ABSTRACT

ENERGY HARVESTING-AWARE DESIGN OF WIRELESS NETWORKS

by
Fabio Iannello

Recent advances in low-power electronics and energy-harvesting (EH) technologies enable the design of self-sustained devices that collect part, or all, of the needed energy from the environment. Several systems can take advantage of EH, ranging from portable devices to wireless sensor networks (WSNs). While conventional design for battery-powered systems is mainly concerned with the battery lifetime, a key advantage of EH is that it enables potential perpetual operation of the devices, without requiring maintenance for battery substitutions. However, the inherent unpredictability regarding the amount of energy that can be collected from the environment might cause temporary energy shortages, which might prevent the devices to operate regularly. This uncertainty calls for the development of energy management techniques that are tailored to the EH dynamics.

While most previous work on EH-capable systems has focused on energy management for single devices, the main contributions of this dissertation is the analysis and design of medium access control (MAC) protocols for WSNs operated by EH-capable devices. In particular, the dissertation first considers random access MAC protocols for single-hop EH networks, in which a fusion center collects data from a set of nodes distributed in its surrounding. MAC protocols commonly used in WSNs, such as time division multiple access (TDMA), framed-ALOHA (FA) and dynamic-FA (DFA) are investigated in the presence of EH-capable devices. A new ALOHA-based MAC protocol tailored to EH-networks, referred to as energy group-DFA (EG-DFA), is then proposed. In EG-DFA nodes with similar energy availability are grouped together and access the channel independently from other groups. It is shown that

EG-DFA significantly outperforms the DFA protocol. Centralized scheduling-based MAC protocols for single-hop EH-networks with communication resource constraints are considered next. Two main scenarios are addressed, namely: i) nodes exclusively powered via EH; ii) nodes powered by a hybrid energy storage system, which is composed by a non-rechargeable battery and a capacitor charged via EH. For the former case the goal is the maximization of the network throughput, while in the latter the aim is maximizing the lifetime of the non-rechargeable batteries. For both scenarios optimal scheduling policies are derived by assuming different levels of information available at the fusion center about the energy availability at the nodes. When optimal policies are not derived explicitly, suboptimal policies are proposed and compared with performance upper bounds.

Energy management policies for single devices have been investigated as well by focusing on radio frequency identification (RFID) systems, when the latter are operated by enhanced RFID tags with energy harvesting capabilities.

ENERGY HARVESTING-AWARE DESIGN OF WIRELESS NETWORKS

by
Fabio Iannello

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

Department of Electrical and Computer Engineering, NJIT

May 2012

APPROVAL PAGE

ENERGY HARVESTING-AWARE DESIGN OF WIRELESS NETWORKS

Fabio Iannello

Dr. Osvaldo Simeone, Dissertation Co-advisor   Date
Assistant Professor, Department of Electrical and Computer Engineering, NJIT

Dr. Umberto Spagnolini, Dissertation Co-advisor   Date
Full Professor, Dipartimento di Elettronica e Informazione, Politecnico di Milano

---

Dr. Mary Ann Ingram, Committee Member   Date
Professor, Department of Electrical and Computer Engineering, Georgia Institute of
Technology

Dr. Yeheskel Bar-Ness, Committee Member   Date
Distinguished Professor, Department of Electrical and Computer Engineering, NJIT

Dr. Alexander M. Haimovich, Committee Member   Date
Professor, Department of Electrical and Computer Engineering, NJIT

BIOGRAPHICAL SKETCH

Author: Fabio Iannello

Degree: Doctor of Philosophy

Date: May 2012

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2012

- Laurea Specialistica (M.Sc.) in Telecommunications Engineering,
  Politecnico di Milano, Milan, Italy, 2008

- Laurea (B.Sc.) in Telecommunications Engineering,
  Politecnico di Milano, Milan, Italy, 2005

Major: Electrical Engineering

Presentations and Publications:

F. Iannello, O. Simeone and U. Spagnolini, "Lifetime maximization for wireless networks with hybrid energy storage systems," in preparation for submission to *IEEE Trans. Commun.*

F. Iannello, O. Simeone and U. Spagnolini, "On the optimal scheduling of independent, symmetric, and time-sensitive tasks," submitted to *IEEE Trans. Autom. Control* (under first revision).

F. Iannello, O. Simeone and U. Spagnolini, "Medium access control protocols for wireless sensor networks with energy harvesting," *IEEE Trans. Commun.*, May 2012 (in press).

F. Iannello, O. Simeone, P. Popovski and U. Spagnolini, "Energy group-based dynamic framed ALOHA for wireless networks with energy harvesting," in *Proc. 46th Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, Mar. 2012.

F. Iannello, O. Simeone and U. Spagnolini, "Optimality of myopic scheduling and whittle indexability for energy harvesting sensors," in *Proc. 46th Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, Mar. 2012.

F. Iannello, O. Simeone and U. Spagnolini, "Dynamic framed-ALOHA for energy-constrained wireless sensor networks with energy harvesting," in *Proc. IEEE GLOBECOM*, Miami, FL, Dec. 2010.

F. Iannello, O. Simeone and U. Spagnolini, "Energy management policies for passive RFID sensors with RF-Energy harvesting," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Cape Town, South Africa, May 2010.

F. Iannello, O. Simeone, "On the throughput region of single and two-way multi-hop fading networks with relay piggybacking," in *Proc. Signal Processing Advances Wireless Commun., (SPAWC)* , Perugia, Italy, Jun. 2009.

F. Iannello, O. Simeone, "Throughput analysis of type-I HARQ strategies in two-way relay channels," in *Proc. 43rd Conf. Inf. Sci. Syst. (CISS)*, Baltimore, MA, Mar. 2009.

To my family

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

Figure   **Page**

CHAPTER 1

INTRODUCTION

In the last decade, the conscience for respecting the environment, reducing pollution and energy consumptions, has tremendously grown in our society, making green one of the most used word in everyday vocabulary. A significant step toward going green is the exploitation of renewable energy sources, which aims on the one hand at reducing pollution and on the other hand at providing alternatives to the finite amount of non-renewable energy sources available on the Earth. Collecting energy from the environment, or *energy-harvesting* (EH), has a long history that dates back to windmills and waterwheels, which represent effective examples on how energy, freely available in the environment, can be efficiently leveraged by human beings.

In the last centuries, several physical effects that convert a form of energy into another have been discovered. Among these, it is worth mentioning the thermoelectric effect, discovered by T. J. Seedback in 1821, where an electric current was shown to deflect a compass needle when inserted into a closed loop between two dissimilar metals subject to different temperatures at the junction. Another milestone was the discovery of the piezoelectric effect by the brothers P. Curie and J. Curie, who realized in 1880 that an electric charge is accumulated in a solid material, such as a crystal, when the latter is subject to mechanical stress. Another fundamental discovery is the photoelectric effect, revealed by H. Hertz in 1887, who realized that, when a surface is exposed to electromagnetic radiation, the radiation can be absorbed and electrons emitted. Effects as the ones listed above, provide the basis for the modern EH technologies.

An important driver for the research on EH technologies was given by the great reduction in the power consumption of electronic circuits. While electronic

devices exclusively powered by EH, such as calculators and watches, have been on the market since the 70s (see Figure 1.1), EH technologies are today applicable to a wider variety of electronic devices. Examples range from cell phones and laptop computers to miniaturized wireless sensors. Furthermore, several energy sources commonly available in the environment, such as sunlight, mechanical, electromagnetic and thermoelectric energy, can now be efficiently converted into electrical energy through energy transducers of suitable sizes and of ever increasing efficiency [1].

One of the main, and perhaps most promising, applications of EH technologies is the deployment of wireless networks with sensing capabilities, also known as wireless sensor networks (WSNs). Such networks are used to monitor phenomena of interest within a prescribed area such the structural monitoring of buildings. The introduction of wireless nodes that are powered via EH not only eases the requirements for battery substitution, but also enables new applications of WSNs by allowing the deployment of battery-less nodes in remote or hazardous areas that are not easily accessible for maintenance. EH is thus expected to play a key role in the near future of WSNs. In fact, the ever increasing demand for a *smart world* [2], i.e., an environment in which objects interacts with each other as well as with human beings, will require an even wider deployment of WSNs.

Figure 1.1 Examples of an older generation of electronic devices powered by solar cells such as calculators and digital watches.

1.1 Overview of Energy Harvesting Technologies

The environment provides multiple sources of energy that can be leveraged for EH. Some are natural sources, such as sunlight and wind, while others are generated by human activities, such as mechanical energy due to machineries movements or electromagnetic energy transmitted by antennas. Regardless of the energy source, a typical architecture for EH-devices consists of three main components [1]:

1. Energy transducer (or converter);

2. Energy conditioning circuitry;

3. Energy storage device (ESD).

The energy transducer is a device that physically converts a given source of energy into electrical energy. Common examples include: photovoltaic cells that convert the energy of light; piezoelectric materials that convert mechanical energy, such as vibrations; thermocouples that convert a temperature gradient; and antennas that convert electromagnetic energy [1]. The energy conditioning circuitry is instead designed in order to efficiently transfer the power from the energy transducer to the device (or to the ESD). The most common examples of conditioning circuits are the maximum power point tracker (MPPT) circuits, which are used (often in photovoltaic cells-based harvesters) to dynamically adjust the working load of the transducer in order to obtain the maximum power transfer to the device [3]. Lastly, the ESD is used to store the surplus of the harvested energy that is not immediately used by the device. The two most important ESDs that are commercially available are rechargeable batteries and capacitors, which are briefly discussed in the next section.

It is worth mentioning that, in some applications, EH-devices are not equipped with ESDs, but they use the harvested power to directly power up their circuitry. One of the most relevant examples is given by passive RFID tags. These are devices

Table 1.1 Typical Values of Power that Can be Harvested from Common Sources
[1].

Energy Source	Power
Light (Indoor - Outdoor)	$10 \ \mu W/cm^2$- $100 \ mW/cm^2$
Mechanical (Human – Machines)	$4 \ \mu W/cm^3$- $800 \ \mu W/cm^3$
Electromagnetic (far from transmitters)	$1 \ \mu W/cm^2$

powered up by an electromagnetic wave generated by a RFID reader that, in their simplest version, are not intended to store energy for later uses [4].

How much energy can be harvested from the environment? Typical values registered through experimental setups are reported in Table 1.1 (see e.g., [1]). As it will be shown in Section 1.2.1, the power that can be harvested from the environment is generally much smaller than the power required for the continuous operation of a wireless node, at least for EH devices of practical dimensions. Therefore, nodes that are powered exclusively by EH can only operate for a small fraction of the time (*duty cycle*). However, this is typically not a limitation, since most WSNs have nodes operating with a very low duty cycle [5].

1.1.1 Batteries and Capacitors

The two most common components that are routinely used as ESDs in electronic systems are rechargeable batteries and capacitors. A battery is an electrochemical component that converts chemical energy into electrical energy, while a capacitor stores energy in the form of an electric field. Due to their distinct nature, the characteristics of batteries and capacitors are quite different [6]. Two of the most important ones being energy density and the component lifetime. In fact, batteries are generally characterized by an energy density higher than that of capacitors, and

are thus able to store more energy in a smaller volume. The component lifetime is often measured as the number of complete charging/discharging cycles before that the ESD suffers a notable loss of nominal capacity. The lifetime of batteries is typically in the order of a few hundreds cycles, while for capacitors is in the order of hundreds of thousands cycles [6]. Other important characteristics include: the rate of self-discharge of the energy stored in the ESD, which is generally smaller for batteries than that of capacitors; the sensitivity to the temperature, which is generally in favor to the capacitors (this is important in outdoor applications where temperature gradient is large); the rates at which the ESD can be charged and discharged, which are generally more flexible for capacitors than those for batteries. The latter aspect is relevant since operating with charging/discharging rates that are not suitable for the ESD at hand might severely degrade its performance. This effects is even more accentuated in EH applications, where the optimal charging rates for batteries cannot be generally guaranteed, and thus the more pronounced flexibility of capacitors might offer a better solution.

1.2 Overview of Wireless Sensor Networks (WSNs)

Recent advances in low-power electronics and wireless communications technologies have enabled the development of low-cost, low-power and multifunctional devices (or *nodes*) that are able to collect information (by sensing) from the surrounding environment and communicate with other devices over short distances [5]. A WSN is composed of several nodes, in order of tens, hundreds or even thousands, which are deployed within the area in which the phenomena of interest are to be monitored. Typical applications of WSNs include monitoring of physical quantities, such as temperature and mechanical vibrations, and object tracking (see e.g., [7]).

An important aspect of WSNs is given by the positioning of the nodes over the area of interest. In particular, the network topology can be engineered or can

be the result of a random deployment. The latter is more suitable when the number of nodes is large and/or the areas to be monitored are hardly accessible [5]. The network topology strongly affects the choice of the communication protocols. Broadly speaking, it is possible to identify three main network topologies (see Figure 1.2): *i)* *point-to-point*; *ii) point-to-multipoint* (or *star topology*); *iii) mesh*. Point-to-point and star networks are generally single-hop, in the sense that nodes only transmit their own data, while mesh networks can be multi-hop as nodes can forward packets belonging to other nodes. It is also possible to add a hierarchical structure to the network such as in cluster-based networks [5] (see Figure 1.2-d)), in which each cluster operates as a star network. Nodes in each cluster generally communicate in a single-hop fashion with the cluster-head, while cluster-heads communicate with each other to guarantee network connectivity.

The next section considers a typical architecture for a node operating in a WSN and discusses the main operations that affect the energy consumption of the nodes.
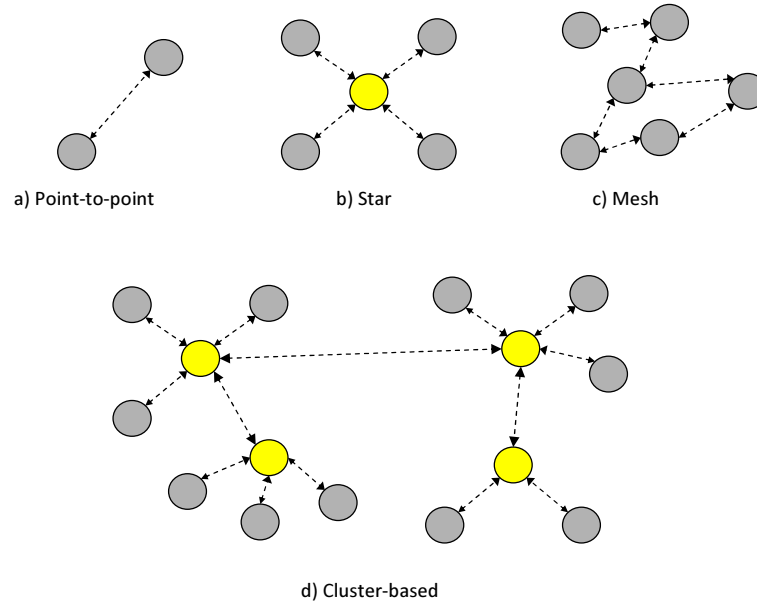
Figure 1.2 Typical network topologies. Dashed arrows indicate wireless links.
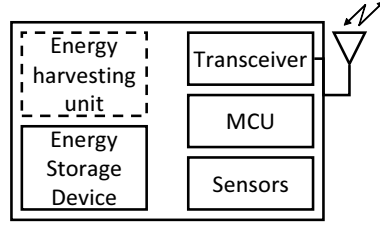
Figure 1.3 Typical architecture of a node employed in a wireless sensor network. An energy harvesting unit might be added.

1.2.1 Architecture of a Sensor and Energy Consumption

A typical architecture of a node employed in a wireless sensor network consists of four main blocks as shown in Figure 1.2 (see e.g., [5]): i) radio transceiver; ii) micro controller unit (MCU); iii) sensors; iv) energy storage device. The node can also be equipped with EH capabilities. Regardless of the application, the energy consumption of a node can be broadly divided into three contributions: sensing; data processing; and data communication [5]. While the contribution of the sensors to the energy budget is strongly application-dependent, some general consideration can be made for the data processing and communication contributions.

To start with, it is interesting to consider the power consumption of typical off-the-shelf MCUs and transceivers that are routinely used in WSNs, such as the ones considered in Table 1.2. In the table, P_{act} and P_{sleep} indicate the power consumption of the component when it is in the active mode and in the sleeping mode, respectively. As shown in Table 1.2, it is not uncommon that, for low-power sensor nodes, the power consumed by the transceiver is the largest one. As it will described in the next section, the impact of the medium access control protocol plays a key role in determining the activity of the transceiver.

As a last remark, it is worth mentioning that, a reduction in the radio frequency (RF) transmitted power might not imply a corresponding reduction in the overall consumed power. This is due to the power consumed for the operation of the transceiver's circuitry, which is not negligible with respect to the power needed for

Table 1.2 Typical Power Consumption for the Micro Controller Units (MCUs) Texas Instruments MSP430 and Microchip PIC24F16, and for the Transceivers (TX/RX) Texas Instruments (TI) CC2500 and Microchip MRF24J40. Such Components are Commonly Used in Wireless Sensor Networks.

Component	Type	P_{act}(typical)	P_{sleep}(typical)
TI MSP430	MCU	$1mW$	$2\mu W$
PIC24F16	MCU	$1.5mW$	$1\mu W$
CC2500	TX/RX	$50mW$	$2\mu W$
MRF24J40	TX/RX	$60mW$	$6\mu W$

Table 1.3 Power Consumption for Different Transmission Powers (TX Power) for the Transceiver Texas Instruments CC2500.

TX Power	P_{act}
$-12\,dBm$	$33.3mW$
$-6\,dBm$	$45mW$
$0\,dBm$	$63.6mW$
$1\,dBm$	$64.5mW$

the RF transmission. Such an example is shown in Table 1.3, where the power consumption of the TI CC2500 transceiver is reported. From Table 1.3 it can be seen that lowering the transmission power of more than an order of magnitude does not imply the same reduction of the overall absorbed power.

1.3 Medium Access Control Protocols for WSNs

One of the main issues in the design of WSNs is the efficient utilization of the radio resources (e.g., frequency bands) when they are shared among multiple nodes. This problem is tackled by designing *medium access control* (MAC) protocols, whose purpose is to regulate the transmissions of the nodes over the shared channel [8].

MAC protocols can be broadly divided into two categories: *random* and *centralized scheduling-based* schemes, which are briefly introduced in the next sections. For an extensive review of MAC protocols see e.g., [9, 10] and reference therein.

1.3.1 Random MAC Protocols

In random access protocols, the nodes access the channel in a random fashion according to a set of rules specified by the MAC. The simplest random access scheme is the pure ALOHA protocol [11], in which any node in the network simply transmits a packet whenever it is generated. Due to the absence of time synchronization and coordination at nodes, the pure ALOHA protocol is severely degraded by the interference that is generated by simultaneously transmitting nodes. In particular, it has been shown in [11] that, under the assumption of a collision channel model (i.e., any packet involved in a simultaneous transmission becomes garbled), the maximum throughput of pure ALOHA is $1/2e$, that is, on average only 18.4% of the time the channel is successfully used.

A simple way to improve the channel utilization of the pure ALOHA protocol is by dividing the time into time-slots, so that nodes can transmit still in random fashion but only within a single time-slot [12]. Note that *slotted*-ALOHA requires synchronization among the nodes. It has been shown in [12] that the slotted-ALOHA protocol can achieve a throughput of $1/e$.

A way to control the transmission of the nodes in the slotted-ALOHA protocol is to have a central controller that organizes time-slots into frames, where each node can transmit only once in each frame [13]. This protocol is referred to as *framed-ALOHA* (FA). A variation of the basic FA protocol, allows the central controller to dynamically adjust the frame size based on the outcomes of nodes transmissions in previous frames, and it is referred to as *dynamic framed-ALOHA* (DFA). It has been shown that DFA have several advantages over the simpler slotted-ALOHA, including

improved data queue stability and reduced acknowledgment overhead [14]. However, as said, it requires a central controller that dynamically selects the frame size.

A way to reduce the chances of packet collisions in ALOHA-based protocols is to consider the *carrier sensing multiple access* (CSMA) protocol [15]. The basic idea of CSMA is that each node listens to the channel before attempting transmission. If no other transmissions are detected, then transmission is performed, while otherwise the node waits and checks the channel later on with the same procedure.

1.3.2   Centralized Scheduling MAC Protocols

In centralized scheduling-based protocols, the nodes are assigned an exclusive channel resource by a central unit (see e.g., [8]). The central unit either pre-assigns the resources to the nodes in a static fashion, such as in the *time division multiple access* (TDMA) protocol, or it dynamically allocates them based on the system conditions (e.g., quality of the radio link), such as in *opportunistic scheduling* schemes (see e.g., [16]). Opportunistic scheduling requires the broadcasting of a scheduling command that specifies when (and for how long) each node is allowed for transmission over the channel. The advantage of scheduling-based protocols is that they prevent the energy wastage due to collisions and that they can often guarantee deterministic performance levels. The disadvantage is that they generally requires tight synchronization and extensive signaling overhead for resource allocation.

1.3.3   MAC Performance Metrics

There are several relevant criteria that measure the performance of a MAC protocol, and the choice of the most appropriate ones depends on the network architecture and on the application requirements. Some of the most important criteria are *throughput*, *transmission delay* and *reliability* [5, 9]. The throughput measures the fraction of the allocated channel resources that are successfully utilized for data transmission.

Instead, the delay measures the average time spent by a packet between the time it is generated and the time it is successfully received by the destination. The reliability is an indicator of the ability of a protocol to correctly deliver data messages.

1.3.4   Energy Consumptions Due to the MAC Protocol

A MAC protocol not only affects the performance of the network in terms of, e.g., throughput, transmission delays and reliability, but it also has a strong impact on the energy consumption of the nodes. In fact, as shown in Section 1.2.1, two of the most power-hungry operations in a wireless node are transmission and reception of data. Therefore, a MAC protocol that parsimoniously utilizes the node's transceiver, and thus the energy resources, is highly favorable.

Depending on the structure of the network, the most common sources of energy wastage due to a MAC protocol are (see e.g., [17]): i) *collisions*; ii) *idle listenings*; iii) *overhearing*; iv) *protocol overhead*. Collisions occur when multiple nodes attempt transmission simultaneously and one or more of the involved messages cannot be correctly decoded by the intended destination(s), due to the interference generated by the other nodes. Depending on the applications, collisions might require that the messages need to be either retransmitted, thus consuming additional energy and increasing delays, or discarded. Idle listenings occur when a node turns its receiver on waiting for other nodes transmissions that do not occur. Overhearing means that a node receives a message that is not intended for it. Protocol overhead includes all the sources of energy consumptions that are related to the exchange of signaling messages required by the MAC.

1.4   Motivation of the Dissertation

The main focus of this dissertation is the study of the impact of EH technologies in the design of wireless networks. Until the last decade, wireless networks have

been conventionally designed by considering that the nodes in the network are either powered by batteries or directly connected to the power grid. Typical examples include cellular networks, where the nodes are *battery-powered* cell phones, or wireless data networks, such as WiFi, in which the nodes can be either powered by batteries or connected to the grid (e.g., laptop or desktop computers, respectively). The main design goals in these networks is either the maximization of the batteries lifetimes or the minimization of the average power consumption while guaranteeing a determined quality of service [5, 18].

However, when the nodes in a network are powered through EH, the energy availability at the terminals might not be guaranteed at any given time. This is due to the fact that the EH-sources are generally unpredictable and highly variable over time [1]. Therefore, despite the energy availability over a long period of time is generally unlimited, the energy available over a short period of time might not be sufficient to guarantee the required operations of the devices. This observation enlightens the fundamental differences between battery-powered and EH devices. The former are equipped with a finite amount of energy that is always available when required within the battery lifetime, while the latter are provided with a theoretically infinite lifetime, but possibly with no guarantee of continuous operations due to temporary energy shortages. Therefore, the design of wireless networks must be restructured to accommodate the novel features introduced by EH. This is the main goal of this dissertation. More specifically, the focus will be on the analysis and design of MAC protocols for EH networks.

Section 1.5 provides an overview of previous work related to the dissertation, while specific contributions of this work are described in detail in Section 1.6.

1.5 State of the Art

General references that describe EH technologies with a focus on wireless networks applications are described in the next section. Previous work that is more directly connected to this dissertation is then discussed by concentrating separately on single-node systems and multi-node systems.

1.5.1 Energy Harvesting Technologies and Principles

An extensive review of EH technologies is given in [6] and [19], while a description more specific to wireless network applications is provided in [1]. Fundamentals of energy neutral operations for EH-capable nodes were established in [20] and reference therein. Reference [6] also provides a discussion of the characteristics of several energy storage devices.

1.5.2 Single-node Systems

Works on single-node systems focus on the problem of trading the energy harvested from the environment with the energy needed by the node to perform the required operations, such as sensing and data transmission. Here, the goal is generally the optimization of the energy usage with the aim of maximizing a given performance criterion such as the data transmission rate.

In [21] a single node equipped with a finite replenishable battery is considered. At any given time, the problem is whether to perform transmission or not based on the current available energy and given that a reward is accrued if transmission is performed. By modeling the evolution of the energy in the battery as a controlled Markov process, where the control action is the decision of whether to transmit or not, the authors found the structure of optimal transmission policies by resorting to theory of Markov decision processes (MDP).

The problem of optimizing transmission policies for a single EH-capable node equipped with infinite battery and data queue is considered in [22]. Here the authors consider random energy and data arrivals and derive throughput-optimal policies as well as delay-minimizing policies. Data queue stability issues are also discussed.

A problem related to [22] is considered in [23], where the node's battery is finite and the times of arrivals of the energy harvested from the environment are assumed to be known in advance at the beginning of the data transmission. The problem is to maximize the amount of data transmitted over a finite horizon of time, by assuming that the node has an unlimited amount of data initially available for transmission. The authors also found an optimal policy for an equivalent problem in which the goal is the minimization of the time needed for the transmission of a given amount of data. A related problem is also considered in [24], where data arrivals are allowed also after the beginning of the transmission but at times known in advance. Extensions of [23] and [24] that include transmission over fading channels and non-idealities in the energy storage devices are considered in [25] and [26].

1.5.3   Multi-node Systems

In multi-node systems several EH-capable nodes interact with each other, and the goal is generally the optimization of either a common performance criterion, such as the total network throughput, or a performance requirement to be satisfied at each single-node, such as data queue stability.

In [27] data queues stability issues are addressed for multiple access problems in single-hop networks, by considering TDMA, CSMA and opportunistic scheduling protocols. Scheduling problems for general mesh networks operated by EH-capable nodes were instead considered in [28], where Lyapunov optimization techniques were leveraged.

A simple MAC problem with two nodes and a single receiver is considered in [29]. Here, the two transmitting nodes receive energy at times that are known in advance, while the data they need to transmit is already available before beginning transmission. The goal is to minimize the overall transmission time by optimally selecting the node transmission powers and data rates. Optimal policies are found explicitly.

There are other previous works for EH networks not strictly related to the objectives of this dissertation that include broadcasting channels [30] and [31], as well as relay networks [32] and routing problems [20, 33].

1.6 Dissertation Outline and Contributions

The main contributions of this dissertation cover the analysis and design of MAC protocols for EH wireless networks. In particular, both random access and centralized scheduling-based MAC protocols are investigated for single-hop wireless networks in Section 1.6.2 and Section 1.6.3, respectively. Energy management techniques for single-node systems are considered as well and are described in Section 1.6.1.

1.6.1 Single-node Systems

Chapter 2 considers a single-node system with EH capabilities where the goal is the maximization of a given performance metric via an optimized energy usage. In particular, a new architecture for enhanced passive radio frequency identification (RFID) tags, equipped with EH capabilities, is proposed jointly with optimal energy management techniques. The new architecture is introduced to tackle the problem of increasing the communication reliability (or the read range) between a passive RFID tag and a RFID reader in a backscatter modulation-based system (see e.g., [34]). It is proposed to introduce a power amplifier (PA) that increases the power of the signal transmitted by the tag to the reader, where the peculiarity is due to

the fact that the PA is exclusively powered via harnessing the electromagnetic energy transmitted by the reader. The architecture proposed in this dissertation is related to the one proposed in [35], where however the PA was powered via a non-rechargeable battery. Whereas, the mathematical modeling developed to establish optimal energy management policies is related to [21], where the authors considered a different energy harvesting model and different performance metrics.

The work in this chapter is based on:

- F. Iannello O. Simeone and U. Spagnolini, "Energy management policies for passive RFID sensors with RF-energy harvesting," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Cape Town, South Africa, May 2010.

1.6.2   Random Access MAC Protocols

In Chapter 3 and Chapter 4 the problem of designing Framed-ALOHA based MAC protocols for single-hop EH networks is investigated. The considered application is a batch resolution problem [36], where data packets are periodically generated at the nodes and need to be collected by a central fusion center in a star-topology network. The EH arrivals at the nodes are described by an arbitrary probability distribution and the energy storage devices are assumed to be finite, while the communication links are subject to random fading.

