solved in each iteration of Algorithm 1 with a polynomial computational complexity in the number of variables and constraints. Accordingly, (4.31) can be transformed into an equivalent optimization problem such that it contains $\mathrm{n_l}=(5N+1)$ linear constraints, $\mathrm{n_p}=(3N)$ posynomial constraints, and $\mathrm{n_b}=(N^2)$ real-valued scalar decision variables. Thus, solving (4.31) requires a complexity of $\mathcal{O}\left(\sqrt{\mathrm{n_l}+\mathrm{n_p}}[\mathrm{n_l}+\mathrm{n_p}+\mathrm{n_b}]\mathrm{n_b}^2\right)$ [108].

It is noted that Algorithm 1 can be straightforwardly modified to solve problems (4.17) and (4.18) by first applying the continuous variables transformation as in (4.23) and (4.24), respectively, and then following the same steps provided in Section 4.3.2 for the binary variables transformation.

4.4 Sub-Optimal Multi-User Scheduling Scheme

In this section, we present the sub-optimal scheme that does not consider multi-user sequencing optimization and the sequence is kept fixed. However, the multi-user scheduling, data compression and transmission rate parameters are jointly

optimized. The proposed optimal and sub-optimal schemes only differ in the consideration of multi-user sequence optimization. Thus, the sub-optimal scheme is proposed to illustrate the impact of multi-user sequencing. In Section 4.5, we will perform a comparative performance analysis of these schemes to show the impact of multi-user sequencing in a multi-user uplink communication system, which is the main contribution of this work. Hence, the sub-optimal scheme serves an important purpose. The three optimization problems are formulated for the sub-optimal scheme considering the three different system objectives previously defined in Section 4.2.2.

For the sub-optimal scheme, the multi-user sequence is fixed and unchanged from one frame to the next. However, the transmission block length of any device is flexible and can be optimized. In this scheme, the transmission rate, compression ratio, and the transmission block length (multi-user scheduling) are jointly optimized for the given energy minimization objective for a fixed multi-user sequence $\big($i.e., $\{x_{n,i},\, \forall\, n, i\}$ is known$\big)$. For the considered system objectives, the optimization problems for the sub-optimal scheduling scheme are given in (4.22), (4.23) and (4.24).

4.5 Numerical Results

This section presents the numerical results to illustrate the performance of the proposed scheme. Unless specified otherwise, the values for the parameters shown in Table 5.1 are adopted.

Remark 10. *Algorithm 1 is implemented in AMPL* [112]*, which is popular for modelling scheduling problems*[1]*. A model for the proposed problem is developed in the AMPL environment and the Couenne (convex over and under envelopes for nonlinear estimation) solver* [113, 114] *is used to solve the problem. The Couenne solver guarantees the globally optimal solution if such a solution exists and we have already proved its existence by Lemmas 1 and 2.*

[1]To cross-check we compared AMPL-Couenne with CVX-sedumi/SDPT3 in geometric programming (GP) mode. The CVX-GP mode requires additional transformations on (4.31) for compliance. Nevertheless, the results match up to 6 decimal points.

Table 4.1: System parameter values.

Name	**Symbol**	**Value**
Amplifier's drain efficiency	μ	0.35
Max. transmit power	$P_{\max}$	0 dB
Scale parameter for channel gain	ς	1
Wavelength	λ	0.333 m
Compression processing power	P_{cp}	24 mW
Number of devices	N	5
Comm. module circuitry power	P_{o}	82.5 mW
Bandwidth	B	1 MHz
Practical modulation power gap	Γ	9.8 dB
Packet size	D_i	{310,500,100,80,200} kbits
Minimum compression ratio	$\frac{D_{\min,i}}{D_i}$	0.4
Distance	d_i	{40,15,31,49,22} m
Per-bit processing time	τ	7.5 ns/b
Noise spectral density	N_0	-174 dBm
Compression cost parameter	β	5
Pathloss exponent	α	4

Let us define system energy cost as the total energy cost of all the devices, i.e., $\sum_{i=1}^{N} E_i$. Moreover, the energy efficiency gain, $\mathbb{G}_{\text{ee}}$, provided by a given scheme A over scheme B be defined as the percentage decrease in the system energy cost of scheme B, $\sum_{i=1}^{N} E_{i,B}$, in comparison to the system energy cost of scheme A, $\sum_{i=1}^{N} E_{i,A}$, and it is given as

$$\mathbb{G}_{\text{ee}} = \frac{\sum_{i=1}^{N} E_{i,B} - \sum_{i=1}^{N} E_{i,A}}{\sum_{i=1}^{N} E_{i,B}}. \tag{4.32}$$

It should be noted that the relative performance of the sum, min-max, and proportionally-fair energy minimization objectives has been well studied in previous studies and thus it is not the focus of this paper. Our focus is rather to evaluate the performance of the proposed joint optimization of multi-user sequencing and scheduling scheme.

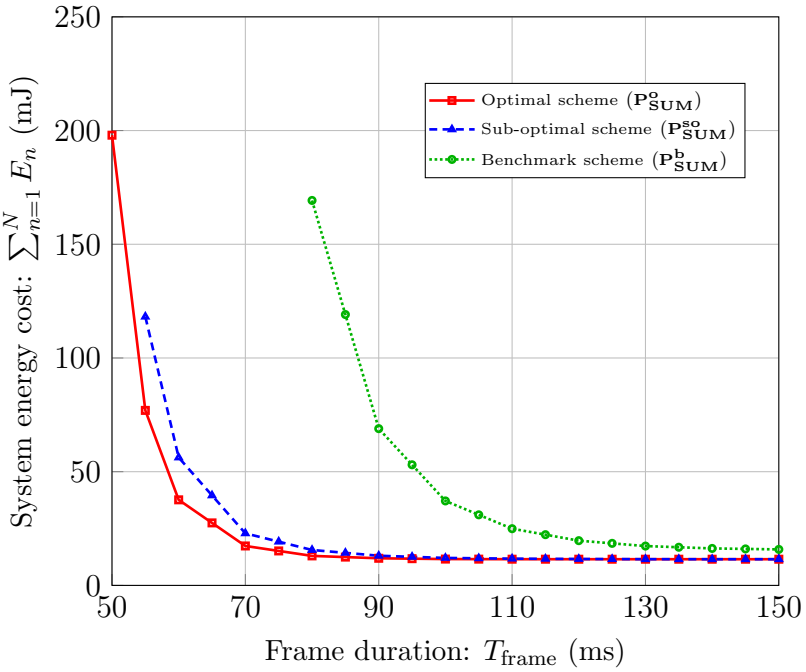
To the best of our knowledge, the recent works [37, 67, 68, 70] are the most relevant to our proposed scheme. Although the system models in these works are based on wireless power transfer, the underlying multi-user scheduling and trans-

mission rate policy designs are similar to our considered system. In this regard, we adopt the multi-user scheduling and transmission rate design policies proposed by these schemes for our considered system model except that data compression and multi-user sequencing are not employed. Moreover, when our considered system is applied, the design problems proposed in [37, 67, 68, 70] can equivalently be represented by the following benchmark scheme.

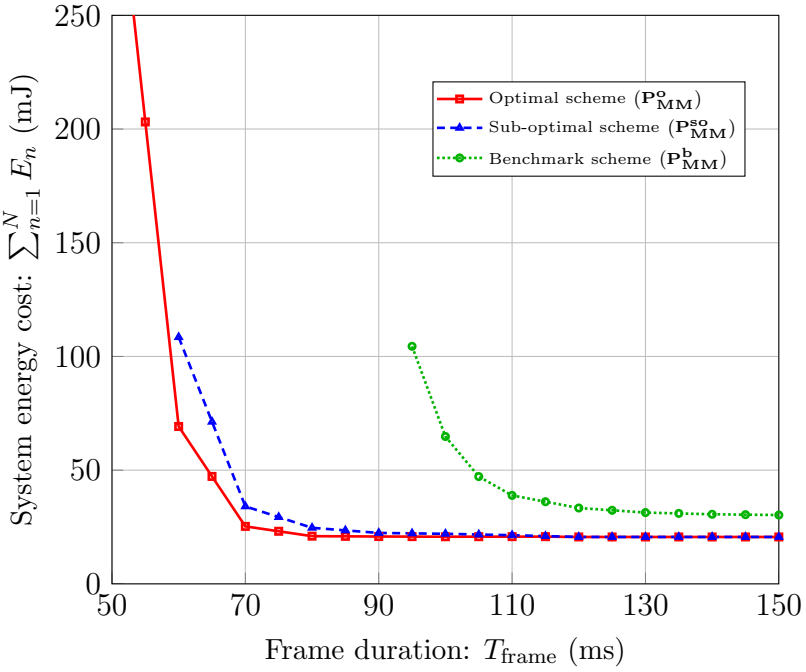
Benchmark scheme: For the benchmark scheme, the multi-user sequence is fixed but the multi-user-scheduling can be optimized, i.e., the transmission block length of each device is flexible. This scheme does not employ data compression. The transmission rate and the transmission block length (scheduling) are jointly optimized for the given energy minimization objective for a fixed sequence and without employing data compression. The strategy followed to optimize the multiuser scheduling and device transmission rate policies for this benchmark scheme is essentially the same as in the state of the art in [37,67,68,70]. The corresponding optimization problems for the benchmark scheme are given in Appendix C.3. For comparison with the proposed scheme, the same energy minimization objectives are considered.

4.5.1 Validation

In this subsection, we carry out a comparative analysis of the proposed scheme with the benchmark scheme (which represents existing state-of-the-art work). Fig. 4.3 plots the system energy cost, $\sum_{i=1}^{N} E_i$, versus the frame duration, T_{frame}, for the system parameters in Table 5.1. The system energy cost is plotted with the proposed optimal scheme and benchmark scheme for the sum, min-max, and proportionally-fair energy minimization objectives in Fig. 4.3. The energy efficiency gain, $\mathbb{G}_{\text{ee}}$, provided by the proposed optimal scheme over benchmark scheme is plotted in Fig. 4.4. The performance for the sub-optimal scheme is also shown in Figs. 4.3 and 4.4 which we will discuss later. It can be seen from Fig. 4.4 that the gains are almost the same irrespective of the considered energy minimization objective. In the following, we will discuss system performance for the sum energy minimization objective. Similar conclusions can be drawn for the min-max and proportionally-fair energy minimization objectives.

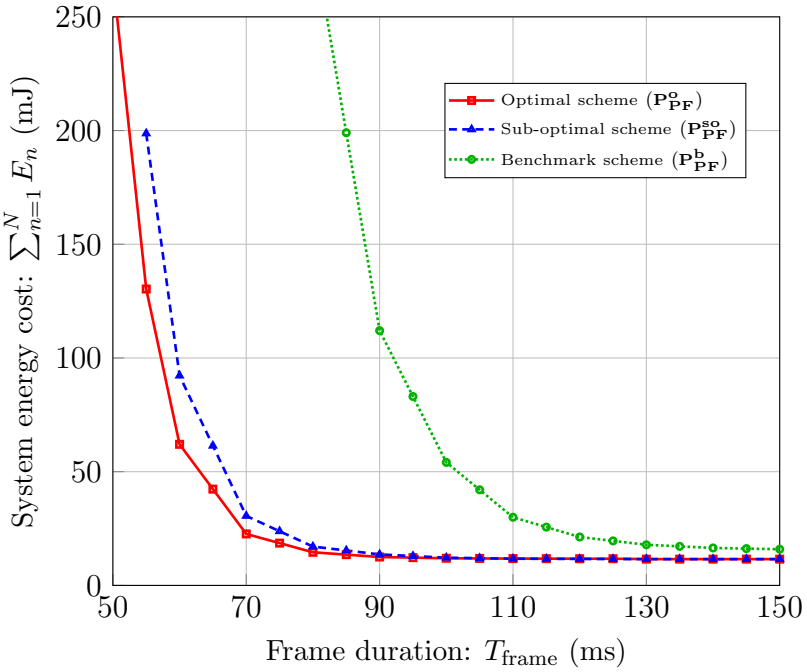


(a) Sum-energy minimization



(b) Min-max energy minimization

Figure 4.3: System energy cost under given power constraints and system objectives.



(c) Proportionally-fair energy minimization

Figure 4.3: System energy cost under given power constraints and system objectives.

When compared with the benchmark scheme, which does not employ data compression, the proposed optimal scheme exhibits significant performance superiority. This shows that employing both multi-user sequence and compression optimization provides notable gains in the energy efficiency, specifically in the lower latency regime. For the sum energy minimization objective, Fig. 4.4 shows that the gain is comparatively significant (between 27% to 92%), for the considered range of delay when the system is feasible for the benchmark scheme (between 150 ms to 80 ms). For the benchmark scheme, the device energy cost is reduced by adapting the minimum required transmit power level under given channel conditions. However, reducing transmission rate through transmit power only helps up to a certain level and any further reduction does not improve energy efficiency. Hence, in general, it is not optimal to transmit at the lowest transmission rate. Note that for the proposed optimal scheme the lower bound delay has a much smaller value as compared to the benchmark scheme.

4.5.2 Impact of Multi-User Sequencing

To illustrate the advantage of the proposed joint multi-user sequencing, we consider the proposed optimal scheme and sub-optimal scheme for comparative analysis. In both schemes, the multi-user scheduling and compression are optimized. However, they differ in an important aspect that the multi-sequencing is employed by the proposed optimal scheme and not by sub-optimal scheme, which uses a fixed multi-user sequence.

From Fig. 4.3, the proposed optimal multi-user sequencing and scheduling scheme clearly outperforms the sub-optimal scheme. Intuitively, it was expected that the multi-user sequencing will always provide non-negative gains. However, the gains are notable, between 13% to 35%, for the considered range of delay when the system is feasible for the sub-optimal scheme as shown in Fig. 4.4, for the sum energy minimization objective. Also, in the lower latency regime the gains are significantly high, between 85 ms to 55 ms. Thus, for a less stringent delay constraint, employing the multi-user sequencing will not pay off. At the same time, it can be conclude that the data compression provides significant gains for all sorts of delay constraints.

In addition, when the proposed optimal scheme is employed, the TDMA-based multi-user transmissions is more likely to be feasible in the lower latency regime subject to the given power constraints. That is, the proposed optimal scheme can support much stringent delay requirements, under given maximum transmit power constraints, as compared to the sub-optimal scheme for the same system parameters. Note that in Fig. 4.3, the system energy cost flattens out as the delay is increased further from a specific value for both schemes. The reason is that for both proposed schemes, with joint data compression and transmission rate strategy, there exists a lower bound on the device energy cost.

4.5.3 Impact of Scheduling Flexibility

Consider a scenario where the transmission blocks are fixed and are not changed from one frame to the other. We illustrate the impact of multi-user scheduling on the overall system performance and multi-user sequencing by allocating the transmission blocks of the same length. That is, the frame duration, T_{frame}, is

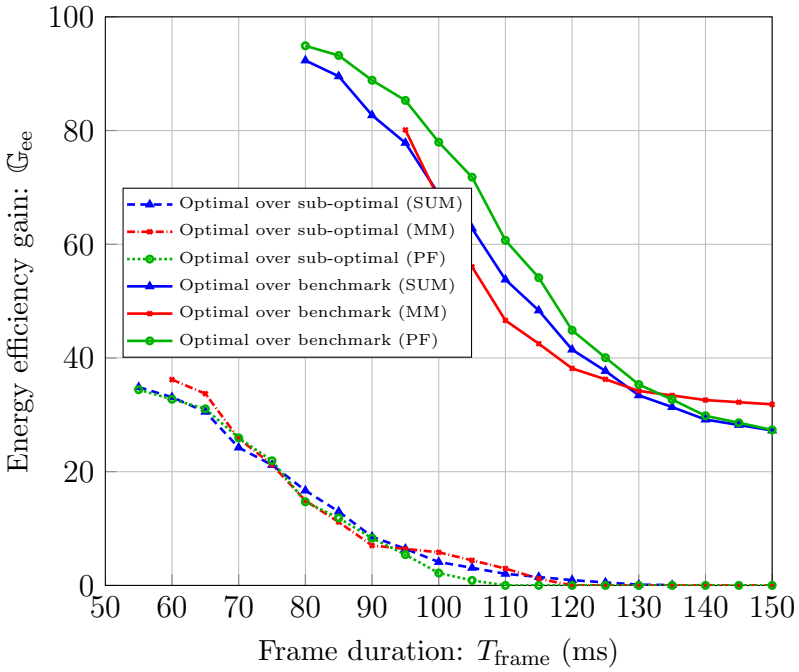


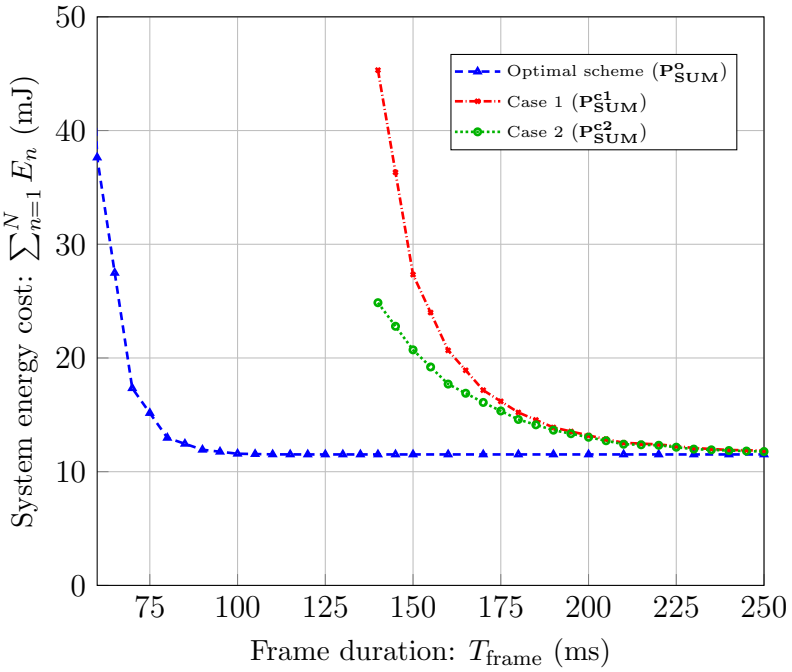
Figure 4.4: Energy efficiency gain performance vs. frame duration for the proposed optimal scheme over the sub-optimal scheme and benchmark scheme, for the considered system objectives.

divided into N equal transmission blocks and assigned to N devices. For this scenario, we consider two cases:

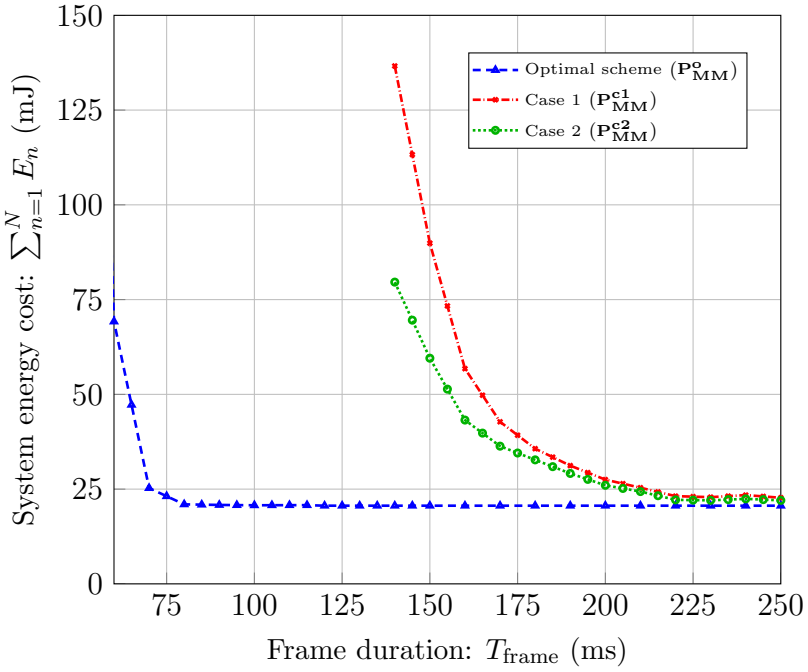
- *Case 1*: the multi-user sequence is fixed,
- *Case 2*: the multi-user sequence can be optimized.