To assess the novel trade-offs in the design of MAC protocols for EH networks, Chapter 3 proposes to utilize two performance metrics. The first metric, referred to as *time efficiency*, measures the data collection rate at the fusion center, while the second metric, referred to as *delivery probability*, accounts for the probability that any packet generated at the nodes is eventually collected by the fusion center. Due to the potential perpetual operations of the nodes enabled by EH, the proposed performance metrics are assessed over a long-term period by developing a mathematical framework based on Markov models, which describes the evolution of the energy availability

at the nodes along time. The critical issue in ALOHA-based scheme of estimating the number of nodes involved in the transmission in each frame is also tackled by proposing a practical reduced-complexity algorithm. This scheme is an extension of the one proposed in [13] that is designed to account for the EH nature of the nodes.

From the analysis of the performance metrics described above, it is inferred that the trade-off between time efficiency and delivery probability is dramatically affected by a design parameter that is used to select the frame size in the framed-ALOHA protocol, which in turns depends on the number of transmitting nodes in each frame. It is shown that the choice of such parameter strongly depends on the probability distribution of the EH processes and on the desired trade-off between time efficiency and delivery probability. Based on this insight, a new protocol, referred to as *energy group dynamic framed-ALOHA* (EG-DFA), is proposed in Chapter 4. The proposed EG-DFA protocol creates groups of nodes according to their energy availability and runs optimized and separated instances of the DFA protocol for each group. It is shown that by judiciously choosing the frame-size parameter for each group of nodes the EG-DFA protocol can remarkably outperform the conventional DFA protocol.

The work in these chapters is based on:

- F. Iannello, O. Simeone, and U. Spagnolini, "Medium access control protocols for wireless sensor networks with energy harvesting," *IEEE Trans. Commun.*, May 2012 (in press).

- F. Iannello, O. Simeone, P. Popovski and U. Spagnolini, "Energy group-based dynamic framed ALOHA for wireless networks with energy harvesting," in *Proc. 46th Conf. Inf. Sci. Syst.* (*CISS*), Princeton, NJ, Mar. 2012.

- F. Iannello, O. Simeone, and U. Spagnolini, "Dynamic framed-ALOHA for energy-constrained wireless sensor networks with energy harvesting," in *Proc. IEEE GLOBECOM*, Miami, USA, Dec. 2010.

1.6.3 Centralized Scheduling MAC Protocols

The third important aspect considered in this dissertation is the design of scheduling-based MAC protocols for EH networks. This issue is addressed in Chapter 5 and Chapter 6. As anticipated in Section 1.5.3, few previous works considered scheduling problems in EH networks. In particular [29] consider a two-nodes system with deterministic energy arrivals, while [28] considers a generally suboptimal Lyapunov optimization approach for a scheduling problems in arbitrarily interconnected networks.

In this dissertation the focus is instead on a star-topology network in which a central fusion center collects data packets that are generated periodically by a set of M nodes, similar to the model considered in Section 1.6.2. The nodes harvest energy from the environment, and their energy storage devices are finite and possibly subject to energy leakage. In each data collection period only a subset of $K \leq M$ nodes is given the chance of transmitting over orthogonal transmission resources, which are allocated by the fusion center.

As mentioned in the previous sections, since the activity of most EH sources is uncertain and unpredictable, nodes that are exclusively powered via EH are possibly subject to temporary energy shortages. Based on this observation, it is possible to distinguish two different scenarios: i) Applications that require continuous operation of the nodes and that do not tolerate temporary energy shortages; ii) applications that tolerate energy shortages. When applications do not tolerate energy shortages, it is not uncommon that EH is used as a secondary energy source that complements the use of a non-rechargeable battery [37]. In this case the nodes are equipped with a so

called *hybrid energy storage system* (HESS), which is composed by a non-rechargeable battery and, e.g., a capacitor that stores the energy harvested from the environment. The network design goal here is to maximize the lifetime of the non-rechargeable batteries. When applications that tolerate temporary energy shortages are instead considered, EH can be used as the unique energy source, and the scheduling policies are designed so as to maximize the network throughput. Scheduling problems for both scenarios are addressed in Chapter 5 and Chapter 6.

In particular, optimal scheduling policies that maximize the battery lifetime of the HESS-nodes are derived under the assumptions that: the fusion center has perfect and instantaneous knowledge of the energy availability at the nodes; the nodes are subject to either energy harvesting only or energy leakage only; the energy harvesting and energy leakage are described by binary random processes, which are assumed symmetric and independent at the nodes and over time. The general case when both energy harvesting and energy leakage processes are non-negligible still remains an open problem.

The scheduling problems above are then addressed under the assumption that the fusion center does not have instantaneous information of the energy availability at the nodes. In this case, the only information available at the fusion center is given by the knowledge of the statistical properties of the energy harvesting and leakage processes at the nodes and by the (observable) history of the system state. The scheduling problem is then formulated as a partially observable Markov decision process (POMDP), which can be seen a restless multiarmed bandit (RMAB) problem [38]. In the scenario in which nodes are equipped with HESS, finding optimal policies explicitly is not straightforward, and thus only heuristic policies are proposed and compared to the full state information scenario.

For the scenario in which the nodes are powered exclusively via EH and under partial state information at the fusion center, optimal scheduling policies are derived

under the assumption that the ESD at the nodes is of capacity one. For this case, it is shown that a myopic, or greedy, policy that operates on the space of the a posteriori probabilities (or beliefs) of the nodes energy levels is optimal. Moreover, it is demonstrated that such policy coincides with the so called Whittle index policy. It is worth mentioning that the derivation of the optimality of the myopic policy and of the Whittle index policy is related to complementary findings in RMAB problems arising in cognitive radio applications [39, 40]. Finally, when the size of the capacitors are arbitrary, a performance upper bound is derived and compared with the performance of the generally suboptimal myopic policy.

The work in these chapters is based on:

- F. Iannello, O. Simeone and U. Spagnolini, "Lifetime maximization for wireless networks with hybrid energy storage systems," in preparation for submission to *IEEE Trans. Commun.*

- F. Iannello, O. Simeone and U. Spagnolini, "On the optimal scheduling of independent, symmetric, and time-sensitive tasks," submitted to *IEEE Trans. Autom. Control* (under first revision).

- F. Iannello, O. Simeone and U. Spagnolini, "Optimality of myopic scheduling and whittle indexability for energy harvesting sensors," in *Proc. 46th Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, Mar. 2012.

Part I

Energy Management Policies for Single-node Systems

This part of the dissertation considers a wireless network in which a single node communicates with a central station, where the latter coordinates the node's transmissions. The node is equipped with energy harvesting (EH) and storage capabilities, so that the use of the harvested energy can be postponed over time. In general, in single node EH networks the design issue is how to trade the energy harvested from the environment with the energy needed by the node to perform the required operations, such as data transmission. Energy management policies are then designed with the aim of optimizing a given performance criterion.

A specific instance of such single node EH networks is considered in the next chapter, where a RFID system operated by enhanced RFID tags is investigated. In particular, in such system, EH is leveraged with the aim of improving the communication reliability between the tag and the central station (or RFID reader). This is done by introducing an additional power amplifier at the tag that is exclusively powered via EH. Energy scheduling policies for the power amplifier are then designed by parsimoniously trading the energy available in the tag's energy storage device and the statistical properties of the EH process.

CHAPTER 2

ENERGY MANAGEMENT POLICIES FOR ENHANCED PASSIVE RFID TAGS WITH ENERGY HARVESTING

2.1   Introduction

Passive radio frequency identification (RFID) technology is finding an ever increasing number of applications, ranging from conventional identification such as supply chain management or toll collections, to wireless sensor networks (WSNs), where identification is provided along with sensed data [41]. A typical far-field passive RFID sensor network consists of one (or more) RFID reader and a number of RFID sensors (also *tags* in the sequel). The tags communicate data to the reader by modulating (possibly amplifying) and transmitting back a continuous wave (CW) that is emitted by the reader itself. This process is referred to as *backscatter modulation* [42].

The RF field emitted by the reader is the only source of energy that allows passive tags to activate their circuitry, while more sophisticated classes of tags, such as semi-active and active, rely on energy storage devices (simply batteries in the sequel) charged at the time of installation [42]. In semi-active tags the onboard battery is used to activate part or all the tag circuitry, but the communication with the RFID reader is still performed via backscatter modulation as in passive tags (i.e., without the use of the on-board battery). Active tags instead do not rely on backscatter modulation, and they use their batteries to activate their circuitry including an on-board transceiver for communication with the reader. Active and semi-active tags enable more sophisticated applications than passive tags, at the price of increasing cost and typically limited lifetime due to the finite energy available in the batteries.
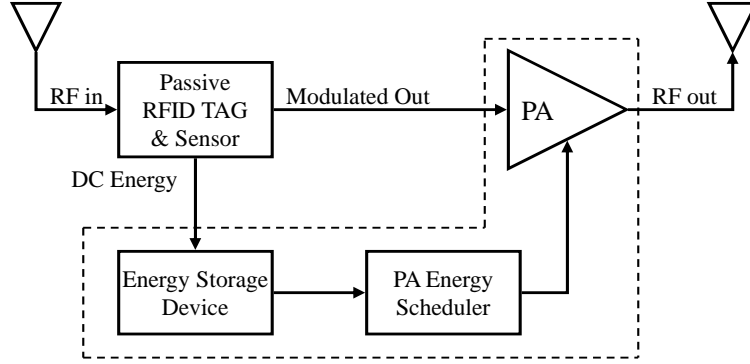
Figure 2.1 Block diagram of an RFID ABEH sensor. The dashed box contains the novel components with respect to classic passive RFID sensors.

One of the most important RFID system performance metric is the *read range*, or equivalently the maximum distance at which the reader can reliably read (or write) the data from (to) the RFID sensors [34]. Two main factors determine the read range: 1) *Tag sensitivity* (tag-limited regime), which is determined by the minimum power received by the tag necessary to activate its circuitry; 2) *Reader sensitivity* (reader-limited regime)*, which is determined by the minimum *signal to noise ratio* (SNR), or alternatively, the minimum power at the reader that enables correct detection of the signal backscattered by the tag.

The new conceptual scheme that is proposed in this chapter aims at addressing the issue of reader-limited regime by introducing two additional components to the hardware architecture of conventional passive tags as shown in Figure 2.1:

- A power amplifier (PA), which is used to amplify the backscatter signal (i.e., the reader's CW processed and transmitted back by the tag);

- An energy storage device (e.g., battery or capacitor), which is charged via energy harvesting.

This enhanced tag architecture, referred to as *amplified backscattering via energy harvesting* (ABEH), is still passive, in the sense that it does not need any initially charged battery (or capacitor). In fact, it exploits the RF-energy transmitted by

the reader, and received by the tag during idle periods, to recharge the onboard battery. The harvested energy is then used by the tags to opportunistically amplify the backscatter signal, with the aim of improving the communication reliability. Notice that RFID tags with ABEH architecture (ABEH tags for short) inherit the theoretically infinite lifetime of passive tags, since in case of depleted battery they can operate as conventional passive tags.

An energy scheduler manages the energy used by the PA to amplify the backscatter signal with the aim of improving the read range of the ABEH tags. This is done by conveniently balancing the instantaneous state of charge of the battery and the energy harvesting rate. The analysis demonstrates that the amplification of the backscatter signal enhances the read range in the reader-limited regime of operation. It is noted that the approach of this chapter could be extended to include the trade-off between energy used for backscatter amplification and for powering the tag circuitry (including the onboard sensor).

2.1.1 Previous Work

A brief overview of previous work related to this chapter is now introduced. In [35] the problem of reader sensitivity is addressed in a similar fashion as ABEH tags by allowing amplified backscatter from the RFID tags. However, in [35] the PA is fed by an external power source (active tags), thus differing from ABEH tags where the energy for amplification is harvested from the CW transmitted by the reader. The problem of tag-limited regime is addressed in [43], where an independent CW source is installed on the tag and acts as an energy pump fed by a battery, while in [44] sleep and wake cycles together with energy harvesting techniques are proposed. Transmission policies optimization for replenishable sensors is addressed in [21] where the authors resort to an analytical model based on Markov decision process (MDP). Battery-free RFID transponders with sensing capability that harvest all the needed energy from

the RF signal emitted by the reader are investigated in [4, and references therein] together with possible applications. Discussion on energy storage architectures, for enhanced RFID tags, can be found in [45]. Measures and statistical characterization of the effect of the fading and path loss in a backscatter modulation-based system are presented in [46].

The chapter is organized as follows. Section 2.2 introduces the signal and system models used throughout the chapter, while Section 2.3 describes the working principle of ABEH tags. The energy scheduling problem is formalized as a MDP in Section 2.4 (see [47] for an overview of MDP), while optimal scheduling policies are derived in Section 2.5. Numerical results are then presented in Section 2.6 and finally some conclusions are drawn in Section 2.7 together with possible extensions.

2.2   System Model

The focus is on a far-field RFID system, with a single-reader and multiple-tags [42, 48]. The operation of the considered RFID network in the presence of passive tags can be generally summarized with the following phases (a commercial example is the Gen-2 standard [48]). The reader transmits a CW to energize the entire population of tags [44]. After a time period long enough for the tags to activate their circuitry (by accumulating energy from the CW), the reader starts transmitting a modulated signal containing a selection command to choose a subset of tags. After this phase, the reader transmits a sequence of *query* commands (Q) of T_q seconds each, to request information from the selected tags. Data transmission from the tags take place during a subsequent period of duration T_c, in which the selected tags perform backscatter modulation. The combination of a query command and CW forms a *time-slot* of duration $T = T_q + T_c$ (see Figure 2.2).

A collision protocol is generally necessary to arbitrate the access of the (possibly multiple) selected tags. In order to simplify the problem and focusing

on the energy management of ABEH tags, it is assumed here that in every slot one *single* tag is selected by the reader's query to respond via backscatter modulation, independently from previous and future queries. Notice that, the impact of collisions, due to the multiple access, could be taken into account by conveniently modifying the probability of successful transmission that will be defined in (2.9). However, this collision-free assumption is reasonable in scenarios where RFID tags are selected according to their unique identifiers (known at the reader) as possibly for RFID-based sensor networks (see [48]).

Because of both collision-free and independent queries assumptions, one can focus on a simplified single-reader single-tag scenario, where the downlink (DL) frame structure transmitted by the reader is composed by successive slots, each one containing a query command and a CW as shown in Figure 2.2. In each slot, the unique tag in this scenario (simply *the tag* in the sequel) after having decoded the query, can assume two different states (see Figure 2.2):
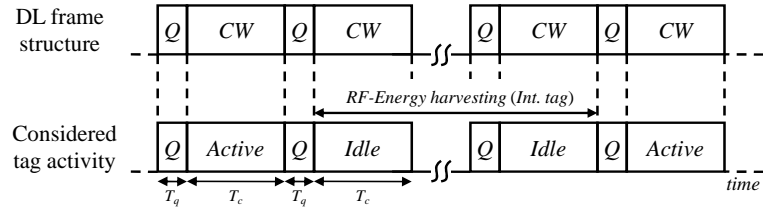
Figure 2.2 Reader DL frame structure and interrogated tag activity. A single time-slot is composed by two parts: Query command (Q) and continuous wave (CW). During the CW period a tag can be either *active* (transmitting data) or *idle* (harvesting energy).

- *Active time-slot* for the tag, with probability p it switches its state to active and performs backscatter modulation to transmit the required data to the reader (the tag is interrogated).

- *Idle time-slots* for the tag, with probability $1-p$, it switches its state to idle and harvests the RF-energy transmitted by the reader (the tag is not interrogated).

Notice that, in a general multiple-tags scenario, the interrogation probability p depends on the number of tags and on the rate at which the reader needs to collect information from each tag. Furthermore, the probability p can also take into account tag collisions at the reader and demodulation errors of the query commands (not explicitly modeled here).

This chapter consider bistatic RFID readers that use two antennas, one for transmission (TX antenna) and one for reception (RX antenna) (see [34] and [46]). The links TX antenna to tag and tag to RX antenna are referred to as downlink (DL) and uplink (UL), respectively. It is assumed the same distance d from tag to reader RX and TX antenna, which is also fixed for the entire operations. During slot k, the DL (UL) channel $h_{dl}(k)$ $(h_{ul}(k))$ is subject to frequency-flat fading, which is assumed being constant over the entire slot. However, the fading in each slot is modeled as an independent and identically distributed (i.i.d.) random variable. Assuming that the duration T_q of the query command is much shorter than the duration T_c of the CW (i.e., $T_q \ll T_c \simeq T$), the signal impinging on the tag is

$$y(t;k) = \sqrt{L}h_{dl}(k)x(t) + w(t;k), \tag{2.1}$$

where $kT \leq t < (k+1)T$ runs over the kth slot (of duration T), and the energy per slot available for the transmission of the CW is E_0. The propagation loss between the reader and the tag is denoted by L and it is assumed constant since the distance d between tag and reader is fixed. The CW transmitted by the reader, of energy E_0, is $x(t) = \sqrt{2E_0/T}\cos 2\pi f_0 t$, where f_0 is the carrier frequency and $w(t;k)$ is an additive white Gaussian noise (AWGN) in the band of interest, with $w(t;k) \sim \mathcal{N}(0, \sigma_t^2)$.

2.3 ABEH Functionality

An ABEH tag is characterized by the following operations: 1) It harvests and stores energy during idle slots; 2) it opportunistically amplifies the backscatter signal during

active slots, as controlled by the energy scheduler. In principle, the energy $E_b(k)$ drawn from the battery by the energy scheduler in slot k may depend on a number of factors, such as the current state of charge of the battery $S(k)$, the energy evolution over the past slots, the interrogation probability p, the DL and UL channels quality (channel state information) and the path loss L. In practice, all this information cannot be dynamically tracked by simple devices like RFID tags and some simpler policies must be used. Specifically, scheduling policies (pre-determined and possibly stored into the tag memory) that do not depend on the entire history of previous observations, i.e., *stationary policies* (see, e.g., [47]) are considered. These policies depend on the following static system parameters, assumed to be time-invariant and known at the tag (or possibly communicated by the reader queries): interrogation probability p, path loss L and DL and UL channel statistics. The only quantity that needs to be measured by the tag is the *state* of the battery $S(k)$.

Optimal policies need to balance the energy harvesting rate, which is out of the tag's control, and the probability of successful transmission, which can be controlled by the energy scheduler by varying the energy drawn from the battery for backscatter amplification. The goal of the energy scheduler is to maximize the performance (read range) of ABEH tags. The next section characterizes the energy harvesting process (during idle slots) and then introduces the effects of the backscatter signal amplification on the backscatter SNR at the reader (during active slots).

2.3.1 Idle Time-Slots: RF-Energy Harvesting

The energy received by the tag during slot k, can be easily derived from (2.1) as

$$E(k) = \int_{kT}^{(k+1)T} |y(t;k)|^2 \, dt \simeq LE_0 \, |h_{dl}(k)|^2 \,, \tag{2.2}$$

where the energy of the noise is negligible compared to the signal energy, i.e., $LE_0 \, |h_{dl}(k)|^2 \gg \sigma_t^2 T$. In order to make the RF-energy available for storage, the signal

(2.1) received by the tag passes through a RF-to-DC converter, with a conversion efficiency $\eta_{DC} \in [0,1)$, which is assumed being constant for all the RF input power levels (see [49] for a more detailed treatment). The energy available for storage during slot k is

$$\mathcal{E}(k) = \eta_{DC}E(k) = \eta_{DC}LE_0 \left|h_{dl}(k)\right|^2 . \tag{2.3}$$

Notice that the randomness of the available energy $\mathcal{E}(k)$ is due to DL fading channel $\left|h_{dl}(k)\right|^2$.

2.3.2 Active Time-Slots: Backscatter SNR

During active slots, the interrogated tag replies to the reader queries by transmitting back information through backscatter modulation. With an ABEH tag, the backscattered signals can be amplified by feeding the PA with an amount of energy $E_b(k)$ that is drawn from the tag's on-board battery (see Section 2.4). The instantaneous SNR at the RFID reader during active slots can thus be written as (derivation is omitted here, see [34] and [50])

$$\gamma(E_b(k);k) \;\;=\;\; \frac{L^2 E_0 \left|h_{ul}(k)\right|^2 \left|h_{dl}(k)\right|^2}{\sigma_r^2 T}\eta_{mod} + \tag{2.4}$$

$$\frac{L \left|h_{ul}(k)\right|^2 E_b(k)}{\sigma_r^2 T}\eta_{amp}, \tag{2.5}$$

where σ_r^2 is the power of the AWGN at the reader, while $h_{dl}(k)$ and $h_{ul}(k)$ are the DL and UL fading channels, respectively. Furthermore, $\eta_{mod} \in (0,1)$ is the tag transmission efficiency accounting for the effects of the backscattering process [34], and $\eta_{amp} \in (0,1)$ is the efficiency of the PA. The first term in (2.4) is the SNR that one would have when using conventional passive tags, which are not equipped with amplification capabilities for the backscatter signal (see [46] and [50]). The second term is due to the amplification performed by the ABEH tag, and depends only on the UL channel.

2.4   Battery Evolution: A Markov Chain Model

The evolution of the energy stored in the battery is modeled by resorting to a discrete Markov chain model (e.g., [51]). The battery is of size $E_{\max}$ [J] and is uniformly divided into N states, representing different energy levels, where the *energy-unit* is $\delta_E = E_{\max}/(N-1)$. The state of the battery is $S(k) \in \{0, ..., N-1\}$. It is noted that the discrete model at hand is an approximation of a continuous quantity (the harvested energy). Therefore, making δ_E as small as possible insures that the state of the battery can be modeled more accurately, at the cost of increasing the complexity of the model.

A stationary policy $\lambda = [\lambda_0, ..., \lambda_{N-1}]^T$ can be defined as the set of actions that the energy scheduler takes for every possible value of the state variable $S(k)$, regardless of the time slot k, and fixed the system parameters as described in Section 2.3. More specifically, action λ_n, for $n \in \{0, ..., N-1\}$, is a non-negative integer $\lambda_n \in \{0, ..., n\}$ that corresponds to the number of energy-units δ_E (or equivalently $E_b(k) = \delta_E \lambda_n$) drawn from the battery for amplification when the tag is in state $S(k) = n$. Notice that, at state $S(k) = n$, the energy scheduler of the ABEH tag has $n + 1$ possible choices for λ_n, so that the total number of available stationary policies for N levels is $1 \cdot 2 \cdot ... \cdot N = N!$. This makes an exhaustive search of the optimal policies an highly complex task. The simplest policy that can be used as a reference is the *draw-all* policy (or greedy), where all the energy currently stored in the battery is used to amplify the backscatter signal (i.e., $\lambda_n = n$). The numerical results presented in Section 2.6, also consider strategies that are limited to schedule energy in steps larger than δ_E due to possible technological constraints.

2.4.1   Transition Probabilities

The evolution of the energy stored by the ABEH tag, depends on tag interrogation probability p, and on the statistical properties of the wireless channel. Specifically,

energy harvesting during idle slots may determine transitions toward higher energy levels, depending on the channel quality (see Section 2.3.1). Conversely, during active slots the energy scheduler draws some energy-units from the battery, thus determining a transition toward a lower energy level (see Section 2.3.2).

For any stationary energy scheduling policies, the state of the battery $S(k)$ evolves over the slots as an irreducible and aperiodic time-homogeneous Markov chain (see Figure 2.3) (the Markov chain is thus ergodic). The transitions toward higher energy levels depend on the probability $q = 1 - p$ of having an idle slot, and on the probability that the harvested energy $\mathcal{E}(k)$ (see (2.3)) allows the ABEH tag to store some energy-units δ_E. The conditional probability β_{nl} that, during an idle slot, there is a transition from state $S(k) = n$ to $S(k+1) = l$, can be obtained as follows

$$
\beta_{nl} = \Pr\left[S(k+1) = l | S(k) = n, idle\right] =
$$
$$
\begin{cases}
\Pr\left[(l-n)\delta_E \leq \mathcal{E}(k) < (l-n+1)\delta_E\right] & l \leq N - 2 \\
\Pr\left[\mathcal{E}(k) \geq (l-n)\delta_E\right] & l = N - 1 \quad , \\
0 & 0 \leq l < n
\end{cases} \tag{2.6}
$$

where the second row of the right-hand side of (2.6) accounts for the highest energy level, while the third row indicates that there is no energy leakage during idle slots. Notice that $\sum_{l=0}^{N-1} \beta_{nl} = 1$, for all $n \in \{0, ..., N-1\}$. Once again energy quantization δ_E should be small enough to capture small variation of the harvested energy $\mathcal{E}(k)$ when modeling the system. Conversely, during active slots the transition toward a lower, or at least the same, energy level, is deterministically defined by the policy λ. To sum up, by resorting to the law of total probability, the nth entry of the lth row $[\mathbf{P}]_{nl}$ of the transition probability matrix $\mathbf{P}$ for the Markov chain in Figure 2.3 can be written as the sum of two contributions, one from idle-time slots with probability

$(1-p)\beta_{nl}$, and one from active slots with probability p if and only if $\lambda_n = n - l$

$$
\begin{aligned}
[\mathbf{P}]_{nl} &= \Pr\left[S(k+1) = l | S(k) = n\right] && (2.7) \\
&= \begin{cases} (1-p)\beta_{nl} & l \neq n - \lambda_n \\ (1-p)\beta_{nl} + p & l = n - \lambda_n \end{cases}. && (2.8)
\end{aligned}
$$

As it will be shown below, the problem of finding optimal stationary policies can be classified as a MDP.

2.5 Optimal Energy Scheduling Policies

The aim of the ABEH tag is to improve the read range in the reader-limited regime (see Section 2.1). Given the randomness induced by the fading channels and the noise at the reader, the read range is evaluated in terms of the probability that the reader correctly decodes the tag signal, referred to as read probability, for a given distance tag-reader. The read probability is defined as follows

$$
r(\lambda_n) = \Pr\left[\gamma(\lambda_n \delta_E; k) \geq \gamma_{th}\right], \tag{2.9}
$$

where $\gamma(\lambda_n \delta_E; k)$ is the instantaneous SNR (see (2.4) with $E_b(k) = \lambda_n \delta_E$) at the reader given that the ABEH tag battery is in state n and λ_n energy-units are drawn from the battery for backscatter amplification, while threshold γ_{th} is the minimum SNR that allows correct decoding. Notice that, the dependence of the read probability (2.9) on the distance d, and thus the relation with the read range, is implicitly contained in the definition of the instantaneous SNR (2.4). In order to evaluate the average performance of the ABEH tag over an increasing number of active slots, it is possible to define the long-term average read probability as follows

$$
g(\lambda) = \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{v}_0^T \mathbf{P}^k(\lambda) \mathbf{r}(\lambda) = \pi^T(\lambda) \mathbf{r}(\lambda), \tag{2.10}
$$

where the product $\mathbf{v}_0^T \mathbf{P}^k(\lambda)$ indicates the probability distribution of the energy in the tag battery after k slots, given the transition matrix $\mathbf{P}(\lambda)$ defined in (2.7), while $\mathbf{v}_0$ is an arbitrary initial distribution vector. The vector $\mathbf{r}(\lambda) = [r_{\lambda_0}, ..., r_{\lambda_{N-1}}]^T$ contains the read probability as a function of the policy λ with entries defined in (2.9). The right-hand side of (2.10) follows from the fact that the Markov chain at hand is ergodic (see Section 2.4.1), and thus one can calculate the steady state stationary distribution vector $\pi(\lambda) = [\pi_0(\lambda), ..., \pi_{N-1}(\lambda)]^T$. Due to the fact that the Markov chain is ergodic, it also follows that $g(\lambda)$ does not depend on the initial vector $\mathbf{v}_0$, but it is uniquely defined by the policy λ.

The optimal stationary policy $\lambda^* = [\lambda_0^*, ..., \lambda_{N-1}^*]$, is defined as the stationary policy that maximizes the long-term average read probability (2.10), such that $g(\lambda^*) \geq g(\lambda)$, for all λ. Notice that, the use of stationary policies is not a restriction for the considered system setting (described in Section 2.3), since it can be proved that they are optimal for the MDP at hand (see [52]). Notice that, the dependence of (2.10) on the channel statistics and on system parameters is embedded in the definition of $\mathbf{P}(\lambda)$ and $\mathbf{r}(\lambda)$.
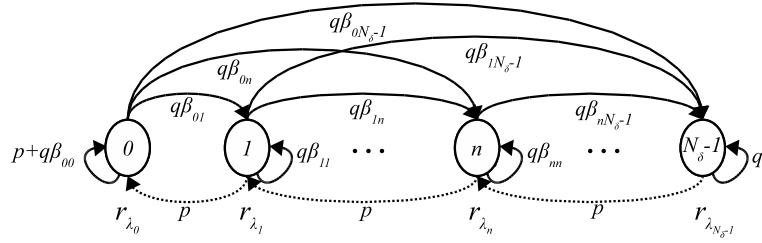
Figure 2.3 Markov chain describing the ABEH tag battery state. Dashed lines indicate policy-dependent transitions.