In both cases, the transmission rate and compression ratio are optimized for each device under a given energy minimization objective. The corresponding optimization problems for both these cases are given in Appendices C.4 and C.5, respectively.

Fig. 4.5 plots the system energy cost, $\sum_{i=1}^{N} E_i$, versus the frame duration, T_{frame}, for the system parameters in Table 5.1. The system energy cost is plotted with the proposed optimal scheme and Case 1 and Case 2 for the sum, min-max, and proportionally-fair energy minimization objectives in Fig. 4.5. As can be seen, both Case 1 and Case 2 perform similar. However, when the delay is stringent Case 2 performs significantly better than Case 1 due to multi-user sequencing. The energy

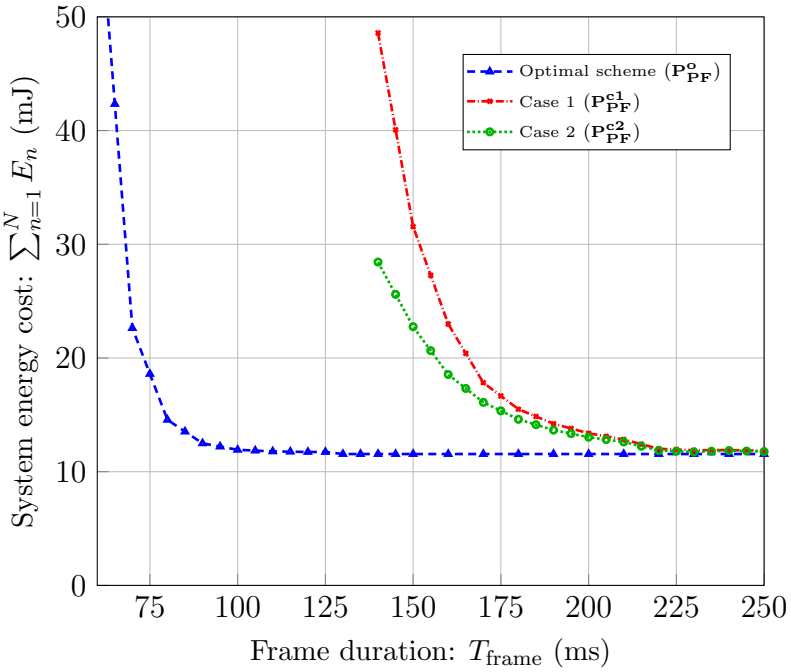


(a) Sum-energy minimization



(b) Min-max energy minimization

Figure 4.5: System energy cost under given power constraints and system objectives.



(c) Proportionally-fair energy minimization

Figure 4.5: System energy cost under given power constraints and system objectives.

efficiency gain is between 11% to 45%, for the considered range of delay from 165 ms to 140 ms for the sum energy minimization objective.

When compared with the optimal scheme, a large performance degradation is observed for the considered scenario when the transmission block lengths are fixed. Thus, optimizing the multi-user sequencing provides a large performance gain even in a restricted scheduling scenario. In addition, multi-user scheduling flexibility when combined with multi-user sequencing has a significant impact on the overall system performance. Similar conclusions can be drawn for the min-max and proportionally-fair energy minimization objectives.

4.6 Summary

In this chapter, we have investigated the joint optimization of sequencing and scheduling in a multi-user uplink machine-type communication scenario, consider-

ing adaptive compression and transmission rate control design. The energy efficiency performance is evaluated for three energy-minimization system objectives, which differ in terms of the overall system performance and fairness among the devices. Our results have showed that the proposed optimal scheme outperforms the schemes without multi-user sequencing. The improvement in energy efficiency observed is up to 35% when multi-user sequencing is optimized, under given maximum transmit power and delay constraints. In an alternate scenario, when the length of the transmission blocks is fixed and equal for each device, multi-user sequence optimization still provides a performance gain of up to 45%. However, the overall system performance degrades quite significantly. The energy efficiency gain of multi-user sequence optimization is paramount under a stringent delay bound. Thus, multi-user sequence optimization makes the TDMA-based multi-user transmissions more likely to be feasible in the lower latency regime subject to the given power constraints.

In this chapter, we assumed that the instantaneous channel gain for each device is perfectly estimated by the BS. In the more practical case of imperfect channel estimation, outage occurs at the BS, and hence, an additional outage probability constraint needs to be introduced into the optimization problem. Nevertheless, the overall principle of the proposed multi-user sequencing and scheduling scheme remains the same.

The proposed solution in this chapter is scalable to networks with massive number of devices in two possible ways: (i) The TDMA based large networks with massive number of devices are extensively handled in the prior studies through clustering of devices that are spatially located in close proximity. Therein, all devices within a cluster use a separate channel or sub-channel for data transmission and the number of devices per frame are kept under control. (ii) For a different scenario where a large number of devices need to share a single frame. The frame can be divided into sub-frames each serving a controlled number of devices. The sequencing and scheduling optimization can then be performed on sub-frames instead of the whole frame. This will significantly reduce the complexity at the cost of sub-optimal performance. It would be interesting to investigate this trade-off in future works.

Chapter 5

Multi-User Transmit Power Diversity in Contention Resolution Diversity Slotted ALOHA Systems

In this chapter, we consider a network scenario that is evolved from the network scenario considered in chapter 4. However, due to lack of coordination among devices, the machine-type devices (MTDs) randomly transmit short data-packets on a shared channel. We consider a network scenario is applicable to a multi-user single-channel massive machine type communication (MTC) system. In such a system, all MTDs share a single channel and randomly transmit short data-packets on this channel. The data may be generated or sensed by the MTD periodically, on demand, or due to the occurrence of an event of interest that is to be delivered under given reliability and delay conditions. This network scenario cover highly reliability and dense MTC systems. For this network scenario, the communication policies for all MTDs are jointly devised for given system parameters.

Motivated by the critical need to better support the MTC in future wireless networks [3], contention resolution diversity slotted ALOHA (CRDSA) [39] and its variants [115–117] have been proposed to enhance the performance of random access schemes [118]. The well known idea of CRDSA is to allow devices to transmit

multiple copies (burst) of the packet in randomly selected slots within a frame and perform iterative successive interference cancellation (SIC) when attempting to resolve collisions. However, most of these variants are based on the clean packet model [39, 117] in which only interference-free packets are recoverable.

Contention resolution diversity slotted ALOHA (CRDSA) is a promising solution to meet the challenge of designing efficient random access in future wireless networks. In this chapter, we consider CRDSA with transmit power diversity where each packet copy from a device is transmitted at a randomly selected power level. This results in inter-slot received power diversity, which is exploited by employing a signal-to-interference-plus-noise ratio based successive interference cancellation (SIC) receiver. Leveraging edge-weighted bipartite graph representation, we propose a novel graph-based message passing algorithm to model the SIC decoding. We derive an expression to characterise the recovery-error probability of the scheme. We also formulate and solve an optimization problem to determine the optimal transmit power distribution.

This chapter is organized as follows. The system model is presented in Section 5.1. The proposed random access mechanism is proposed in Section 5.2. The recovery-error probability is derived in Section 5.3. The optimization problem for the proposed scheme is formulated in Section 5.4. Numerical results are presented in Section 5.5. Finally, Section 5.6 concludes the chapter.

5.1 System Model

We consider an uncoded slotted random access system with M machine-type communication devices (MTDs), which contend to access a single base station (BS). The time is divided into frames and each frame is divided into N equal duration slots. The MTDs are frame and slot synchronized, e.g., using global positioning system (GPS) or using periodic beacons transmitted by the BS [115]. Similar to existing works [73, 74, 80], we assume that (i) both the large-scale fading and small-scale fading coefficients are known perfectly to the BS. Hence, we do not consider fading in this chapter, (ii) the BS knows the number of MTDs M [39, 117] and (iii) each MTD generates a single packet, which fits in exactly one slot, for transmission in each frame. The system load G is defined as the normalised number of packets

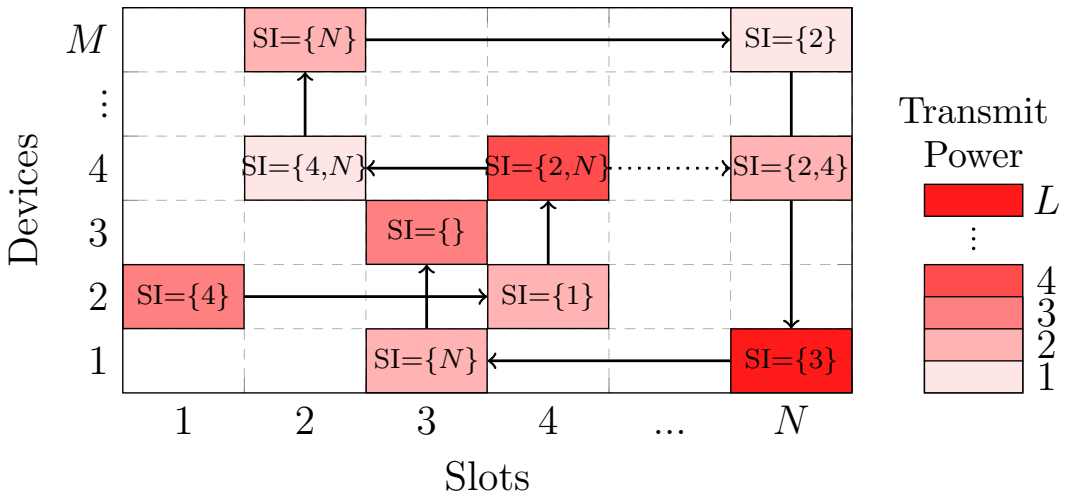


Figure 5.1: Illustration of proposed random access mechanism (SI = slot index). The horizontal arrows represent the SIC iteration and the vertical arrows depict the burst recovery in the next iteration. The dotted horizontal arrow shows that SIC alleviated the interference power, yet the collision is not resolved.

per slot. Since, each device generates only one packet, $G = \frac{M}{N}$.

5.2 Proposed Random Access Mechanism

5.2.1 Transmission Scheme

Similar to previous works on CRDSA, we assume that each MTD transmits d copies (i.e., bursts) of its single packet in randomly and uniformly selected slots within one frame, where $1 \leq d \leq D$ and D is the maximum number of allowable copies, as shown in Fig. 5.1.

Different from existing works on CRDSA, in addition to the packet diversity due to the copies, we allow *transmit power diversity*. Hence, each MTD randomly chooses a transmit power level, E_l, for each of its copies from the set of equally spaced L available power levels, denoted by $\mathcal{E} = \{E_1, E_2, \ldots, E_L\}$, where $1 \leq l \leq L$. We allow selection of transmit power levels with replacement, i.e., two copies from a MTD may be transmitted on the same power level. Note in Fig. 5.1 the transmit power level is indicated with a color gradient.

Due to slot-level synchronization, copies of packets from different MTDs either

completely overlap (i.e., collide) or not at all. Therefore, a single slot may contain collided copies of packets from k MTDs, where $1 \leq k \leq M$. A collision free slot is called a singleton slot and the packet is called a clean packet.

Let Λ_d denote the probability that a MTD transmits d copies, Ψ_k denote the probability that k MTDs choose to transmit a copy on any given single slot and Γ_l denote the probability that a MTD chooses power level E_l for a given copy. According to [117], the polynomial representations for the probabilities Ψ_k, Λ_d, and Γ_l are

$$\begin{aligned} \Psi(z) &\triangleq \sum_k \Psi_k z^k, \\ \Lambda(z) &\triangleq \sum_d \Lambda_d z^d, \\ \Gamma(z) &\triangleq \sum_l \Gamma_l z^l. \end{aligned} \tag{5.1}$$

5.2.2 Recovery Mechanism

The received signal at the BS in the nth slot is given by

$$y_n = \mathbf{e_n}\mathbf{x_n} + \omega_n, \tag{5.2}$$

where $n = 1, 2, \ldots, N$ is the slot index, $m = 1, 2, \ldots, M$ is the user index, $\mathbf{e_n} = \left[\sqrt{e_{1,n}}, \ldots, \sqrt{e_{m,n}}, \ldots, \sqrt{e_{M,n}}\right]$ is the $1 \times M$ vector of transmit power levels and $e_{m,n} \in \mathcal{E}$, $\mathbf{x_n} = [x_{1,n}, \ldots, x_{m,n}, \ldots, x_{M,n}]^T$ is the $M \times 1$ vector of transmitted signals, where $x_{m,n}$ refers to one of the copies of the packet transmitted by the mth MTD on the nth slot, and ω_n is the additive white gaussian noise (AWGN) with zero mean and variance σ^2.

We assume that the BS employs SIC to iteratively recover each MTD's packet. For a given iteration of the SIC process, the BS first checks to see if any packet in any slot can be successfully decoded. In this chapter, we consider that a packet is successfully decoded if, at least in one of its copies, the received signal-to-interference-plus-noise ratio (SINR) is above a threshold γ^{th}. We assume that each copy contains the slot indices to all its copies so that once a packet is successfully decoded, full information about the location of its copies is available. Hence, once a MTD's packet is decoded at the BS, the interference caused by that packet can

be completely removed from all slots. This in turn increases the SINR of the remaining packets in the following iterations. The process is repeated until all MTD packets are recovered or a maximum number of iterations $i_{\max}$ is reached. This repetition and SIC based decoding process at the BS is depicted with the help of arrows in Fig. 5.1.

Let $\mathcal{M}_n$ denote the set of MTDs transmitting on the nth slot. Let $\mathcal{N}_m$ denote the set of slots selected by the mth MTD for transmitting of its copies. We are interested in the recovery-error probability, which is defined as the probability that the BS fails to recover the mth MTD's packet in the ith iteration of the SIC process. It can be expressed as

$$q_m^{(i)} = 1 - \mathbb{P}\Big\{ \max\big\{\gamma_{m,n}^{(i)}\big\} \geqslant \gamma^{\text{th}}\Big\}, \tag{5.3}$$

where $\mathbb{P}\{\cdot\}$ denotes the probability, $\max\{\cdot\}$ denotes the max operation, $n \in \mathcal{N}_m$, $|\mathcal{N}_m| = d$, and $\gamma_{m,n}^{(i)}$ is the SINR of the mth MTD in the ith iteration for its copy transmitted with power level E_l in the nth slot, which is given by

$$\gamma_{m,n}^{(i)}(E_l, u) = \frac{E_l}{\Theta_{m,n}^{(i)}(u) + \sigma^2}, \tag{5.4}$$

where σ^2 is the noise power and $\Theta_{m,n}^{(i)}(u)$ represents the cumulative interference power from u unresolved MTDs. Note that $q_m^{(i)}$ in (5.3) is the same for each MTD device since the probabilities of selecting a transmission slot, power level and number of copies, respectively, are the same for all nodes.

Due to the $\max\{\cdot\}$ operation, (5.3) is not tractable across iterations of the SIC process. Hence, we employ graph based analysis to analyze $\gamma^{(i)}$. In this regard, by defining an interference threshold $\varphi^{\text{th}}(E_l) = \frac{E_l}{\gamma^{\text{th}}} - \sigma^2$, we can get the following relationship which will be used in the subsequent analysis,

$$\mathbb{P}\Big\{\gamma_{m,n}^{(i)}(E_l, u) \geqslant \gamma^{\text{th}}\Big\} = \mathbb{P}\Big\{\Theta_{m,n}^{(i)}(u) \leqslant \varphi^{\text{th}}(E_l)\Big\}. \tag{5.5}$$

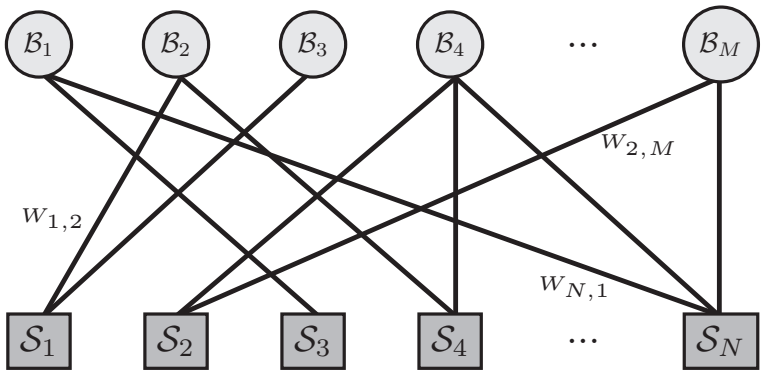


Figure 5.2: Edge-weighted bipartite graph representation of the proposed scheme.

5.3 Recovery-Error Probability Analysis

5.3.1 Graph Representation

The system under consideration can be represented as an edge-weighted bipartite graph and can be analysed using the theory of codes on graph [76, 117]. An example graph, relative to Fig. 5.1, is illustrated in Fig. 5.2, where the MTDs and slots are shown by circles and rectangles, representing burst nodes and sum nodes, respectively.

An edge-weighted bipartite graph is defined by $\mathcal{G} = \{\mathcal{B},\ \mathcal{S},\ \mathcal{E}_W\}$, where $\mathcal{B}$, $\mathcal{S}$, and $\mathcal{E}_W$ represent the sets of burst nodes (MTDs), sum nodes (slots), and edges with weights, respectively. The edge weight $W_{m,n}$ corresponds to the transmit power level $e_{m,n} \in \mathcal{E}$. The number of edges connected to a node represents the node degree.

Let λ_d denote the probability that an edge is connected to a degree-d burst node and ρ_k denote the probability that an edge is connected to a degree-k sum node. According to [117], the definitions of λ_d and ρ_k can be given as

$$\begin{aligned} \lambda_d &\triangleq \frac{\Lambda_d d}{\sum_r \Lambda_r r}, \\ \rho_k &\triangleq \frac{\Psi_k k}{\sum_j \Psi_j j}, \end{aligned} \tag{5.6}$$

where the probabilities Λ_d and Ψ_k are represented in (5.1).

5.3.2 Message Passing Algorithm

We develop a graph-based message passing algorithm representation to track the iterative SIC process in the proposed scheme.

The set of burst nodes transmitting on sum node $\mathcal{S}_n$ is equal to $\mathcal{M}_n$ and the set of sum nodes selected by burst node $\mathcal{B}_m$ for burst transmission is equal to $\mathcal{N}_m$. The message $\mathbf{S}_{n\to m}^{(i)}$ is sent by the sum node $\mathcal{S}_n$ to burst node $\mathcal{B}_m$ to advertise the cumulative interference power $\Theta_{m,n}^{(i)}$ (defined below (5.4)) experienced in the ith iteration due to the unresolved edges connected to it. On the graph, this cumulative interference power is the sum of weights of all unresolved edges on a given sum node. In response to the message $\mathbf{S}_{n\to m}^{(i)}$, a burst node $\mathcal{B}_m$ replies with a message $\mathbf{B}_{m\to n}^{(i)}$. This message tells the sum node $\mathcal{S}_n$ whether the edge connected to the burst node $\mathcal{B}_m$ should be removed or not. If the interference power is below the interference threshold $\varphi^{\text{th}}(W_{m,n})$ (defined above (5.5)), which implies that the edge can be removed, then the message $\mathbf{B}_{m\to n}^{(i)}$ equals the interference power contributed by $\mathcal{B}_m$ on sum node $\mathcal{S}_n$, i.e., the weight $W_{m,n}$, is subtracted (edge is removed) from the cumulative interference power on $\mathcal{S}_n$ and the updated interference power is reported to the remaining unresolved burst nodes in the next iteration. Otherwise, if the message $\mathbf{B}_{m\to n}^{(i)}$ is zero then it implies that the edge cannot be removed in the current iteration. This process is repeated until the maximum number of iterations $i_{\max}$ is reached or all burst nodes are recovered. Algorithm 1 summarizes the proposed graph-based message passing algorithm representation of the iterative SIC-based decoding process. Note that the developed message passing algorithm is different from [40] due to the consideration of transmit power diversity and slot-wise SINR based recovery in our work.

5.3.3 Derivation of Recovery-Error Probability

We derive the recovery-error probability using ANR-OR tree analysis technique. The basic principle of this approach is to represent the graph as a tree and then formulate $q_m^{(i)}$. In the ANR-OR process, an unknown edge of a burst node is recovered if at least one of its edges are revealed (*OR rule*). Similarly, for a sum node an edge is recovered in the given iteration if all other edges have been recovered in prior iterations (*AND rule*).

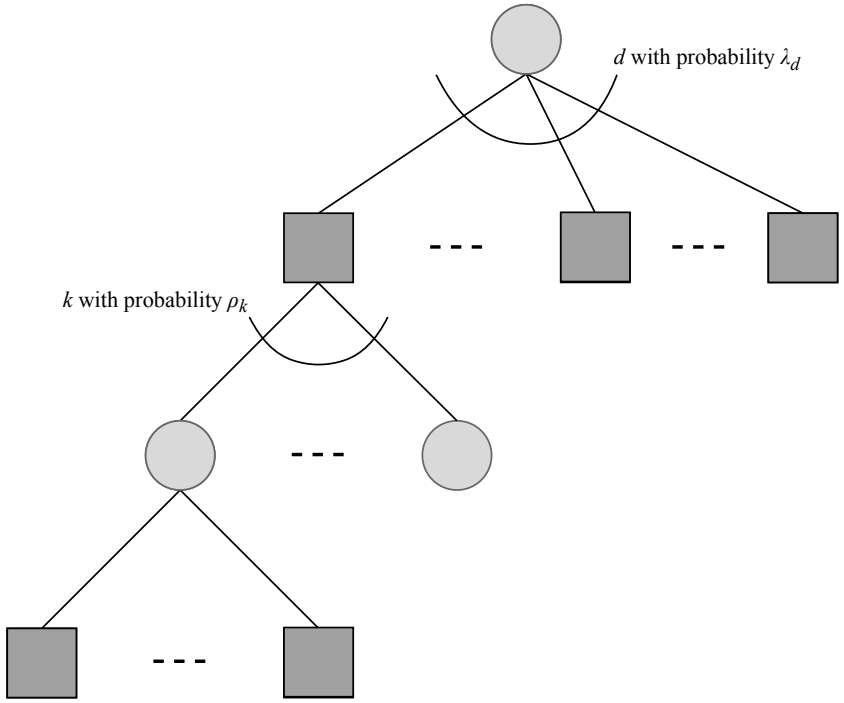


Figure 5.3: AND-OR tree.

We construct a tree using the edge-weighted bipartite graph $\mathcal{G}$, as illustrated in Fig. 5.3. The depth of the tree is twice the maximum allowed number of iterations $i_{\max}$. A burst node is the root of the tree at depth 0. The children of a burst node are those sum nodes which were chosen for the burst transmissions. Similarly, the sum nodes have those burst nodes as children who chose them for burst transmissions. A node at depth ζ has children in depth $\zeta + 1$. Thereby, for the proposed random access mechanism, burst nodes and sum nodes are located at depths $0, 2, 4, .., 2\,i_{\max}$ and $1, 3, 5, ..., 2\,i_{\max} - 1$, respectively. According to Algorithm 1, in each iteration the burst nodes at depth ζ send the messages to their parent sum nodes at depth $\zeta - 1$ and the sum nodes send messages to their parent burst nodes at depth $\zeta - 2$. *Note that a packet is lost if in the final iteration (i_{max}) the message received by the root burst node does not satisfy the SINR criterion. The probability of this event is denoted by* $q^{(i_{\max})}$.

The following proposition gives the main result for $q_m^{(i)}$.

Proposition 5. *In the asymptotic setting, as the number of slots $N \to \infty$, the recovery-error probability for a burst node $\mathcal{B}_m$ in the ith iteration, denoted by $q_m^{(i)}$,*

Algorithm 2 Iterative SIC as a message passing algorithm

1: Initialization: i=0, $\mathbf{B}_{n\to m}^{(0)}=0$, $\mathbf{S}_{n\to m}^{(0)}=\Theta_{m,n}^{(0)}$ $\forall m,n$.
2: **for** $i \geqslant 0$ && $i \leqslant i_{\max}$ && $\mathbf{S}_{n\to m}^{(i)}=0$ **do**
3: **for** $n \geqslant 1$ && $n \leqslant N$ **do**
4: **for** $\mathcal{M}_n \neq \{\}$ **do**
5: In ith iteration sum node $\mathcal{S}_n$ sends a message $\mathbf{S}_{n\to m}^{(i)}$ to burst node $\mathcal{B}_m$, $\forall m \in \mathcal{M}_n$

$$\mathbf{S}_{n\to m}^{(i)} = \mathbf{S}_{n\to m}^{(i-1)} - \sum_{m'\in\mathcal{M}_n\setminus\mathcal{B}_m} \mathbf{B}_{m'\to n}^{(i-1)}$$

6: Within ith iteration, burst node $\mathcal{B}_m$ sends a message $\mathbf{B}_{m\to n}^{(i)}$ to sum node $\mathcal{S}_n$, $\forall n \in \mathcal{N}_m$

$$\mathbf{B}_{m\to n}^{(i)} = \begin{cases} W_{m,n}, & \text{if } \mathbf{S}_{n\to m}^{(i)} - W_{m,n} \leqslant \varphi^{\text{th}} \\ 0, & \text{otherwise.} \end{cases}$$

7: **end for**
8: **end for**
9: **end for**

is given by

$$q_m^{(i)} = \sum_{d=1}^{D} \Lambda_d \left(1 - \sum_{l=1}^{L} \sum_{u=0}^{M-1} p^{(i)}(u) \mathbb{P}\left\{ \Theta_m^{(i)}(u) \leqslant \varphi^{th}(E_l) \right\} \Gamma_l \right)^d \tag{5.7}$$

where Λ_d *and* Γ_l *are defined in* (5.1), $\mathbb{P}\left\{\Theta_m^{(i)}(u) \leqslant \varphi^{th}(E_l)\right\}$ *is defined in* (5.5), $\Theta_m^{(i)} = \Theta_{m,n}^{(i)}$, *for* $n \in \mathcal{N}_m$, *since experiencing given interference power on any sum node is equally probable, and* $p^{(i)}(u)$ *is given by*

$$p^{(i)}(u) = \sum_{k=u}^{M-1} \Psi_k \binom{k}{u} \left(1 - q_m^{(i-1)}(k,u)\right)^{k-u} \left(q_m^{(i-1)}(k,u)\right)^u \tag{5.8}$$

where Ψ_k is defined in (5.1) and $q_m^{(i-1)} = 1$ for $i \leqslant 1$.

Proof. The proof is provided in Appendix D.1. Note that (5.5) cannot be expressed

in closed form but it can be found numerically using the following equation

$$\Theta_{m,n}^{(i)}(u) = \sum_{c=0}^{u} W_{c,n}, \quad 0 \leqslant u \leqslant M-1, \tag{5.9}$$

as discussed in the Appendix D.1. ■

5.4 Performance Optimization

We define the system load G as the normalised number of packets per slots. Since each device generates only one packet, we have $G = \frac{M}{N}$.

Based on the condition $q^{(i)} < q^{(i-1)}$ (note that $q^{(i)} = q_m^{(i)}$ since the recovery-error probability is the same for each device), we can define the maximum achievable load G^* such that recovery-error probability $q^{(i_{\max})}$ after $i_{\max}$ iterations will be almost zero for system load $G \leqslant G^*$ in the asymptotic setting when $N \to \infty$.

Mathematically, it is given as follows:

$$\begin{aligned} & G^* = \max \\ & \text{such that} \\ & q^{(i_{\max})} = 0. \end{aligned} \tag{5.10}$$

In this chapter, we are interested in finding the optimal transmit power probability distribution which will maximize G^* for a given degree distribution. This optimization problem can be formulated as follows.

$$\begin{aligned} \underset{\Gamma_l}{\text{maximize}} \quad & G^* \\ \text{subject to} \quad & E_1 = \sigma^2 \gamma^{\text{th}}, \\ & \overline{E\Lambda} = E_{\text{tot}}, \\ & \sum_l \Gamma_l = 1, \ \forall\, l, \\ & 0 \leqslant \Gamma_l \leqslant 1, \ \ \forall\, l, \end{aligned} \tag{5.11}$$

where the first constraint defines the criterion for the minimum power level, the second constraint mandates that the average power consumption must be equal

Table 5.1: System parameter values.

Name	**Symbol**	**Value**
Number of slots	N	1000
Max. number of iterations	$i_{\max}$	10000
Number of power levels	L	2
Probability metric	Γ_1	0.5
Probability metric	Γ_2	0.5
Power budget per frame	E_{tot}	3.52 dB
Transmit power level	E_1	0 dB
Transmit power level	E_2	6.02 dB
SINR threshold	γ^{th}	0 dB
Noise power	σ^2	0 dB

to the power budget per frame per device denoted by E_{tot}, where $\overline{E} = \sum_l \Gamma_l E_l$ represents the average power level and $\overline{\Lambda}$ is the average degree of repetition given as $\overline{\Lambda} = \sum_d \Lambda_d d$, the third constraint mandates that the sum of all transmit power probability levels is equal to one and the last constraint mandates that each probability level needs to be between zero and one.

5.5 Numerical Results

In this section, first we present the results to illustrate the advantage of employing transmit power diversity in conjunction with SINR based packet recovery in CRDSA, referred to as the *proposed scheme*. For comparison, we consider a *benchmark scheme* in which the random access transmission scheme does not employ transmit power diversity, however the rest of the transmission strategy and the recovery mechanism is the same as in Section 5.3.

5.5.1 Impact of Transmit Power Diversity

We compare the performance of the proposed scheme with the benchmark scheme by considering a simple transmit power diversity model, i.e., only two equally probable power levels. For both schemes the SINR based recovery model is used which employs SIC. Unless specified otherwise, the values for the parameters shown

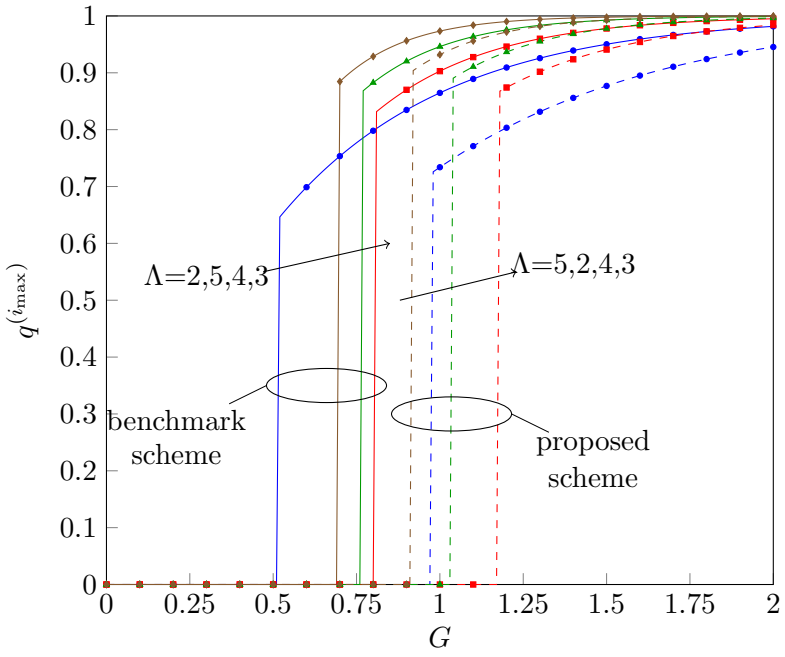


Figure 5.4: Recovery-error probability $q^{(i_{\max})}$ vs. system load G, for the clean packet model and SINR based model, in asymptotic setting, with $\Lambda = \{2, 3, 4, 5\}$.

in Table 5.1 are adopted. The transmit power level for each packet copy for the benchmark scheme is the same, i.e., $E_2 = 6.02$ dB.

Fig. 5.4 plots the recovery-error probability $q^{(i_{\max})}$ using Proposition 5 for both the proposed scheme and the benchmark scheme with repetition degrees $\Lambda = 2, 3, 4, 5$. We can see that for both schemes, the recovery-error probability is zero until a specific load value (G^*). As the load is further increased beyond G^*, the recovery-error probability abruptly escalates. As expected, the proposed scheme significantly outperforms the benchmark scheme for different repetition rates. The maximum achievable load G^* is tabulated in Table 5.2. This performance improvement of the proposed scheme is owing to the enhanced capture effect as a result of the transmit power diversity. We can see that, for the given system parameters, repetition degree 3 can achieve the highest load. This can be explained as follows. Transmitting multiple packets copies increases the probability of success but it also increases the level of interference. This tradeoff results in the best performance for a certain repetition degree (3 in this case) and the performance degrades after that.

Table 5.2: Maximum achievable load, G^*.

Λ	2	3	4	5
Benchmark scheme	0.52	0.81	0.77	0.70
Proposed scheme	0.98	1.18	1.04	0.82

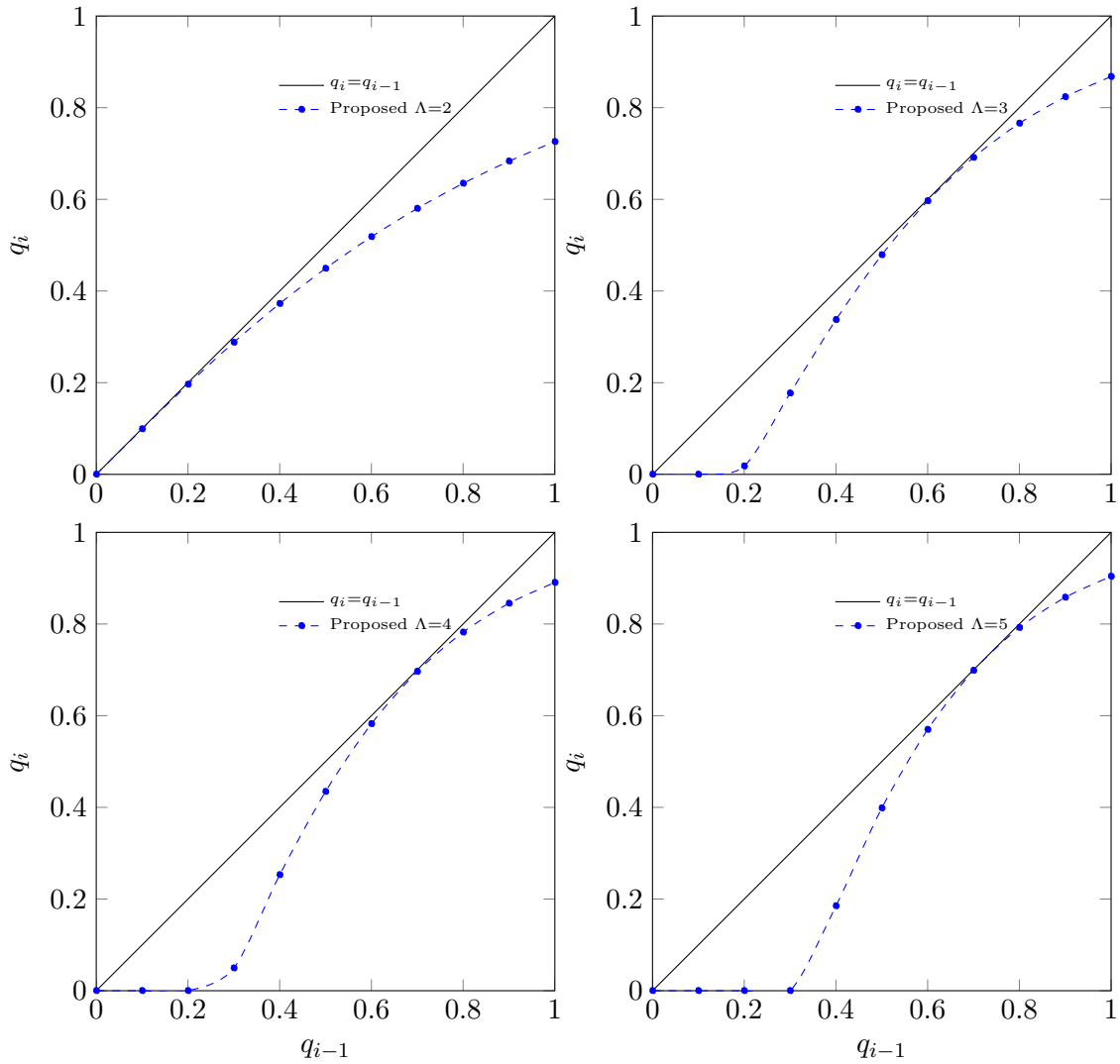


Figure 5.5: Evolution of error probability $q^{(i)}$ when operating at maximum achievable load $G^* = \{0.98, 1.18, 1.04, 0.92\}$ for $\Lambda = \{2, 3, 4, 5\}$.

Table 5.3: G^* for optimal probability distribution vector $\{\Gamma^*\}$.