2.5.1 Howard Policy Improvement Algorithm

The complexity of a brute force approach algorithm, which exhaustively evaluates all the possible $N!$ policies to find the optimal stationary policy λ^*, becomes prohibitive for large N. As alternative approach, it is possible to resort to the Howard Policy

Improvement Algorithm (HPI-A) [47], which allows the optimal policy to be found in a finite number of steps (generally much lower than the exhaustive search). The starting point for HPI-A is the equation relating the long-term average read probability $g(\lambda)$ (2.10) and the *relative gain vector* $w(\lambda) = [0, w_1(\lambda), ..., w_{N-1}(\lambda)]$, whose nth entry is defined as the gain of having the Markov chain starting in state n rather than in state 0. This vector equation is given by: $\mathbf{w}(\lambda) + g(\lambda)\mathbf{1} = \mathbf{r}(\lambda) + \mathbf{P}(\lambda)\mathbf{w}(\lambda)$, where $\mathbf{1} = [1, ..., 1]^T$ and $\mathbf{P}(\lambda)$ is the transition probability matrix defined in Section 2.4.1. Notice that the vector equation above defines a linear system with N equations and N unknowns that are given by $w_1(\lambda), ..., w_{N-1}(\lambda)$, since it is possible to arbitrarily fix $w_0(\lambda) = 0$ as reference. The HPI-A is an iterative algorithm that converges to the exact solution in a finite number of steps. It works as follows: 1) Choose an arbitrary policy $\lambda = [\lambda_0, ..., \lambda_{N-1}]^T$; 2) calculate $\mathbf{w}(\lambda)$ from the linear system above; 3) if $\mathbf{r}(\lambda) + \mathbf{P}(\lambda)\mathbf{w}(\lambda) \geq \mathbf{r}(\theta) + \mathbf{P}(\theta)\mathbf{w}(\lambda)$ for all possible $\theta = [\theta_0, ..., \theta_{N-1}]^T$ (N entry-wise inequalities have to be satisfied), then λ is optimal; 4) otherwise, find θ such that at least one of the inequalities above is not satisfied; 5) update $\lambda = \theta$ and iterate with the new policy steps from 2 to 5 until the algorithm converges (that is, all the N inequalities at step 3 are satisfied). Further details on the HPI-A can be found in [21] and [47].

2.6   Numerical Results

This section provides some numerical results to show the read range improvement of ABEH tag with respect to conventional passive tags. It is assumed that the DL and UL channels are statistically independent Rayleigh channels. The long-term average read probability (2.10) of ABEH tags is compared to the one attainable with passive tags, which using the notation above is $g^{std} = \Pr\left[\gamma(0; k) \geq \gamma_{th}\right]$ (see [46] for a closed-form expression). The value of δ_E is chosen by imposing that $\Pr[\mathcal{E}(k) < \delta_E] \simeq 5\%$ for the maximum distance tag-reader ($d = 16m$). This value provides a reasonable

trade-off between approximation and complexity of the model. Figure 2.4 shows the long-term average read probability versus the tag-reader distance d. Energy-unit is $\delta_E = 0.22\mu J$ while the size of the battery is varied $E_{\max} \in \{14, 56, 224\}\mu J$ by changing the size of the discrete model $N \in \{64, 256, 1024\}$ (notice that, keeping δ_E fixed implies more accuracy of the discrete model for distances smaller than $d = 16m$, see Section 2.4). The duration of a slot is $T = 10ms$, the transmitted power is $E_0/T = 36dBm$, while the product between SNR threshold and noise power is $\gamma_{th}\sigma_r^2 = -67dBm$ (this is equivalent to define the power sensitivity of the reader, see (2.4) for details). The interrogation probability is $p = 0.1$, the CW frequency is $f_0 = 915MHz$ and the efficiencies are $\eta_{mod} = \eta_{amp} = 0.2$ and $\eta_{DC} = 0.4$. ABEH tags provide considerable gains in terms of read range (for the given requirements $\gamma_{th}\sigma_r^2$) with respect to passive tags, especially for sufficiently large batteries.

The effects of the interrogation probability p and the complexity of the energy scheduler on the system performance are now evaluated. Low complexity schedulers can discern only $N_L < N$ battery levels. The advantage is that they require less memory to store policies (λ has N_L elements compared to N) and simpler circuit to measure the battery state. Figure 2.5 shows the average read probability of ABEH tags for different values of $N_L \in \{2, 16, 1024\}$ versus p, along with the performance of the draw-all policy for $d = 16m$ and $E_{\max} = 224\mu J$ (other parameters as above). It is seen that energy schedulers with only $N_L = 16$ states suffers negligible performance penalty with respect to more complex scheduler with $N_L = N$ states. Notice also that, even with $N_L = 2$ states (i.e., a threshold at the half size of the battery), ABEH tags still perform much better than passive tags. Clearly for $p \to 1$ there is no performance gain when using ABEH tags, as no energy can be harvested, while gains of orders of magnitude are possible for smaller p. Finally, the draw-all policy becomes highly suboptimal for moderate-to-high values of p since in this regime, the tag needs to manage accurately the stored energy. The shapes of the policies for a

moderate-high interrogation probability $p = 0.1$ versus battery state and for different distances d, are shown in Figure 2.6. Note that, for increasing distances d, the energy scheduling preserves energy until the battery has stored enough energy.
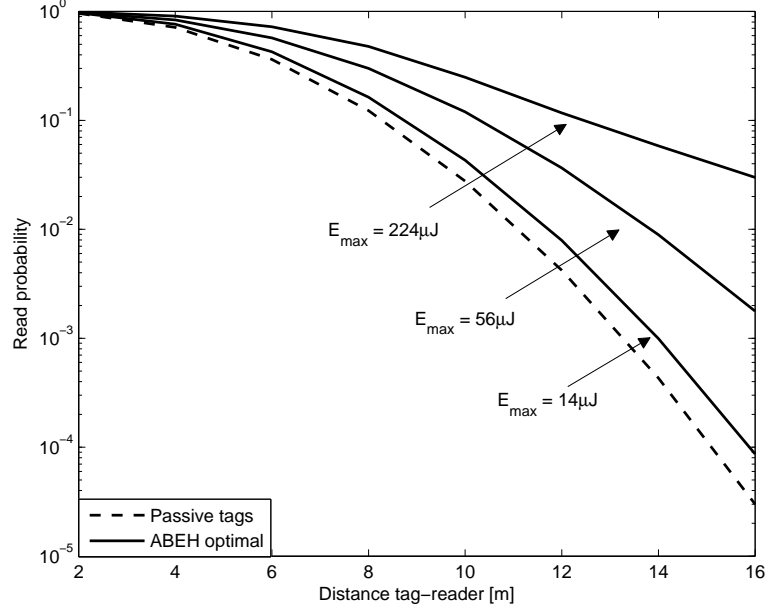
Figure 2.4 Long-term average read probability of ABEH and passive tags versus tag-reader distance for different battery sizes ($\gamma_{th}\sigma_r^2 = -67dBm$, $\delta_E = 0.22\mu J$, $E_0/T = 36dBm$, $T = 10ms$, $p = 0.1$, $\eta_{amp} = \eta_{mod} = 0.2$, $\eta_{DC} = 0.4$).

2.7   Concluding Remarks

The problem of increasing the tag read range for passive RFID-based sensor networks limited by the reader sensitivity has been addressed. An approach that leverages an onboard battery at the tag, recharged exclusively through RF-energy harvesting of the reader signal during tag inactivity period, to opportunistically amplify the backscatter signal (Amplified Backscatter through Energy Harvesting, ABEH) has been proposed. The analysis presented in this chapter shows remarkable performance gains in terms of read range achievable with ABEH tags, even in the presence of moderate-to-large interrogation probabilities, i.e., for a small number of tags and/or high rate of information collection from the tags to reader. Moreover, it points to the
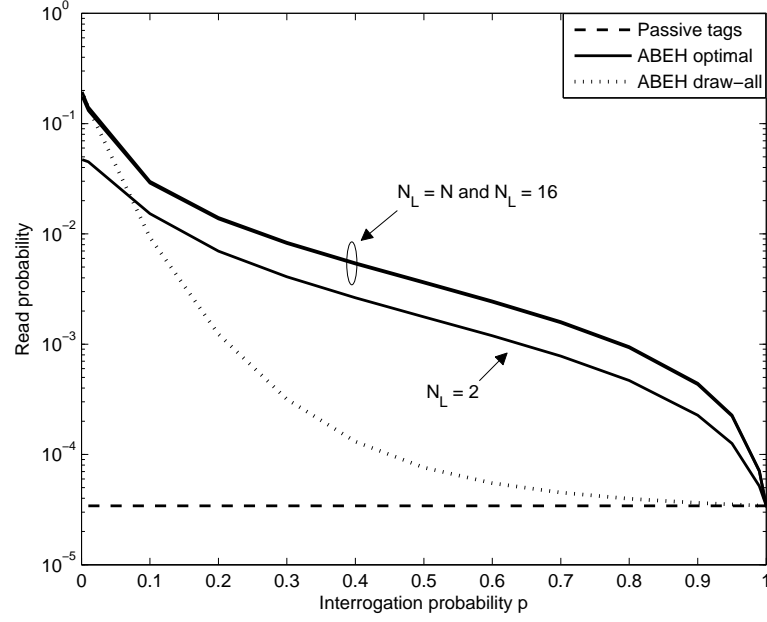
Figure 2.5 Long-term average read probability of ABEH and passive tags versus interrogation probability p for different policy complexities N_L ($E_{max} = 224\mu J$, $d = 16m$, other parameters as in Figure 2.4).
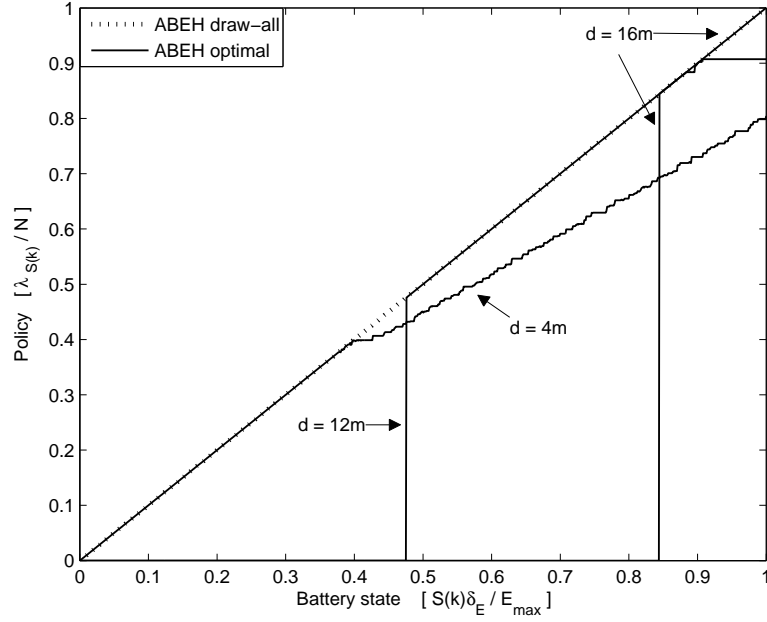
Figure 2.6 Normalized policies λ/N versus normalized battery state $S(k)\delta_E/E_{max}$ for different distances tag-reader d ($E_{max} = 224\mu J$, $p = 0.1$, other parameters as in Figure 2.4).

importance of a well-designed energy scheduling techniques at the tag, especially in the regime of moderate-to-high interrogation probabilities. Low-complexity policies have also been designed and shown to provide small performance loss over optimal strategies. As a final remark, it is noted that the proposed solution and analysis framework can be extended to the case of tag sensitivity-limited systems, by allowing a trade-off between the use of the on-board battery to amplify the backscatter signal and to reduce the tag sensitivity requirement. Finally, more complex propagation scenario can also be analyzed based on this framework with minor modifications.

Part II

Medium Access Control Protocols for Energy Harvesting Wireless

Networks

This part of the dissertation considers the issue of medium access control (MAC) for single-hop wireless networks operated by nodes equipped with energy harvesting (EH) capabilities. In particular, Chapter 3 considers the design and investigates the performance of random access protocols such as ALOHA, while Chapter 4, based on the insights obtained in Chapter 3, proposes a new ALOHA-based protocol that is specifically tailored to EH networks.

Centralized scheduling MAC schemes are instead addressed in Chapter 5 and Chapter 6. Specifically, Chapter 5 considers the design of scheduling policies in a single-hop network where the nodes are powered via a hybrid energy storage system (HESS), while Chapter 6 considers the same setting but with nodes exclusively powered via EH.

CHAPTER 3

RANDOM ACCESS PROTOCOLS FOR ENERGY HARVESTING WIRELESS SENSOR NETWORKS

The design of medium access control (MAC) protocols for wireless sensor networks (WSNs) has been conventionally tackled by assuming battery-powered devices and by adopting the battery lifetime as the main performance criterion [5]. While WSNs operated by energy-harvesting (EH) devices are not limited by network lifetime, they pose new design challenges due to the uncertain amount of energy that can be harvested from the environment. Novel design criteria are thus required to capture the trade-offs between the potentially infinite network lifetime and the uncertain energy availability.

This chapter considers a single-hop WSN in which a fusion center (FC) collects data packets from M wireless nodes distributed in its surrounding (see Figure 3.1). The considered application is a batch resolution problem, in which each node in the network potentially generates a new packet periodically and simultaneously to other nodes. The packets generated at any given time compose the batch of packets that need to be collected by a central station (or *fusion center*, FC). In particular, this chapter investigates the novel performance trade-offs arising due to EH when designing conventional MAC protocols, namely TDMA, framed-ALOHA (FA) and dynamic-FA (DFA). Furthermore, based on the insights obtained through the analysis of ALOHA-based schemes, a novel random access MAC protocol, tailored to EH networks and referred to as Energy Group Dynamic Framed-ALOHA (EG-DFA), will be proposed in Chapter 4.

3.1 Related Work and Systems

While performance analysis of MAC protocols for battery powered wireless networks have been investigated in depth (see e.g., [8, 13, 14]), analyses of MAC protocols with EH devices are hardly available. A notable exceptions are [27], where data queue stability has been studied for TDMA and carrier sense multiple access (CSMA) protocols in EH networks, and [28] where a scheduling problem for EH mesh networks has been considered. Whereas, to the best of this dissertation author's knowledge the first work on random access MACs for EH wireless networks has been addressed in [53] and then in extended in [54]. It is finally remarked that routing for EH networks has instead received more attention, see e.g., [20, 55].

In recent years, wireless networks with EH-capable nodes have attracted a lot of attention also at commercial level. To provide some examples, the Enocean Alliance proposes to use a MAC protocol for EH devices based on pure ALOHA strategies [56], while an enhanced self-powered RFID tag created by Intel, referred to as WISP [4], has been conceived to work with the EPC Gen 2 standard [48] that adopts a FA-like MAC protocol.

3.1.1 Contributions

As introduced above, this chapter considers the design and analysis of TDMA, FA and DFA MAC protocols in the light of the novel challenges introduced by EH. Section 3.3 proposes to measure the system performance in terms of the trade-off between the *delivery probability*, which accounts for the number of sensors' measurements successfully reported to the FC, and the *time efficiency*, which measures the rate of data collection at the FC (formal definitions are in Section 3.3). An analytical framework is then introduced in Section 3.5, while Section 3.6 assesses the performance of the considered MAC protocols in terms of the mentioned trade-off for TDMA, FA and DFA protocols. Section 3.7 tackles the critical issue in ALOHA-based

protocols of estimating the number of EH nodes involved in transmission, referred to as *backlog*, by proposing a practical reduced-complexity algorithm. Extensive numerical simulations are then presented in Section 3.8 to get insights into the MAC protocol design trade-offs, and to validate the analytical derivations.

3.2 System Model

This chapter considers a single-hop WSN with a fusion center (FC) surrounded by M wireless nodes labeled as $U_1, U_2, ..., U_M$ (see Figure 3.1). Each node (or sensor) is equipped with an EH unit (EHU) and an energy storage device (ESD), where the latter is used to store the energy harvested by the EHU. The FC retrieves measurements from nodes via periodic *inventory rounds* (IRs), once every T_{int} seconds [s]. Each IR is started by the FC by transmitting an initial *query command* (Q), which provides both synchronization and instructions to nodes on how to access the channel. Time is slotted, with each *slot* lasting T_s [s]. The effective duration of the nth IR, during which the communication between the FC and the nodes takes place, is denoted by $T_{IR}(n)$. It is assumed that $T_{IR}(n) \ll T_{int}$ for all IR n, and also that the query duration is negligible, so that the ratio $T_{IR}(n)/T_s$ indicates the total number of slots allocated by the FC during the nth IR.

In every IR, each node has a new measurement to transmit with probability (w.p.) α, independently of other nodes and previous IRs. If a new measurement is available, the node will mandatory attempt to report it successfully to the FC as long as enough energy is stored in its ESD (see Section 3.2.2 for details). Each measurement is the payload of a packet, whose transmission fits within the slot duration T_s. nodes' transmissions within each IR are organized into *frames*, each of which is composed of a number of slots that is selected by the FC. Depending on the adopted MAC protocol, any node that needs to (and can) transmit in a frame either chooses or is assigned a single slot within the frame for transmission as it will be detailed below. Moreover,
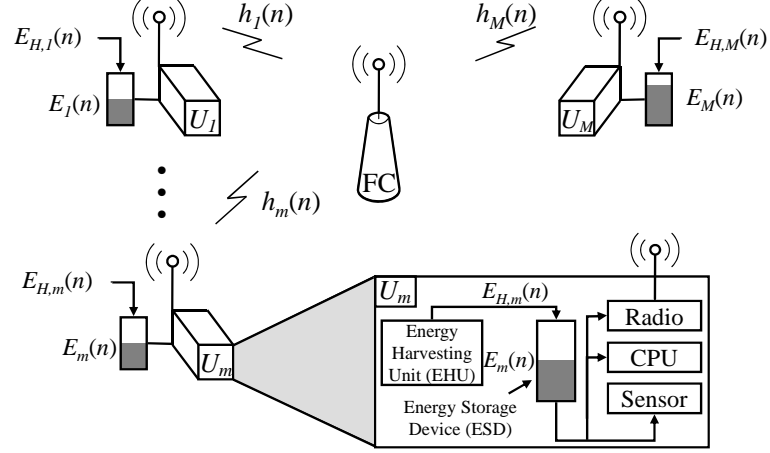
Figure 3.1 A WSN where a fusion center (FC) collects data from M nodes. Each node is equipped with an energy storage device (ESD) and an energy-harvesting unit (EHU).

after a node has successfully transmitted its packet to the FC, it first receives an acknowledge (ACK) of negligible duration by the FC and then it becomes inactive for the remaining of the IR. It is emphasized that the FC knows neither the number of nodes with a new measurement to transmit, nor the state of nodes' ESDs.

3.2.1 Interference Model

This chapter considers *interference-limited* communication systems in which the *downlink* packets transmitted by the FC are always correctly received (error-free) by the nodes, while *uplink* packets transmitted by the nodes to the FC are subject to communication errors due to possible interference arising from collisions with other transmitting nodes. The uplink channel power gain for the mth node during the nth IR is $h_m(n)$. Channel gain $h_m(n)$ is assumed to be constant over the entire IR but subject to random independent and identically distributed (i.i.d.) fading across IRs and nodes, with probability density function (pdf) $f_h(\cdot)$ and normalized such that $E[h_m(n)] = 1$, for all n, m. In the presence of simultaneous transmissions in the same slot during the kth frame of the nth IR, a node U_m is correctly received by the FC if and only if its instantaneous signal-to-interference ratio (SIR) $\gamma_{m,k}(n)$ is larger than

a given threshold γ_{th}, i.e., if

$$\gamma_{m,k}(n) = \frac{h_m(n)}{\sum_{l \in \mathcal{I}_{m,k}(n)} h_l(n)} \geq \gamma_{th}, \tag{3.1}$$

where $\mathcal{I}_{m,k}(n)$ denotes the set of nodes that transmit in the same slot selected by U_m in frame k and IR n. It is assumed $\gamma_{th} > 0dB$ so that, in case a slot is selected by more than one node, at most one of the colliding node can be successfully decoded in the slot.

According to the interference model (3.1), any slot can be: *empty* when it is not selected by any node; *collided* when it is chosen by more than one nodes but none of them transmits successfully; *successful* when a node transmits successfully possibly in the presence of other (interfering) nodes. Successful transmission in the presence of interfering nodes within the same slot is often referred to as *capture effect* [14].

Remark 1: Errors in the decoding of downlink query packets can be accounted for through the parameter α as well. In fact, let α_Q be the probability that a node correctly decodes the downlink packet sent by the FC at the beginning of an IR. Moreover, assume that downlink decoding errors are i.i.d. across nodes and IRs, and let α_N be the probability that a node has a new measurement to transmit in any IR. Then, the probability that any node U_m has a new packet and correctly decodes the FC's query is given by the product $\alpha = \alpha_Q \alpha_N$.

3.2.2 ESD and Energy Consumption Models

This chapter considers a discrete ESD with $N+1$ energy levels in the set $\mathcal{E} = \{0, \delta, 2\delta, ..., N\delta\}$, where δ is referred to as *energy unit*. Let $E_m(n) \in \mathcal{E}$ be the energy stored in the ESD of the mth node at the beginning of the nth IR. Energy $E_m(n)$ is a random variable that is the result of the EH process and the energy consumption of the node across IRs; its probability mass function (pmf) is $p_{E(n)}(\cdot)$ and the corresponding complementary cumulative distribution function (ccdf) is $G_{E(n)}(x) = \Pr[E_m(n) \geq x]$.

Note that, the initial energy distribution $p_{E(1)}(\cdot)$ is given, while the evolution of the pmf $p_{E(n)}(\cdot)$ for $n > 1$ depends on both the MAC protocol and the EH process.

It is assumed that each time a node transmits a packet it consumes an energy ε, which accounts for the energy consumed in the: *a)* reception of the FC's query that starts the frame (see Figure 3.2); *b)* transmission; *c)* reception of FC's ACK or not ACK (NACK) packet, if any. At the beginning of each IR, a node with a new measurement to transmit can participate to the current IR only if the energy stored in its ESD is at least ε. Let $\varepsilon_\delta = \varepsilon/\delta$ be the number of energy units δ required for transmission, where ε_δ is assumed to be an integer value without loss of generality. Let $F_\varepsilon = N\delta/\varepsilon = N/\varepsilon_\delta$ be the (normalized) *capacity* of the ESD, which is assumed to be an integer indicating the maximum number of (re)transmissions allowed by a fully charged ESD.
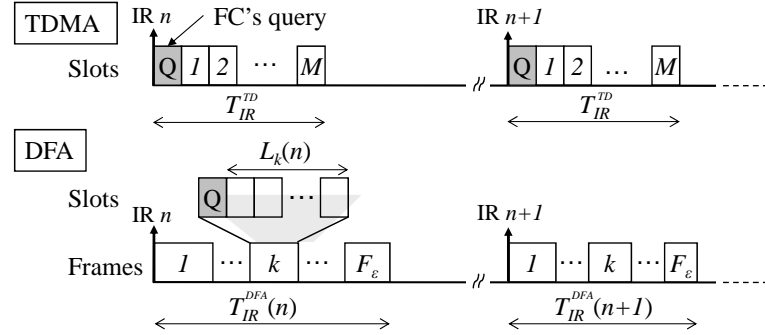
Figure 3.2 Organization of time in slots and frames for TDMA and DFA protocols (FA is a special case of DFA with only one frame).

3.2.3 Energy Harvesting Model

During the time T_{int} between the nth and $(n+1)$th IRs the mth node U_m harvests an energy $E_{H,m}(n)$, which is modeled as a discrete random variable, i.i.d. over IRs and nodes, with pmf $q_i = \Pr[E_{H,m}(n) = i\delta]$, with $i \in \{0, 1, 2, ...\}$. For technical reasons discussed in Section 3.6.2, it is assumed that the probabilities q_0 and q_1 of harvesting

zero and one energy unit, respectively, are both strictly positive, i.e., $q_0 > 0$ and $q_1 > 0$.

It is assumed that the EH dynamics is much slower than the IR duration $T_{IR}(n)$, so that the amount of energy harvested within $T_{IR}(n)$ can be considered as negligible with respect to ε (recall also that $T_{IR}(n) \ll T_{int}$). Hence, the only energy that a node can actually use throughout an IR is the energy initially available at the beginning of the IR itself (i.e., $E_m(n)$).

3.3   MAC Performance Metrics

The next sections introduce the MAC performance metrics that are considered throughout this chapter.

3.3.1   Delivery Probability

The delivery probability $p_d(n)$ measures the capability of the MAC protocol to successfully deliver the packet of any node, say U_m, to the FC in IR n

$$p_d(n) = \Pr\left[U_m \text{ TX successfully in IR } n | U_m \text{ has a new measurement in IR } n\right].$$

$$(3.2)$$

The statistical equivalence of all nodes makes the probability (3.2) independent of the specific node. Notice that a node fails to report its measurement during an IR if either it has an energy shortage before (re)transmitting the packet correctly, or the MAC protocol does not provide the node with sufficient retransmission opportunities. Given the potentially perpetual operation enabled by EH, it is relevant to evaluate the delivery probability when the system is in *steady-state*. The *asymptotic delivery probability* is thus obtained by taking the limit of $p_d(n)$ for large IR index n, provided that it exists, as

$$p_d^{AS} = \lim_{n \to \infty} p_d(n). \qquad (3.3)$$

3.3.2 Time Efficiency

The time efficiency $p_t(n)$ measures the probability that any slot allocated by the MAC within the nth IR is successfully used (see Section 3.2.1)

$$p_t(n) = \Pr\left[\text{The FC correctly retrieves a packet in any slot of the } n\text{th IR}\right]. \quad (3.4)$$

By taking the limit of (3.4) for $n \to \infty$, one can obtain the *asymptotic time efficiency*

$$p_t^{AS} = \lim_{n\to\infty} p_t(n). \quad (3.5)$$

Remark 2: Informally speaking, the time efficiency $p_t(n)$ measures the ratio in IR n between the total number of packets successfully received by the FC and the total number of slots allocated by the MAC protocol (i.e., $T_{IR}(n)/T_s$, see Section 3.2). As it will be shown in Section 3.4, the IR duration $T_{IR}(n)$ is in general a random variable, and consequently, time efficiency $p_t(n)$ differs from more conventional definitions of throughput that measure the number of packets delivered over the interval T_{int} between two successive IRs (see [13]), instead of $T_{IR}(n)$. The rationale for this definition of time efficiency is that it actually captures more effectively the rate of data collection at the FC. Whereas, the delivery probability accounts for the fraction of nodes, with a new measurement to transmit at the beginning of the current IR, which are able to successfully report their payload to the FC within the IR, where delivery failures are due to collisions and energy shortages.

In contention based MACs (e.g., ALOHA), there is a trade-off between delivery probability and time efficiency. In fact, increasing the former generally requires the FC to allocate a larger number of slots in an IR to reduce packet collisions, which in turn decreases the time efficiency.

3.4   MAC Protocols

The next sections review the MAC protocols that are considered in this chapter.

3.4.1 TDMA

With the TDMA protocol, each node is pre-assigned an exclusive slot that it can use in every IR, irrespective of whether it has a packet to deliver or enough energy to transmit. Recall that such information is not available at the FC. Any IR is thus composed by one frame with M slots and has fixed duration $T_{IR}^{TD} = MT_s$ (see Figure 3.2). Since TDMA is free of communication errors in the considered interference-limited scenario, its delivery probability $p_d(n)$ is only limited by energy availability and it is thus an upper bound for ALOHA-based MACs. However, TDMA might not be time efficient due to the many empty slots when the probability of having a new measurement α and/or the EH rate are small.