Λ	G^* Uniform	G^* Optimal	$\{\Gamma^*\} = \{\Gamma_1\ \Gamma_2\ \Gamma_3\ \Gamma_4\ \Gamma_5\ \}$
2	1.62	1.68	{ 0.35 0.20 0.15 0.15 0.15 }
3	1.46	1.58	{ 0.55 0.05 0.10 0.05 0.25 }
4	1.07	1.33	{ 0.72 0.00 0.00 0.00 0.28 }
5	0.92	1.06	{ 0.55 0.00 0.00 0.00 0.45 }

5.5.2 Convergence Analysis

We investigate the convergence of the iterative SIC process at the maximum load G^*. Fig. 5.5 plots the evolution of the recovery-error probability in the asymptotic setting, when the system is operated at the maximum load G^*. It can be seen from Fig. 5.5 that the convergence condition $q^{(i)} < q^{(i-1)}$, defined above (5.10), is fulfilled.

Finally, we investigate the tightness of the result in Proposition 5 when the number of slots N is finite. Figs. 5.6 and 5.7 plot the recovery-error probability for $\Lambda = 2$ and $\Lambda = 3$ using Proposition 5 and simulations for N=100, 500, 1000, 5000, respectively. We can see that as N increases, the simulation results quickly approach the asymptotic performance.

5.5.3 Optimal Transmit Power Distribution

We solve (5.11) to determine the optimal transmit power distribution and investigate the improvement in performance. (5.11) can be efficiently solved using differential evolution optimization technique [119]. The parameter values used are as follows: $N = 1000, i_{\max} = 10000, \Lambda = 3, L = 5, \gamma^{\text{th}} = 0$ dB, $\sigma^2 = 0$ dB, $E_1 = 0$ dB, $E_2 = 3.77$ dB, $E_3 = 7.53$ dB, $E_4 = 11.30$ dB, $E_5 = 15.06$ dB, $E_{\text{tot}} = 20$ dB. For comparison, two probability models are considered:

- An uniform probability model in which each power level has equal probability of being selected.
- An optimized probability model in which each power level is assigned an optimal probability.

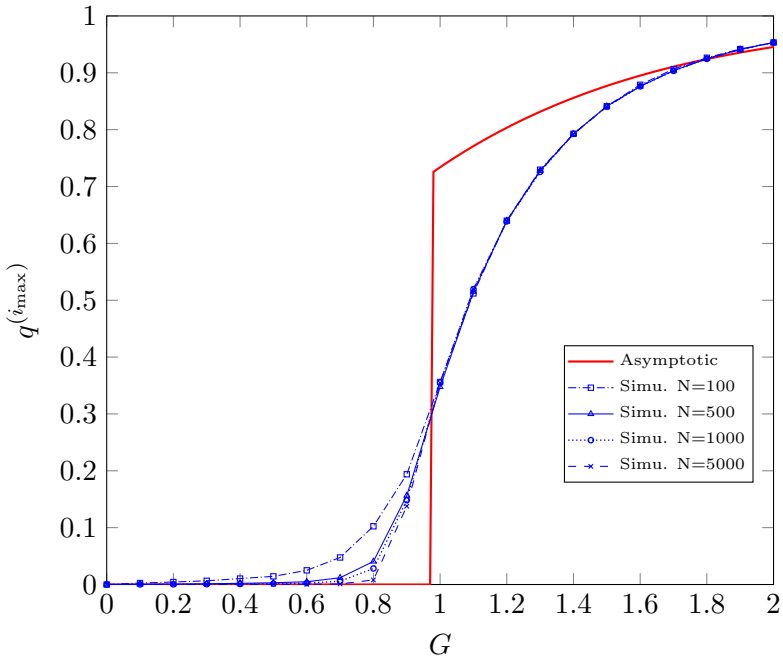


Figure 5.6: Recovery-error probability $q^{(i_{\max})}$ vs. system load G, for the proposed scheme with asymptotic setting, compared with simulations for $\Lambda = 2$.

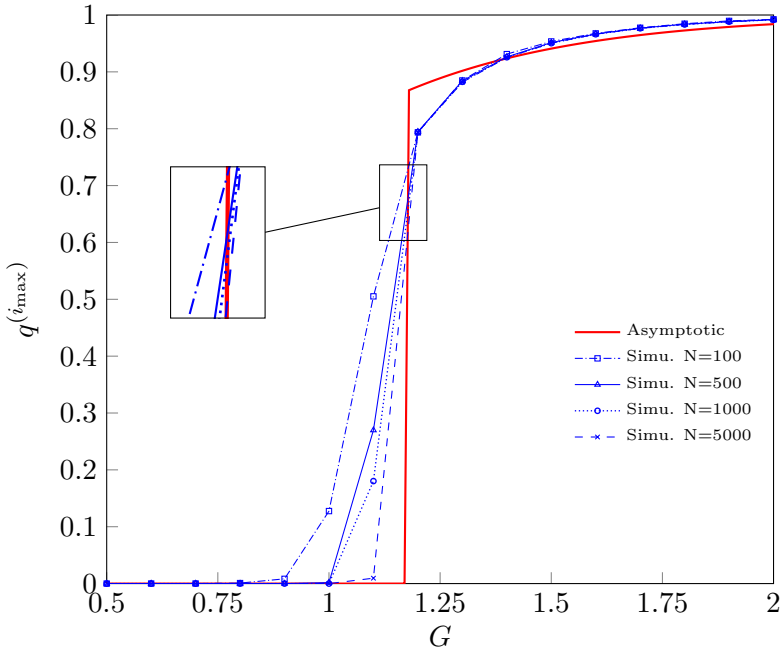


Figure 5.7: Recovery-error probability $q^{(i_{\max})}$ vs. system load G, for the proposed scheme with asymptotic setting, compared with simulations for $\Lambda = 3$.

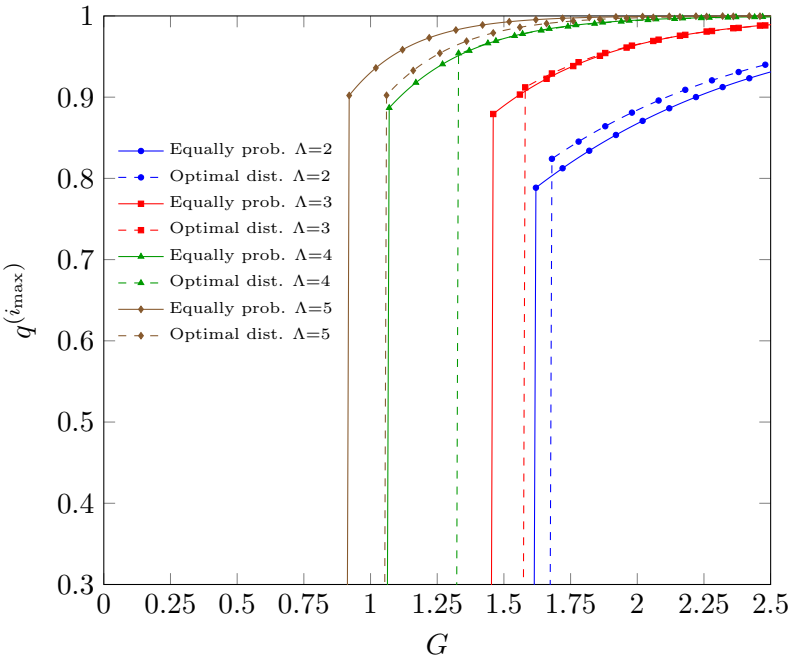


Figure 5.8: Recovery-error probability $q^{(i_{\max})}$ vs. system load G, for the proposed scheme with uniform and optimal distribution $\{\Gamma^*\}$, compared with simulations for $\Lambda = \{2, 3, 4, 5\}$.

The maximum achievable load G^* for both models is presented in Fig. 5.8 and Table 5.3. It can be seen that the optimized probability distribution achieves better performance as compared to uniform probability distribution. This is due to the fact that the optimized transmit power distributions maximize the probability of the gap between the power levels of two copies transmitted in a given slot. This has a greater impact on maximizing the capture effect. Hence, for the considered system parameters, as the number of copies Λ increases, from 2 to 5, intermediate power levels may not necessarily be used. Accordingly, their probability is zero in the optimized transmit power distribution.

Finally, we investigate the impact of the total power budget per device per frame (E_{tot}), which is an important constraint in (5.11). Fig. 5.9 plots the maximum achievable load G^* vs. E_{tot} for repetition rate $\Lambda = 2$ and other parameters same as before. It can be seen that a larger E_{tot} gives more room to exploit the transmit power diversity and thus boosts the performance. This leads to further significant performance improvement.

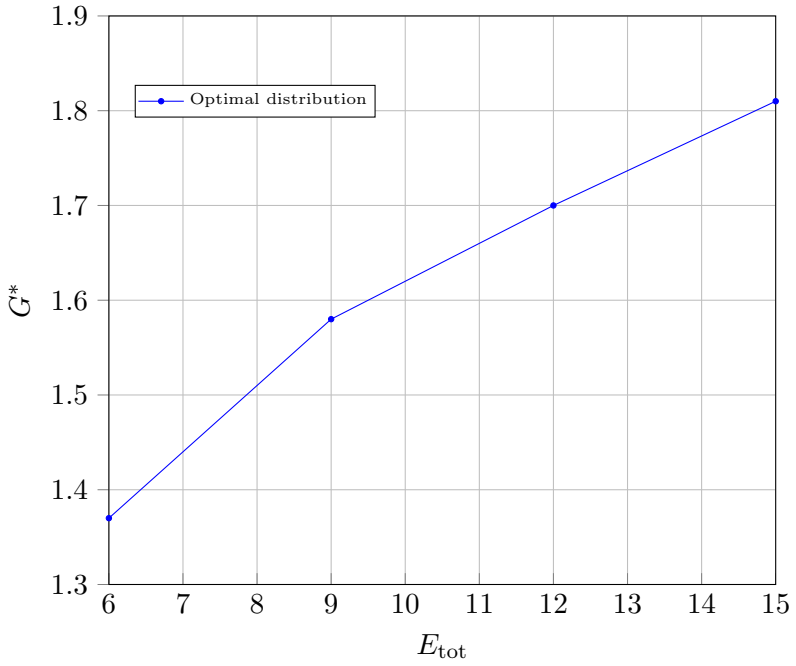


Figure 5.9: Maximum achievable load of CRDSA with transmit power diversity when the frame budget is varied.

5.6 Summary

In this correspondence, we have considered CRDSA with transmit power diversity and SINR based iterative SIC decoding, in the asymptotic frame length regime. We have proposed a framework for accurately predicting the performance of the system. The proposed framework is validated by simulation results. The results show promising improvements in maximum achievable system load, which is relevant to efficient random access for MTC scenario.

In this work, fading is not considered since fading cannot provide *inter-slot* receive power diversity. This is because two packet copies transmitted by a given MTD will under go the same fading experience. The recovery-error probability will be same for given slots if two MTDs happen to transmit at the same set of slots. It is due to the fact that the BS can receive different received power levels from the MTDs located at different locations, even if MTDs transmit their packet copies at the same power level. In this case, if two or more slots contain packet copies from the same MTDs then we might recover some packets from these slots, since

there is a received power diversity between the packet copies from different MTDs. However, in this scenario, there is no *inter-slot* received power diversity, i.e., the received power level for each copy from a given MTD will be same. Nevertheless, a fading power gain parameter can easily be included in the Proposition 5. However, it will substantially increase the overhead of simulations required to do the SIC based iterative convergence analysis of the optimization problem in (5.10). Due to this complexity, existing works [73,74,80] involving SIC based iterative convergence analysis of CRDSA, using the differential evolution optimization technique, do not consider fading.

Chapter 6

Conclusions and Future Research Directions

In this chapter, we summarize the general conclusions drawn from this thesis. We also outline some future research directions arising from this thesis.

In this thesis, we classify machine-type communication (MTC) use-cases in the following major challenges and requirements: deployment diversity, traffic heterogeneity, energy constraint, and operational efficiency. The intensity and further categorization of these major issues vary from one use-case to the other as discussed in Chapter 1. One the most crucial requirements and challenges for wireless MTC systems is the energy-efficient operation of the machine-type devices (MTDs). This is because in most cases MTDs are unattended and battery operated. In addition, MTC systems require low-latency, perpetual operation, massive-access, etc. This thesis focuses on the modelling, analysis and design of novel communication strategies for wireless MTC systems to realize the notion of Internet of things (IoT).

We consider sensor based MTDs which acquire physical information from the environment and transmit it to a base station (BS) while satisfying application specific quality-of-service (QoS) requirements. In particular, our overall objective is to minimize the energy cost of the MTDs. The considered network scenarios constitute fairly generic MTC system settings. As such, the design policies proposed and the insights and observations given in this thesis provide novel wireless solutions for a wide range of MTC use-cases and applications. The proposed novel

techniques and insights gained from this thesis aim to better utilize the limited energy resources of MTDs in order to effectively serve the future wireless networks.

Single-User With Limited Battery Capacity: We firstly consider a multi-user multi-channel sparse MTC network scenario which emulates practical use-cases requiring high reliability and low-latency. In such a system, each MTD is assigned a channel and transmit data on dedicated. The data may be generated or sensed by the MTD periodically, real-time, on demand (only when requested by the receiver), or due to the occurrence of an event of interest that is to be delivered under given reliability and delay conditions. For this model, designing a generic single-user communication policy would suffice and the same policy can be employed by all other users for corresponding parameters.

In Chapter 2, we investigated the joint optimization of compression and transmission strategy for an energy-constrained sensor MTD, and illustrated their trade-off. We showed that the joint optimization performs much better than only optimizing transmission without compression under any bit error rate (BER) and delay constraints. The performance gain observed ranges from 90% to 2000% and is most profound when the delay constraint is stringent. Overall, it is best to reduce compression and increase the transmission rate when the delay constraint gets more stringent and vice versa. The optimal level of compression is insensitive to the change in the BER requirement.

Single-User With Limited Battery Capacity: The second network scenario is evolved from the first scenario and considers an important aspect of energy-harvesting for a multi-user multi-channel sparse MTC system. Each MTD is battery powered which is recharged by an energy-harvesting source for perpetual operation. This network scenario cover loss tolerant and high reliability demanding MTC systems. As before, designing a generic single-user communication policy would suffice for this network scenario.

In Chapter 3, we investigated the joint wireless power transfer (WPT) and wireless information transfer (WIT) policies, employing data compression, to minimize the energy transferred by the harvesting access point (HAP) under given system constraints. For a set of practical parameter values, the joint optimization performs significantly better than only optimizing harvesting-time ratio and transmission rate without compression. Specifically, the gain is relatively large, up to

19%, when the delay constraint is stringent.

Multi-User Sequential Channel Access: The third network scenario is applicable to a multi-user single-channel sparse or massive MTC system. In such a system, all MTDs share a single channel and sequentially transmit data on this channel for interference free transmission. The data may be generated or sensed by the MTD periodically or due to the occurrence of an event of interest that is to be delivered under given reliability and delay conditions. This network scenario covers high reliability and low-latency MTC systems. For this network scenario, the communication policies for all MTDs are jointly devised for the given system parameters.

In Chapter 4, we investigated the joint optimization of sequencing and scheduling in a multi-user uplink machine-type communication scenario, considering adaptive compression and transmission rate control design. The energy efficiency performance is evaluated for three energy-minimization system objectives, which differ in terms of the overall system performance and fairness among the devices. Our results have showed that the proposed optimal scheme outperforms the schemes without multi-user sequencing. The improvement in energy efficiency observed is up to 45% when multi-user sequencing is optimized, under given maximum transmit power and delay constraints.

Multi-User Random Channel Access: The fourth network scenario is evolved from the third scenario. However, due to lack of coordination among devices, the MTDs randomly transmit short data-packets on a shared channel. This network scenario covers highly reliability and dense MTC systems. For this network scenario, the communication policies for all MTDs are jointly devised for given system parameters.

In Chapter 5, we investigated transmit power diversity and signal-to-interference-plus-noise ratio (SINR) based iterative successive interference cancellation (SIC) decoding for contention resolution diversity slotted ALOHA (CRDSA), in the asymptotic frame length regime. A mathematical framework is devised to carry out the performance analysis of the proposed scheme. The results show promising improvements in maximum achievable system load, which is relevant to efficient random access for massive MTC scenario. We show that the proposed strategy provides up to 88% system load performance improvement for massive-MTC.

In this thesis, we considered Rayleigh fading channel for network scenarios modelled in Chapters 2-4. Our focus in these network scenarios is on the benefit of data compression for the overall data transmission energy cost minimization and its relation with the availability of the channel state information at the transmitter side. Therefore, in our case a randomly distributed channel coefficient, i.e., Rayleigh distributed, served the purpose.

In industrial sensor network scenario, different fading models can be considered, e.g., Nakagami fading. Nevertheless, the overall principle of the proposed schemes for the considered network scenarios in Chapters 2-4 remains the same. The results for a different channel fading model can be derived following the same steps given in Chapters 2-4. Note that the insights and the conclusions drawn would still remain the same.

The data communication in the network scenarios considered in Chapter 2 are uncoded QAM transmission. An alternate way to achieve energy efficiency is to employ channel coding and increasing the bandwidth. Accordingly, a constraint on bandwidth can be defined and the computational cost can be balanced with the transmission cost. In future this approach to reduce energy cost can be investigated.

6.1 Future Research Directions

A number of emerging practical MTC use-cases call for the design and analysis of further novel communication solutions for energy-efficient operation of constrained MTDs. In this regard, a few promising research directions are briefly discussed in the following for future considerations.

Energy-efficient operation of MTDs under the coexistence of the ultra-reliable MTC (uMTC) and massive MTC (mMTC) systems with drastically diverse traffic types and QoS requirements is a critical research challenge. In addition, for many smart-city MTC use-cases the same network MTDs may generate sporadic data traffic and pose context-based QoS requirements. One way forward for such scenarios could be based on two-fold operation. First step could be to identify the dynamics of the data traffic and QoS requirements on-the-go. Second step could be to adapt the context-based communication policy design.

The short-packet communication in mMTC systems, under different low-latency

and limited computational resources at MTDs, poses crucial challenges for the implementation of cryptographic methods. The tradeoff between throughput and reliability for short-packet transmission has been extensively studied recently. However, the tradeoff between latency and energy efficiency has been given little attention for the implementation of security and privacy methods in mMTC systems. Therefore, we believe that this is an important research challenge for the future MTC systems calling for novel solutions.

Cooperative communication is an effective solution for multi-hop communication in terms of energy efficiency, spectrum efficiency reliability, scalability and coverage. Mostly, the routing technique for cooperation among devices and the transmission design are optimized separately because of modularity. For energy-constrained MTDs, a joint cross-layer optimization design for multi-hop MTC may provide advantageous operation. However, deploying too many relays, though increases diversity, reduces the energy efficiency. A future direction in this regard is to devise a joint cooperation and communication design to determine the optimal number of relays and transmission policy exploiting the diversity and energy efficiency tradeoff. Moreover, an incentive and fairness based relay selection design can also be employed for appropriate relay selection.