3.4.2 Framed-ALOHA (FA) and Dynamic-FA (DFA)

Hereafter only the DFA protocol is described, since FA follows as a special case of DFA with no retransmissions capabilities as discussed below. The nth IR, of duration $T_{IR}^{DFA}(n)$, is organized into a set of frames as shown in Figure 3.2. The *backlog* $\mathcal{B}_k(n)$ for the kth frame is the set composed of all nodes that simultaneously satisfy the following three conditions: i) have a new measurement to transmit in the nth IR; ii) have transmitted unsuccessfully (because of collisions) in the previous $k-1$ frames (this condition does not apply for frame $k=1$); iii) have enough energy left in the ESD to transmit in the kth frame. All the nodes in the set $\mathcal{B}_k(n)$, whose cardinality $|\mathcal{B}_k(n)| = B_k(n)$ is referred to as *backlog size*, thus attempt transmission during frame k. To make this possible, the FC allocates a frame of $L_k(n)$ slots, where $L_k(n)$ is selected based on the estimate $\hat{B}_k(n)$ of the backlog size $B_k(n)$ (estimation of $B_k(n)$ is discussed in Section 3.7) as

$$L_k(n) = \left\lceil \rho \hat{B}_k(n) \right\rceil, \tag{3.6}$$

where $\lceil \cdot \rceil$ is the upper nearest integer operator, and ρ is a design parameter. Note that, if the backlog size is B, the probability $\beta(j, B, L)$ that $j \leq B$ nodes transmit in the same slot in a frame of length L is binomial [57]

$$\beta(j, B, L) = \binom{B}{j} \left(\frac{1}{L}\right)^{j} \left(1 - \frac{1}{L}\right)^{B-j}. \tag{3.7}$$

Finally, FA is a special case of DFA where only one single frame of size $L_1(n)$ is announced as retransmission of collided packets is not allowed.

3.5 Analysis of the MAC Performance Metrics

This section derives the performance metrics defined in Section 3.3 for TDMA, FA and DFA. The analysis is based on two simplifying assumptions:

- $\mathcal{A}.1$ *Known backlog*: the FC knows the backlog size $B_k(n) = |\mathcal{B}_k(n)|$ before each kth frame;

- $\mathcal{A}.2$ *Large backlog*: the backlog size $B_k(n)$, in any IR n and any frame k of size $L_k(n) = \lceil \rho B_k(n) \rceil$, is large enough to let the probability (3.7) be approximated by the Poisson distribution [57]:

$$\beta(j, B_k(n), L_k(n)) \simeq \frac{e^{-\frac{1}{\rho}}}{\rho^j j!}. \tag{3.8}$$

Assumption $\mathcal{A}.1$ simplifies the analysis as in reality the backlog can only be estimated by the FC (see Section 3.7 and Section 3.8 for the impact of backlog estimation). Assumption $\mathcal{A}.2$ is standard and analytically convenient, as it makes the probability $\beta(j, B_k(n), L_k(n))$ dependent only on the ratio ρ between the frame length $L_k(n)$ and the backlog size $B_k(n)$. The assumptions above are validated numerically in Section 3.8.

The next sections derive the delivery probability (3.2) and the time efficiency (3.4) for the MAC protocols considered in this chapter, under the assumptions $\mathcal{A}.1$ and $\mathcal{A}.2$ introduced above. The IR index n is dropped to simplify the notation.

3.5.1   Delivery Probability for TDMA

Since the TDMA protocol is free of collisions, each node U_m that has a new measurement to report in the current IR cannot deliver its payload to the FC only when it is in energy shortage, namely if $E_m < \varepsilon$. Provided that node U_m has a new packet to transmit, the delivery probability (3.2) reduces to

$$p_d^{TD} = \Pr\left[E_m \geq \varepsilon\right] = G_E^{TD}\left(\varepsilon\right), \tag{3.9}$$

which is independent of the node index m and dependent only on the ccdf $G_E^{TD}(\cdot)$ of the energy stored in node ESD at the beginning of the considered IR. The ESD energy distribution for any arbitrary nth IR is derived in Section 3.6.

3.5.2   Delivery Probability for FA

In the FA protocol, each node U_m that has a new measurement to report in the current IR is able to correctly deliver its payload to the FC only if: a) it transmits successfully in the selected slot, possibly in the presence of interfering nodes provided that its SIR is $\gamma_{m,1} \geq \gamma_{th}$; and b) it has enough energy to transmit. From (3.1), the probability that node U_m, with $U_m \in \mathcal{B}_1$, transmits successfully in the selected slot, given that $|\mathcal{I}_{m,1}| = j$ nodes select the same slot of U_m (thus colliding), is given by

$$p_c(j) = \Pr\left[h_m \geq \gamma_{th} \sum_{l=1}^{j} h_l\right], \tag{3.10}$$

where, without loss of generality, it is assumed that $\mathcal{I}_{m,1} = \{U_1, ..., U_j\}$, and $U_m \notin \mathcal{I}_{m,1}$, as nodes are stochastically equivalent. Under the large backlog assumption $\mathcal{A}.2$, the probability that there are j interfering nodes is Poisson-distributed (see (3.8)),

and thus the unconditional probability p_c that U_m captures the selected slot can be approximated as

$$p_c \simeq e^{-\frac{1}{\rho}} \sum_{j=0}^{\infty} \frac{1}{\rho^j j!} p_c(j).$$

(3.11)

Note that, in (3.11) the number of possible interfering nodes have been extended up to infinity as $p_c(j)$ rapidly vanishes for increasing j. Moreover, depending on the channel gain pdf $f_h(\cdot)$, probabilities (3.10) can be calculated either analytically (e.g., when $f_h(\cdot)$ is exponential, see [58]) or numerically.

Finally, under assumption $\mathcal{A}.2$, the successful transmission event is independent of the ESD energy levels (which in principle determine the actual backlog size in (3.7)), and thus the delivery probability (3.2) for the FA protocol can be calculated as the product between the probability $G_E^{FA}(\varepsilon) = \Pr[E_m \geq \varepsilon]$ that node U_m has enough energy to transmit and the (approximated) capture probability (3.11) as

$$p_d^{FA} \simeq G_E^{FA}(\varepsilon)\, e^{-\frac{1}{\rho}} \sum_{j=0}^{\infty} \frac{1}{\rho^j j!} p_c(j),$$

(3.12)

where the ESD energy ccdf $G_E^{FA}(\varepsilon)$ for any arbitrary nth IR is derived in Section 3.6.

3.5.3 Delivery Probability for DFA

The DFA protocol is composed of several instances of FA, one for each kth frame of the current IR. As DFA allows retransmissions, one needs to calculate the probability $p_{c,k}(j)$ that any node active during frame k, say $U_m \in \mathcal{B}_k$, transmits successfully in the selected slot given that there are $|\mathcal{I}_{m,k}| = j$ nodes that transmit in the same slot, with $\mathcal{I}_{m,k} \subseteq \mathcal{B}_k$. The computation of $p_{c,k}(j)$, for $k > 1$, is more involved than (3.10). In fact, packets collisions introduce correlation among the channel gains of collided nodes, as any node in the backlog $\mathcal{B}_k$, for $k > 1$, might have collided with some other

nodes in the set $\mathcal{B}_k$. It is recalled that, even though the channel gains are i.i.d. at the beginning of the IR, they remain fixed for the entire IR.

While the exact computation of probabilities $p_{c,k}(j)$ is generally cumbersome, the large backlog assumption $\mathcal{A}.2$ enables some simplifications. Specifically, correlation among channel gains can be neglected, since for large backlogs it is unlikely that two nodes collide more than once within the same IR. By assuming independence among the channel gains at any frame, calculation of $p_{c,k}(j)$ requires only to evaluate the channel gain pdf $f_h^{(k)}(\cdot)$ at the kth frame for any node within $\mathcal{B}_k$, which is the same for all nodes by symmetry. The computation of pdf $f_h^{(k)}(\cdot)$ can be done recursively, starting from frame $k = 1$, so that at frame k one can condition on the event that the SIR (3.1) was $\gamma_{m,k-1} < \gamma_{th}$. Under assumption $\mathcal{A}.2$, this can be done numerically (see Appendix A and [59] for more details).

Now, let $\tilde{h}_m^{(k)}$, for $m \in \{1, ..., M\}$ and $k \in \{1, ..., F_\varepsilon\}$, be random variables with pdf $f_h^{(k)}(\cdot)$ independent over m, where $\tilde{h}_m^{(1)} = h_m$. The conditional capture probabilities $p_{c,k}(j)$ can then be approximated as (compare to (3.10))

$$p_{c,k}(j) \simeq \Pr\left[\tilde{h}_m^{(k)} \geq \gamma_{th} \sum_{l=1}^{j} \tilde{h}_l^{(k)}\right], \tag{3.13}$$

for any $m \notin \{1, ..., j\}$ as nodes are stochastically equivalent. By exploiting the Poisson approximation similarly to (3.11), the unconditional probability that any node within the backlog successfully transmits in the selected slot in the kth frame is

$$p_{c,k} \simeq e^{-\frac{1}{\rho}} \sum_{j=0}^{\infty} \frac{1}{\rho^j j!} p_{c,k}(j). \tag{3.14}$$

Recalling that a node keeps retransmitting its packet until it is successfully delivered to the FC, then the successful delivery of a packet in a frame is a mutually exclusive event with respect to the delivery in previous frames. Therefore, the probability of transmitting successfully in the kth frame, given that enough energy is available, is $p_{c,k} \prod_{i=1}^{k-1} (1 - p_{c,i})$. Finally, by accounting for the probability

$G_E^{DFA}(k\varepsilon) = \Pr[E_m \geq k\varepsilon]$ of having enough energy in each kth frame, the DFA delivery probability can be obtained, under assumption $\mathcal{A}.2$, as[1]

$$p_d^{DFA} \simeq \sum_{k=1}^{F_\varepsilon} G_E^{DFA}(k\varepsilon) \, p_{c,k} \prod_{i=1}^{k-1} (1 - p_{c,i}), \qquad (3.15)$$

where the ESD energy ccdf $G_E^{DFA}(k\varepsilon)$ for any arbitrary nth IR is derived in Section 3.6.

3.5.4 Time Efficiency for TDMA

Let $\mathcal{M}_m$ be the event indicating that node U_m has a new measurement to report in the current IR, with $\Pr[\mathcal{M}_m] = \alpha$, then the TDMA time efficiency (3.4) is given by the probability that the mth node has enough energy to transmit and a packet to report as

$$
\begin{aligned}
p_t^{TD} &= \Pr[E_m \geq \varepsilon, \mathcal{M}_m] = \Pr[E_m \geq \varepsilon]\Pr[\mathcal{M}_m] \\
&= \alpha G_E^{TD}(\varepsilon), \qquad (3.16)
\end{aligned}
$$

where the independence between the energy availability E_m and the event $\mathcal{M}_m$ has been exploited.

3.5.5 Time Efficiency for FA

Since it has been assumed $\gamma_{th} > 0dB$, then when more than one node transmits within the same slot, only one of them can be decoded successfully (i.e., successful transmissions of different nodes within the same slot are disjoint events). Hence, the probability that a slot, simultaneously selected by j nodes, is successfully used by any of them is given by $jp_c(j-1)$, where $p_c(j-1)$ is (3.10) by recalling that any node have $(j-1)$ interfering nodes. Furthermore, under assumption $\mathcal{A}.2$, the probability

---

[1]Note that in principle the backlogs $\mathcal{B}_1, \mathcal{B}_2...$ are correlated, and therefore the exact p_d^{DFA} should be obtained by averaging over the joint distribution of the backlog sizes. However, the assumption $\mathcal{A}.2$ removes the dependence on the backlog size.

that exactly j nodes select the same slot is $e^{-\frac{1}{\rho}}/(\rho^j j!)$, and by summing up over the number of simultaneously transmitting nodes j one gets

$$p_t^{FA} \simeq e^{-\frac{1}{\rho}} \sum_{j=1}^{\infty} \frac{1}{\rho^j j!} j p_c(j-1) = e^{-\frac{1}{\rho}} \sum_{j=0}^{\infty} \frac{1}{\rho^{(j+1)} j!} p_c(j). \qquad (3.17)$$

Note that, a consequence of assumption $\mathcal{A}.2$ is to make the FA time efficiency (3.17) independent of the ESD energy pmf. Moreover it is remarked that, when $\rho = 1$, $p_c(j) = 1$ for $j = 0$ and $p_c(j) = 0$ for $j > 0$ (i.e., no capture), then one has $p_t^{FA} = e^{-1}$, which is the throughput of slotted ALOHA [13].

3.5.6 Time Efficiency for DFA

The DFA time efficiency p_t^{DFA} follows from the FA time efficiency by accounting for the presence of multiple frames within an IR similarly to Section 3.5.3. Since the time efficiency is defined over multiple frames, it is possible to first derive the time efficiency in the kth frame, similarly to (3.17) but considering (3.13) instead of (3.10), as

$$p_{t,k}^{DFA} \simeq e^{-\frac{1}{\rho}} \sum_{j=0}^{\infty} \frac{1}{\rho^{(j+1)} j!} p_{c,k}(j). \qquad (3.18)$$

It is then possible to calculate p_t^{DFA} by summing (3.18) up, for all $k \in \{1, ..., F_\varepsilon\}$, weighted by the (random) length of the corresponding frame L_k normalized to the total number of slots in the IR $\sum_{k=1}^{F_\varepsilon} L_k$. Note that, under assumption $\mathcal{A}.2$ the random frame length L_k is well-represented by its (deterministic) average value $L_k \simeq E[L_k] = \rho E[B_k]$ and thus the DFA time efficiency results

$$p_t^{DFA} \simeq \frac{\sum_{k=1}^{F_\varepsilon} p_{t,k}^{DFA} E[B_k]}{\sum_{k=1}^{F_\varepsilon} E[B_k]}, \qquad (3.19)$$

where the average backlog size $E[B_k]$ in frame k, can be computed, under assumption $\mathcal{A}.2$, as $E[B_k] = M\alpha G_E^{DFA}(k\varepsilon) \prod_{i=1}^{k-1} (1 - p_{c,i})$. In fact, $M\alpha$ is the average number of nodes with a new measure to report in the current IR, $G(k\varepsilon)$ is the probability that

$k\varepsilon$ energy units are stored in the ESD at the beginning of the IR, thus allowing k consecutive transmissions, and $\prod_{i=1}^{k-1}(1-p_{c,i})$ is the probability that a node collides in all of the first $(k-1)$ frames.

3.6 ESD Energy Evolution

Section 3.5 shown that the performance metrics for the nth IR depend on the ESD energy distribution at the beginning of the IR. The goal of this section is to derive the ccdf $G_{E(n)}(\cdot)$, in any IR n, to obtain the asymptotic performance metrics (3.3) and (3.5) described in Section 3.5.

In general, in DFA, the evolution of node ESDs across IRs are correlated with each other due to the retransmission opportunities after collisions. However, under the large backlog assumption $\mathcal{A}.2$, similarly to the discussion in Section 3.5.3, the evolution of node ESDs become decoupled and can thus be studied separately. Accordingly, this section develops a stochastic model, based on a discrete Markov chain (DMC) that focuses on a single node ESD as shown in Figure 3.3. In addition, the focus is on the DFA protocol since the ESD evolutions for TDMA and FA follow as special cases. Note that, in TDMA (or FA), the evolution of node ESDs are actually independent with each other as retransmissions are not present.

3.6.1 States of a Node

The state of a node is uniquely characterized by: i) node activity or idleness (see below); ii) the amount of energy in its ESD; iii) the current frame index if the node is active. A node is *active* if it has a new packet still to be delivered to the FC in the current IR and enough energy in its ESD, while it is *idle* otherwise. States in which a node is active (or *active states*), are denoted by A_j^k and they are characterized by: a) the current frame index $k \in \{1, ..., F_\varepsilon\}$; and b) the number $j \in \{0, ..., N\}$ of energy units δ stored in the node ESD.
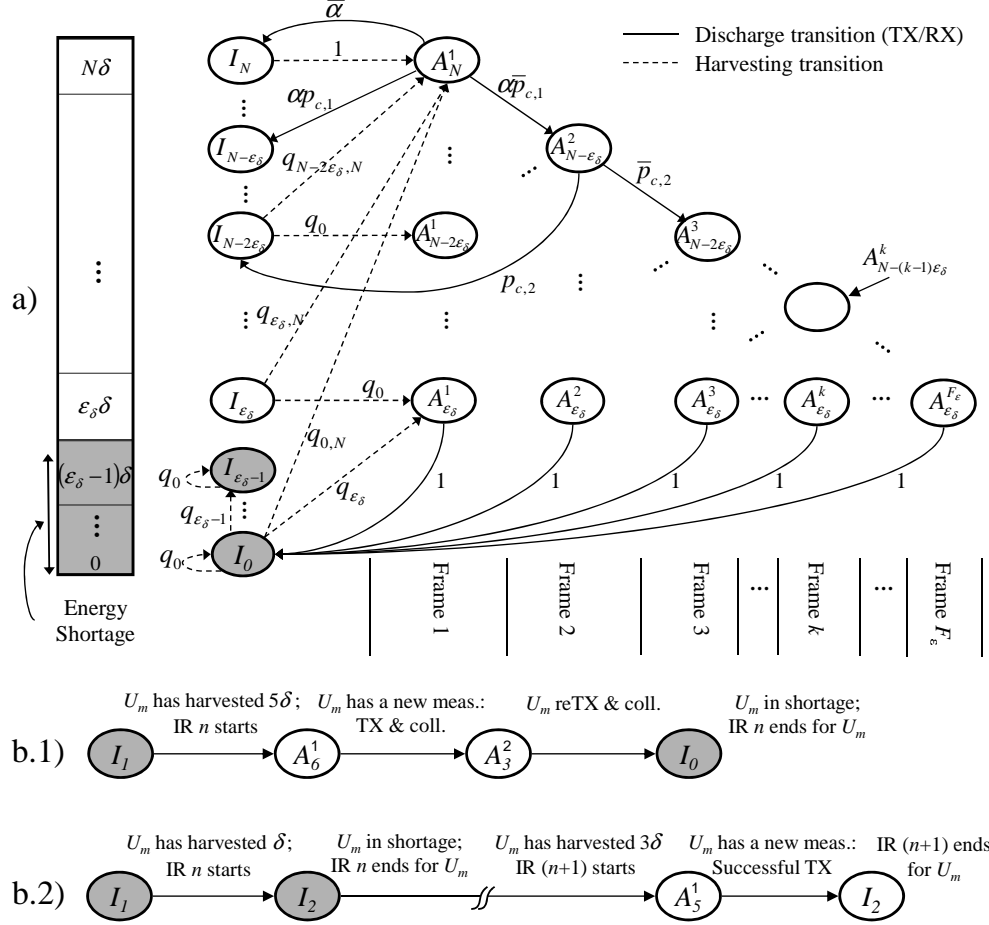
Figure 3.3 a) Discrete Markov chain used to model the evolution of the energy stored in the discrete ESD of a node in terms of the energy unit δ. In b.1) and b.2) there are two outcomes of possible state transition chains for $\varepsilon_\delta = 3$. Grey shaded states indicate energy shortage condition. Some transitions are not depicted to simplify representation. ($\bar{\alpha} = 1 - \alpha$ and $\bar{p}_{c,k} = 1 - p_{c,k}$).

States in which the node is idle (or *idle states*) are denoted by I_j and they are only characterized by the number $j \in \{0, ..., N\}$ of energy units stored in the ESD. EH is associated to idle states given the assumption that energy harvested in the current IR can only be used in the next IR (see Section 3.2.3).

3.6.2 Discrete Markov Chain (DMC) Model

Operations of a node across IRs are as follows. When node U_m is not involved in an IR, it is in an idle state, say I_j, waiting for the next IR. When a new IR begins, the energy harvested in the last interval T_{int} is added, so that, if the ESD is not in energy shortage, the state makes a transition $I_j \rightarrow A_l^1$ toward an active state, with $l \geq \varepsilon_\delta \geq j$. Otherwise, if it is in energy shortage, it makes a transition $I_j \rightarrow I_l$ toward another idle state, with $j \leq l < \varepsilon_\delta$. If node U_m is not in energy shortage, it remains in state A_j^1 at the beginning of the IR only if it has a new packet to transmit, which happens w.p. α. Instead, w.p. $\bar{\alpha} = 1 - \alpha$ the state makes a transition toward an idle state as $A_j^1 \rightarrow I_j$. If there is a new packet, the node keeps transmitting it in successive frames until either the packet is correctly delivered to the FC, or its ESD falls in energy shortage, or both. A collision in frame k happens w.p. $\bar{p}_{c,k} = 1 - p_{c,k}$ (see Section 3.5.3) and leads to a transition either $A_j^k \rightarrow A_{j-\varepsilon_\delta}^{k+1}$, for $j \geq 2\varepsilon_\delta$ (no shortage after collision) or $A_j^k \rightarrow I_{j-\varepsilon_\delta}$, for $j < 2\varepsilon_\delta$ (shortage after collision). Successful transmission in frame k, which happens w.p. $p_{c,k}$, instead leads to a transition $A_j^k \rightarrow I_{j-\varepsilon_\delta}$. Transition probabilities are summarized in Figure 3.4, where it has been defined $q_{j,N} = \Pr[E_{H,m} \geq (N-j)\delta] = 1 - \sum_{i=0}^{N-j-1} q_i$. Note that, the probability α of having a new measurement is only accounted for in active states in the first frame (i.e., in states A_j^1, for $j \in \{0, ..., N\}$, see Figure 3.4-b)). In fact, being in any state A_j^k for $k > 1$ already implies that a new measurement was available at the beginning of the IR. Note that, state transitions in the DMC at hand are event-driven and do not happen at fixed time intervals. A sketch of the proposed DMC is shown

in Figure 3.3-a), while two outcomes of possible state transition chains are shown in Figure 3.3-b.1) and 3.3-b.2).

From Figure 3.3-a), it can be seen that, when $q_0 > 0$, $q_1 > 0$ and $p_{c,k} > 0$, for $k \in \{1, ..., F_\varepsilon\}$, the DMC at hand is irreducible and aperiodic and thus, by definition, ergodic (see [47]). In fact, if $q_1 > 0$, any state of the DMC can be reached from any other state with non-zero probability, and therefore the DMC is irreducible. Moreover, the probability of having a self-transition from state I_0 to itself is $q_0 > 0$, and therefore state I_0 is aperiodic. The presence of an aperiodic state in a finite state irreducible DMC is enough to conclude that the chain is aperiodic [47, Ch. 4, Th. 1]. Since the DMC is ergodic it admits a unique steady-state probability distribution $\phi = [\phi_{I_0}, ..., \phi_{I_N}, \phi_{A^1_{\varepsilon_\delta}}, ..., \phi_{A^{F_\varepsilon}_N}]$, regardless of the initial distribution, which can be calculated by resorting to conventional techniques [47]. This also guarantees the existence of limits (3.3) and (3.5). Vector ϕ represents the steady-state distribution in any discrete time instant of the interrogation period (i.e., during either a frame of an IR or an idle period). However, to calculate (3.3) and (3.5) one needs the DMC steady-state distribution ϕ^+ conditioned on being at the beginning of the IR. This can be calculated by recalling that between the end of the last issued IR and the beginning of a new one, node U_m can only be in an idle state I_j, with $j \in \{0, ..., N\}$, and thus its state conditional distribution $\phi^- = [\phi^-_{I_0}, ..., \phi^-_{I_N}, \phi^-_{A^1_{\varepsilon_\delta}}, ..., \phi^-_{A^{F_\varepsilon}_N}]$, is given by $\phi^-_{I_j} = \phi_{I_j} / \sum_{i=0}^{N} \phi_{I_i}$, $\forall j \in \{0, ..., N\}$ and $\phi^-_{A^k_j} = 0$, for all j, k. The desired distribution ϕ^+ of the state at the beginning of the next IR can be obtained as $\phi^+ = \phi^- \mathbf{P}$, where $\mathbf{P}$ is the DMC probability transition matrix of the DMC in Figure 3.3-a) that can be obtained through Figure 3.4. Note that, according to the transition probabilities in Figure 3.4, starting from any state I_j, with $j \in \{0, ..., N\}$, only states I_j, with $j \in \{0, ..., \varepsilon_\delta - 1\}$ and states A^1_j, with $j \in \{\varepsilon_\delta, ..., N\}$ can be reached. Therefore, the only possible non-zero entries of distribution ϕ^+ are $\phi^+_{I_j}$ for $j \in \{0, ..., \varepsilon_\delta - 1\}$ and $\phi^+_{A^1_j}$ for $j \in \{\varepsilon_\delta, ..., N\}$.

Once the DMC steady-state distribution ϕ^+ at the beginning of any (steady-state) IR is obtained, it is possible to calculate the corresponding ESD steady-state distribution $p_{E(n\to\infty)}(\cdot)$ by mapping the DMC states into the energy level set $\mathcal{E}$ as follows

$$p_{E(n\to\infty)}(j) = \begin{cases} \phi^+_{I_j} & \text{for } j \in \{0, ..., \varepsilon_\delta - 1\} \\ \phi^+_{A^1_j} & \text{for } j \in \{\varepsilon_\delta, ..., N\} \end{cases}. \qquad (3.20)$$

The ccdf $G_{E(n\to\infty)}(\cdot)$ is immediately derived from $p_{E(n\to\infty)}(\cdot)$. Finally, it is remarked that analysis of FA and TDMA can be done by limiting the set of active states to $A^1_{\varepsilon_\delta}, ..., A^1_N$ (i.e., no retransmission), since any node after transmission returns idle regardless the transmission outcome.

From / To	I_l; $l \in \{j, ..., \varepsilon_\delta - 1\}$	A^1_l; $l \in \{\varepsilon_\delta, ..., N-1\}$	A^1_l; $l = N$
a) I_j	q_{l-j}	q_{l-j}	$q_{j,N} = 1 - \sum_{i=0}^{N-j-1} q_i$

From / To	I_j	$I_{j-\varepsilon_\delta}$	$A^{k+1}_{j-\varepsilon_\delta}$
A^k_j; $j \in \{\varepsilon_\delta, ..., 2\varepsilon_\delta - 1\}$, for $k = 1$	$\overline{\alpha}$	α	0
A^k_j; $j \in \{2\varepsilon_\delta, ..., N\}$, for $k = 1$	$\overline{\alpha}$	$\alpha p_{c,1}$	$\alpha \overline{p}_{c,1}$
A^k_j; $j \in \{\varepsilon_\delta, ..., 2\varepsilon_\delta - 1\}$, for $k > 1$	0	1	0
A^k_j; $j \in \{2\varepsilon_\delta, ..., N\}$, for $k > 1$	0	$p_{c,k}$	$\overline{p}_{c,k}$

Figure 3.4 State transition probabilities for the DMC model in Section 3.6.2 due to: a) energy harvesting; and b) the bidirectional communication with the FC. The transition matrix $\mathbf{P}$ can be derived according to the probabilities in a) and b) for all the values of $k \in \{1, ..., F_\varepsilon\}$ and $j \in \{0, ..., N\}$.

3.7   Backlog Estimation

This section proposes a backlog estimation algorithm for the DFA protocol (extension to FA is straightforward). Unlike previous work on the subject [57, 60], here backlog estimation is designed by accounting for the interplay of EH, capture effect and multiple access. Computational complexity of optimal estimators is generally

intractable for a large number of nodes even for non-EH systems [60]. This section thus proposes a low-complexity two-steps backlog estimation algorithm that, neglecting the IR index, operates in every IR as follows: i) the FC estimates the initial backlog size B_1 based on the ccdf $G_E(\varepsilon)$ of the ESD energy at the beginning of the current IR; ii) the backlog estimates for the next frames are updated based on the channel outcomes and the residual ESD energy.

For the first frame, the backlog size estimate and the frame length are $\hat{B}_1 = M\alpha G_E(\varepsilon)$ and $L_1 = \lceil \rho\hat{B}_1 \rceil$, respectively. For subsequent frames, assume that the FC announced a frame of $L_k = \lceil \rho\hat{B}_k \rceil$ slots. The FC estimates the backlog size for frame $k+1$ by counting the number of slots that are successful ($N_{D,k}$) and collided ($N_{C,k}$) within the kth frame of length L_k slots. Since the FC cannot discern exactly how many nodes transmitted in each successful slot, the estimate of the total number $C_{D,k}$ of nodes that collided in $N_{D,k}$ successful slots is $\hat{C}_{D,k} = (\beta_{D,k} - 1)N_{D,k}$, with $\beta_{D,k}$ being the conditional average number of nodes that transmit in a slot given that the slot is successful (with no capture $\beta_{D,k} = 1$). Similarly, for the collided slots one can obtain $\hat{C}_{C,k} = \beta_{C,k}N_{C,k}$, where $\beta_{C,k}$ is now conditioned on observing a collided slot. Derivations of $\beta_{D,k}$ and $\beta_{C,k}$ are in Section 3.7.1. Since the estimate of the total number of nodes that unsuccessfully transmitted is $\hat{C}_k = \hat{C}_{C,k} + \hat{C}_{D,k}$, the backlog size estimate $\hat{B}_{k+1}$ for the $(k+1)$th frame is obtained by accounting for the fraction of nodes within $\hat{C}_k$ that are not in energy shortage: $\hat{B}_{k+1} = \hat{C}_k G_E((k+1)\varepsilon|k\varepsilon)$, where $G_E((k+1)\varepsilon|k\varepsilon) = \Pr[E_m \geq (k+1)\varepsilon|E_m \geq k\varepsilon]$. The proposed backlog estimation scheme thus works as follows:

$$\hat{B}_k = \begin{cases} M\alpha G_E(\varepsilon) & \text{if } k = 1 \\ \hat{C}_{k-1}G_E(k\varepsilon|(k-1)\varepsilon) & \text{if } k > 1 \end{cases}. \tag{3.21}$$

Algorithm (3.21) can be applied to any IR n by deriving the ESD distribution $p_{E(n)}(\cdot)$ (or $G_{E(n)}(\cdot)$) from any initial distribution $p_{E(1)}(\cdot)$ by exploiting the DMC model in Section 3.6.2.