Appendix A

This appendix contains the proofs needed in Chapter 2.

A.1 Proof of Proposition

Substituting BER constraint (2.15) and SNR expression (2.16) in second constraint in (2.17) yields

$$\omega_2 \exp\left(- \frac{\omega_1}{(M-1)} \frac{\kappa P_{\mathrm{t}} |h|^2}{\sigma^2 d^{\alpha}} \right) \leqslant \phi. \tag{A.1.1}$$

For given instantaneous CGI, $|h|^2$, a lower bound on the transmit power, P_{t}, can be obtained by rearranging (A.1.1) as follows

$$P_{\mathrm{t}} \geqslant (1-M) \frac{\sigma^2 d^{\alpha} \ln\left(\phi/\omega_2\right)}{\omega_1 \kappa |h|^2}. \tag{A.1.2}$$

By substituting T_{cp}, T_{tx}, P_{tx}, r and ε from (2.1), (2.3), (2.9), (2.4) and (2.8), respectively, in (2.13) yields Ψ as a function of M, D_{cp} and P_{t}, which can be expressed as follows

$$\Psi \;=\; \tau D^{\beta+1} D_{\mathrm{cp}}^{-\beta} P_{\mathrm{cp}} \;-\; \tau D P_{\mathrm{cp}} \;+\; \frac{D_{\mathrm{cp}} T_{\mathrm{s}} \ln(2)}{\ln(M)} \left(\frac{3P_{\mathrm{t}}(M^{\frac{1}{2}}-1)}{\mu(M^{\frac{1}{2}}+1)} \;+\; P_{\mathrm{o}} \right). \tag{A.1.3}$$

For given values of M and D_{cp}, Ψ is an increasing function of P_{t}. Hence, the best choice of P_{t} to minimize Ψ while satisfying the constraint in (A.1.2) is the minimum value obtained by setting (A.1.2) with equality. Thus, P_{t} can be expressed as a function of M, as given in (2.19).

A.2 Proof of Theorem

It can be shown that (2.21) is not convex in M. By substitution of variable $M = \exp(z)$ in (2.21), $\widetilde{\Psi}$ can be equivalently defined as

$$\widetilde{\Psi}(z, D_{\mathrm{cp}}) \;=\; \tau D^{\beta+1} D_{\mathrm{cp}}^{-\beta} P_{\mathrm{cp}} - \tau D P_{\mathrm{cp}} - \frac{T_{\mathrm{s}} \ln(2)}{z D_{\mathrm{cp}}^{-1}} \left(\frac{3\Omega(\exp(z/2)-1)^2}{\mu |h|^2} - P_{\mathrm{o}} \right). \tag{A.2.1}$$

Accordingly, the problem defined in (2.17) can be equivalently given as follows

$$\begin{aligned} \underset{z, D_{\mathrm{cp}}}{\text{minimize}} \quad & \widetilde{\Psi}(z, D_{\mathrm{cp}}) \\ \text{subject to} \quad & \tau D^{\beta+1} D_{\mathrm{cp}}^{-\beta} - \tau D + z^{-1} D_{\mathrm{cp}} T_{\mathrm{s}} \ln(2) - T \leqslant 0, \\ & 2 - \exp(z) \leqslant 0, \\ & \exp(z) - \ln(M_{\max}) \leqslant 0, \\ & D_{\min} - D_{\mathrm{cp}} \leqslant 0, \\ & D_{\mathrm{cp}} - D \leqslant 0. \end{aligned} \tag{A.2.2}$$

For brevity we omit the proof, however using basic calculus and with some algebraic manipulation, it can be shown that the problem in (A.2.2) is a convex optimization problem. Lagrangian function for (A.2.2) can be given as in (A.2.3) shown at the top of the next page, where $\Lambda_i \in \mathbf{\Lambda} = \{\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5\}$ is the Lagrangian multiplier associated with the ith constraint.

$$\begin{aligned} \mathcal{L}(z, D_{\mathrm{cp}}, \mathbf{\Lambda}) \;=\; & \tau D P_{\mathrm{cp}} \Big(\frac{D^\beta}{D_{\mathrm{cp}}^\beta} - 1 \Big) - \frac{T_{\mathrm{s}} \ln(2)}{z D_{\mathrm{cp}}^{-1}} \left(\frac{3\Omega(\exp(z/2)-1)^2}{\mu |h|^2} - P_{\mathrm{o}} \right) \\ & + \Lambda_1 \left(\tau D \Big(\frac{D^\beta}{D_{\mathrm{cp}}^\beta} - 1 \Big) + \frac{T_{\mathrm{s}} \ln(2)}{z D_{\mathrm{cp}}^{-1}} - T \right) \\ & + \Lambda_2 \big(2 - \exp(z) \big) \\ & + \Lambda_3 \big(\exp(z) - \ln(M_{\max}) \big) \\ & + \Lambda_4 (D_{\min} - D_{\mathrm{cp}}) \\ & + \Lambda_5 (D_{\mathrm{cp}} - D), \end{aligned} \tag{A.2.3}$$

The Karush-Kuhn-Tucker (KKT) conditions for (A.2.2) are:

$$\begin{aligned}
&\tau D^{\beta+1} D_{\mathrm{cp}}^{-\beta} - \tau D + z^{-1} D_{\mathrm{cp}} T_{\mathrm{s}} \ln(2) - T \leqslant 0, \\
&2 - \exp(z) \leqslant 0, \quad \exp(z) - \ln(M_{\max}) \leqslant 0, \\
&D_{\min} - D_{\mathrm{cp}} \leqslant 0, \quad D_{\mathrm{cp}} - D \leqslant 0,
\end{aligned} \tag{A.2.4a}$$

$$\Lambda_1 \geqslant 0, \;\; \Lambda_2 \geqslant 0, \;\; \Lambda_3 \geqslant 0, \;\; \Lambda_4 \geqslant 0, \;\; \Lambda_5 \geqslant 0, \tag{A.2.4b}$$

$$\begin{aligned}
&\Lambda_1 \Big(\tau D^{\beta+1} D_{\mathrm{cp}}^{-\beta} - \tau D + z^{-1} D_{\mathrm{cp}} T_{\mathrm{s}} \ln(2) - T \Big) = 0, \\
&\Lambda_2 (2 - \exp(z)) = 0, \quad \Lambda_3 (\exp(z) - \ln(M_{\max})) = 0, \\
&\Lambda_4 (D_{\min} - D_{\mathrm{cp}}) = 0, \quad \Lambda_5 (D_{\mathrm{cp}} - D) = 0,
\end{aligned} \tag{A.2.4c}$$

$$\nabla_{z, D_{\mathrm{cp}}} \mathcal{L}(z, D_{\mathrm{cp}}, \Lambda) = \left[\frac{\partial \mathcal{L}}{\partial z} \;\; \frac{\partial \mathcal{L}}{\partial D_{\mathrm{cp}}} \right]^{\top} = [0 \;\; 0]^{\top}. \tag{A.2.4d}$$

where ∇ is the gradient operator and $[\cdot]^{\top}$ is the transpose operator.

We first determine the optimal constellation size which will minimize (A.2.1) and then use it to get the optimal transmit power and compression ratio. It can be shown that $\widetilde{\Psi}$ in (A.2.1) is convex in z and there lies a global minima. From (A.2.4d) we have $\left[\frac{\partial \mathcal{L}}{\partial z} \;\; \frac{\partial \mathcal{L}}{\partial D_{\mathrm{cp}}} \right]^{\top} = [0 \;\; 0]^{\top}$. Taking partial derivative of (A.2.3) with respect to z and setting $\frac{\partial \mathcal{L}}{\partial z} = 0$ and after simplification we get

$$\frac{3\Omega T_{\mathrm{s}}}{\mu |h|^2} \Big(\exp\big(z/2\big) - 1 \Big) \Big((z-1) \exp\big(z/2\big) + 1 \Big) + T_{\mathrm{s}} P_{\mathrm{o}} + \Lambda_1 T_{\mathrm{s}} + \frac{(\Lambda_2 - \Lambda_3) z^2 \exp(z)}{D_{\mathrm{cp}} \ln(2)} = 0. \tag{A.2.5}$$

Similarly, it can also be shown that $\widetilde{\Psi}$ in (A.2.1) is convex in D_{cp} and there lies a global minima. Taking partial derivative of (A.2.3) with respect to D_{cp} and setting $\frac{\partial \mathcal{L}}{\partial D_{\mathrm{cp}}} = 0$ and after simplification we get

$$\begin{aligned}
\frac{3\Omega T_{\mathrm{s}}}{\mu |h|^2} \big(\exp\big(z/2\big) - 1 \big)^2 + T_{\mathrm{s}} P_{\mathrm{o}} + \Lambda_1 \left(T_{\mathrm{s}} - \frac{z \tau \beta D^{\beta+1}}{\ln(2) D_{\mathrm{cp}}^{\beta+1}} \right) \\
- \frac{\Lambda_4 z}{\ln(2)} + \frac{\Lambda_5 z}{\ln(2)} = \frac{z \tau \beta D^{\beta+1} P_{\mathrm{cp}}}{\ln(2) D_{\mathrm{cp}}^{\beta+1}}.
\end{aligned} \tag{A.2.6}$$

From complimentary slackness condition (A.2.4c) we know either Λ_i is zero or the associated constraint function is zero for any given i. First we consider one of the possible cases that is Λ_1, $\Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5$ do not exist, i.e., unconstrained minimization. Accordingly, plugging in Λ_1=0, Λ_2=0, Λ_3=0, Λ_4=0, Λ_5=0 in (A.2.5), (A.2.6) yields following expressions, respectively,

$$\frac{3\Omega}{\mu|h|^2}\Big(\exp\big(z/2\big)-1\Big)\Big((z-1)\exp\big(z/2\big)+1\Big) + P_{\mathrm{o}} = 0, \tag{A.2.7}$$

$$-\frac{z\tau\beta D^{\beta+1}P_{\mathrm{cp}}}{T_{\mathrm{s}}D_{\mathrm{cp}}^{\beta+1}\ln(2)} + \frac{3\Omega}{\mu|h|^2}\Big(\exp\big(z/2\big)-1\Big)^2 + P_{\mathrm{o}} = 0. \tag{A.2.8}$$

Solving (A.2.8) for D_{cp} yields

$$D_{\mathrm{cp}} = D\left(\frac{z\tau\beta P_{\mathrm{cp}}}{\frac{3\Omega T_{\mathrm{s}}\ln(2)}{\mu|h|^2}\big(\exp(z/2)-1\big)^2 + P_{\mathrm{o}}T_{\mathrm{s}}\ln(2)}\right)^{\frac{1}{\beta+1}}. \tag{A.2.9}$$

Numerically solving (A.2.7) for z yields its value $\widetilde{z}$. Substituting this value of z in (A.2.9) and solving for D_{cp} yields its value $\widetilde{D}_{\mathrm{cp}}$. $\widetilde{z}$ and $\widetilde{D}_{\mathrm{cp}}$ provide a lower bound on optimization problem in (A.2.2) for given instantaneous CGI, $|h|^2$. It can be shown that $\widetilde{z}$ and $\widetilde{D}_{\mathrm{cp}}$ satisfy all the KKT conditions when the first constraint in (A.2.2) is slack, i.e.,

$$\tau D^{\beta+1}\widetilde{D}_{\mathrm{cp}}^{-\beta} - \tau D + \widetilde{z}^{-1}\widetilde{D}_{\mathrm{cp}}T_{\mathrm{s}}\ln(2) - T < 0, \tag{A.2.10}$$

and other constraints are also slack. Thus, the optimal lagrange multiplier Λ_1, $\Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5$ are zero. Hence, the derived solution in (A.2.7) and (A.2.9) is the optimal solution for the optimization problem in (A.2.2) for given instantaneous CGI, $|h|^2$, when all constraints in (A.2.2) are slack.

Now consider another possible case when the first constraint in (A.2.2) is not slack, i.e., Λ_1 exits and $\Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5$ do not exist. Thus, plugging in $\Lambda_1 \neq 0$, Λ_2=0, Λ_3=0, Λ_4=0, Λ_5=0 in (A.2.5), (A.2.6) and the complimentary slackness conditions (A.2.4c) yields following expressions, respectively,

$$\frac{3\Omega}{\mu|h|^2}\Big(\exp(z/2)-1\Big)\Big((z-1)\exp\big(z/2\big)+1\Big) + P_{\mathrm{o}} + \Lambda_1 = 0, \tag{A.2.11}$$

$$\Lambda_1\Big(1 - \frac{z\tau\beta D^{\beta+1}}{T_{\rm s}D_{\rm cp}^{\beta+1}\ln(2)}\Big) - \frac{z\tau\beta D^{\beta+1}P_{\rm cp}}{T_{\rm s}D_{\rm cp}^{\beta+1}\ln(2)} + \frac{3\Omega}{\mu|h|^2}\Big(\exp\big(z/2\big)-1\Big)^2 + P_{\rm o} = 0. \tag{A.2.12}$$

Solving (A.2.11) for Λ_1 and substituting its value in (A.2.12) yields (A.2.13) shown at the top of the next page.

$$\Big(\frac{3\Omega}{\mu|h|^2}\Big(\exp(z/2)-1\Big)\Big((z-1)\exp(z/2)+1\Big)+P_{\rm o}\Big)\Big(1-\frac{z\tau\beta D^{\beta+1}}{T_{\rm s}D_{\rm cp}^{\beta+1}\ln(2)}\Big) + \frac{z\tau\beta D^{\beta+1}P_{\rm cp}}{T_{\rm s}D_{\rm cp}^{\beta+1}\ln(2)} - \frac{3\Omega}{\mu|h|^2}\big(\exp(z/2)-1\big)^2 = P_{\rm o}. \tag{A.2.13}$$

Solving (A.2.13) for $D_{\rm cp}$ yields

$$D_{\rm cp} = D\upsilon^{\frac{1}{\beta+1}}, \tag{A.2.14}$$

where

$$\upsilon = \frac{P_{\rm cp}-P_{\rm o}-\frac{3\Omega}{\mu|h|^2}\big(\exp(z/2)-1\big)\big((z-1)\exp(z/2)+1\big)}{\frac{3\Omega}{\tau\beta\mu|h|^2}\big(\exp(z/2)-\exp(z)\big)}. \tag{A.2.15}$$

Substituting $D_{\rm cp}$ from (A.2.14) in (A.2.13) yields

$$\frac{T}{D}+\tau-\tau\upsilon^{\frac{-\beta}{\beta+1}} = \frac{T_{\rm s}\ln(2)}{z}\upsilon^{\frac{1}{\beta+1}}. \tag{A.2.16}$$

Numerically solving (A.2.12) for z yields its value $\widehat{z}$. Substituting this value of z in (A.2.11) and solving for $D_{\rm cp}$ yields its value $\widehat{D}_{\rm cp}$. The value of z can be obtained by numerically solving (A.2.12). Substituting value of z in (A.2.11) and solving for $D_{\rm cp}$ yields its value. It can be shown that $\widehat{z}$ and $\widehat{D}_{\rm cp}$ satisfy all the KKT conditions when all constraints in (A.2.2) are slack except the first constraint, thus the optimal lagrange multiplier Λ_1 is positive and $\Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5$ are zero. Hence, the derived solution in (A.2.14) and (A.2.16) is the optimal solution for the optimization problem in (A.2.2), when all constraints in (A.2.2) are slack except the first constraint. For all other cases, similar steps can be followed and it can be shown that these cases violate one or more constraints.

The problem in (A.2.2) is equivalent to (2.17), thus the optimal values of M and

D_{cp} for both cases can be obtained by substituting $z = \ln(M)$ in (A.2.7), (A.2.9) and (A.2.16), (A.2.14), respectively, which will minimize the objective function in (2.17). Finally, by substituting the optimal value of M in (2.19) we can determine the optimal P_{t} which will minimize Ψ for given instantaneous CGI, $|h|^2$.

A.3 Proof of Proposition

Substituting BER constraint (2.15) and SNR expression (2.16) in second constraint in (2.32) yields

$$\mathbb{P}\left\{\omega_2 \exp\left(- \frac{\omega_1}{(M-1)} \frac{\kappa P_{\mathrm{t}} |h|^2}{\sigma^2 d^{\alpha}} \right) \leqslant \phi \right\} \geqslant \vartheta. \tag{A.3.1}$$

Solving (A.3.1) for fading power gain, $|h|^2$, yields

$$\mathbb{P}\left\{ |h|^2 \geqslant (1-M) \frac{\sigma^2 d^{\alpha} \ln(\phi/\omega_2)}{\omega_1 \kappa P_{\mathrm{t}}} \right\} \geqslant \vartheta. \tag{A.3.2}$$

The left hand side of (A.3.2) represents the complimentary cumulative distribution function (CCDF) for $|h|^2$. For our considered Rayleigh fading channel, the fading power gain, $|h|^2$, is exponentially distributed. (A.3.2) can be given as follows

$$1 - \left[1 - \exp\left((M-1) \frac{\sigma^2 d^{\alpha} \ln(\phi/\omega_2)}{\varsigma \omega_1 \kappa P_{\mathrm{t}}} \right) \right] \geqslant \vartheta, \tag{A.3.3}$$

where ς represents the scale parameter of the probability distribution. Solving (A.3.3) for P_{t} yields

$$P_{\mathrm{t}} \geqslant (M-1) \frac{\sigma^2 d^{\alpha} \ln(\phi/\omega_2)}{\varsigma \omega_1 \kappa \ln(\vartheta)}. \tag{A.3.4}$$

By substituting T_{cp}, T_{tx}, P_{tx}, r and ε from (2.1), (2.3), (2.9), (2.4) and (2.8), respectively, in (2.13) yields Ψ as a function of M, D_{cp} and P_{t}, which can be expressed as follows

$$\Psi \;=\; \tau D^{\beta+1} D_{\mathrm{cp}}^{-\beta} P_{\mathrm{cp}} \;-\; \tau D P_{\mathrm{cp}} \;+\; \frac{D_{\mathrm{cp}} T_{\mathrm{s}} \ln(2)}{\ln(M)} \left(\frac{3 P_{\mathrm{t}} (M^{\frac{1}{2}} - 1)}{\mu (M^{\frac{1}{2}} + 1)} \;+\; P_{\mathrm{o}} \right). \tag{A.3.5}$$

For given values of M and D_{cp}, Ψ is an increasing function of P_{t}. Hence, the best choice of P_{t} to minimize Ψ while satisfying the constraint in (A.3.4) is the minimum value obtained by setting (A.3.4) with equality. Thus, P_{t} can be expressed as a function of M, as given in (2.33).

Appendix B

This appendix contains the proof needed in Chapter 3.