3.7.1 Average Number of Node Transmissions per Slot

The conditional averages $\beta_{D,k}$ and $\beta_{C,k}$ are calculated similarly to [13] by accounting for the capture effect and an arbitrary ρ. Let Y be the number of simultaneous transmissions in the same slot, and let $\mathcal{U}_k$ and $\mathcal{C}_k$ be the event of successful and collided slot in frame k, respectively. The average number of nodes per successful and collided slot are respectively

$$\beta_{D,k} = \sum_{j=1}^{\infty} j \Pr\left[Y = j | \mathcal{U}_k\right]; and \ \beta_{C,k} = \sum_{j=2}^{\infty} j \Pr\left[Y = j | \mathcal{C}_k\right]. \tag{3.22}$$

To calculate $\beta_{D,k}$ consider $\mathcal{A}.1$ and $\mathcal{A}.2$ and allow the number of possible interfering nodes up to infinity as in Section 3.5.2. By exploiting the Bayes rule, one can get $\Pr\left[Y = j | \mathcal{U}_k\right] = \Pr\left[\mathcal{U}_k | Y = j\right] \frac{\Pr[Y=j]}{\Pr[\mathcal{U}_k]}$, where $\Pr\left[\mathcal{U}_k | Y = j\right] = j p_{c,k}(j-1)$, $\Pr\left[Y = j\right] = e^{-\frac{1}{\rho}}/(\rho^j j!)$ and $\Pr\left[\mathcal{U}_k\right] = p_{t,k}^{DFA}$ (see 3.18). Similarly, one can obtain $\beta_{C,k}$ given that $\Pr\left[\mathcal{C}_k\right] = 1 - \Pr\left[\mathcal{U}_k\right] - \beta\left(0, B, L\right)$, where $\beta\left(0, B, L\right) \simeq e^{-\frac{1}{\rho}}$ is the probability of an empty slot, and $\Pr\left[\mathcal{C}_k | Y = j\right] = 1 - \Pr\left[\mathcal{U}_k | Y = j\right]$ for $j \geq 1$.

3.8 Numerical Results

This section presents extensive numerical results to get insight into the MAC protocols design. Moreover, to validate the analysis proposed in Section 3.5 and Section 3.6, the analytical results derived therein are compared with a simulated system that does not rely on simplifying assumptions $\mathcal{A}.1$ and $\mathcal{A}.2$. The performances of the backlog estimation algorithm proposed in Section 3.7 are also assessed through a comparison with the ideal case of perfectly known backlog at the FC.

3.8.1 MAC Performance Metrics Trade-offs

The energy $E_{H,m}(n)$ harvested between two successive IRs is assumed as geometrically-distributed so that $q_i = \Pr[E_{H,m}(n) = i\delta] = \xi(1 - \xi)^i$, with $\xi = \delta/(\delta + \mu_H)$, where it has been defined the *harvesting rate* μ_H as the average harvested energy normalized by ε as $\mu_H = E[E_{H,m}(n)/\varepsilon]$.

The asymptotic time efficiencies (3.5) for TDMA, FA and DFA protocols, are shown in Figure 3.5 versus design parameter ρ (recall (3.6)). System performance is evaluated by considering: $\mu_H \in \{0.15, 0.35\}$, $M = 400$, $\gamma_{th} = 3dB$, $\alpha = 0.3$; ε is normalized to one, energy unit is $\delta = 1/50$ so that $\varepsilon_\delta = 50$ and $F_\varepsilon = 10$. Figure 3.5 compares the analytical performance metrics derived in Section 3.5 with simulated scenarios for both known and estimated backlog. TDMA's performance is clearly independent of ρ, while in FA and DFA there is a time efficiency-maximizing ρ that is close to one (in [13] the optimal value was $\rho = 1$ since the capture effect was not considered). The effect of decreasing (or increasing) the harvesting rate μ_H on the TDMA time efficiency is due to the larger (or smaller) number of nodes that are in energy shortage and whose slots are not used, while it is negligible for FA and DFA due to their ability to dynamically adjust the frame size according to backlog estimates $\hat{B}_k$. The tight match between analytical and simulated results also validates assumptions $\mathcal{A}.1$ and $\mathcal{A}.2$ and the efficacy of the backlog estimation algorithm.

The asymptotic delivery probability (3.3), for harvesting rate $\mu_H \in \{0.05, 0.15, 0.35\}$, versus parameter ρ is shown in Figure 3.6 with the same system parameters as for Figure 3.5. Unlike for the time efficiency, TDMA always outperforms FA and DFA in terms of delivery probability. In fact, nodes operating with TDMA and FA have the same energy consumption since they transmit at most once per IR, while possibly more than once in DFA. However, TDMA does not suffer collisions and thus it is able to eventually deliver more packets to the FC. The delivery probability strongly depends on the harvesting rate μ_H, which influences

the ESD energy distribution and thus the energy shortage probability. Moreover, DFA outperforms FA thanks to the retransmission capability when the harvesting rate is relatively high (e.g., $\mu_H = 0.35$). Whereas, for low harvesting rate (e.g., $\mu_H \in \{0.05, 0.15\}$) DFA and FA perform similarly since most of the nodes are either in energy shortage or have very low energy in their ESDs, thus being unable to fully exploit the retransmission opportunities provided by DFA.

The trade-off between asymptotic delivery probability (3.3) and asymptotic time efficiency (3.5) is shown in Figure 3.7 for different values of the harvesting rate $\mu_H \in \{0.05, 0.15, 0.35\}$. System parameters are the same as for Figure 3.5. For TDMA, the trade-off consists of a single point on the plane, whereas FA and DFA allow for more flexibility via the selection of parameter ρ. When increasing ρ more nodes might eventually report their measurements to the FC, thus increasing the delivery probability to the cost of lowering time efficiency (see Figure 3.5 and 3.6). For FA and DFA, the trade-off curves are obtained as $\max_\rho \{p_d^{AS}\}$, s.t. $p_t^{AS} = \lambda$ for each achievable λ.

The impact of the capture effect on the performance metrics trade-offs is shown in Figure 3.8, where the SIR threshold $\gamma_{th} \in \{0.01, 3, 10\}dB$ is varied while the harvesting rate $\mu_H = 0.15$ is kept fixed (other parameters are as in Figure 3.5). As expected, the lower the SIR threshold γ_{th} the higher the probability that the SIR of any of the colliding nodes is above γ_{th}, and thus the higher the performance obtained with ALOHA-based protocols. TDMA is insensitive to γ_{th}.

3.9 Conclusions

The design of medium access control (MAC) protocols for single-hop wireless node networks (WSNs) with energy-harvesting (EH) devices offers new challenges as compared to the standard scenario with battery-powered (BP) nodes. New performance criteria are called for, along with new design solutions. This chapter
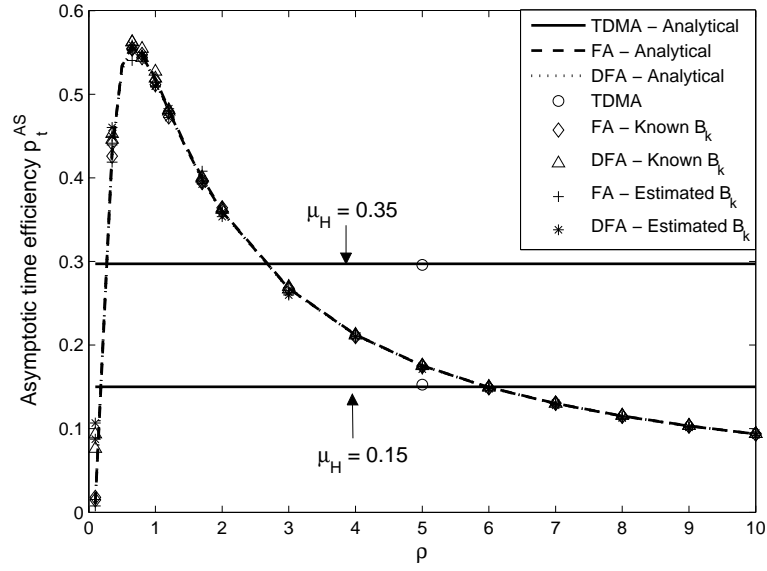
Figure 3.5 Asymptotic time efficiency (3.5) versus ρ, for different harvesting rates $\mu_H \in \{0.15, 0.35\}$. Comparisons are between analytical and simulated results with both known (B_k) and estimated backlog $(\hat{B}_k$, see (3.21)), ($M = 400$, $\gamma_{th} = 3dB$, $\alpha = 0.3$, $F_\varepsilon = 10$, $\varepsilon = 1$, $\delta = 1/50$).
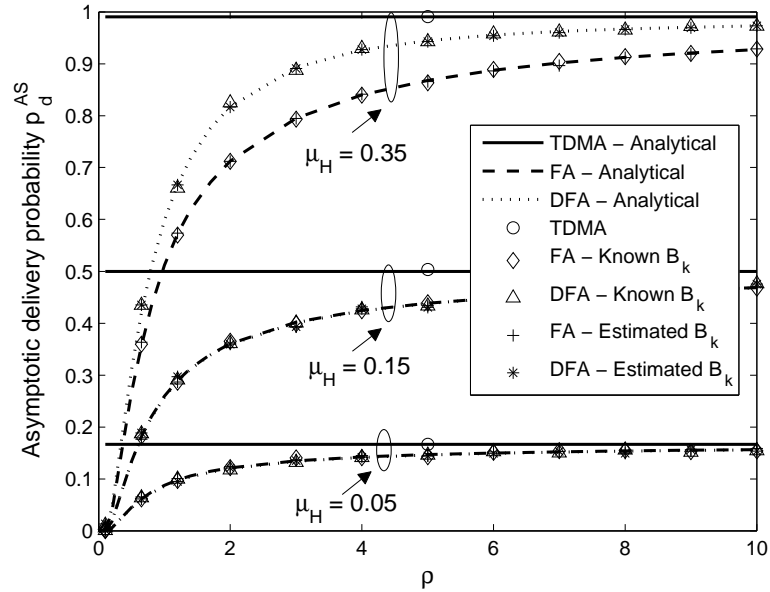
Figure 3.6 Asymptotic delivery probability (3.3) versus ρ, for different harvesting rate $\mu_H \in \{0.05, 0.15, 0.35\}$. Comparisons are between analytical and simulated results with both known (B_k) and estimated backlog $(\hat{B}_k$, see (3.21)), ($M = 400$, $\gamma_{th} = 3dB$, $\alpha = 0.3$, $F_\varepsilon = 10$, $\varepsilon = 1$, $\delta = 1/50$).
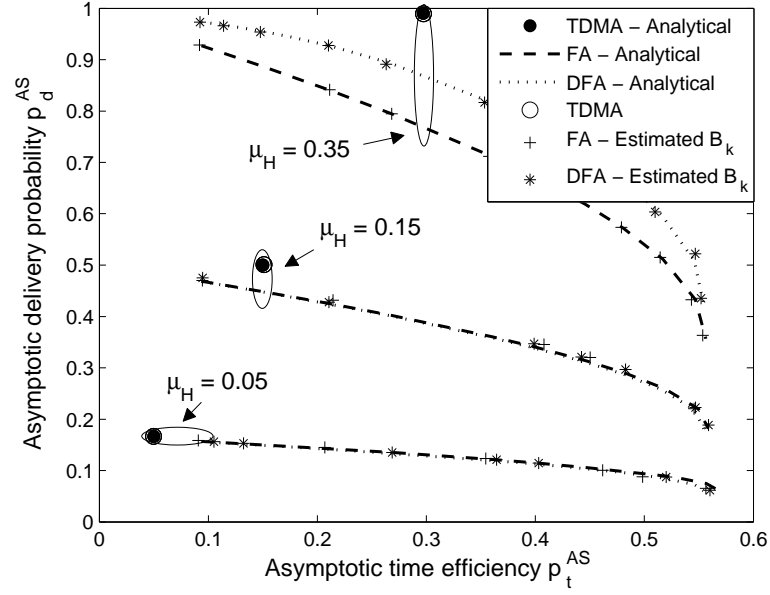
Figure 3.7 Trade-off between asymptotic delivery probability (3.3) and asymptotic time efficiency (3.5) for different harvesting rate $\mu_H \in \{0.05, 0.15, 0.35\}$. Comparisons are between analytical and simulated results with estimated backlog ($\hat{B}_k$, see (3.21)), ($M = 400$, $\gamma_{th} = 3dB$, $\alpha = 0.3$, $F_\varepsilon = 10$, $\varepsilon = 1$, $\delta = 1/50$).
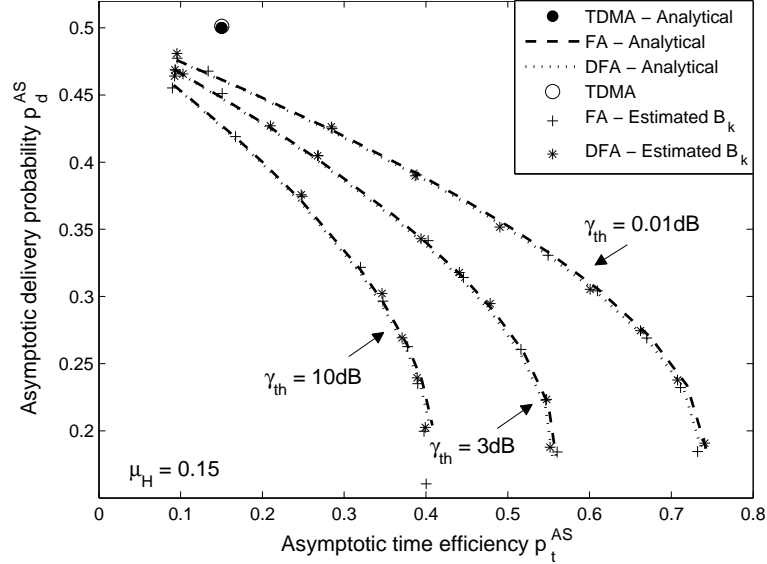
Figure 3.8 Trade-off between asymptotic delivery probability (3.3) and asymptotic time efficiency (3.5) for different SIR threshold $\gamma_{th} \in \{0.01, 3, 10\}dB$ values and fixed energy harvesting rate $\mu_H = 0.15$. Comparisons are between analytical derivations and simulated results with estimated backlog ($\hat{B}_k$, see (3.21)), ($M = 400$, $\alpha = 0.3$, $F_\varepsilon = 10$, $\varepsilon = 1$, $\delta = 1/50$).

addressed these issues by investigating the novel trade-off between the *delivery probability,* which measures the capability of a MAC protocol to deliver the measure of any node in the network to the intended destination (i.e., fusion center, FC) and the *time efficiency*, which measures the data collection rate at the FC. The analysis is focused on standard MAC protocols, such as TDMA, Framed-ALOHA (FA) and Dynamic-FA (DFA). Novel design issues are also discussed, such as backlog estimation and frame length selection. Extensive numerical results and discussions validate the proposed analytical framework and provide insight into the design of EH-WSNs.

CHAPTER 4

ENERGY GROUP DYNAMIC FRAMED-ALOHA PROTOCOL

This chapter proposes a novel random access protocol for data collection from a set of energy harvesting (EH) capable wireless nodes. The scheme is a variant of the dynamic framed-ALOHA (DFA) protocol and it is tailored to EH networks. The proposed scheme, referred to as energy group-DFA (EG-DFA), is based on the observation that, when DFA is operated with EH-capable nodes, the optimal number of slots in a frame (i.e., the frame size) must balance two conflicting performance requirements as shown in Section 3.8. First, in a perfect collision channel (i.e., no capture effect at the FC) it is well-known that the time efficiency (i.e., the data collection rate see Section 3.5.5) is maximized when the frame size is equal to the backlog, namely the number of transmitting nodes. Second, since each node can store and harvest a finite energy, the number of (re)transmissions attempts that each node can perform during the channel contention process is limited. Thus, to reduce the probability of packets collisions and thus the energy wastage due to retransmissions, the frame size should be selected as large as possible, so that the delivery probability is increased. Therefore, the choice of the frame size is crucial in determining the trade-off between time efficiency and delivery probability, where their trade-off strongly depends on the energy harvesting rate and thus the energy availability at nodes as shown in Section 3.8.

Based on the insights above, the key idea of EG-DFA is to divide nodes in groups according to their energy availability, and let each group access the channel via a separate instance of DFA, whereby different values of ρ can be selected for each group. A grouping technique for ALOHA-based MACs was proposed in [61], via a protocol that here it is referred to as *Group-DFA* (G-DFA). G-DFA divides

nodes in groups, and each groups' data packets are collected by the FC through separated instances of DFA all with the same parameter ρ. G-DFA improves DFA's time efficiency, as decreasing the number of nodes competing for the same frame increases the chance of a successful transmission. However, the G-DFA protocol was developed for nodes with no energy constraints, thus without considering the crucial trade-off between time efficiency and delivery probability. This trade-off is instead tackled in EG-DFA, which combines the *grouping gain* of G-DFA with the ability to tune the design parameter ρ to the group's energy availability. In fact, as discussed above, obtaining an high delivery probability in groups with small energies requires large ρ values to decrease energy-wastage due to collisions, whereas ρ close to one is expected to be optimal for groups with large energies.

To analyze the performance of the EG-DFA protocol, here it is considered a simplified system model with respect to the one considered in Section 3.2.2. Furthermore, to denote time, this section will generally use a double index (n,i), which denotes the beginning of the ith frame, $i = 1, 2, ...,$ in the nth IR, $n = 1, 2, ...$ (slots are not indexed).

4.1   Energy Model for EG-DFA

To simplify the presentation of the EG-DFA protocol, here it is considered that the energy unit δ, used to describe the granularity of the discrete ESD (see Section *3.2.2*), is equal to the energy per frame ε, so that $\varepsilon_\delta = 1$. Therefore, let $E_m(n, i) \in \{0, 1, ..., C\}$ be the energy stored in the ESD of the mth node at the beginning of the ith frame during the nth IR, where C is the ESD capacity. The energy $E_m(1, 1)$ initially stored in the mth node's ESD is a random variable independent and identically distributed (i.i.d.) among nodes.

The EHU of the mth node harvests energy $e_m(n)$ during the time T_{int} between the beginning of the nth and $(n+1)$th IRs. The harvested energy $e_m(n)$ is a random
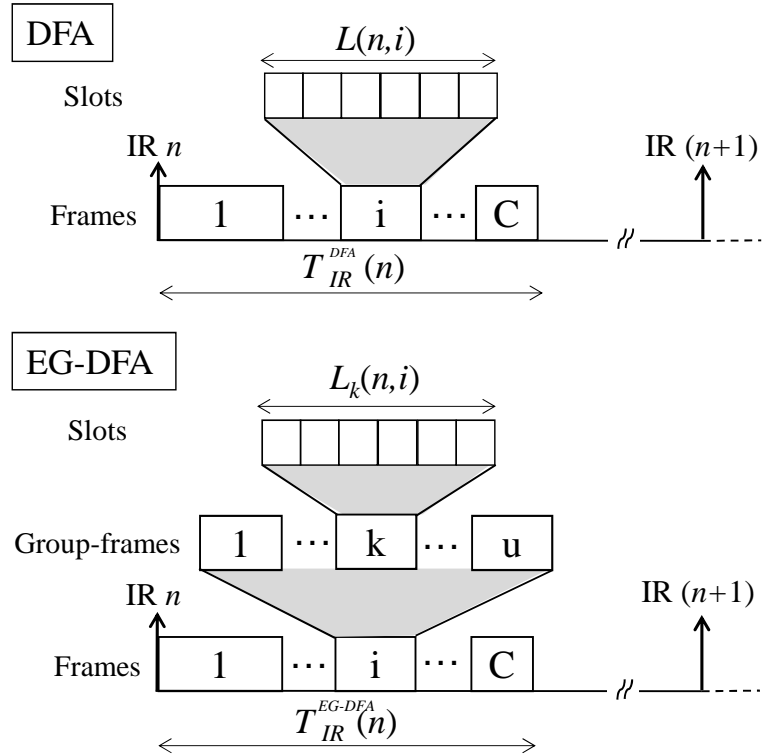
Figure 4.1 Organization of slots into frames in the dynamic framed aloha (DFA) protocol, and into group-frames and frames in the energy group-DFA (EG-DFA) protocol. The same structure is repeated every T_{int} [s] for each IR. Frames in DFA and group-frames in EG-DFA are designed according to Section 4.2.1 and Section 4.2.2, respectively. Group-DFA (G-DFA) uses a structure similar to EG-DFA (see Section 4.2.1).

variable, i.i.d. across nodes and IRs, independent on the IR duration $T_{IR}(n)$, and with probability mass function (pmf) $p_e(k) = \Pr[e_m(n) = k]$. Note that, as the ESD is finite the energy harvested when the ESD is fully charged is wasted. This section assumes that each node operates in each nth IR using only the energy stored in its ESD at time $(n, 1)$, while the energy harvested during the current IR can only be used in the next IRs. The energy in the mth node's ESD is a random variable that evolves across IRs as $E_m(n+1, 1) = \min\{C,\ E_m(n, 1) - \sum_i T_m(n, i) + e_m(n)\}$, where the indicator $T_m(n, i)$ equals one if node m transmits in the ith frame of the nth IR, and zero otherwise. It follows $\sum_i T_m(n, i) \le E_m(n, 1)$. Moreover, the energy in the mth node's ESD evolves across successive frames of any nth IR as $E_m(n, i) = E_m(n, 1) - \sum_{k=1}^{i-1} T_m(n, k)$.

At the beginning of the nth IR at time $(n, 1)$, the mth node is assumed to have a new data packet to transmit with probability α, and no packet with probability $(1 - \alpha)$, independently from the other nodes and on previously generated packets and IRs (i.e., there is no data buffer). The mth node with a new packet is *active* at time $(n, 1)$, if it has enough energy to transmit, i.e., if $E_m(n, 1) \ge 1$. At the ith frame at time (n, i), with $i > 1$, the mth node is active if: i) it was active at time $(n, 1)$; ii) its energy is $E_m(n, i) \ge 1$; iii) its packet still has to be received correctly by the FC (i.e., all previous attempts, if any, were unsuccessful).

4.2 Energy-Group Based DFA

This section first reviews the DFA and G-DFA protocol with the notation adapted to this chapter and then introduces the EG-DFA protocol. Let $M_k(n, i)$ be the number of nodes with energy $E_m(n, i) = k$ at time (n, i), and $B_k(n, i) \le M_k(n, i)$ be the number of *active* nodes, within the $M_k(n, i)$ with energy k. Let

$$B(n, i) = \sum_{k=1}^{C} B_k(n, i) \le \sum_{k=0}^{C} M_k(n, i) = M, \qquad (4.1)$$

be the overall *backlog*, i.e., the total number of active nodes, at time (n, i). To simplify protocols' description, hereafter it is assumed that the FC exactly knows the backlogs $B_k(n, i)$ at any time. Backlog estimation algorithms for DFA and G-DFA protocols have been investigated in previous works (see e.g., [53, 60]). For the sake of completeness, a simple backlog estimation algorithm specifically designed for the EG-DFA protocol is proposed in Section 4.3.

4.2.1   DFA and G-DFA

In DFA, the number of slots in each frame at time (n, i) is selected as

$$L(n, i) = \lceil \rho B(n, i) \rceil, \tag{4.2}$$

where $\lceil \rceil$ is the nearest upper integer operator, and the design parameter ρ is selected such that $\rho \in [1, \rho_{\max}]$. Parameter ρ is chosen greater than one since for $\rho < 1$ both time efficiency and delivery probability are simultaneously penalized, while choosing $\rho \leq \rho_{\max}$ is to consider frame sizes of practical values. Each of the $B(n, i)$ active nodes randomly and uniformly selects one slot for transmissions in the current frame. After the end of the ith frame, the FC updates the backlog size for the next $(i+1)$th frame as $B(n, i+1) = B(n, i) - D(n, i) - S(n, i)$, where $D(n, i)$ denotes the number of packets successfully decoded and $S(n, i)$ indicates the number of nodes that collided in frame i and that have no energy left in the ESD for transmitting in frame $(i + 1)$. The FC keeps announcing frames until no more nodes are available for transmission so that the ith is the last frame if $B(n, i + 1) = 0$. Clearly, since the ESD is finite there cannot be more than C frames in an IR.

G-DFA is characterized by grouping, namely at the beginning of the nth IR, each active node randomly and uniformly selects one out of G groups to belong to. Each group of nodes then accesses the channel by resorting to G separate instances of DFA (through time-division over the same channel), one for each group. Specifically,

in each frame of the IR, G subframes, referred to as *group-frames*, are allocated (see Figure 4.1). Each group-frame contains slots intended only for nodes belonging to the specific group. Note that, only one group-frame per group is allowed within a frame and that all the G instances of DFA are operated with the same ρ.

4.2.2 Energy-Group DFA

Similarly to G-DFA, the EG-DFA protocol divides the nodes into groups as shown in Figure 4.1. However, in EG-DFA each active node selects its own group in each frame (say at time (n, i)) *based on the energy currently available in its ESD*. Specifically, the kth group at time (n, i) contains all the active nodes with energy k at time (n, i). Accordingly, those active nodes that are initially in the kth group at time $(n, 1)$ and that collide for j consecutive times $(j < k)$ will belong to group $k - j$ in frame j. Note that, even if (colliding) active nodes change group index across frames, they always compete with the same set of nodes that were in the same group at time $(n, 1)$. The EG-DFA thus has C parallel instances of the DFA protocol (one for each energy level in the ESD), similarly to G-DFA (where $C = G$), but here the kth instance of DFA resolves only nodes with equal initial energy level k. Furthermore, the instance of DFA for each energy level k is operated with a different parameter ρ_k, so that the trade-off between time efficiency and DER can be addressed according to the energy availability at nodes.

To elaborate, in the ith frame the FC announces $(C - i + 1)$ group-frames since no active nodes can have energy greater than $(C - i + 1)$ at time (n, i). Recall in fact that the energy harvested during an IR will be available only in the next IR. Let $B_k(n, i)$ be the backlog for group k at time (n, i), then the number of slots in the kth group-frame, for $(1 \leq k \leq C - i + 1)$, is

$$L_k(n, i) = \lceil \rho_k B_k(n, i) \rceil, \tag{4.3}$$

with ρ_k chosen as $\rho_1 \geq ... \geq \rho_C$, since a larger ρ_k is generally preferable for low-energy group as it decreases energy wastage due to collisions. Consequently, active nodes that collided in the current frame will transmit in the next frame with a generally larger ρ_k.

The backlogs $B_k(n,i)$ are updated at the end of each ith frame as

$$B_k(n, i+1) = \begin{cases} B_{k+1}(n,i) - D_{k+1}(n,i) & \text{for } 1 \leq k \leq C - i \\ 0 & \text{for } C - i < k \leq C \end{cases}, \qquad (4.4)$$

where $D_k(n,i) \leq B_k(n,i)$ is the number of nodes in group k at time (n,i) that successfully transmitted in frame i. Eq. (4.4) holds as active nodes with energy $(k+1)$ at time (n,i) (i.e., $B_{k+1}(n,i)$), which collide in frame i, will be the only $B_k(n, i+1)$ active nodes in the $(i+1)$th frame with energy k (for $k \geq 1$). The procedure repeats until the overall backlog (4.1) becomes empty, i.e., $B(n, i+1) = 0$.

4.2.3 Performance Metrics

This section focuses on the performance of the data collection process in a high delivery probability regime, so that it is more convenient to consider the fraction of the backlog that is not correctly retrieved by the FC, which is referred to as delivery error rate (DER). The DER can be seen as the complement of the delivery probability, and it is defined as

$$\nu = 1 - \frac{\lim_{n \to \infty} \frac{1}{n} \sum_{l=1}^{n} E[D_{IR}(l)]}{\alpha M}, \qquad (4.5)$$

where αM is the average number of nodes with a new packet to transmit at the beginning of an IR if there were no energy limitations (recall that α is the probability that a node has a new measure to transmit in an IR). The DER counts as lost both the packets of active nodes that end up in energy shortage during the IR, and the potential packets of nodes that have no energy since the IR's beginning. This is

relevant for EH systems as protocols are expected to be able to collect a large number of packets in the given IR while saving energy for next IRs.

Clearly, there is a critical trade-off between the (asymptotic) time efficiency and the DER. The time efficiency (defined as in (3.5)) accounts for the speed of the collection process, while the DER indicates how many packets of the (average) potential overall batch of αM transmitting nodes are not retrieved due to energy shortages. A reasonable design criterion is thus to maximize the time efficiency while constraining the DER ν to be smaller than a threshold value $\bar{\nu}$ (i.e., $\nu \leq \bar{\nu}$) as

$$p_t^* = \max_{\rho_1,...,\rho_C} p_t \text{ s.t. } \nu \leq \bar{\nu}, \tag{4.6}$$

with the goal of optimizing parameters $\rho_1, ..., \rho_C$. In this regard, [54] shows that by judiciously selecting parameter ρ in DFA, small DER values (or high delivery probability) can be achieved with limited losses on time efficiency.

4.3   Backlog Estimation Algorithm for EG-DFA

Since optimal backlog estimation algorithms are computational expensive even for DFA [60], this section proposes a low-complexity two-phases scheme [54] tailored to the EG-DFA protocol. The first phase is operated by the FC within each nth IR, and it is based on the observations of the channel outcomes (e.g., collided slots) [13]. In the second phase the FC accounts for the EH process.

Phase 1. Let $\hat{M}_k(n, 1)$ be the estimated number of nodes in the kth group at time $(n, 1)$. The estimate at time $(1, 1)$ is $\hat{M}_k(1, 1) = M \Pr[E_m(1, 1) = k]$ (i.e., the expected number of nodes with energy k). The kth group's backlog estimation at time $(n, 1)$ is $\hat{B}_k(n, 1) = \alpha \hat{M}_k(n, 1)$. When the first frame ends, the FC counts the number of successfully received packets $D_k(n, 1)$ and collided slots $Z_k(n, 1)$ in each kth group-frame. According to (4.4), the nodes that transmitted in the $Z_k(n, 1)$ collided slots will form the backlog $B_{k-1}(n, 2)$ for the $(k-1)$th group in the next frame.

However, the FC cannot discern how many nodes were involved in the collision, and thus an estimation of $B_{k-1}(n,2)$ can be obtained as $\hat{B}_{k-1}(n,2) = Z_k(n,1)\beta(\rho_k)$, where $\beta(\rho_k)$ is the average number of nodes per observed-collided slot when the frame is dimensioned as $L = \lceil \rho B \rceil$. This estimator, first proposed in [13] for $\rho = 1$ with $\beta(1) \simeq 2.39$, was then extended in [54], where $\beta(\rho)$ was computed, under a large backlog approximation for any ρ, as $\beta(\rho) \simeq (1 - e^{-1/\rho})/(\rho - \rho e^{-1/\rho} - e^{-1/\rho})$. By iterating this procedure, the backlog estimate at time (n,i) for the kth group (with $k \geq 1$) is $\hat{B}_k(n,i) = \alpha \hat{M}_k(n,i)$, for $i = 1$ and $\hat{B}_k(n,i) = Z_{k+1}(n,i-1)\beta(\rho_{k+1})$ for $i > 1$.

Phase 2. Let $M'_k(n)$ be the number of nodes in the kth group after the nth IR ends and before accounting for the EH process. $M'_k(n)$ is given by the sum of the number of: *i*) nodes $\sum_{i=1}^{C-k} D_{k+1}(n,i)$ that transmitted successfully within the $(k+1)$th group in the ith frame (known by the FC); *ii*) idle nodes $M_k(n,1) - B_k(n,1)$ that were initially in the kth group at time $(n,1)$ and that did not have a new measure to transmit, this is estimated (packet generation is random) as $\hat{M}_k(n,1) - \hat{B}_k(n,1) = \hat{M}_k(n,1)(1-\alpha)$. Accordingly, $M'_k(n)$ is estimated as $\hat{M}'_k(n) = \sum_{i=1}^{C-k} D_{k+1}(n,i) + \hat{M}_k(n,1)(1-\alpha)$, which might need to be conveniently normalized so that $\sum_{k=1}^{C} \hat{M}'_k(n) = M$. The number of nodes $M_k(n+1,1)$ at the $(n+1)$th IR's beginning can be obtained from $\hat{M}'_k(n)$ by using the expectation over the EH pmf $p_e(\cdot)$ as $\hat{M}_k(n+1,1) = \sum_{j=0}^{k} \hat{M}'_j(n) p_e(k-j)$ if $0 \leq k < C$, while $\hat{M}_k(n+1,1) = M - \sum_{k=0}^{C-1} \hat{M}_k(n+1,1)$ if $k = C$.

4.4 Numerical Results and Discussion

This section presents extensive numerical results to get insights into EG-DFA's design and performance by numerically solving the constrained optimization problem (4.6) through a grid search.

Figure 4.2 illustrates the (asymptotic) time efficiency p_t^* versus the DER constraint $\bar{\nu}$ for the DFA, G-DFA and EG-DFA protocols. For reference, DFA and G-DFA's performances are shown by assuming that the FC perfectly knows the backlog at all times, while EG-DFA's performance are shown with both known and estimated backlog (see algorithm in Section 4.3). Figure 4.2 also shows the performance of the EG-DFA protocol when the solution of (4.6) is restricted by setting $\rho_k = \rho$ for each group $k \in [1, C]$, thus only exploiting nodes' grouping gain. System parameters are: $M = 100$ nodes; ESD's capacity $C = 8$; number of G-DFA's groups $G = C = 8$; $\alpha = 0.5$; EH's pmf $p_e(\cdot)$ exponential with mean $E[e_m(n)] = 2$.

From Figure 4.2, it can be seen that EG-DFA with known backlog outperforms DFA, in terms of time efficiency, for any DER constraints $\bar{\nu}$, and also G-DFA for moderate-to-low DER values (here $\bar{\nu} \leq 4 \cdot 10^{-1}$). For higher DER constraints ($\bar{\nu} > 4 \cdot 10^{-1}$) G-DFA outperforms EG-DFA. This is because, one can decrease the design parameters ρ_k towards one, as collisions and thus energy wastage are less penalized when increasing the DER threshold $\bar{\nu}$. This implies that, when the EH rate is limited, most of the nodes have a small stored energy and only few groups in EG-DFA will have non-zero backlogs, thus drastically reducing grouping gain. Conversely, in G-DFA, groups are occupied uniformly (and randomly) regardless of the nodes' energy, and hence grouping gain is still fully exploited. Notice that, even if backlog estimation reduces EG-DFA's performance, it still allows to outperform both DFA and G-DFA with known backlog for a wide range of DER constraint $\bar{\nu}$. The results in this section also suggest (not shown) that the optimal ρ_k^* values increases as DER decreases and they increase more as the energy availability (i.e., group index k) gets smaller, consistently with the intuition in Section 4.2.2, while they approach unity for each k for large DER values, as in this regime the time efficiency is the relevant metric.

The effects of the ESD capacity C are shown in Figure 4.3 for DER constraints $\bar{\nu} \le \{2 \cdot 10^{-1}, 5 \cdot 10^{-3}\}$. System parameters are: $M = 100$; $G = C$; $\alpha = 0.5$; $E[e_m(n)] = 3$. For small ESD's capacity C, the energy harvested when the ESD is full cannot be stored, and thus nodes can easily get in energy shortage even when the harvesting rate is large (i.e., $E[e_m(n)] \gg 1$). This causes a significant performance loss and it imposes constraints on the achievable values of DER. For instance, if $C < 6$ a DER smaller than $\bar{\nu} \le 5 \cdot 10^{-3}$ is not achievable by any technique. Moreover, small C values reduce the capability of grouping nodes and thus enabling small grouping gains only.

Figure 4.4 shows the effects of varying the average harvesting rate $E[e_m(n)]$ on the time efficiency for DER constraints $\bar{\nu} \le \{5 \cdot 10^{-2}, 5 \cdot 10^{-3}\}$. Parameters are as above with $C = G = 8$. When the harvesting rate is small (e.g., $E[e_m(n)] \le 3$), EG-DFA outperforms both G-DFA and DFA for both DER constraints. However, the gap between EG-DFA and G-DFA gets smaller as the harvesting rate increases. In fact, most nodes have full ESDs, and this causes only high-energy availability groups to have non-zero backlog, thus reducing EG-DFA's grouping gain. G-DFA's grouping gain is instead preserved as groups are uniformly occupied as described above.

As a final remark, note that for large C values EG-DFA can be operated by bundling close energy groups together without increasing the protocol complexity (i.e., number of groups).

4.5   Conclusions

The design of protocols for wireless networks with Energy-Harvesting (EH) calls for novel approaches that address the unique requirements imposed by the variability of the energy available at the nodes. This chapter proposed a variant of dynamic framed ALOHA (DFA) that is tailored to the problem of periodic data collection from a set of EH nodes. The proposed scheme, termed energy group-DFA (EG-DFA), improves
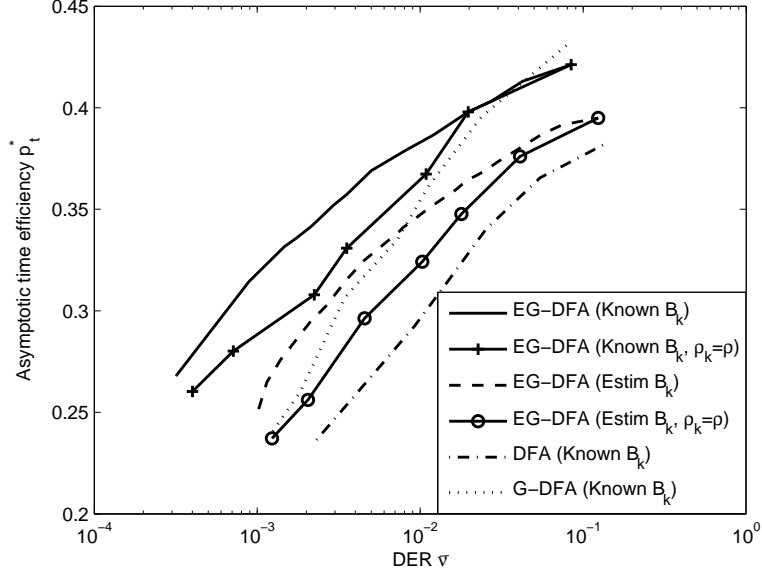
Figure 4.2 Asymptotic time efficiency p_t^* versus DER $\bar{\nu}$ for the DFA and G-DFA protocols with known backlog, and for EG-DFA with both known and estimated backlog ($M = 100$, $\alpha = 0.5$, $C = G = 8$, $E\left[e_m(n)\right] = 2$).
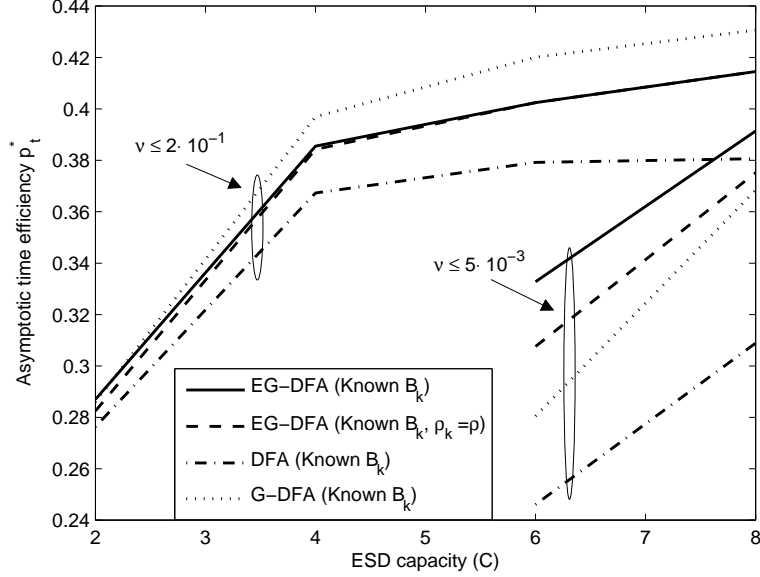
Figure 4.3 Asymptotic time efficiency p_t^* versus ESD capacity C for the EG-DFA, G-DFA and DFA protocols, assuming perfect knowledge of the backlog. The DER is constrained to be $\nu \leq \{5 \cdot 10^{-3}, 2 \cdot 10^{-1}\}$ ($M = 100$, $\alpha = 0.5$, $G = C$, $E\left[e_m(n)\right] = 3$).

the performance of DFA by leveraging the observation that the optimal size of the frame in DFA, when implemented over EH nodes, depends critically on the energy levels at the nodes and on the harvesting rate. Performance is evaluated in terms of
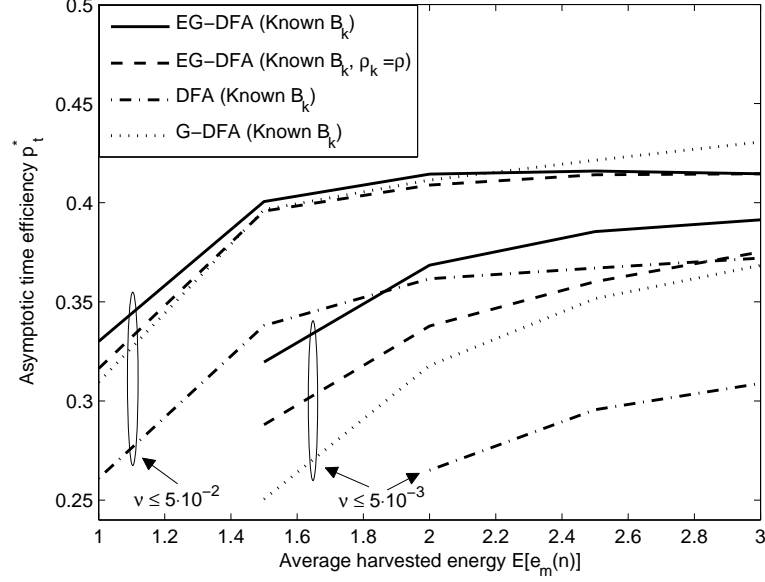
Figure 4.4 Asymptotic time efficiency p_t^* versus average harvested (normalized) energy per IR $E[e_m(n)/\varepsilon]$ for the EG-DFA, G-DFA and DFA protocols, assuming perfect knowledge of the backlog. The DER is constrained to be $\nu \leq \{5 \cdot 10^{-3}, 5 \cdot 10^{-2}\}$ ($M = 100$, $\alpha = 0.5$, $G = C = 8$).

the trade-off between the time efficiency and the delivery error rate (DER), where the latter measures the capability of collecting data from the nodes before they run out of energy. EG-DFA is shown via simulations to outperform known strategies in terms of time efficiency in the low DER regime. Impacts of the size of the energy storage device and of the harvesting rate are investigated as well. Extensions to this work can include the development of analytical tools for the design of the EG-DFA's optimal frame sizes and to derive performance in closed form.

CHAPTER 5

LIFETIME MAXIMIZATION FOR WIRELESS NETWORKS WITH HYBRID ENERGY STORAGE SYSTEMS

5.1 Introduction

As discussed in the previous chapters, energy harvesting (EH) technologies potentially enable perpetual operations of electronic devices without requiring maintenance for battery substitution. However, the inherent unpredictability regarding the presence (or the dynamics) of ambient energy sources might lead devices that are exclusively powered via EH to operate discontinuously due to temporary energy shortages. Therefore, for those applications that do not tolerate temporary energy shortages, the use of EH as the unique source of energy is generally not sufficient to guarantee the required level of activity. To overcome this problem, a promising solution is to equip the devices with a *hybrid energy storage system* (HESS). A HESS typically includes a *battery*, either rechargeable or not, which operates as the primary energy source, and a *capacitor*, which is recharged via EH and is intended to provide support to the battery, thus extending its lifetime [62].

As a specific instance of a system operated by devices equipped with a HESS, this chapter considers a single-hop wireless network, in which a central controller, referred to as *fusion center* (FC), periodically collects information from M nodes distributed in its surrounding as shown in Figure 5.1. Time is slotted, and, in each slot, the FC schedules $K \leq M$ nodes for transmission over K orthogonal transmission resources (e.g., frequencies), as shown in Figure 5.2. Each node is equipped with a HESS composed by a non-rechargeable battery and a capacitor charged via EH. In particular, this chapter considers the design of scheduling policies such that K nodes in each slot are selected, with the aim of maximizing the *network lifetime* (to be

rigorously defined below). It is emphasized that the number K of nodes scheduled in each slot is specified by the considered application and is thus assumed to be fixed and given.

The scheduling problem in this chapter is tackled in two different scenarios: *i) full state information*, in which the FC knows the state of the HESSs, i.e., the states of the capacitors and of the batteries, of all the nodes at any *decision epoch* (i.e., the beginning of each slot); *ii) partial state information*, in which the FC does not have direct access to the instantaneous state of the HESSs, but it only knows the statistical properties of the EH and leakage processes and the outcomes of previous scheduling commands. Note that, unlike the setting with partial state information, the set-up with full state information requires overhead uplink transmissions in order for each node to transmit the state of its HESS to the FC before each decision epoch.
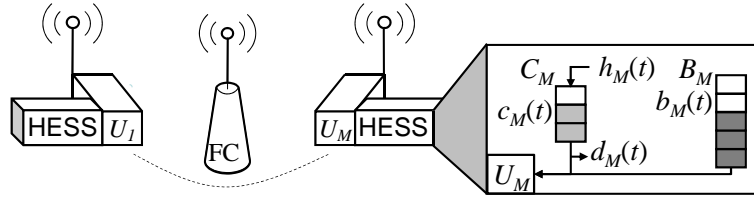
Figure 5.1 Wireless network with a single fusion center (FC) that collects packets from a set of M nodes equipped with a hybrid energy storage system (HESS). Any ith node U_i is equipped with a battery B_i and a capacitor C_i that contain energy $b_i(t)$ and $c_i(t)$ at the beginning of slot t, respectively. The energy harvesting (EH) and leakage processes of node U_i at slot t are denoted by $h_i(t)$ and $d_i(t)$, respectively.
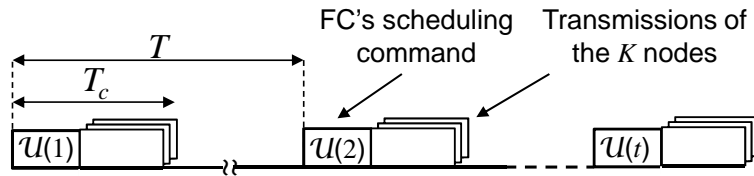
Figure 5.2 Overview of the periodic data collection. Time is organized into slots of duration T each, while the transmission time in each slot (including the scheduling command and transmissions of the nodes) lasts T_c, with $T_c \ll T$. The transmission resources are allocated in each slot by the FC that broadcasts a scheduling command $\mathcal{U}(\cdot)$.

5.1.1 Related Work and Contribution

As introduced above, this chapter considers a centralized scheduling problem for a wireless network with nodes equipped with HESSs, as illustrated in Figure 5.1. To simplify the discussion, it is assumed that the batteries are non-rechargeable and that they do not suffer from energy leakage in the time scale of interest. Instead, the capacitors are rechargeable and can potentially leak energy. This a reasonable approximation of practical scenarios, since non-rechargeable batteries typically suffer from self-discharge over a time-scale much larger than that of capacitors (see, e.g., [63]). It is also assumed that the EH and leakage processes are independent and identically distributed (i.i.d.) across nodes and time-slots, and that they are modeled as binary random processes. That is, in each slot a node either harvests a unit of energy (to be defined below) or not, and similarly for the leakage process.

Main contributions of this chapter. For the *full state information* scenario, considered in Section 5.3, the scheduling problem is formulated as the maximization of the network lifetime, and it is shown to reduce to a stochastic shortest path (SSP) problem [64], which is a special instance of a Markov decision process (MDP). Under the assumption that the system is *symmetric*, so that the statistics of the EH and leakage processes at the nodes are equivalent, optimal scheduling policies are obtained in two limiting cases: *a) harvesting-only* model, in which the amount of energy leaked is negligible; and *b) leakage-only* model, in which the amount of energy harvested is negligible. As it will be discussed, these two limiting scenarios are useful approximations of situations in which the capacitors tend to be close to full or close to empty most of the time, respectively.

An optimal policy in the harvesting-only scenario is shown to select in each slot the K nodes with the largest energy stored in their capacitors (when available). Instead, for the leakage-only model, the optimal policy selects in each slot the K nodes with the smallest non-zero energy in their capacitors. An easily computable

performance upper bound on the network lifetime is also proposed for the general scenario, which can be used as a performance reference when the size of the network makes the numerical computation of the optimal policy intractable. It is then shown that, when the FC schedules only one node in each slot (i.e., $K = 1$), the computation of the network lifetime can be decomposed into separate contributions due to batteries and due to capacitors, and, based on this result, an algorithm that enables the computation of the network lifetime with reduced complexity is proposed.

In the *partial state information* case, considered in Section 5.4, finding the optimal scheduling policies explicitly is more challenging than in the full state information case, and the numerical computation of optimal policies is generally intractable. Therefore, based on the insights obtained from the analysis of the full state information scenario, two heuristic policies that can be easily implemented in practical systems are proposed. Moreover, to improve on these policies, opportunistic feedback schemes are considered, in which each node with a sufficiently large energy in its capacitor opportunistically provides additional information to the FC over a dedicated transmission resource. It is then shown in the numerical results in Section 5.5 that this limited-feedback approach has the potential to greatly improve the lifetime performance.

Related work: The lifetime of *battery-powered* wireless networks was studied in [65], where the problem of scheduling a subset of battery-powered nodes in a wireless network, subject to fading channels, was tackled by resorting to a SSP formulation. In [65] the nodes are equipped with non-rechargeable batteries (i.e., they have no EH capabilities), while the transmission power of each node is adapted to the channel quality in each slot. A similar system setting is considered in [66], where the emphasis is instead on the development of distributed access protocols based on the channel state information and the residual energy information at each node. The work in this chapter differs from [65, 66] in that the energy availability at

the nodes keeps changing even when nodes are not scheduled due to EH and energy leakage. However, the impact of fading is not considered here. A relevant reference for HESS systems is [62] (see also references therein), where the problem of routing in wireless networks operated by nodes equipped by a HESS is considered. Reference [62] also provides a review of the properties of batteries and capacitors and their trade-offs.

5.2   System Model

This chapter considers a wireless network in which a FC is tasked with collecting data packets from a set of M nodes, labeled as $U_1, U_2, ..., U_M$, under the constraint that K packets must be collected in each time-slot (see Figure 5.1 and Figure 5.2). To this end, in each slot t, of duration T, a subset $\mathcal{U}(t) \subseteq \{U_1, ..., U_M\}$ of $|\mathcal{U}(t)| = K$ nodes is selected for transmission. Each node has a new packet to transmit at each slot. It is assumed that nodes' transmissions take place over orthogonal communication resources (e.g., frequencies) so that they do not interfere with each other. It is also assumed that the total duration of the communication between the nodes and the FC in each slot is fixed and equal to T_c (see Figure 5.2), where T_c is generally assumed to be much smaller than the slot duration T, i.e., $T_c \ll T$ for reasons that will be clarified below (see Figure 5.2). Moreover, the FC's scheduling commands and the nodes' packets are considered to be received without error by the intended destinations.

5.2.1   HESS Model

Each node is powered by a HESS, which is composed by a non-rechargeable battery and a capacitor that is charged via EH. It is assumed that each transmission consumes a given energy amount, referred to as *energy unit*, which is normalized to one for simplicity. This energy can be drawn by a node either from the capacitor or from the battery. No energy is consumed by non-scheduled nodes, except for possible energy

leakages. Batteries and capacitors are assumed to be finite and have capacities $1 \leq E_b < \infty$ and $1 \leq E_c < \infty$ energy units, respectively. The capacitor and the battery of node U_i are denoted by C_i and B_i, respectively, while their energies at the beginning of any slot t are denoted by $c_i(t) \in \{0, ..., E_c\}$ and $b_i(t) \in \{0, ..., E_b\}$, respectively. Let $(c_i(t), b_i(t)) \in \mathcal{S}$ be the state of node U_i at slot t, where $\mathcal{S} = \{0, ..., E_c\} \times \{0, ..., E_b\}$ is the single-node state space. Finally, let $(\mathbf{c}(t), \mathbf{b}(t)) \in \mathcal{S}^M$ be the *system state* at slot t, where $\mathbf{c}(t) = [c_1(t), ..., c_M(t)]$ and $\mathbf{b}(t) = [b_1(t), ..., b_M(t)]$ are two $(1 \times M)$ vectors.

5.2.2   Energy Harvesting and Leakage Models

The energy harvested at node U_i during any slot t is modeled as a binary random variable, which is denoted by $h_i(t) \in \{0, 1\}$. This random variable has probability mass function (pmf) $\Pr[h_i(t) = 1] = p_h$ and $\Pr[h_i(t) = 0] = 1 - p_h$. In other words, the node harvests one energy unit in each slot with probability p_h. This harvested energy unit can be stored by node U_i only when its capacitor C_i is not full (i.e., if $c_i(t) < E_c$), while otherwise an *energy overflow* occurs and the harvested energy unit is lost. Similarly, the energy lost by a non-empty capacitor C_i due to leakage is modeled as a binary random variable $d_i(t) \in \{0, 1\}$ with pmf $\Pr[d_i(t) = 1] = p_d$ and $\Pr[d_i(t) = 0] = 1 - p_d$. Clearly, only non-empty capacitor can leak energy. Both the EH and leakage processes are independent across nodes and also i.i.d. across slots.

It is assumed that the energy in the capacitor that can be used for transmission by node U_i during the tth slot is given by $c(t)$, which is the energy initially available at the beginning of the tth slot, while the energies $h_i(t)$ and $d_i(t)$ (potentially) harvested and lost during the tth slot do not affect the energy availability during slot t. This is typically a good approximation of reality, especially if the slot duration T is much longer than the transmission duration T_c, as assumed here. The energy $c_i(t)$ in the capacitor C_i thus evolves as

$$c_i(t+1) = \min\left\{ \left((c_i(t) - 1[U_i \in \mathcal{U}(t)])^+ - d_i(t)\right)^+ + h_i(t), \; E_c \right\}, \tag{5.1}$$

where $1[x]$ is an indicator function such that $1[x] = 1$ if event x is true and zero otherwise, while $(x)^+ = x$ if $x \geq 0$ and $(x)^+ = 0$ if $x < 0$. Note that in (5.1), it is assumed that the leakage process $d_i(t)$ affects the energy in the capacitor before the newly harvested energy $h_i(t)$ is added. This assumption does not affect the main results in this chapter and can be easily modified. Moreover, equation (5.1) implicitly assumes that, when any node U_i is scheduled for transmission (i.e., if $1[U_i \in \mathcal{U}(t)] = 1$), it draws energy from its capacitor first, and thus uses the battery only when its capacitor is empty. As it is intuitive, and it will be further argued below, this choice maximizes the lifetime of the network, since the battery cannot be replenished and thus any energy unit drawn from the battery is irreparably lost.

Finally, by recalling that the batteries are non-rechargeable and that do not suffer from energy leakage, the state of the battery B_i is updated across each slot t as

$$b_i(t+1) = (b_i(t) - 1[U_i \in \mathcal{U}(t) \text{ and } c_i(t) = 0])^+. \tag{5.2}$$

5.3 Full State Information Scenario

This section considers the full state information scenario, in which the FC has perfect knowledge of the state of the HESS of each node at the each slot. The section starts by introducing some useful definitions and the problem formulation, and it continues by proposing two scheduling policies that are proved to be optimal for the harvesting-only and leakage-only scenarios (to be defined exactly below), respectively.

5.3.1 Preliminary Definitions

Definition 1. A state $(\mathbf{c}, \mathbf{b})$ is terminal if at least one node has an empty battery. The set $\mathcal{T}$ of terminal states is thus defined as

$$\mathcal{T} = \left\{ (\mathbf{c}, \mathbf{b}) \in \mathcal{S}^M : \sum_{i=1}^{M} 1[b_i = 0] \geq 1 \right\}. \tag{5.3}$$

As formalized below, the network is assumed to be active for as long as the state $(\mathbf{c}, \mathbf{b})$ does not enter the set $\mathcal{T}$ [1].

Definition 2. I) *Scheduling policy:* A scheduling policy $\pi = \{\pi_1, \pi_2, ...\}$ is a sequence of functions π_t, for $t \in \{1, 2, ...\}$, that map the history $\mathcal{H}(t)$ of the system states up to slot t, with $\mathcal{H}(t) = \{(\mathbf{c}(1), \mathbf{b}(1)), ..., (\mathbf{c}(t-1), \mathbf{b}(t-1))\}$, into a scheduling decision $\mathcal{U}^\pi(t) \in \{U_1, ..., U_M\}$. $\mathcal{U}^\pi(t)$ is the set of nodes that are scheduled for transmission in slot t under policy π, with the constraint $|\mathcal{U}^\pi(t)| = K$, where the notation $|\mathcal{A}|$ indicates the cardinality of the set $\mathcal{A}$.