B.1 Proof of Theorem

Lagrangian function for (3.14) can be given as

$$\begin{aligned}\mathcal{L}(\rho_{\mathrm{MTC}}, z, D_{\mathrm{cp}}, \mathbf{\Lambda}) = \rho_{\mathrm{MTC}} \\ &+ \Lambda_1\Big(P_{\mathrm{sen}}T_{\mathrm{sen}}+\frac{\tau D^{\beta+1}P_{\mathrm{cp}}}{D_{\mathrm{cp}}^{\beta}}-\tau DP_{\mathrm{cp}}+\frac{b(\exp(z)+c)}{z\widehat{D}_{\mathrm{cp}}^{-1}}+\eta\rho_{\mathrm{MTC}}Td\Big) \\ &+ \Lambda_2\Big(T_{\mathrm{sen}}+\frac{\tau D^{\beta+1}}{D_{\mathrm{cp}}^{\beta}}-\tau D+\frac{\ln(2)}{BzD_{\mathrm{cp}}^{-1}}-T+\rho_{\mathrm{MTC}}T\Big) \\ &- \Lambda_3\rho_{\mathrm{MTC}} \\ &+ \Lambda_4(\rho_{\mathrm{MTC}}-1) \\ &+ \Lambda_5\Big(\frac{2}{\ln(2)}-z\Big) \\ &+ \Lambda_6(D_{\min}-D_{\mathrm{cp}}) \\ &+ \Lambda_7(D_{\mathrm{cp}}-D),\end{aligned} \tag{B.1.1}$$

where $\Lambda_i \in \mathbf{\Lambda} = \{\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4\}$ is the Lagrangian multiplier associated with the ith constraint.

The Karush-Kuhn-Tucker (KKT) complementary slackness and optimality con-

ditions for (3.14) are as follows:

$$\begin{aligned}
&\Lambda_1\Big(\frac{D_{\rm cp}b}{z}\big(\exp(z)+c\big)+\frac{\tau D^{\beta+1}}{D_{\rm cp}^{\beta}P_{\rm cp}^{-1}}-\tau DP_{\rm cp}+\eta\rho_{\rm MTC}Td+P_{\rm sen}T_{\rm sen}\Big)=0,\\
&\Lambda_2\Big(T_{\rm sen}+\frac{\tau D^{\beta+1}}{D_{\rm cp}^{\beta}}-\tau D+\frac{\ln(2)}{BzD_{\rm cp}^{-1}}-T+\rho_{\rm MTC}T\Big)=0,\\
&\Lambda_3(-\rho_{\rm MTC})=0,\ \ \Lambda_4(\rho_{\rm MTC}-1)=0,\ \ \Lambda_5\Big(\frac{2}{\ln(2)}-z\Big)=0,\\
&\Lambda_6(D_{\rm min}-D_{\rm cp})=0,\ \ \Lambda_7(D_{\rm cp}-D)=0,
\end{aligned}\tag{B.1.2a}$$

$$\Big[\frac{\partial\mathcal{L}}{\partial\rho_{\rm MTC}}\ \frac{\partial\mathcal{L}}{\partial z}\ \frac{\partial\mathcal{L}}{\partial D_{\rm cp}}\Big]^{\top}=[0\ 0\ 0]^{\top}.\tag{B.1.2b}$$

where $[\cdot]^{\top}$ is the transpose operator.

Taking the partial derivative of (B.1.1) with respect to $\rho_{\rm MTC}$ and by setting $\frac{\partial\mathcal{L}}{\partial\rho_{\rm MTC}}=0$ we get

$$1+\Lambda_1\eta Td+\Lambda_2T-\Lambda_3=0.\tag{B.1.3}$$

Taking partial derivative of (B.1.1) with respect to z and setting $\frac{\partial\mathcal{L}}{\partial z}=0$ and after some simplification we get

$$\frac{\Lambda_1D_{\rm cp}b}{z^2}\Big((z-1)\exp(z)-c\Big)-\frac{\Lambda_2D_{\rm cp}\ln(2)}{Bz^2}-\Lambda_5=0.\tag{B.1.4}$$

Taking partial derivative of (B.1.1) with respect to $D_{\rm cp}$ and setting $\frac{\partial\mathcal{L}}{\partial D_{\rm cp}}=0$ and after some simplification we get

$$\begin{aligned}
\Lambda_1\Big(-\frac{\tau\beta D^{\beta+1}P_{\rm cp}}{D_{\rm cp}^{\beta+1}}+\frac{b}{z}\big(\exp(z)+c\big)\Big)+&\\
\Lambda_2\Big(-\frac{\tau\beta D^{\beta+1}}{D_{\rm cp}^{\beta+1}}+\frac{\ln(2)}{Bz}\Big)-\Lambda_6+\Lambda_7=0.&
\end{aligned}\tag{B.1.5}$$

From (B.1.2a) we know either Λ_i is zero or the associated constraint function is zero for any given i. First we consider one of the possible cases that is Λ_1 exists and $\Lambda_2,\Lambda_3,\Lambda_4,\Lambda_5,\Lambda_6,\Lambda_7$ do not exist. Accordingly, solving (B.1.3) for Λ_1 yields its value

$$\widetilde{\Lambda}_1=-\frac{1}{\eta Td}.\tag{B.1.6}$$

Substituting the value of Λ_1 from (B.1.6) into (B.1.4), yields

$$-\frac{D_{\text{cp}}b}{z^2\eta T d}\big((z-1)\exp(z)-c\big)=0. \tag{B.1.7}$$

With some algebraic manipulation, solving (B.1.7) for z yields its value given as follows

$$\widetilde{z}=\mathrm{W}_0\Big(\exp\big(\ln(c)-1\big)\Big)+1, \tag{B.1.8}$$

where $\mathrm{W}_0(\cdot)$ is the principle branch of the Lambert W function. Substituting the value of z from (B.1.8) into (B.1.5), and solving for D_{cp} yields its value given as follows

$$\widetilde{D}_{\text{cp}}=D\Big(\frac{\tau\beta P_{\text{cp}}\widetilde{z}}{b\exp(\widetilde{z})+bc}\Big)^{\frac{1}{\beta+1}}. \tag{B.1.9}$$

Since, $\Lambda_1\neq 0$, thereby, from (B.1.2a), we have

$$\frac{\eta T d}{\rho_{\text{MTC}}^{-1}}=\frac{\tau D}{P_{\text{cp}}^{-1}}-\frac{T_{\text{sen}}}{P_{\text{sen}}^{-1}}-\frac{\tau D^{\beta+1}}{P_{\text{cp}}^{-1}\widetilde{D}_{\text{cp}}^{\beta}}-\frac{\widetilde{D}_{\text{cp}}b}{\widetilde{z}}\big(\exp(\widetilde{z})+c\big). \tag{B.1.10}$$

Solving (B.1.10) for ρ_{MTC} yields its value given as follows

$$\widetilde{\rho}_{\text{MTC}}=\frac{1}{\eta T d}\left(\frac{\tau D}{P_{\text{cp}}^{-1}}-\frac{T_{\text{sen}}}{P_{\text{sen}}^{-1}}-\frac{\tau D^{\beta+1}}{P_{\text{cp}}^{-1}\widetilde{D}_{\text{cp}}^{\beta}}-\frac{\exp(\widetilde{z})+c}{\widetilde{z}b^{-1}\widetilde{D}_{\text{cp}}^{-1}}\right). \tag{B.1.11}$$

It can be shown that $\widetilde{z}$, $\widetilde{D}_{\text{cp}}$, $\widetilde{\rho}_{\text{MTC}}$ satisfy all the KKT conditions and thus are optimal for the problem in (3.14), when all constraints in (3.14) are slack, except the first constraint.

Now consider another possible case, when both Λ_1 and Λ_2 exist and $\Lambda_3,\Lambda_4,\Lambda_5$, Λ_6,Λ_7 do not exist. Accordingly, solving (B.1.3) for Λ_2 yields its value

$$\widehat{\Lambda}_2=-\frac{1}{T}-\Lambda_1\eta d. \tag{B.1.12}$$

Substituting the value of Λ_2 from (B.1.12) into (B.1.4) and solving for Λ_1, yields its value

$$\widehat{\Lambda}_1=\frac{\frac{\ln(2)}{TB}}{b(z-1)\exp(z)-bc-\frac{\eta d\ln(2)}{B}}. \tag{B.1.13}$$

Substituting the value of Λ_2 from (B.1.12) into (B.1.5) and solving for D_{cp} yields its value given as follows

$$\widehat{D}_{\text{cp}} = D\left(\frac{z\tau\beta\left(T^{-1}\widehat{\Lambda}_1^{-1}+\eta d-P_{\text{cp}}\right)}{\frac{\widehat{\Lambda}_1^{-1}\ln(2)}{TB}+\frac{\eta d\ln(2)}{B}-b\exp(z)-bc}\right)^{\frac{1}{\beta+1}}, \tag{B.1.14}$$

where $\widehat{\Lambda}_1$ is a function of z and is defined in (B.1.13).

Since, $\Lambda_2 \neq 0$, accordingly, from (B.1.2a), we have

$$T_{\text{sen}} + \tau D^{\beta+1}\widehat{D}_{\text{cp}}^{-\beta} - \tau D + \frac{\widehat{D}_{\text{cp}}\ln(2)}{Bz} - T + \rho_{\text{MTC}}T = 0. \tag{B.1.15}$$

Solving (B.1.15) for ρ_{MTC} yields its value given as follows

$$\widehat{\rho}_{\text{MTC}} = \frac{1}{T}\left(T - T_{\text{sen}} - \tau D^{\beta+1}\widehat{D}_{\text{cp}}^{-\beta} + \tau D - \frac{\widehat{D}_{\text{cp}}\ln(2)}{Bz}\right). \tag{B.1.16}$$

Also, $\Lambda_1 \neq 0$, thereby, from (B.1.2a), we have

$$\frac{T_{\text{sen}}}{P_{\text{sen}}^{-1}}+\frac{\tau D^{\beta+1}}{P_{\text{cp}}^{-1}\widehat{D}_{\text{cp}}^{\beta}}-\frac{\tau D}{P_{\text{cp}}^{-1}}+\frac{b(\exp(z)+c)}{z\widehat{D}_{\text{cp}}^{-1}}+\frac{\eta\widehat{\rho}_{\text{MTC}}}{(Td)^{-1}} = 0. \tag{B.1.17}$$

Numerically solving (B.1.17) for z yields its value $\widehat{z}$. Substituting this value of z in (B.1.12), (B.1.13), (B.1.14), (B.1.16) yields the values of $\widehat{\Lambda}_2$, $\widehat{\Lambda}_1$, $\widehat{D}_{\text{cp}}$ and $\widehat{\rho}_{\text{MTC}}$, respectively. It can be shown that $\widehat{D}_{\text{cp}}$, $\widehat{\rho}_{\text{MTC}}$, $\widehat{z}$ satisfy all the KKT conditions and thus are optimal for (3.14), when all constraints are slack except the first and second constraint. It can be shown that all other cases for the lagrange multipliers violate one or more KKT conditions.

The problem in (3.14) is equivalent to (3.12), thus the optimal values of D_{cp}, ρ_{MTC}, P_{IT}, for both cases can be obtained by substituting $z = \ln\left(1-\frac{\kappa\lambda\ln(\delta)P_{\text{IT}}}{\sigma^2 r^{\alpha}\Gamma}\right)$ in (B.1.12), (B.1.13), (B.1.14), (B.1.16), which will minimize the objective function in (3.12).

Appendix C

This appendix contains the proofs needed in Chapter 4.

C.1 Proof of Lemma 1

It can be shown that the energy cost of the ith device, E_i, defined in (4.14), is non-convex in P_i. By substitution of variable $Z_i = \ln\big(1+\kappa\frac{P_i|h_i|^2}{\Gamma\sigma^2 d_i^\alpha}\big)$ in (4.14), E_i can equivalently be expressed as

$$E_i\big(Z_i, D_{\mathrm{cp},i}\big) = \tau D_i P_{\mathrm{cp}}\big(\big(\frac{D_i}{D_{\mathrm{cp},i}}\big)^\beta - 1\big) + \frac{D_{\mathrm{cp},i} b_i}{Z_i}\big(\exp(Z_i) + c_i\big), \tag{C.1.1}$$

where $b_i = \frac{\Gamma\sigma^2 d_i^\alpha \ln(2)}{\mu B \kappa |h_i|^2}$, $c_i = \frac{\mu\kappa|h_i|^2 P_{\mathrm{o}}}{\Gamma\sigma^2 d_i^\alpha} - 1$. Substituting Z_i in constraints (4.15c), (4.15e) and (4.15h) yields

$$x_{1,i}\tau D_i\big(\big(\frac{D_i}{D_{\mathrm{cp},i}}\big)^\beta - 1\big) + x_{1,i}\frac{D_{\mathrm{cp},i}}{BZ_i} \leqslant T_1, \quad \forall\, i, \tag{C.1.2a}$$

$$\sum_{n=2}^{N} x_{n,i}\frac{D_{\mathrm{cp},i}}{BZ_i} \leqslant \sum_{n=2}^{N} x_{n,i} T_n, \quad \forall\, i, \tag{C.1.2b}$$

$$0 \leqslant Z_i \leqslant Z_{\max}, \quad \forall\, i, \tag{C.1.2c}$$

where $Z_{\max} = \ln\big(1+\kappa\frac{P_{\max}|h_i|^2}{\Gamma\sigma^2 d_i^{\alpha}}\big)$. Accordingly, problems (4.16) and (4.17) respectively become as

$$\begin{aligned} \widetilde{\mathbf{P}}_{\mathbf{SUM}}: \quad & \underset{Z_i,\, D_{\mathrm{cp},i},\, T_n,\, \forall\, n,i}{\text{minimize}} \quad && \sum_{i=1}^{N} E_i(Z_i, D_{\mathrm{cp},i}) \\ & \text{subject to} && \text{(4.15b), (4.15d), (C.1.2a), (C.1.2b), (C.1.2c),} \\ & && \text{(4.15i), (4.15j),} \end{aligned} \tag{C.1.3}$$

$$\begin{aligned} \widetilde{\mathbf{P}}_{\mathbf{MM}}: \quad & \underset{Z_i,\, D_{\mathrm{cp},i},\, T_n,\, \forall\, n,i}{\text{minimize}} \quad && \max_{1\leqslant i\leqslant N}\big\{E_i(Z_i, D_{\mathrm{cp},i})\big\} \\ & \text{subject to} && \text{(4.15b), (4.15d), (C.1.2a), (C.1.2b), (C.1.2c),} \\ & && \text{(4.15i), (4.15j).} \end{aligned} \tag{C.1.4}$$

Using basic calculus and with some algebraic manipulation, it can be shown that $E_i\big(Z_i, D_{\mathrm{cp},i}\big)$ in (C.1.1) and constraint functions in (4.15d), (C.1.2a) and (C.1.2b), respectively, are jointly convex in Z_i and $D_{\mathrm{cp},i}$, $\forall\, i$. For brevity, we omit the detailed proof of this result here. Also, because the sum of convex functions is convex and the maximum of the convex functions is also convex [120]. $\sum_{i=1}^{N} E_i(Z_i, D_{\mathrm{cp},i})$ and $\max\limits_{1\leqslant i\leqslant N}\big\{E_i(Z_i, D_{\mathrm{cp},i})\big\}$ both are jointly convex in Z_i and $D_{\mathrm{cp},i}$, $\forall\, i$. Hence, for a given sequence, $\widetilde{\mathbf{P}}_{\mathbf{SUM}}$ and $\widetilde{\mathbf{P}}_{\mathbf{MM}}$ both are convex optimization problems.

Now consider problem $\widehat{\mathbf{P}}_{\mathbf{PF}}$ in (4.21). It can be shown that the objective function in (4.21) is jointly non-convex in P_i and $D_{\mathrm{cp},i}$, $\forall\, i$. We propose substituting $Z_i = \ln\big(1+\kappa\frac{P_i|h_i|^2}{\Gamma\sigma^2 d_i^{\alpha}}\big)$ and $V_i = \ln\big(D_{\mathrm{cp},i}\big)$ in (4.14) to arrive at

$$\log\Big(E_i\big(Z_i, V_i\big)\Big) = \log\Big(\tau D_i P_{\mathrm{cp}}\Big(\frac{D_i^{\beta}}{\exp(\beta V_i)}-1\Big) + \frac{b_i\exp(V_i)}{Z_i}\Big(\exp\big(Z_i\big)+c_i\Big)\Big). \tag{C.1.5}$$

Substituting Z_i and V_i in the constraint functions (4.15c), (4.15d), (4.15e), (4.15h) and (4.15i) yields

$$x_{1,i}\tau D_i\big(\frac{D_i^{\beta}}{\exp\big(\beta V_i\big)}-1\big) + x_{1,i}\frac{\exp\big(V_i\big)\ln(2)}{BZ_i} \leqslant T_1, \quad \forall\, i, \tag{C.1.6a}$$

$$\sum_{n=2}^{N} x_{n,i}\tau D_i\big(\frac{D_i^{\beta}}{\exp\big(\beta V_i\big)}-1\big) \leqslant \sum_{n=2}^{N}\sum_{k=1}^{n-1} x_{n,i}T_k, \quad \forall\, i, \tag{C.1.6b}$$

$$\sum_{n=2}^{N} x_{n,i} \frac{\exp\left(V_i\right) \ln(2)}{B Z_i} \leqslant \sum_{n=2}^{N} x_{n,i} T_n, \quad \forall\, i, \tag{C.1.6c}$$

$$0 \leqslant Z_i \leqslant Z_{\max}, \quad \forall\, i, \tag{C.1.6d}$$

$$\ln(D_{\min,i}) \leqslant V_i \leqslant \ln(D_i), \quad \forall\, i, \tag{C.1.6e}$$

For a known sequence, problem $\widehat{\mathbf{P}}_{\mathbf{PF}}$ in (4.21) can equivalently be written as

$$\begin{aligned} \widetilde{\mathbf{P}}_{\mathbf{PF}}: \quad & \underset{Z_i, V_i, T_n, \forall\, n,i}{\text{minimize}} \quad && \sum_{i=1}^{N} \log\Big(E_i(Z_i, V_i)\Big) \\ & \text{subject to} && \text{(4.15b), (C.1.6a), (C.1.6b), (C.1.6c), (C.1.6d),} \\ & && \text{(C.1.6e), (4.15j).} \end{aligned} \tag{C.1.7}$$

It can also be shown that (C.1.5) is jointly convex in Z_i and V_i. Since the sum of convex functions is convex [120], $\sum_{i=1}^{N} \log\big(E_i(Z_i, V_i)\big)$ is jointly convex in both Z_i and V_i, $\forall\, i$. Similarly, it can be shown that (C.1.6b) is convex in V_i, $\forall\, i$, (C.1.6c) is jointly convex in Z_i and V_i, $\forall\, i$, and (C.1.6a) is jointly convex in Z_1 and V_1. Hence, for a given sequence, $\widetilde{\mathbf{P}}_{\mathbf{PF}}$ is convex.

C.2 Proof of Lemma 3

This proof is based on [109–111]. Let $\mathcal{S}^{(j+1)} = (Z_i^{(j+1)}, D_{\text{cp},i}^{(j+1)}, T_n^{(j+1)}, x_{n,i}^{(j+1)}, \forall\, n, i)$ be the solution to problem (4.31) at the $(j+1)$th iteration at a given point $x_{n,i}^{(j)}$, $\forall\, n, i$, which yields

$$\begin{aligned} \mathrm{OF}^{(j+1)}\big(\mathcal{S}^{(j+1)}\big) &= \sum_{i=1}^{N} E_i^{(j+1)}(Z_i^{(j+1)}, D_{\text{cp},i}^{(j+1)}) + \Lambda \sum_{n=1}^{N} \sum_{i=1}^{N} \Big(x_{n,i}^{(j+1)}\big(1-2x_{n,i}^{(j)}\big) + \big(x_{n,i}^{(j)}\big)^2\Big) \\ &\leqslant \sum_{i=1}^{N} E_i^{(j)}(Z_i^{(j)}, D_{\text{cp},i}^{(j)}) + \Lambda \sum_{n=1}^{N} \sum_{i=1}^{N} \Big(x_{n,i}^{(j)}\big(1-2x_{n,i}^{(j-1)}\big) + \big(x_{n,i}^{(j-1)}\big)^2\Big) = \mathrm{OF}^{(j)}\big(\mathcal{S}^{(j)}\big) \end{aligned} \tag{C.2.1}$$

Thus, the Algorithm 1 produces a monotone sequence of improved solutions, $\big\{\mathrm{OF}^{(j)}\big(\mathcal{S}^{(j)}\big)\big\}$, for problem (4.16). Moreover, $\big\{\mathrm{OF}^{(j)}\big(\mathcal{S}^{(j)}\big)\big\}$ is bounded by constraint functions (4.15h), (4.15i) and (4.15j), and therefore convergence is guaran-

teed, i.e., $\Big(\mathrm{OF}^{(j+1)}\big(\mathcal{S}^{(j+1)}\big) - \mathrm{OF}^{(j)}\big(\mathcal{S}^{(j)}\big)\Big) \to 0$ as $j \to \infty$.

Following [110], problem (4.16), which is solved by Algorithm 1, can be represented as

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \varphi_{\mathrm{o}}(\mathbf{s}) && \text{(C.2.2a)} \\ \text{subject to} \quad & \psi_u(\mathbf{s}) \leqslant 0, \quad \forall\, u \in \{1, 2, \cdots, U\}, && \text{(C.2.2b)} \\ & \varphi_v(\mathbf{s}) \leqslant 0, \quad \forall\, v, \in \{1, 2, \cdots, V\}, && \text{(C.2.2c)} \end{aligned}$$

where $\mathbf{s}$ are design parameters, φ_{o} is the objective function, ψ_u are convex constraints and φ_v are non-convex constraints. The convex approximation of (4.16) in (4.31) can be rewritten as

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \widetilde{\varphi}_{\mathrm{o}}(\mathbf{s}, \mathbf{s}^{(j)}) && \text{(C.2.3a)} \\ \text{subject to} \quad & \psi_u(\mathbf{s}, \mathbf{s}^{(j)}) \leqslant 0, \quad \forall\, u \in \{1, 2, \cdots, U\}, && \text{(C.2.3b)} \\ & \widetilde{\varphi}_v(\mathbf{s}, \mathbf{s}^{(j)}) \leqslant 0, \quad \forall\, v, \in \{1, 2, \cdots, V\}, && \text{(C.2.3c)} \end{aligned}$$

where $\widetilde{\varphi}_{\mathrm{o}}$ and $\widetilde{\varphi}_v$ are the convex approximations of objective function, φ_{o}, and the non-convex constraints, $\widetilde{\varphi}_v$, respectively, at a given point $\mathbf{s}^{(j)}$. From (4.30), (4.31) and (C.2.1), we have

$$\begin{aligned} \varphi_v(\mathbf{s}) &\leqslant \widetilde{\varphi}_v(\mathbf{s}, \mathbf{s}^{(j)}), \quad \forall\, v, && \text{(C.2.4a)} \\ \varphi_v(\mathbf{s}^{(j)}) &= \widetilde{\varphi}_v(\mathbf{s}^{(j)}, \mathbf{s}^{(j)}), \quad \forall\, v, && \text{(C.2.4b)} \\ \nabla\varphi_v(\mathbf{s}^{(j)}) &= \widetilde{\varphi}_v(\mathbf{s}^{(j)}, \mathbf{s}^{(j)}), \quad \forall\, v. && \text{(C.2.4c)} \end{aligned}$$

Let $\mathbf{s}^{(j)}$ be the solution for Algorithm 1 at convergence. It is the optimal and the Fritz John point for problem (C.2.3) satisfying the following conditions [111]

$$\lambda_{\mathrm{o}} \nabla\widetilde{\varphi}_{\mathrm{o}}(\mathbf{s}^{(j)}, \mathbf{s}^{(j)}) + \sum_{u=1}^{U} \lambda_u \nabla\psi_u(\mathbf{s}^{(j)}) + \sum_{v=1}^{V} \lambda_v \nabla\widetilde{\varphi}_v(\mathbf{s}^{(j)}, \mathbf{s}^{(j)}) = 0, \qquad \text{(C.2.5a)}$$

$$\lambda_u \psi_u(\mathbf{s}^{(j)}) = 0, \quad \forall\, u, \qquad \text{(C.2.5b)}$$

$$\lambda_v \widetilde{\varphi}_v(\mathbf{s}^{(j)}, \mathbf{s}^{(j)}) = 0, \quad \forall\, v, \qquad \text{(C.2.5c)}$$

where λ_u and λ_v are Lagrange variables for the uth and the vth convex and non-convex constraints, respectively. Substituting (C.2.4b) and (C.2.4c) in (C.2.5) yields

$$\lambda_{\mathrm{o}} \nabla \varphi_{\mathrm{o}}(\mathbf{s}^{(j)}) + \sum_{u=1}^{U} \lambda_u \nabla \psi_u(\mathbf{s}^{(j)}) + \sum_{v=1}^{V} \lambda_v \nabla \varphi_v(\mathbf{s}^{(j)}) = 0, \tag{C.2.6a}$$

$$\lambda_u \psi_u(\mathbf{s}^{(j)}) = 0, \quad \forall\, u, \tag{C.2.6b}$$

$$\lambda_v \varphi_v(\mathbf{s}^{(j)}) = 0, \quad \forall\, v. \tag{C.2.6c}$$

The above results hold and thus satisfy the conditions (C.2.5), implying that $\mathbf{s}^{(j)}$ is a Fritz John solution of (C.2.2). Since (C.2.2) represents (4.16) the same conclusion can be drawn for (4.16). ■

C.3 Optimization Problems for the Benchmark Scheme

For the benchmark scheme, the problems for the considered objectives can be expressed as

$$\mathbf{P}^{\mathbf{b}}_{\mathbf{SUM}}: \quad \underset{Z_i,\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \sum_{i=1}^{N} \Big(\frac{D_i b_i}{Z_i} \big(\exp(Z_i) + c_i \big) \Big) \tag{C.3.1a}$$

$$\text{subject to} \quad (4.15\text{b}), (4.15\text{j}), (4.22\text{e}),$$

$$\sum_{n=1}^{N} x_{n,i} \frac{D_i}{B Z_i} \leqslant \sum_{n=1}^{N} x_{n,i} T_n, \quad \forall\, i, \tag{C.3.1b}$$

$$\mathbf{P}^{\mathbf{b}}_{\mathbf{MM}}: \quad \underset{Z_i,\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \max_{1 \leqslant i \leqslant N} \Big\{ \frac{D_i b_i}{Z_i} \big(\exp(Z_i) + c_i \big) \Big\} \tag{C.3.2}$$

$$\text{subject to} \quad (4.15\text{b}), (4.15\text{j}), (4.22\text{e}), (\text{C.3.1b}),$$

$$\mathbf{P}^{\mathbf{b}}_{\mathbf{PF}}: \quad \underset{Z_i,\, T_n,\, \forall\, n,i}{\text{minimize}} \quad \sum_{i=1}^{N} \log \Big(\frac{D_i b_i}{Z_i} \big(\exp(Z_i) + c_i \big) \Big) \tag{C.3.3}$$

$$\text{subject to} \quad (4.15\text{b}), (4.15\text{j}), (4.22\text{e}), (\text{C.3.1b}), .$$

C.4 Optimization Problems for Case 1

For Case 1, the problems for the considered objectives can be expressed as

$$\mathbf{P^{c1}_{SUM}}: \underset{Z_i,\, D_{\mathrm{cp},i},\, \forall\, i}{\text{minimize}} \quad \sum_{i=1}^{N} \Big(\tau D_n P_{\mathrm{cp}} \big((\frac{D_n}{D_{\mathrm{cp},n}})^{\beta} - 1 \big) + \frac{D_{\mathrm{cp},n} b_n}{Z_n} \big(\exp(Z_n) + c_n \big) \Big) \tag{C.4.1a}$$

$$\text{subject to} \quad (4.15\text{i}), (4.22\text{e}),$$

$$x_{1,i} \tau D_i \Big((\frac{D_i}{D_{\mathrm{cp},i}})^{\beta} - 1 \Big) + x_{1,i} \frac{D_{\mathrm{cp},i}}{B Z_i} \leqslant \frac{T_{\mathrm{frame}}}{N}, \quad \forall\, i, \tag{C.4.1b}$$

$$\sum_{n=2}^{N} x_{n,i} \tau D_i \Big((\frac{D_i}{D_{\mathrm{cp},i}})^{\beta} - 1 \Big) \leqslant \sum_{n=2}^{N} \sum_{k=1}^{n-1} x_{n,i} \frac{T_{\mathrm{frame}}}{N}, \quad \forall\, i, \tag{C.4.1c}$$

$$\sum_{n=2}^{N} x_{n,i} \frac{D_{\mathrm{cp},i}}{B Z_i} \leqslant \sum_{n=2}^{N} x_{n,i} \frac{T_{\mathrm{frame}}}{N}, \quad \forall\, i, \tag{C.4.1d}$$

$$\mathbf{P^{c1}_{MM}}: \underset{Z_i,\, D_{\mathrm{cp},i},\, \forall\, i}{\text{minimize}} \quad \max_{1 \leqslant i \leqslant N} \Big\{ \tau D_i P_{\mathrm{cp}} \Big((\frac{D_i}{D_{\mathrm{cp},i}})^{\beta} - 1 \Big) + D_{\mathrm{cp},i} b_i Z_i^{-1} \big(\exp(Z_i) + c_i \big) \Big\}$$

$$\text{subject to} \quad (4.15\text{i}), (4.22\text{e}), (\text{C.4.1b}) - (\text{C.4.1d}), \tag{C.4.2}$$

$$\mathbf{P^{c1}_{PF}}: \underset{Z_i,\, V_i,\, \forall\, i}{\text{minimize}} \quad \sum_{i=1}^{N} \log \Big(\tau P_{\mathrm{cp}} \Big(\frac{D_i D_i^{\beta}}{\exp(\beta V_i)} - D_i \Big) + \frac{b_i \exp\big(V_i\big)}{Z_i} \Big(\exp\big(Z_i\big) + c_i \Big) \Big) \tag{C.4.3a}$$

$$\text{subject to} \quad (4.22\text{e}), (4.24\text{e}),$$

$$x_{1,i} \tau D_i \Big(\frac{D_1^{\beta}}{\exp(\beta V_1)} - 1 \Big) + x_{1,i} \frac{\exp(V_1) \ln(2)}{B Z_i} \leqslant \frac{T_{\mathrm{frame}}}{N}, \quad \forall\, i, \tag{C.4.3b}$$

$$\sum_{n=2}^{N} x_{n,i} \tau D_i \Big((\frac{D_i^{\beta}}{\exp(\beta V_1)}) - 1 \Big) \leqslant \sum_{n=2}^{N} \sum_{k=1}^{n-1} x_{n,i} \frac{T_{\mathrm{frame}}}{N}, \quad \forall\, i, \tag{C.4.3c}$$

$$\sum_{n=2}^{N} x_{n,i} \frac{\exp(V_n)\ln(2)}{BZ_i} \leqslant \sum_{n=2}^{N} x_{n,i} \frac{T_{\text{frame}}}{N}, \quad \forall\, i, \tag{C.4.3d}$$

C.5 Optimization Problems for Case 2

For Case 2, the problems for the considered objectives can be expressed as

$$\begin{aligned} \mathbf{P}^{\mathbf{c2}}_{\mathbf{SUM}}: \quad & \underset{\substack{Z_i,\, D_{\text{cp},i},\\ x_{n,i},\forall\, n,i}}{\text{minimize}} \quad \sum_{i=1}^{N} \Big(\tau D_i P_{\text{cp}} \Big(\big(\frac{D_i}{D_{\text{cp},i}}\big)^{\beta} - 1\Big) + \frac{D_{\text{cp},i} b_i}{Z_i} (\exp(Z_i) + c_i)\Big) \\ & \text{subject to} \quad (4.15\text{f}), (4.15\text{g}), (4.15\text{i}), (4.15\text{k}), (4.22\text{e}), \\ & \qquad\qquad (\text{C.4.1b}) - (\text{C.4.1d}), \end{aligned} \tag{C.5.1}$$

$$\begin{aligned} \mathbf{P}^{\mathbf{c2}}_{\mathbf{MM}}: \quad & \underset{\substack{Z_i,\, D_{\text{cp},i},\\ x_{n,i},\forall\, n,i}}{\text{minimize}} \quad \max_{1 \leqslant i \leqslant N} \Big\{\tau D_i P_{\text{cp}} \Big(\big(\frac{D_i}{D_{\text{cp},i}}\big)^{\beta} - 1\Big) + \frac{D_{\text{cp},i} b_i}{Z_i} (\exp(Z_i) + c_i)\Big\} \\ & \text{subject to} \quad (4.15\text{f}), (4.15\text{g}), (4.15\text{i}), (4.15\text{k}), (4.22\text{e}), \\ & \qquad\qquad (\text{C.4.1b}) - (\text{C.4.1d}), \end{aligned} \tag{C.5.2}$$

$$\begin{aligned} \mathbf{P}^{\mathbf{c2}}_{\mathbf{PF}}: \quad & \underset{\substack{Z_i,\, V_i,\\ x_{n,i},\forall\, n,i}}{\text{minimize}} \quad \sum_{i=1}^{N} \log\Big(\tau P_{\text{cp}} \Big(\frac{D_i D_i^{\beta}}{\exp(\beta V_i)} - D_i\Big) + \frac{b_i \exp\big(V_i\big)}{Z_i} \Big(\exp\big(Z_i\big) + c_i\Big)\Big) \\ & \text{subject to} \quad (4.15\text{f}), (4.15\text{g}), (4.15\text{k}), (4.22\text{e}), (4.24\text{e}), \\ & \qquad\qquad (\text{C.4.3b}) - (\text{C.4.3d}). \end{aligned} \tag{C.5.3}$$

Appendix D

This appendix contains the proof needed in Chapter 5.

D.1 Proof of Proposition

The probability $\mathbb{P}\big(\max\big\{\gamma_{m,n}^{(i)}\big\} \geqslant \gamma^{\text{th}}\big)$, defined in (5.3), is not tractable due to the max $\{\cdot\}$ function, thus we use bipartite graph to find $q_m^{(i)}$. First we formulate $p_{m,n}^{(i)}$ which denotes the probability that the burst node $\mathcal{B}_m$ selects edge weight equal to E_l and the cumulative weight of u other edges $\Theta_{m,n}^{(i)}(u)$, connected to the sum node $\mathcal{S}_n$, is less than or equal to $\varphi^{\text{th}}(E_l)$. It can be given as

$$p_{m,n}^{(i)} = \sum_{l=1}^{L} \sum_{u=0}^{M-1} p^{(i)}(u)\, \mathbb{P}\left\{\Theta_{m,n}^{(i)}(u) \leqslant \varphi^{\text{th}}(E_l)\right\} \Gamma_l, \tag{D.1.1}$$

where $p^{(i)}(u)$, defined in (5.8), represents the probability of k edges connected to a given sum node, out of which $k-u$ edges are resolved in the previous iterations, but u edges are still connected to burst node $\mathcal{B}_m$'s edge in the ith iteration. To compute $p^{(i)}(u)$, we need to know Ψ_k which in the asymptotic setting, when $N \to \infty$, can be given as [121]

$$\Psi_k = e^{-\frac{\alpha^{k+1}}{k!}}, \tag{D.1.2}$$

where $\alpha = \overline{\Lambda}\frac{M}{N}$ is the average degree of a sum node, and $\overline{\Lambda}$ is the average degree of a burst node given as $\overline{\Lambda} = \sum_d \Lambda_d d$.

Next, we need $\Theta_{m,n}^{(i)}(u)$, in order to know the cumulative weight of edges on a sum node $\mathcal{S}_n$, which may vary from iteration to iteration depending upon the

number of connected edges u. It is given as

$$\Theta_{m,n}^{(i)}(u) = \sum_{c=0}^{u} W_{c,n}, \quad 0 \leqslant u \leqslant M-1, \tag{D.1.3}$$

where $W_{c,n}$ represents the edge weight of burst node $\mathcal{B}_c$, $\forall c \in \mathcal{M}_n \setminus \mathcal{B}_m$, connected with sum node $\mathcal{S}_n$. A burst node selects weight equal to E_l with probability Γ_l which is defined in (5.1). Using (D.1.3), the probability in (5.5) can be numerically evaluated.

Since, the selection of any sum node is equally probable, it implies that $p_{m,n}^{(i)} = p_{m,n'}^{(i)}$, for $n \neq n'$, where $n, n' \in \mathcal{N}_m$. Thus, $p_m^{(i)} = p_{m,n}^{(i)}$, which is given by (D.1.1). Recall, the *OR rule* for success and *AND rule* for error explained in Section 5.3. Thus, the recovery-error probability for burst node $\mathcal{B}_m$ with node degree $d = |\mathcal{N}_m|$, in the ith iteration, is equal to $(1 - p_m^{(i)})^d$. By averaging it over the degree distribution represented by $\{\Lambda_d\}$, defined in (5.1), yields

$$q_m^{(i)} = \sum_d \Lambda_d \big(1 - p_m^{(i)}\big)^d \tag{D.1.4}$$

Substituting the value of $p_m^{(i)}$ from (D.1.1) in (D.1.4), yields (5.7).

Bibliography

[1] M. R. Bhalla and A. V. Bhalla, "Generations of mobile wireless technology: A survey," *International Journal of Computer Applications*, vol. 5, no. 4, pp. 26–32, Aug. 2010.

[2] International Telecommunication Union*, Measuring digital development facts and figures 2019*.

[3] Z. Dawy, W. Saad, A. Ghosh, J. G. Andrews, and E. Yaacoub, "Toward massive machine type cellular communications," *IEEE Wireless Commun. Mag.*, vol. 24, no. 1, pp. 120–128, Feb. 2017.

[4] H. Shariatmadari, R. Ratasuk, S. Iraji, A. Laya, T. Taleb, R. Jntti, and A. Ghosh, "Machine-type communications: current status and future perspectives toward 5G systems," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 10–17, Sep. 2015.

[5] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, Jun. 2015.

[6] R. Ratasuk, S. Iraji, K. Hugl, L. Wang, and A. Ghosh, "Performance of low-cost LTE devices for advanced metering infrastructure," in *Proc. IEEE VTC*, Jun. 2013, pp. 1–5.