II) *Stationary policy:* A policy π is said to be *stationary* if the mapping π_t is independent of t and instead depends only on the current system state $(\mathbf{c}(t), \mathbf{b}(t))$. A stationary policy can be thus characterized by a function $\mathcal{U}^\pi(\mathbf{c}, \mathbf{b}) \in \{U_1, ..., U_M\}$.

III) *Proper policy:* A stationary policy is said to be *proper* if the system reaches a terminal state $(\mathbf{c}(t), \mathbf{b}(t)) \in \mathcal{T}$ with probability one, regardless of the initial state $(\mathbf{c}(1), \mathbf{b}(1))$, that is, if $\lim_{t \to \infty} \Pr[(\mathbf{c}(t), \mathbf{b}(t)) \in \mathcal{T} | \mathbf{c}(1), \mathbf{b}(1)] = 1$ for all initial states $(\mathbf{c}(1), \mathbf{b}(1))$ [64].

The network lifetime is defined as follows.

Definition 3. Let $(\mathbf{c}(1), \mathbf{b}(1)) \in \mathcal{S}^M$ be the initial state at slot $t = 1$. The network lifetime $L^\pi(\mathbf{c}(1), \mathbf{b}(1))$ under a scheduling policy π is the average number of slots in which the network is active before the terminal set $\mathcal{T}$ is entered:

$$L^\pi(\mathbf{c}(1), \mathbf{b}(1)) = \lim_{T \to \infty} E^\pi \left[\sum_{t=1}^{T} 1\left[(\mathbf{c}(t), \mathbf{b}(t)) \notin \mathcal{T}\right] \bigg| \mathbf{c}(1), \mathbf{b}(1) \right]. \qquad (5.4)$$

It will be shown in Lemma 4 that the limit (5.4) always exists and it is also finite when the harvesting probability is strictly smaller than one, i.e., $p_h < 1$.

---

[1]Other definitions of the set $\mathcal{T}$ of terminal states can be considered as well. For instance, a relevant setting is one in which the terminal set $\mathcal{T}$ includes only the states in which *all* of the batteries are empty. This can be assumed without requiring substantial modifications to the derivations in this chapter.

The optimization goal is the maximization of the lifetime $L^\pi\left(\mathbf{c}(1), \mathbf{b}(1)\right)$ over the set of all (not necessarily stationary) policies π, where the optimal network lifetime and the corresponding optimal policy are given, respectively, by

$$L^*(\mathbf{c}(1), \mathbf{b}(1)) = \max_\pi L^\pi\left(\mathbf{c}(1), \mathbf{b}(1)\right) \tag{5.5}$$

$$\text{and} \pi^*(\mathbf{c}(1), \mathbf{b}(1)) = \operatorname*{argmax}_\pi L^\pi(\mathbf{c}(1), \mathbf{b}(1)). \tag{5.6}$$

5.3.2 Controlled Markov Process Formulation

According to the model described in Section 5.2, the state $(c_i(t), b_i(t))$ of any node U_i evolves as a controlled Markov chain. Specifically, the transition probabilities for the state $(c_i(t), b_i(t))$ of node U_i are obtained as follows (see Figure 5.3). Let $p_{kj}^{(0)} = \Pr\left[c_i(t+1) = j \middle| c_i(t) = k, U_i \notin \mathcal{U}(t)\right]$ and $p_{kj}^{(1)} = \Pr\left[c_i(t+1) = j \middle| c_i(t) = k, U_i \in \mathcal{U}(t)\right]$ be the probability that energy stored in the capacitor C_i at slot $t+1$ is $c_i(t+1) = j$, given that the energy at slot t is $c_i(t) = k$ and that node U_i is either not scheduled (i.e., $U_i \notin \mathcal{U}(t)$) or scheduled (i.e., $U_i \in \mathcal{U}(t)$), respectively. To simplify the notation, let

$$\lambda = p_h\left(1 - p_d\right) \text{ and } \mu = p_d\left(1 - p_h\right). \tag{5.7}$$

The transition probabilities for the energy in the capacitor can be easily calculated as follows. For the case $U_i \notin \mathcal{U}(t)$, it results $p_{00}^{(0)} = 1 - p_h$ and $p_{01}^{(0)} = p_h$ (recall that empty capacitors do not lose energy, see Section 5.2.2); for $1 \leq k < E_c$, it is possible to write $p_{kk+1}^{(0)} = \lambda$, $p_{kk}^{(0)} = 1 - \lambda - \mu$ and $p_{kk-1}^{(0)} = \mu$; and finally, it follows that $p_{E_c E_c}^{(0)} = 1 - \mu$ and $p_{E_c E_c-1}^{(0)} = \mu$. Instead, for the case $U_i \in \mathcal{U}(t)$ one has $p_{00}^{(1)} = 1 - p_h$ and $p_{01}^{(1)} = p_h$, and $p_{10}^{(1)} = 1 - p_h$ and $p_{11}^{(1)} = p_h$; and for $2 \leq k \leq E_c$, one can write $p_{kk-2}^{(1)} = \mu$, $p_{kk-1}^{(1)} = 1 - \mu - \lambda$, $p_{kk}^{(1)} = \lambda$. For the battery, the energy $b_i(t)$ of a scheduled node $U_i \in \mathcal{U}(t)$, is decremented by one energy unit only when the capacitor is empty, so that $b_i(t+1) = b_i(t) - 1\left[U_i \in \mathcal{U}(t) \text{ and } c_i(t) = 0\right]$.

Based on the above and on the independence of the harvesting and leakage processes at different nodes, the probability that the next overall system state is $(\mathbf{c}(t+1), \mathbf{b}(t+1)) = (\mathbf{c}', \mathbf{b}')$, given the current state $(\mathbf{c}(t), \mathbf{b}(t)) = (\mathbf{c}, \mathbf{b})$ and the scheduling decision $\mathcal{U}(t) = \mathcal{U}$, is

$$\Pr[(\mathbf{c}(t+1), \mathbf{b}(t+1)) = (\mathbf{c}', \mathbf{b}') \,|\, (\mathbf{c}(t), \mathbf{b}(t)) = (\mathbf{c}, \mathbf{b}), \mathcal{U}(t) = \mathcal{U}] =$$

$$\begin{cases} \prod_{U_i \in \mathcal{U}} p^{(1)}_{c_i c'_i} \prod_{U_i \notin \mathcal{U}} p^{(0)}_{c_i c'_i} & \text{if } b'_i = b_i - 1 \,[U_i \in \mathcal{U} \text{ and } c_i = 0]\,, \text{ for all } i \in \{1, ..., M\} \\ 0 & \text{otherwise} \end{cases}.$$

$$(5.8)$$

The following lemma is instrumental in deriving the properties of the optimal policies
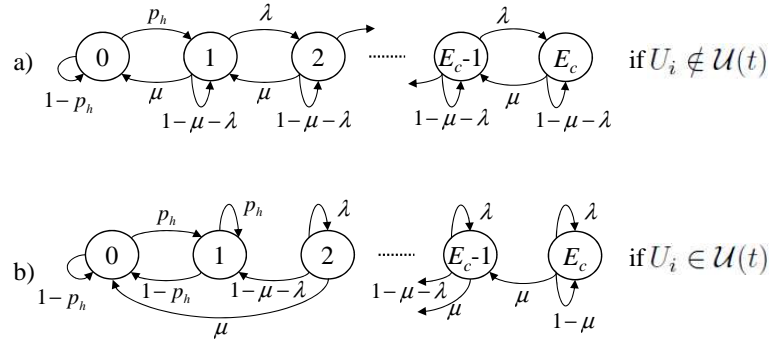
Figure 5.3 Markov chains that describe the evolution of the energy in the capacitor C_i of node U_i when U_i is: b) not scheduled ($U_i \notin \mathcal{U}(t)$) b) scheduled ($U_i \in \mathcal{U}(t)$).

for problem (5.5)-(5.6).

Lemma 4. I) *Finite lifetime:* If the probability of harvesting p_h is strictly smaller than one, i.e., $p_h < 1$, then all the policies π are proper and the maximum network lifetime is finite.

II) *Optimality of stationary policies:* There exists an optimal stationary policy π^* (5.6).

Proof. Part **I)** As seen above, for any fixed policy π, the system state $(\mathbf{c}(t), \mathbf{b}(t))$ evolves as a finite-state Markov chain over the state space $\mathcal{S}^M$, which contains some

absorbing states. It is not hard to see that, regardless of the action taken in any non-terminal state $s_i = (\mathbf{c}(t), \mathbf{b}(t)) \notin \mathcal{T}$, there always exists a path over the space $\mathcal{S}^{\mathcal{M}}$ that connects (with non-zero probability) the state s_i to a terminal state $s_j \in \mathcal{T}$. For a finite-state Markov chain this is sufficient to conclude that the average time before absorption is finite and so is the network lifetime [47]. Therefore, all the policies are proper according to Definition 2. Part **II)** The existence of an optimal stationary policy when all the policies are proper is a well known result (see e.g., [64]). $\qquad\square$

5.3.3 Dynamic Programming Equations

To simplify the analysis below, this section introduces a dynamic programming formulation of problem (5.5)-(5.6). Specifically, the lifetime of any stationary policy π, when the initial (non-terminal) state is $(\mathbf{c}, \mathbf{b})$, can be calculated by solving the following *Bellman equation*

$$L^{\pi}(\mathbf{c}, \mathbf{b}) = 1 + \sum_{(\mathbf{c}', \mathbf{b}')} L^{\pi}(\mathbf{c}', \mathbf{b}') \Pr\left[\mathbf{c}', \mathbf{b}' | (\mathbf{c}, \mathbf{b}), \mathcal{U}^{\pi}(\mathbf{c}, \mathbf{b})\right], \tag{5.9}$$

where the distribution $\Pr\left[\mathbf{c}', \mathbf{b}' | \mathbf{c}, \mathbf{b}, \mathcal{U}^{\pi}(\mathbf{c}, \mathbf{b})\right]$ is given by (5.8) (the time-dependence here is dropped since only stationary policies are considered). Note that, when the initial state is terminal, i.e., $(\mathbf{c}, \mathbf{b}) \in \mathcal{T}$, it follows $L^{\pi}(\mathbf{c}, \mathbf{b}) = 0$ for any π. Moreover, any optimal stationary policy π^*, with optimal actions $\mathcal{U}^*(\mathbf{c}, \mathbf{b})$, satisfies the following *optimality equations*

$$L^*(\mathbf{c}, \mathbf{b}) = \max_{\mathcal{U} \in \{U_1, \ldots, U_M\}} L(\mathbf{c}, \mathbf{b} | \mathcal{U}), \tag{5.10}$$

$$\text{where } \mathcal{U}^*(\mathbf{c}, \mathbf{b}) = \operatorname*{argmax}_{\mathcal{U} \in \{U_1, \ldots, U_M\}} L(\mathbf{c}, \mathbf{b} | \mathcal{U}), \tag{5.11}$$

where

$$L(\mathbf{c}, \mathbf{b} | \mathcal{U}) = 1 + \sum_{(\mathbf{c}', \mathbf{b}')} L^*(\mathbf{c}', \mathbf{b}') \Pr\left[\mathbf{c}', \mathbf{b}' | (\mathbf{c}, \mathbf{b}), \mathcal{U}\right] \tag{5.12}$$

is defined as the lifetime of a policy that selects action $\mathcal{U} \in \{U_1, ..., U_M\}$ at the current slot and then proceeds optimally from the next slot onward.

The following lemma introduces some useful properties of the optimal lifetime $L^*(\cdot)$.

Lemma 5. I) *Monotonicity:* $L^*(\mathbf{c}, \mathbf{b}) \geq L^*(\mathbf{c}', \mathbf{b}')$ if $c_i \geq c_i'$ and $b_i \geq b_i'$ for all $i \in \{1, ..., M\}$.

II) *Dependence on total battery energy:* $L^*(\mathbf{c}, \mathbf{b}) = L^*(\mathbf{c}, \mathbf{b}')$ for all $\mathbf{b}$ and $\mathbf{b}'$ such that $\sum_{i=1}^{M} b_i = \sum_{i=1}^{M} b_i'$, when the states $(\mathbf{c}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{b}')$ are not terminal.

III) *Symmetry:* $L^*(\mathbf{c}, \mathbf{b}) = L^*(\mathcal{P}(\mathbf{c}), \mathbf{b})$ for any permutations $\mathcal{P}$ of the node indices.

IV) *Schedule capacitor first:* Any optimal stationary policy always schedules nodes with a non-empty capacitor first, while batteries are used only when there are less than K nodes with a non-empty capacitor.

Proof. The proofs of parts **I)** **II)** and **III)** are omitted since they are trivial consequences of the symmetry of the nodes. The formal proof of part **IV)** is omitted for brevity and follows by the same techniques used below in Section 5.6. ☐

5.3.4 Optimal Scheduling Policies

This section proposes two stationary scheduling policies for the full state information scenario, which are referred to as *Most Charged capacitor First* (MCF) and *Least Charged capacitor First* (LCF). It is then shown that these policies are optimal for problem (5.5)-(5.6) in the harvesting-only (i.e., $p_d = 0$) and leakage-only (i.e., $p_d = 0$) scenarios, respectively. The optimal policy for the general scenario is instead addressed numerically in Section 5.5 by using standard dynamic programming tools.

Definition 6. The MCF policy π^{MCF}, with lifetime $L^{MCF}(\cdot)$, schedules K nodes in each slot according to the following priority rules: 1) nodes with the largest energy in

the capacitors; 2) nodes with empty capacitors and at least two energy units in their batteries; 3) nodes with empty capacitors and one energy unit in their batteries.

The LCF policy is instead defined as follows.

Definition 7. The LCF policy π^{LCF}, with lifetime $L^{LCF}(\cdot)$, schedules K nodes in each slot according to the following priority rules: 1) nodes with the smallest (non-zero) energy in the capacitors; 2) and 3) as for the MCF policy above.

The optimality of the LCF and MCF policies above is summarized in the following proposition.

Proposition 8. a) The MCF policy π^{MCF} is optimal for problem (5.5)-(5.6) when $p_d = 0$; **b)** The LCF policy π^{LCF} is optimal for problem (5.5)-(5.6) when $p_h = 0$. Namely

$$L^{MCF}(\mathbf{c}, \mathbf{b}) = L^*(\mathbf{c}, \mathbf{b}), \ \text{ for } p_d = 0 \tag{5.13}$$

$$L^{LCF}(\mathbf{c}, \mathbf{b}) = L^*(\mathbf{c}, \mathbf{b}), \ \text{ for } p_h = 0. \tag{5.14}$$

Proof. See Section 5.6. □

The intuition behind the optimality of the MCF policy in the harvesting-only scenario is that, when the energy in the capacitor is subject to harvesting only, then it is better to schedule nodes that are fully charged in order to reduce the chance of energy overflows. Instead, the optimality of the LCF policy is due to the fact that, when the energy in the capacitor is subject to leakage only, then it is better to concentrate all the energy in a small number of capacitors, so that there are fewer chances that the capacitors lose energy. The full proof of the Proposition 8 is given in Section 5.6.

It is remarked that the two limiting scenarios addressed above represent approximation of the following two practical situations. The harvesting-only case

$(p_d = 0)$ approximates a situations in which the capacitors tend to be full, and thus the main contribution to energy loss is given by energy overflows. Instead, the leakage-only case $(p_h = 0)$ provides an approximation for a system in which the capacitors tend to be non-full most of the time, so that the most prominent source of energy loss is energy leakage rather than energy overflow. These intuitions are confirmed by the numerical results shown in Section 5.5.

The lifetime of both MCF and LCF when only one node is scheduled in each slot (i.e., $K = 1$) can be calculated efficiently as described in Appendix B.

5.4   Partial State Information

This section tackles the scheduling problem of Section 5.2 by assuming that the FC does not have full state information about the instantaneous states of the HESS at the nodes. Specifically, beside the statistical properties of the EH and leakage processes, it is assumed that the FC learns the state of the capacitors only for the scheduled nodes upon reception of their packets. Packets are thus assumed to contain information about the state of the capacitor of the transmitting node. It follows that the observations available at the $(t + 1)$th decision epoch at the FC are given by the state $c_i(t)$ of the capacitor of node U_i of all the nodes that were scheduled at slot t, i.e., for all $U_i \in \mathcal{U}(t)$. Note that the state $c_i(t + 1)$ at slot $t + 1$ is affected also by the energies harvested $h_i(t)$ and lost $d_i(t)$ by node U_i during slot t, whose values are not available at the FC. Lastly, it is assumed that the state of the batteries is perfectly known by the FC, since the latter can be easily tracked by the FC without any additional communication overhead as the batteries are ideal and do not leak energy.

5.4.1 Problem Formulation

When the FC has only partial information about the system state $(\mathbf{c}(t), \mathbf{b}(t))$, the scheduling problem can be formalized as a Partially Observable Markov Decision Process (POMDP) [67]. For such problems, it is well-known that the a posteriori probability distribution of the system state, typically referred to as the *belief*, is a sufficient statistics [67]. Specifically, let $\omega_{i,k}(t)$ be the a posteriori probability that the energy $c_i(t)$ stored in the capacitor C_i at slot t is $c_i(t) = k$, where conditioning is done over the observations accrued by the FC up to time t. The combinations of the vectors of the a posteriori probabilities $\boldsymbol{\omega}_i(t) = [\omega_{i,1}(t), ..., \omega_{i,E_c}(t)]$ and of the known batteries' values $\mathbf{b}(t)$ for all nodes $\{U_1, ..., U_M\}$, can be then used as the system state for the POMDP at hand [67], and the problem can be formalized similarly to Section 5.3. In particular, the network lifetime is still defined as the right hand side of (5.4), where the average is taken over the a posteriori probabilities of the state for any given policy. Note that this definition of lifetime is meaningful since, even though the FC has no direct access to the actual system state $(\mathbf{c}(t), \mathbf{b}(t))$, the first slot t in which at least one node gets its battery depleted can be easily recognized by the FC as discussed above.

Solving a POMDP is notoriously complex [67]. Therefore, the next Section derives two simple heuristic policies that are inspired by the results described above for the full state information case. A simple opportunistic feedback scheme, in which a limited amount of overhead is transmitted by the nodes to the FC, is then considered in Section 5.4.3.

5.4.2 Index-based Heuristic Policies

The first heuristic policy that is proposed in this section is inspired by the fact that, with full state information, it is always optimal that a scheduled node transmit by drawing energy from its capacitor if not empty (see Lemma 5-4)). The proposed

policy, referred to as *Capacitor-Greedy* (CG) is defined as the policy that schedules the K nodes in each slot that have the largest probability of having a non-empty capacitor, i.e.,

$$\mathcal{U}^{CG}(t) = \underset{\mathcal{U}:\,|\mathcal{U}|=K}{\operatorname{argmin}} \sum_{U_i \in \mathcal{U}} \omega_{i,0}(t), \tag{5.15}$$

where $\omega_{i,0}(t) = 1 - \sum_{k=1}^{E_c} \omega_{i,k}(t)$. It is assumed that when two or more nodes have the same probability $\omega_{i,0}(t)$ of having an empty capacitor then the node with the largest energy in the battery is scheduled.

The drawback of the CG policy is that a node with a relatively small probability $\omega_{i,0}(t)$ can be scheduled regardless of the state of its battery. This could produce a lifetime termination even when there are potentially other nodes with a larger energy in the batteries. Therefore, the second proposed policy, which is referred to as *Largest Average total Energy* (LAE) policy schedules in each slot the K nodes with the largest average sum energy stored in their battery and capacitor, i.e.,

$$\mathcal{U}^{LAE}(t) = \underset{\mathcal{U},\,|\mathcal{U}|=K}{\operatorname{argmax}} \sum_{U_i \in \mathcal{U}} \left(b_i(t) + \sum_{j=1}^{E_c} j\omega_{i,j}(t) \right). \tag{5.16}$$

The lifetimes for these two heuristic policies are compared in Section 5.5 with the performance for the full state information scenario.

5.4.3   Partial State Information with Opportunistic Feedback

This section briefly investigates a set-up in which additional feedback from the nodes to the FC is allowed. Specifically, it is assumed that, immediately before the beginning of each slot, any node with an energy in the capacitor greater than or equal to a threshold $\lambda_{FB} \in \{1, ..., E_c\}$ sends one bit of information to the FC. This bit is used to inform the FC about the fact that the transmitting node has an energy in the capacitor larger than the threshold λ_{FB}. The FC collects the received feedback and updates the belief state accordingly (not detailed here).

More specifically, the *one-bit feedback policy* proposed in this section works as follows. If there are at least K nodes with an energy in the capacitor larger than the threshold λ_{FB} (and thus that transmit the feedback to the FC), then the FC selects the K nodes among them with the largest average energy in the capacitor, where the average is calculated based on the current belief. If instead there are only $K_1 < K$ nodes that have energy in the capacitor larger than λ_{FB}, then the FC selects such K_1 nodes, and the remaining $K - K_1$ ones are selected according to the LAE policy defined above. The performance of the one-bit feedback policy is investigated numerically in Section 5.5.

5.5   Numerical Results

This section presents extensive numerical results to get insight into the system design and to validate the analytical derivations of the previous sections. The performance criterion that is considered hereafter is the network lifetime (5.4) normalized by the optimal lifetime of a *battery-only* system. The latter is simply given by $\sum_{i=1}^{M} 1[b_i \geq 1] (b_i - 1) + 1$ for an initial battery state $[b_1, b_2, ..., b_M]$. This normalization enables the performance advantages of adopting HESS to be more clearly highlighted.

Numerical results for the full state information scenario. Figure 5.4 shows the normalized lifetime versus the size of the capacitor for the optimal policy, which is evaluated numerically using standard dynamic programming tools, and for the MCF and LCF policies. As a reference, a policy is considered that *randomly* selects K nodes for transmission among the ones with non-empty capacitors, if available, and otherwise follows the steps 2) and 3) of the MCF policy in Definition 6. The system parameters are $M = 5$, $K = 1$, $Mp_h/K = 0.9$, $p_d = 0.1$ and $E_b = 5$. Note that the quantity Mp_h/K captures the ratio between the average cumulative energies harvested and consumed at the M nodes.

In Figure 5.4 it can be seen that, when the capacitors are of capacity one, i.e., $E_c = 1$, the optimal performance can be achieved by any policy that schedules a capacitor first, and thus, as special cases, MCF, LCF and the random policies are all optimal. Instead, when the capacitors get larger, the LCF policy approaches the optimal performance, while the MCF policy does not perform well. This follows from the discussion in Section 5.3.4: when the leakage probability is non-zero (here $p_d = 0.01$) and the ratio Mp/K is smaller than one (here $Mp_h/K = 0.9$), then the capacitors are non-full most of the time, and thus the LCF policy, which reduces the chances of energy loss due to leakage, is almost optimal.

The effects of the energy leakage probability on the network lifetime is instead shown in Figure 5.5 for system parameters $M = 5$, $K = 1$, $Mp_h/K = 0.9$, $E_b = 5$ and $E_c = 6$. It can be seen that for small leakage probability, e.g., $p_d \leq 10^{-3}$, the MCF policy approaches the optimal performance, while the LCF policy does not perform satisfactorily. Instead, as the leakage probability p_d gets larger, the LCF policy approaches the optimal performance.

Numerical results for the partial state information scenario. Figure 5.6 shows the performance of the heuristic policies CG in (5.15) and LAE in (5.16) that are proposed in Section 5.4.2. These heuristic policies are compared with the optimal network lifetime for the full state information case. Note that the numerical computation of the optimal network lifetime for the partial state information scenario is prohibitively complex. Figure 5.6 also shows the performance of the opportunistic feedback scheme described in Section 5.4.3, where the threshold λ_{FB} is optimized for each capacitor size in order to maximize the network lifetime. It turns out that, for the considered systems parameters, the optimal threshold is $\lambda_{FB} = 1$ for all the capacitor values. The other system parameters are $M = 5$, $K = 1$, $Mp_h/K = 0.9$, $p_d = 0.01$ and $E_b = 5$.

As it can be seen in Figure 5.6, when no feedback in considered, the lifetimes of the LAE and CG policies are considerably far from the optimal performance of the full state information scenario. Moreover, as expected, the CG policy performs poorly since it does not account for the batteries. Instead, the presence of opportunistic feedback enables the performance to be significantly improved.

5.6 Proof of Proposition 8

This section provides the main steps of the proof of Proposition 8 regarding the optimality of the MCF policy in the harvesting-only scenario. The remaining details and the optimality of the LCF policy in the leakage-only setting are addressed in Appendix D.

Optimality of the MCF policy in the harvesting-only scenario. The proof leverages sample path arguments of the EH processes and the stochastic symmetry of the nodes. To elaborate, recall from Lemma 4 that, in order to prove the optimality of the MCF policy, it is sufficient to show that the action $\mathcal{U}^{MCF}(\mathbf{c}, \mathbf{b})$ of the MCF policy satisfies (5.11). This amounts to showing that the inequality

$$L(\mathbf{c}, \mathbf{b}|\mathcal{U}^{MCF}(\mathbf{c}, \mathbf{b})) \geq L(\mathbf{c}, \mathbf{b}|\mathcal{U}), \tag{5.17}$$

holds for all $(\mathbf{c}, \mathbf{b}) \in \mathcal{S}^M$ and all actions $\mathcal{U} \subseteq \{U_1, ..., U_M\}$.

Regarding inequality (5.17), it is first noted that, by the definition (5.12), the quantity $L(\mathbf{c}, \mathbf{b}|\mathcal{U})$ is the lifetime of a policy π that selects action $\mathcal{U}$ at slot $t = 1$ (so that $\mathcal{U}^\pi(1) = \mathcal{U}$) and then operates optimally from slot $t = 2$ onward (the time index t is started from one only for reference and it is dropped below when unnecessary). The lifetime of policy π can thus be written as $L(\mathbf{c}, \mathbf{b}|\mathcal{U}) = L^\pi(\mathbf{c}, \mathbf{b})$. The inequality (5.17) can be proved by showing that it is possible to construct an auxiliary policy γ, which acts as the MCF policy at slot $t = 1$, so that: *a)* $\mathcal{U}^\gamma(1) = \mathcal{U}^{MCF}(\mathbf{c}, \mathbf{b})$; *b)*

the lifetime is no smaller than that of π, i.e.,

$$L^\gamma(\mathbf{c}, \mathbf{b}) \geq L^\pi(\mathbf{c}, \mathbf{b}), \tag{5.18}$$

for all $(\mathbf{c}, \mathbf{b}) \in \mathcal{S}^M$ and all $\mathcal{U}^\pi(1) = \mathcal{U} \subseteq \{U_1, ..., U_M\}$. If such a policy γ exists, then the inequality (5.17) is automatically implied since the inequality $L(\mathbf{c}, \mathbf{b}|\mathcal{U}^{MCF}(\mathbf{c}, \mathbf{b})) \geq L^\gamma(\mathbf{c}, \mathbf{b})$ holds by the fact that policy γ is not required to operate optimally from slot $t = 2$ onward. The remaining of this section is devoted to prove inequality (5.18).

Let $c_i{}^\gamma(t)$ be the energy stored in the capacitor of node U_i at the beginning of slot t when policy γ is implemented, and similarly, define $c_i^\pi(t)$ when policy π is followed. Moreover, define the vectors $\mathbf{c}^\gamma(t) = [c_1{}^\gamma(t), ..., c_M{}^\gamma(t)]$ and $\mathbf{c}^\pi(t) = [c_1{}^\pi(t), ..., c_M{}^\pi(t)]$, where $\mathbf{c}^\gamma(1) = \mathbf{c}^\pi(1) = \mathbf{c}(1)$. The following definition is instrumental for the proof.