[7] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Computer networks*, vol. 38, no. 4, pp. 393–422, Mar. 2002.

[8] L. Atzori, A. Iera, and G. Morabito, "The Internet of things: A survey," *Computer networks*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010.

[9] Lu Tan and Neng Wang, "Future Internet: The Internet of things," in *Proc. IEEE ICACTE*, vol. 5, Aug. 2010, pp. 376–380.

[10] H.-D. Ma, "Internet of things: Objectives and scientific challenges," *Journal of Computer Science and Technology*, vol. 26, no. 6, pp. 919–924, Nov. 2011.

[11] L. Mainetti, L. Patrono, and A. Vilei, "Evolution of wireless sensor networks towards the Internet of things: A survey," in *Proc. IEEE SoftCOM*. IEEE, Sep. 2011, pp. 1–6.

[12] C. C. Aggarwal, N. Ashish, and A. Sheth, "The Internet of things: A survey from the data-centric perspective," in *Managing and mining sensor data*. Springer, 2013, pp. 383–428.

[13] M. Zorzi, A. Gluhak, S. Lange, and A. Bassi, "From today's intranet of things to a future Internet of things: A wireless-and mobility-related view," *IEEE Wireless Commun. Mag.*, vol. 17, no. 6, pp. 44–51, Dec. 2010.

[14] J.-P. Vasseur and A. Dunkels, *Interconnecting smart objects with IP: The next Internet*. Morgan Kaufmann, 2010.

[15] O. Hersent, D. Boswarthick, and O. Elloumi, *The Internet of things: Key applications and protocols*. John Wiley & Sons, 2011.

[16] Ericsson, *Ericsson Mobility Report*, Nov. 2019.

[17] T. Taleb and A. Kunz, "Machine type communications in 3GPP networks: Potential, challenges, and solutions," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 178–184, Mar. 2012.

[18] K.-C. Chen and S.-Y. Lien, "Machine-to-machine communications: Technologies and challenges," *Ad Hoc Networks*, vol. 18, pp. 3–23, Jul. 2014.

[19] C. Anton-Haro and M. Dohler, *Machine-to-machine (M2M) communications: Architecture, performance and applications*. Elsevier, 2014.

[20] "IEEE standard for information technology - Telecommunications and information exchange between systems local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," *IEEE Std 802.11-2016 (revision of IEEE Std 802.11-2012)*, pp. 1–3534, Dec. 2016.

[21] "IEEE standard for low-rate wireless networks," *IEEE Std 802.15.4-2015 (revision of IEEE Std 802.15.4-2011)*, pp. 1–709, Apr. 2016.

[22] *IEEE Std 802.15.1-2005 (Revision of IEEE Std 802.15.1-2002)*, pp. 1–700, Jun. 2005.

[23] N. Kouzayha, M. Jaber, and Z. Dawy, "M2M data aggregation over cellular networks: signaling-delay trade-offs," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2014, pp. 1155–1160.

[24] M. D6.6, *Final Report on the METIS System Concept and Technology Roadmap*, 2015.

[25] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.

[26] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. H. Hanzo, "A survey of network lifetime maximization techniques in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 828–854, Jan. 2017.

[27] S. A. Alvi, B. Afzal, G. A. Shah, L. Atzori, and W. Mahmood, "Internet of multimedia things: Vision and challenges," *Ad Hoc Networks*, vol. 33, pp. 87–111, Oct. 2015.

[28] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1658–1686, Jan. 2014.

[29] N. Navet, Y. Song, F. Simonot-Lion, and C. Wilwert, "Trends in automotive communication systems," *Proc. IEEE*, vol. 93, no. 6, pp. 1204–1223, Jun. 2005.

[30] J. Qiao, X. S. Shen, J. W. Mark, Q. Shen, Y. He, and L. Lei, "Enabling device-to-device communications in millimeter-wave 5G cellular networks," *IEEE Trans. Mobile Comput.*, vol. 53, no. 1, pp. 209–215, Jan. 2015.

[31] A. Zanella, M. Zorzi, A. F. dos Santos, P. Popovski, N. Pratas, C. Stefanovic, A. Dekorsy, C. Bockelmann, B. Busropan, and T. A. H. J. Norp, "M2M massive wireless access: Challenges, research issues, and ways forward," in *Proc. IEEE Globecom Workshops*, Dec. 2013, pp. 151–156.

[32] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 117–125, Apr. 2015.

[33] D. N. K. Jayakody, J. Thompson, S. Chatzinotas, and S. Durrani, *Wireless information and power transfer: A new paradigm for green communications.* Springer International Publishing AG, 2017.

[34] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, Nov. 2015.

[35] D. Niyato, P. Wang, and D. I. Kim, "Performance analysis and optimization of TDMA network with wireless energy transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4205–4219, Aug. 2014.

[36] V. Cionca, T. Newe, and V. Dadrlat, "TDMA protocol requirements for wireless sensor networks," in *Proc. IEEE SENSORCOMM*, Aug. 2008, pp. 30–35.

[37] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.

[38] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: Physical and mac-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, Sep. 2016.

[39] E. Casini, R. D. Gaudenzi, and O. D. R. Herrero, "Contention resolution diversity slotted ALOHA (CRDSA): An enhanced random access scheme for satellite access packet networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1408–1419, Apr. 2007.

[40] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Random access for M2M communications with QoS guarantees," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2889–2903, Jul. 2017.

[41] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Computer Networks*, vol. 51, no. 4, pp. 921 – 960, Mar. 2007.

[42] T. Srisooksai, K. Keamarungsi, P. Lamsrichan, and K. Araki, "Practical data compression in wireless sensor networks: A survey," *Journal of Network and Computer Applications*, vol. 35, no. 1, pp. 37 – 59, Jan. 2012.

[43] P. Kasirajan, C. Larsen, and S. Jagannathan, "A new data aggregation scheme via adaptive compression for wireless sensor networks," *ACM Trans. Sen. Netw.*, vol. 9, no. 1, pp. 5:1–5:26, Nov. 2012.

[44] X. Deng and Y. Yang, "Online adaptive compression in delay sensitive wireless sensor networks," *IEEE Trans. Comput.*, vol. 61, no. 10, pp. 1429–1442, Oct. 2012.

[45] Y. Wang, D. Wang, X. Zhang, J. Chen, and Y. Li, "Energy-efficient image compressive transmission for wireless camera networks," *IEEE Sensors J.*, vol. 16, no. 10, pp. 3875–3886, May 2016.

[46] D. Jung, T. Teixeira, and A. Savvides, "Sensor node lifetime analysis: Models and tools," *ACM Trans. Sen. Netw.*, vol. 5, no. 1, pp. 3:1–3:33, Feb. 2009.

[47] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 40–50, Mar. 2002.

[48] M. Tahir and R. Farrell, "A cross-layer framework for optimal delay-margin, network lifetime and utility tradeoff in wireless visual sensor networks," *Ad Hoc Networks*, vol. 11, no. 2, pp. 701–711, Mar. 2013.

[49] C. M. Sadler and M. Martonosi, "Data compression algorithms for energy-constrained devices in delay tolerant networks," in *Proc. SENSYS, ACM*, Nov. 2006, pp. 265–278.

[50] Z. Qin, Y. Liu, Y. Gao, M. Elkashlan, and A. Nallanathan, "Wireless powered cognitive radio networks with compressive sensing and matrix completion," vol. 65, no. 4, pp. 1464–1476, Apr 2017.

[51] Z. Qin, J. Fan, Y. Liu, Y. Gao, and G. Y. Li, "Sparse representation for wireless communications: A compressive sensing approach," *IEEE Signal Process. Mag.*, vol. 35, no. 3, pp. 40–58, May 2018.

[52] E. Uysal-Biyikoglu, B. Prabhakar, and A. E. Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 487–499, Aug. 2002.

[53] E. Uysal-Biyikoglu and A. E. Gamal, "On adaptive transmission for energy efficiency in wireless data networks," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3081–3094, Dec. 2004.

[54] M. Zafer and E. Modiano, "Optimal rate control for delay-constrained data transmission over a wireless channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4020–4039, Sep. 2008.

[55] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1135–1149, May 2002.

[56] M. A. Zafer and E. Modiano, "A calculus approach to energy-efficient data transmission with quality-of-service constraints," *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 898–911, Jun. 2009.

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

[58] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: Tutorial, survey, and open issues," *IEEE Wireless Commun. Mag.*, vol. 18, no. 6, pp. 28–35, Dec. 2011.

[59] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.

[60] Q. Wu, M. Tao, D. W. K. Ng, W. Chen, and R. Schober, "Energy-efficient resource allocation for wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2312–2327, Mar. 2016.

[61] X. Lin, L. Huang, C. Guo, P. Zhang, M. Huang, and J. Zhang, "Energy-efficient resource allocation in TDMS-based wireless powered communication networks," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 861–864, Apr. 2017.

[62] S. Hemour and K. Wu, "Radio-frequency rectifier for electromagnetic energy harvesting: Development path and future outlook," *Proc. of the IEEE,*, vol. 102, no. 11, pp. 1667–1691, Nov. 2014.

[63] W. Liu, X. Zhou, S. Durrani, H. Mehrpouyan, and S. D. Blostein, "Energy harvesting wireless sensor networks: Delay analysis considering energy costs of sensing and transmission," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4635–4650, Jul. 2016.

[64] M. A. Antepli, E. Uysal-Biyikoglu, and H. Erkal, "Optimal packet scheduling on an energy harvesting broadcast link," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1721–1731, Sep. 2011.

[65] M. Gregori and M. Payaro, "Energy-efficient transmission for wireless energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1244–1254, Mar. 2013.

[66] Z. Hadzi-Velkov, I. Nikoloska, G. K. Karagiannidis, and T. Q. Duong, "Wireless networks with energy harvesting and power transfer: Joint power and time allocation," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 50–54, Jan. 2016.

[67] P. D. Diamantoulakis and G. K. Karagiannidis, "Maximizing proportional fairness in wireless powered communications," *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 202–205, Apr. 2017.

[68] C. Guo, B. Liao, L. Huang, Q. Li, and X. Lin, "Convexity of fairness-aware resource allocation in wireless powered communication networks," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 474–477, Mar. 2016.

[69] X. Kang, C. K. Ho, and S. Sun, "Optimal time allocation for dynamic-TDMA-based wireless powered communication networks," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 3157–3161.

[70] Q. Wu, W. Chen, and J. Li, "Wireless powered communications with initial energy: QoS guaranteed energy-efficient resource allocation," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2278–2281, Dec. 2015.

[71] S. A. Alvi, X. Zhou, and S. Durrani, "Optimal compression and transmission rate control for node-lifetime maximization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7774–7788, Nov. 2018.

[72] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted ALOHA," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 477–487, Feb. 2011.

[73] A. Mengali, R. D. Gaudenzi, and P. D. Arapoglou, "Enhancing the physical layer of contention resolution diversity slotted ALOHA," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4295 – 4308, Oct. 2017.

[74] F. Clazzer, E. Paolini, I. Mambelli, and Č. Stefanović, "Irregular repetition slotted ALOHA over the Rayleigh block fading channel with capture," in *Proc. IEEE ICC*, May 2017, pp. 1–6.

[75] Č. Stefanović, M. Momoda, and P. Popovski, "Exploiting capture effect in frameless ALOHA for massive wireless random access," in *Proc. IEEE WCNC*, Apr. 2014, pp. 1762–1767.

[76] M. Luby, M. Mitzenmacher, and M. A. Shokrollahi, "Analysis of random processes via and-or tree evaluation." in *Proc. SODA*, vol. 98, Jan. 1998, pp. 364–373.

[77] S. Fan, L. Zhang, W. Feng, W. Zhang, and Y. Ren, "Optimization-based design of wireless link scheduling with physical interference model," *IEEE Trans. Veh. Technol.*, vol. 61, no. 8, pp. 3705–3717, Oct. 2012.

[78] M. H. Cheung and V. W. S. Wong, "Interference pricing for SINR-based random access game," vol. 12, no. 5, pp. 2292–2301, May 2013.

[79] A. Mohsenian-Rad, V. W. S. Wong, and R. Schober, "Optimal SINR-based random access," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.

[80] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "On SINR-based random multiple access using codes on graph," in *Proc. IEEE GLOBECOM*, Dec. 2015, pp. 1–6.

[81] H. Lin, K. Ishibashi, W. Shin, and T. Fujii, "Decentralized power allocation for secondary random access in cognitive radio networks with successive interference cancellation," in *Proc. IEEE ICC*, May 2016, pp. 1–6.

[82] A. Goldsmith, *Wireless communications*. Cambridge Univ. Press, 2005.

[83] S. A. Alvi, X. Zhou, and S. Durrani, "A lifetime maximization scheme for a sensor based MTC device," in *Proc. IEEE GLOBECOM*, Dec. 2018.

[84] S. A. Alvi, X. Zhou, and S. Durrani, "Wireless powered machine-type communication: Energy minimization via compressed transmission," in *Proc. IEEE PIMRC*, Sep. 2019, pp. 1–7.

[85] S. A. Alvi, X. Zhou, S. Durrani, and D. T. Ngo, "Sequencing and scheduling for multi-user machine-type communication," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2459–2473, Apr. 2020.

[86] S. A. Alvi, X. Zhou, S. Durrani, and D. T. Ngo, "Proportionally-fair sequencing and scheduling for machine-type communication," in *Proc. IEEE ICC*, Jun. 2020.

[87] S. A. Alvi, S. Durrani, and X. Zhou, "Enhancing CRDSA with transmit power diversity for machine-type communication," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7790–7794, Aug. 2018.

[88] W. Liu, X. Zhou, S. Durrani, H. Mehrpouyan, and S. D. Blostein, "Energy harvesting wireless sensor networks: Delay analysis considering energy costs of sensing and transmission," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4635–4650, Jul. 2016.

[89] S. Mao, M. H. Cheung, and V. W. S. Wong, "Joint energy allocation for sensing and transmission in rechargeable wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 6, pp. 2862–2875, Jul. 2014.

[90] T. H. Lee, *The design of CMOS radio-frequency integrated circuits.* Cambridge University Press, Dec. 2003.

[91] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital communication receivers: Synchronization, channel estimation, and signal processing.* New York, NY, USA: John Wiley & Sons, Inc., 1997.

[92] M. K. Banavar, C. Tepedelenlioglu, and A. Spanias, "Estimation over fading channels with limited feedback using distributed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 414–425, Jan. 2010.

[93] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, Sep. 2007.

[94] G. Foschini and J. Salz, "Digital communications over fading radio channels," *The Bell System Technical Journal*, vol. 62, no. 2, pp. 429–456, Feb. 1983.

[95] P. Nuggehalli, V. Srinivasan, and R. R. Rao, "Delay constrained energy efficient transmission strategies for wireless devices," in *Proc. IEEE INFOCOM*, vol. 3, Nov. 2002, pp. 1765–1772.

[96] W. Xu, Q. Shi, X. Wei, Z. Ma, X. Zhu, and Y. Wang, "Distributed optimal rate-reliability-lifetime tradeoff in time-varying wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4836–4847, Sep. 2014.

[97] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[98] T. Q. Wu and H. C. Yang, "On the performance of overlaid wireless sensor transmission with RF energy harvesting," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1693–1705, Aug. 2015.

[99] S. C. Liew and Y. J. Zhang, "Proportional fairness in multi-channel multi-rate wireless networks-part I: The case of deterministic channels with application to AP association problem in large-scale WLAN," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3446–3456, Sep. 2008.

[100] A. Laya, L. Alonso, and J. Alonso-Zarate, "Is the random access channel of LTE and LTE-A suitable for M2M communications? A survey of alternatives," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 4–16, Dec. 2014.

[101] T. David and P. Viswanath, *Fundamentals of wireless communication.* Cambridge University Press, May 2005.

[102] Q. Wu, G. Y. Li, W. Chen, and D. W. K. Ng, "Energy-efficient D2D overlaying communications with spectrum-power trading," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4404–4419, Jul. 2017.

[103] T. Bonald, L. Massoulié, A. Proutière, and J. Virtamo, "A queueing analysis of max-min fairness, proportional fairness and balanced fairness," *Queueing Systems*, vol. 53, no. 1, pp. 65–84, Jun. 2006.

[104] F. P. Kelly, "Charging and rate control for elastic traffic," *Euro. Trans. Telecommun.*, vol. 8, no. 1, pp. 7–20, Jan. 1997.

[105] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. Sagastizábal, *Numerical optimization theoretical and practical aspects, 2nd ed.* Springer, Aug. 2006.

[106] U. Rashid, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks," *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 488–499, Feb. 2014.

[107] E. Che, H. D. Tuan, and H. H. Nguyen, "Joint optimization of cooperative beamforming and relay assignment in multi-user wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5481–5495, Oct. 2014.

[108] H. H. M. Tam, H. D. Tuan, D. T. Ngo, T. Q. Duong, and H. V. Poor, "Joint load balancing and interference management for small-cell heterogeneous networks with limited backhaul capacity," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 872–884, Feb. 2017.

[109] T. T. Vu, D. T. Ngo, M. N. Dao, S. Durrani, and R. H. Middleton, "Spectral and energy efficiency maximization for content-centric C-RANs with edge caching," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6628–6642, Dec. 2018.

[110] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations research*, vol. 26, no. 4, pp. 681–683, Aug. 1978.

[111] M. C. Ferris and O. L. Mangasarian, "Parallel variable distribution," *SIAM Journal on Optimization*, vol. 4, no. 4, pp. 815–832, Nov. 1994.

[112] R. Fourer, D. M. Gay, and B. W. Kernighan, *AMPL: A mathematical programing language. In: Wallace s.w. (eds) algorithms and model formulations in mathematical programming. NATO ASI series (series F: Computer and systems sciences).* Springer, Berlin, Heidelberg, Apr. 1989, vol. 51.

[113] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Wchter, "Branching and bounds tightening techniques for non-convex MINLP," *Optimization Methods and Software*, vol. 24, no. 4-5, pp. 597–634, Aug. 2009.

[114] "Couenne, an exact solver for nonconvex MINLPs," 2006. [Online]. Available: https://projects.coin-or.org/Couenne/

[115] M. Ivanov, F. Brännström, A. G. i Amat, and P. Popovski, "Broadcast coded slotted ALOHA: A finite frame length analysis," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 651–662, Feb. 2017.

[116] Z. Sun, Y. Xie, J. Yuan, and T. Yang, "Coded slotted ALOHA for erasure channels: Design and throughput analysis," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–1, Aug. 2017.

[117] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted ALOHA," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 477–487, Feb. 2011.

[118] M. S. Ali, E. Hossain, and D. I. Kim, "LTE/LTE-A random access for massive machine-type communications in smart cities," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 76–83, Jan. 2017.

[119] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, Dec. 1997.

[120] S. Boyd and V. Lieven, *Convex optimization*. Cambridge University Press, Mar. 2004.

[121] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, Jun. 2006.