Definition 9. Let $\Omega^\gamma(t)$, $\Omega^\gamma(t)$ and $\Omega^c(t)$, for $t = \{1, 2, ...\}$, be three sequences of sets defined as follows. For $t = 1$ let

$$\Omega^\gamma(1) = \mathcal{U}^\gamma(1) \setminus \mathcal{U}^\pi(1) \setminus \{U_i \in \mathcal{U}^\gamma(1) \text{ and } U_j \in \mathcal{U}^\pi(1) : c_i(1) = c_j(1)\} \tag{5.19}$$

$$\Omega^\pi(1) = \mathcal{U}^\pi(1) \setminus \mathcal{U}^\gamma(1) \setminus \{U_i \in \mathcal{U}^\gamma(1) \text{ and } U_j \in \mathcal{U}^\pi(1) : c_i(1) = c_j(1)\} \tag{5.20}$$

$$\Omega^c(1) = \{U_1, ..., U_M\} \setminus \Omega^\gamma(1) \setminus \Omega^\pi(1); \tag{5.21}$$

while, for any $t \geq 2$, the set $\Omega^c(t)$ is updated as

$$\Omega^c(t) = \Omega^c(t-1) \cup \{U_k\} \cup \{U_i, U_j\}, \tag{5.22}$$

for all $U_k \in \Omega^\gamma(t-1)$ or $U_k \in \Omega^\pi(t-1)$ such that $c_k^\gamma(t) = c_k^\pi(t)$; and

for all $U_i \in \Omega^\gamma(t-1)$ and $U_j \in \Omega^\pi(t-1)$ such that $c_i^\gamma(t) = c_j^\pi(t)$ and $c_j^\gamma(t) = c_i^\pi(t)$,

while sets $\Omega^\gamma(t)$, $\Omega^\gamma(t)$, for any $t \geq 2$, are updated as

$$\Omega^\gamma(t) = \Omega^\gamma(1) \setminus \Omega^c(t) \tag{5.23}$$

$$\Omega^\pi(t) = \Omega^\pi(1) \setminus \Omega^c(t). \tag{5.24}$$

The set $\Omega^\gamma(1)$ contains all the nodes that are scheduled by policy γ but not by policy π at slot $t = 1$, excluding the nodes $U_i \in \mathcal{U}^\gamma(1)$ for which there exists a node $U_j \in \mathcal{U}^\pi(1)$ with $c_i(1) = c_j(1)$. The set $\Omega^\pi(1)$ is similarly defined for policy π. The *common set* $\Omega^c(1)$ contains all the nodes that are neither in $\Omega^\gamma(1)$ nor in $\Omega^\pi(1)$. This set can be interpreted as consisting of the nodes that, as seen below from the definition of policy γ, evolve in the same fashion under the two policies π and γ, possibly upon an index permutation (see (5.22)). More specifically, for any node $U_i \in \mathcal{U}^c(t)$, either the energy stored in its capacitor is the same under the two policies (i.e., $c_i^\gamma(t) = c_i^\pi(t)$) or it is possible to find another node $U_j \in \mathcal{U}^c(t)$ such that, upon an index permutation between i and j, their capacitors have the same stored energies (i.e., $c_i^\gamma(t) = c_j^\pi(t)$ and $c_j^\gamma(t) = c_i^\pi(t)$). The sets evolve as per (5.22) so that nodes from sets $\Omega^\gamma(t)$ and $\Omega^\pi(t)$ are removed to be added to the common set $\Omega^c(t)$ when appropriate conditions apply. In other word, the set $\Omega^c(1)$ contains the nodes that are equivalent under the two policies.

Using the definitions above, it is possible to define the operations of policy γ from slot $t = 2$ onward. In particular, policy γ selects the same nodes that are scheduled by policy π in all the slots $t \geq 2$ with the following exception, which is referred to as *switch*: When policy π schedules some nodes that are in the set $\Omega^\gamma(t)$, then policy γ schedules the same number of nodes (arbitrarily selected) from the set $\Omega^\pi(t)$. Moreover, at the first slot t in which the set $\Omega^\gamma(t)$ becomes empty, policy γ keeps selecting the same nodes as for policy π with no more exceptions.

Having defined policy γ, the proof is turned to show that inequality (5.18) holds. The key idea of the proof is the following. At the first slot t in which at least one of the two sets $\Omega^\gamma(\cdot)$ and $\Omega^\pi(\cdot)$ becomes empty three scenarios can occur: $i)$ $|\Omega^\gamma(t)| = |\Omega^\pi(t)| = 0$; $ii)$ $|\Omega^\pi(t)| > |\Omega^\gamma(t)| = 0$; and $iii)$ $|\Omega^\gamma(t)| > |\Omega^\pi(t)| = 0$. If case $i)$ occurs, then the two policies γ and π have the same lifetime, since all the

capacitors in the system are equivalent. If case $ii)$ occurs, policy γ starts selecting the same nodes of policy π with no switch exceptions. However, as proved in Appendix D, it holds that $c_i^\gamma(t) > c_i^\pi(t)$ for all $U_i \in \Omega^\pi(t)$, while all the other nodes are equivalent under the two policies, and thus policy γ has a lifetime no smaller than policy π. Finally, based on sample path arguments of the EH process, it is shown in Appendix D that case $iii)$ can never occur. This is sufficient to conclude that policy γ has a lifetime no smaller than policy π, which implies that inequality (5.18) holds, thus completing the proof.
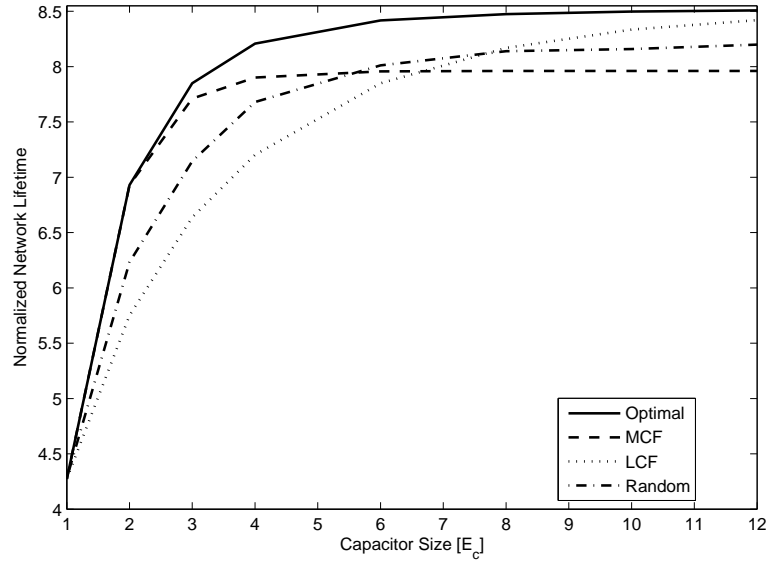
Figure 5.4 Normalized lifetime (5.4) versus the capacitor size E_c for the full state information scenario. The system parameters are $M = 5$, $K = 1$, $Mp_h/K = 0.9$, $p_d = 0.01$ and $E_b = 5$.

5.7 Conclusions

This chapters considered the design of centralized scheduling policies for a single-hop wireless network, in which a fusion center (FC) collects data packet periodically from a set of nodes powered via hybrid energy storage systems (HESSs). Each HESS is composed by a non-rechargeable battery and a capacitor that is recharged via energy harvesting (EH) and that is subject to energy leakage. The capacitors are aimed to
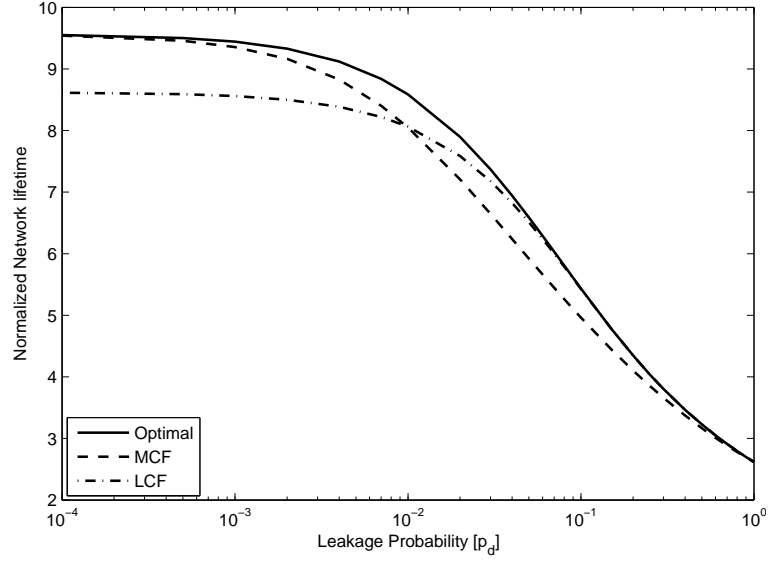
Figure 5.5 Normalized lifetime (5.4) versus the leakage probability p_d. The system parameters are $M = 5$, $K = 1$, $Mp_h/K = 0.9$, $E_b = 5$ and $E_b = 6$.
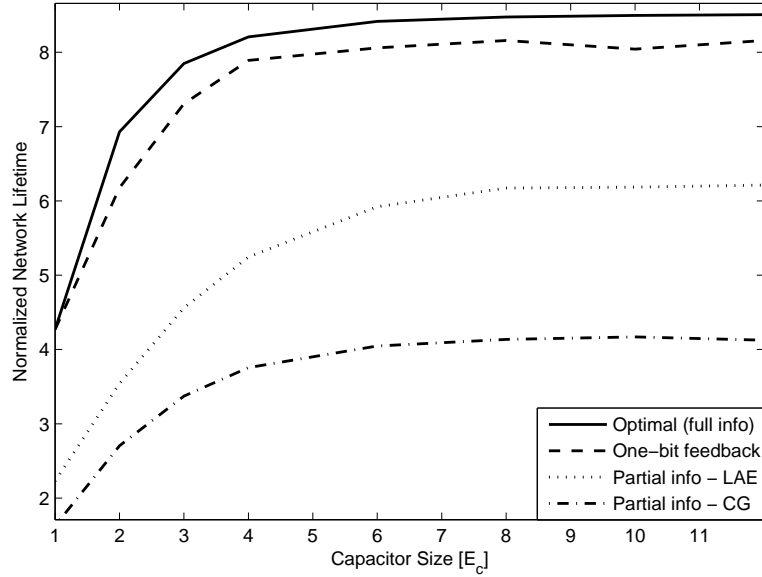
Figure 5.6 Normalized lifetime (5.4) versus the capacitor size for the partial state information scenario. The system parameters are $M = 5$, $K = 1$, $Mp_h/K = 0.9$, $p_d = 0.01$, $E_b = 5$ and $E_c = 6$.

support the batteries and thus to extend their lifetime. The FC's scheduling policies are thus designed with the aim of maximizing the network lifetime.

In particular, the scheduling policies are designed with different level of information available at the FC. When the FC knows the instantaneous energy availability at the HESS of each node (or full state information) then two optimal policies for two limiting case are derived explicitly. Specifically, when the energy leakage is negligible (or harvesting-only case), a policy that selects in each slot the nodes with the largest energy in the capacitors, referred to as most charged capacitor first (MCF), is shown to be optimal. When the energy harvesting is negligible instead (or leakage-only case), it is shown that a policy that selects the nodes with the smallest non-zero energy in the capacitors, referred to as least charged capacitor first (LCF), is optimal.

For the scenario in which the FC does not know the instantaneous state of the HESSs of the nodes (or partial state information case), then heuristic scheduling policies that take decision based only on the statistical properties of the EH and energy leakage processes are proposed. Furthermore, opportunistic feedback schemes, where the nodes with a sufficiently large amount of energy in the capacitor transmit a bit of information to the FC, are considered as well and shown to approach the performance of the full state information scenario.

Overall, it has been shown that a careful design of the scheduling policies can improve the network lifetime remarkably.

CHAPTER 6

OPTIMALITY OF MYOPIC SCHEDULING AND WHITTLE INDEXABILITY FOR ENERGY HARVESTING NODES

6.1   Introduction and System Model

This chapter considers a single-hop WSN, where a central node, referred to as fusion center (FC), collects data from a set of M energy harvesting (EH) nodes, labeled as $U_1, ..., U_M$, deployed in its surrounding as shown in Figure 6.1. Each node, is equipped with an energy harvesting unit (EHU) that converts a given source of energy available in the environment into electrical energy. The harvested energy is stored in a rechargeable battery (or a capacitor). The battery is also subject to energy leakage (or self-discharge). Nodes perform the monitoring of a given phenomenon of interest and the task of the FC is to collect as many measurements (packets) as possible. To this end, in each time slot the FC schedules transmission of a subset $\mathcal{U}(t) \subseteq \{U_1, ..., U_M\}$ of K nodes, where each of the K scheduled nodes is allocated an orthogonal transmission resource, e.g., frequency. The problem is thus similar to the one considered in Chapter 5 with the difference that in this chapter the nodes are equipped with a single replenishable energy storage device.

One of the main challenge to be addressed when designing a scheduling algorithm for EH-nodes is that the energy availability at the nodes keeps changing due to EH and energy leakage. Therefore, in general, the FC cannot be aware of the exact energy level in the battery of each node anytime, unless dedicated transmission resources (overhead) are arranged with the aim of collecting information regarding the energy availability at the nodes (see also the discussion in Chapter 5). It is worth mentioning that nodes exclusively powered via EH are potentially subject to temporary energy shortages. Hence, when a node is scheduled while in energy

shortage, then the allocated transmission resource remains unused for the entire slot (i.e., it is wasted).

In this chapter the the focus is on the design of scheduling policies that do not require any communication overhead from the nodes. In particular, it is assumed that the FC can only take scheduling decisions based on the outcomes of previous node transmissions and on the knowledge of the stochastic properties of the EH and leakage processes. The scheduling policies will thus be designed with the aim of maximizing the average *throughput*, i.e., the number of packets collected in each slot within a given amount of slots (*horizon*), as it will be formally defined below.
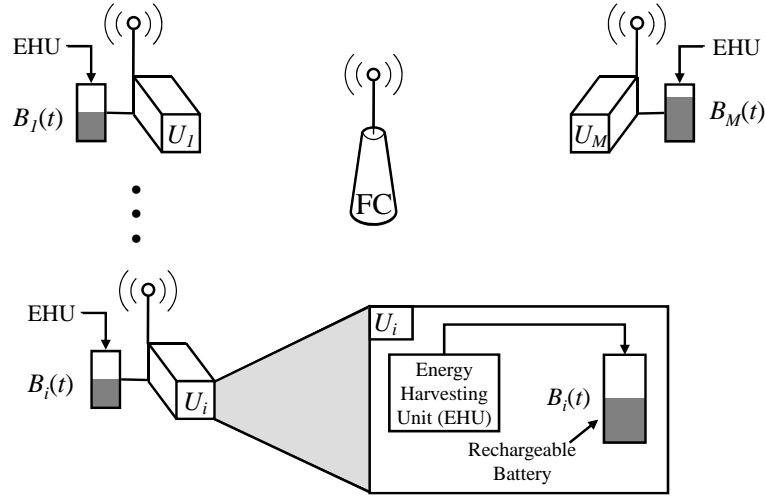
Figure 6.1 A WSN where a fusion center (FC) collects data from M energy-harvesting (EH) nodes. Each node U_i is equipped with a rechargeable battery with energy $B_i(t)$ at time-slot t.

6.1.1 Markov Formulation

To elaborate on the contribution of this chapter and on related works, this section introduces the model that is considered throughout this chapter for the evolution of the energy in the battery. The energy in the battery is considered discrete, where the granularity of the discrete model is referred to as *energy unit*. Let $B_i(t) \in \{0, ..., C\}$ be the number of energy units stored in the battery of node U_i at slot t, for $i \in \{1, ..., M\}$,

where C is the capacity of the battery and where the energy unit is normalized to one for simplicity. An energy unit is consumed for the transmission of a data packet. As explained above, the scheduling decisions consist in the assignment at each slot t of the K communication resources to a subset $\mathcal{U}(t) \subseteq \{U_1, ..., U_M\}$ of K nodes, with $|\mathcal{U}(t)| = K$, where the operator $|\mathcal{A}|$ indicates the cardinality of set $\mathcal{A}$.
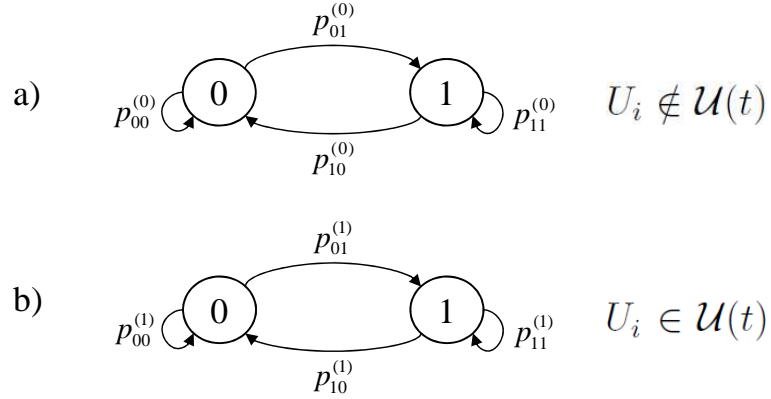
Figure 6.2 Markov model for the evolution of the state of the battery $B_i(t) \in \{0, 1\}$, of capacity $C = 1$, when the node U_i: a) is not scheduled in slot t (i.e., $U_i \notin \mathcal{U}(t)$); b) is scheduled in slot t (i.e., $U_i \in \mathcal{U}(t)$).

The evolution of the battery $B_i(t)$ given the scheduling decision $\mathcal{U}(t)$ is illustrated in Figure 6.2 for the case of capacity $C = 1$. At each slot, node U_i can be either scheduled ($U_i \in \mathcal{U}(t)$) or not ($U_i \notin \mathcal{U}(t)$). If U_i is not scheduled (i.e., $U_i \notin \mathcal{U}(t)$, see Figure 6.2-a)) and an energy unit is stored in its battery (i.e., $B_i(t) = 1$), then the battery will be empty in the next slot with probability (w.p.) $p_{10}^{(0)} = \Pr[B_i(t+1) = 0 | B_i(t) = 1, U_i \notin \mathcal{U}(t)]$, while it will remain full w.p. $p_{11}^{(0)} = 1 - p_{10}^{(0)}$ (this accounts for possible self-discharge of the battery). Instead, if node U_i is scheduled (i.e., $U_i \in \mathcal{U}(t)$, see Figure 6.2-b)) and $B_i(t) = 1$, the node transmits successfully and its battery in the next slot will be either empty or full w.p. $p_{10}^{(1)} = \Pr[B_i(t+1) = 0 | B_i(t) = 1, U_i \in \mathcal{U}(t)]$ and $p_{11}^{(1)} = 1 - p_{10}^{(1)}$, respectively. Probability $p_{11}^{(1)}$ accounts for the possible arrival of a new energy unit due to EH. If $B_i(t) = 0$ the probabilities of receiving an energy unit due to EH when U_i is

not scheduled and scheduled are $p_{01}^{(0)} = \Pr[B_i(t+1) = 1 | B_i(t) = 0, U_i \notin \mathcal{U}(t)]$ and $p_{01}^{(1)} = \Pr[B_i(t+1) = 1 | B_i(t) = 0, U_i \in \mathcal{U}(t)]$, respectively, while the probabilities of not receiving any energy unit are $p_{00}^{(0)} = 1 - p_{01}^{(0)}$ and $p_{00}^{(1)} = 1 - p_{01}^{(1)}$, respectively.

Given the model in Figure 6.2, the FC aims at scheduling a subset $\mathcal{U}(t)$ of nodes so as to maximize the throughput. The scheduling problem can thus be formalized as a partially observable Markov decision process (POMDP) [68]. In fact, the current state of the system, which amounts to the number of energy units $B_1(t), ..., B_M(t)$ in the batteries, is not directly accessible by the FC, which in turns is only aware of the transitions probabilities $p_{xy}^{(u)}$, with $x, y, u \in \{0, 1\}$ (the same for all the nodes), and of the outcomes of previously scheduled transmissions. This is unlike standard Markov decision processes (MDP) where the FC has full access to the system state [69]. By following standard steps to be discussed below, the scheduling problem discussed above can be cast into the framework of restless multi-armed bandit (RMAB) problems [70]. In a RMAB problem, there are M independent controlled Markov chains referred to as *arm*s (i.e., nodes in the formulation in this chapter). Any arm that is selected by the FC provides a reward that depends only on its current state. The goal of the FC is to select, at each slot, K out of the M arms, so as to maximize an average reward criterion. Note that the formulation of the problem as a RMAB does not solve by itself the complexity issue of POMDPs [71], as finding solutions of general RMABs is known to be prohibitively complex as well [72].

6.1.2 Related Work and Contributions

In this work, the scheduling problem above is tackled by assuming that the transition probabilities of the Markov chains in Figure 6.2, the number of nodes M and communication resources K are such that

$$M = Km, \text{ for } m \text{ integer, and} \tag{6.1a}$$

$$p_{11}^{(1)} \le p_{01}^{(1)} \le p_{01}^{(0)} \le p_{11}^{(0)}, \tag{6.1b}$$

Condition (6.1a) states that the number of nodes is proportional to the number of communication resources, generalizing the single-resource ($K = 1$) case. Conditions (6.1b) are motivated as follows. Inequality $p_{11}^{(1)} \leq p_{01}^{(1)}$ imposes that the probability that a new energy unit is harvested when the battery is full and the node is scheduled ($p_{11}^{(1)}$) is no larger than when the battery is empty ($p_{01}^{(1)}$). This is the case, for instance, in the relevant setting in which the probability of the arrival of an energy unit is independent on the battery state and the scheduling decisions taken by the FC, so that one has $p_{11}^{(1)} = p_{01}^{(1)}$, or in the setting in which energy units arriving when the battery is full are discarded, so that $p_{11}^{(1)} = 0$. The second inequality $p_{01}^{(1)} \leq p_{01}^{(0)}$ imposes that the probability $p_{01}^{(1)}$ that a new energy unit is harvested when the battery is empty and the node is scheduled is no larger than when it is not scheduled ($p_{01}^{(0)}$). Similarly to the discussion above, this is true, for instance, if the EH probability does not dependent on the battery state and on the scheduling decisions, so that $p_{01}^{(1)} = p_{01}^{(0)}$. Finally, the last inequality $p_{01}^{(0)} \leq p_{11}^{(0)}$ or equivalently $p_{00}^{(0)} \geq p_{10}^{(0)}$ indicates that, when a node is not scheduled, the probability $p_{10}^{(0)}$ that an energy unit is lost due to leakage is no larger than the probability of not harvesting any energy unit ($p_{00}^{(0)}$). This is the case, for instance, if the probability of energy leakage is sufficiently small.

Main Contributions: The contribution of this chapter are as follows. It is first shown that a myopic policy (MP) under assumptions (6.1) is a round robin (RR) strategy that: *i*) re-numbers the nodes in a decreasing order according to the initial probability that their respective battery is full; and then *ii*) schedules the nodes periodically in group of K by exploiting the initial ordering. The MP is proved to be throughput-optimal. It is then shown that, for the special case in which $p_{01}^{(0)} = p_{01}^{(1)}$ and $p_{10}^{(0)} = p_{11}^{(1)} = 0$, the MP coincides with the Whittle index policy, which is a generally suboptimal index strategy for RMAB problems [38]. Finally, the model of Section 6.1.1 is extended to batteries with an arbitrary capacity C. Characterizing optimal policies for $C > 1$ is significantly more complicated than the case of $C = 1$.

Hence, inspired by the optimality of the MP for $C = 1$, the performance of the MP for $C > 1$ are compared with a upper bound based on a relaxation of the scheduling constraints of the original RMAB problem [38].

Related Work: The derivations in this chapter are inspired by the works [39, 73, 40], in which a RMAB problem is studied by assuming that the evolution of the battery is not affected by the scheduling decision. This is equivalent to the setting $p_{01}^{(0)} = p_{01}^{(1)}$ and $p_{11}^{(0)} = p_{11}^{(1)}$ in the Markov chains of Figure 6.2. In [39] it is shown that the MP is optimal for $p_{11}^{(0)} = p_{11}^{(1)} \leq p_{11}^{(0)} = p_{11}^{(1)}$ with $K = 1$, while [73] extends this result to an arbitrary K. The work [39] also demonstrates that the MP in not generally optimal in the case $p_{01}^{(0)} = p_{01}^{(1)} \geq p_{11}^{(0)} = p_{11}^{(1)}$. Finally, paper [40] proves the optimality of the Whittle index policy for $p_{11}^{(0)} = p_{11}^{(1)} \leq p_{11}^{(0)} = p_{11}^{(1)}$. It is emphasized that neither the model considered in this chapter nor the one considered in [39, 40, 73] subsumes the other, and the results here and in the mentioned previous works should be considered as complementary.

Notation: Vectors are denoted in bold, while the corresponding unbolded letters denote the vectors components. Given a vector $\mathbf{x} = [x_1, ..., x_M]$ and a set $\mathcal{S} = \{i_1, ..., i_K\} \subseteq \{1, ..., M\}$ of cardinality $K \leq M$, it is defined vector $\mathbf{x}_{\mathcal{S}} = [x_{i_1}, ..., x_{i_K}]$, where $i_1 \leq ... \leq i_K$. A function $f(\mathbf{x})$ of vector $\mathbf{x}$ is also denoted as $f(x_1, ..., x_M)$ or as $f(x_1, ..., x_l, \mathbf{x}_{\{l+1,...,M\}})$ for some $1 \leq l \leq M$, or similar notations depending on the context. Notation $\mathbf{1}_K$ indicates a vector of K components, all equal to one. Given a set $\mathcal{A}$ and a subset $\mathcal{B} \subseteq \mathcal{A}$, $\mathcal{B}^c$ represents the complement of $\mathcal{B}$ in $\mathcal{A}$.

6.2 Problem Formulation

This section formalizes the scheduling problem of Section 6.1 (see Figure 6.1), in which the EH and energy leakage processes are modeled, independently at each node, by the Markov models of Section 6.1.1 with battery capacity $C = 1$ (see Figure 6.2). Extension to batteries of arbitrary capacity is addressed in Section 6.6.

6.2.1 Problem Definition

The scheduling problem at the FC is addressed in a finite-horizon scenario in slots $t \in \{1, ..., T\}$. Let $\mathbf{B}(t) = [B_1(t), ..., B_M(t)]$ be the vector collecting the states of the batteries at slot t. At slot $t = 1$, the FC is only aware of the initial probability distribution $\boldsymbol{\omega}(1) = [\omega_1(1), ..., \omega_M(1)]$ of $\mathbf{B}(1)$, whose ith entry is $\omega_i(1) = \Pr[B_i(1) = 1]$. The subset $\mathcal{U}(1)$ of $|\mathcal{U}(1)| = K$ nodes scheduled at slot $t = 1$ is chosen as a function of the initial distribution $\boldsymbol{\omega}(1)$ only. For any node $U_i \in \mathcal{U}(t)$ scheduled at slot t, an *observation* is made available to the FC at the end of the slot (or equivalently before the scheduling decision is taken at slot $t+1$), while no observations are available for non-scheduled nodes $U_i \notin \mathcal{U}(t)$. Specifically, if $B_i(t) = 1$ and $U_i \in \mathcal{U}(t)$, the packet of U_i is collected successfully within slot t, and the FC observes that $B_i(t) = 1$. Conversely, if $B_i(t) = 0$ and $U_i \in \mathcal{U}(t)$, no packets are collected and the FC observes that $B_i(t) = 0$. The set $\mathcal{O}(t)$ contains the (new) observations available at the FC before the scheduling decision is taken at slot $(t + 1)$, which include the states of the batteries of the nodes scheduled at slot t, i.e., $\mathcal{O}(t) = \{B_i(t) : U_i \in \mathcal{U}(t)\}$. At time t, the FC thus knows the history of all decision (or actions) and previous observations along with the initial distribution $\boldsymbol{\omega}(1)$, namely $\mathcal{H}(t) = \{\mathcal{U}(1), ..., \mathcal{U}(t-1), \mathcal{O}(1), ..., \mathcal{O}(t-1), \boldsymbol{\omega}(1)\}$, with $\mathcal{H}(1) = \{\boldsymbol{\omega}(1)\}$. In general, the scheduling decision $\mathcal{U}(t)$ is a function of the history $\mathcal{H}(t)$.

A policy $\pi = [\mathcal{U}^\pi(1), ..., \mathcal{U}^\pi(T)]$ is a collection of functions $\mathcal{U}^\pi(t)$ that map the history $\mathcal{H}(t)$ to a subset $\mathcal{U}(t)$ of $|\mathcal{U}(t)| = K$ nodes, $\mathcal{U}^\pi(t) : \mathcal{H}(t) \to \mathcal{U}(t)$. Note that, strictly speaking $\mathcal{U}^\pi(t)$ is a mapping function, however it is also referred as the subset of scheduled nodes throughout this chapter. In designing the policy π, the FC aims at maximizing the throughput in terms of the average number of packets collected over the finite horizon $t \in \{1, ..., T\}$. For generality, the throughput includes a discount factor β [39], while the infinite horizon scenario (i.e., $T \to \infty$) will be discussed in Section 6.4.

To measure the throughput it is first introduced the *immediate reward* $R(\mathbf{B},\mathcal{U})$, accrued by the FC, as the number of packets collected by the FC in a slot where the state of the batteries is $\mathbf{B}$ and the scheduled set is $\mathcal{U}$:

$$R(\mathbf{Q},\mathcal{U}) = \sum_{i=1}^{M} 1[B_i = 1] 1[U_i \in \mathcal{U}], \tag{6.2}$$

where $1[A]$ is the indicator function of event A, with $1[A] = 1$ if event A is true and zero otherwise. Notice that $R(\mathbf{B},\mathcal{U}) \leq K$ since there are only K transmission resources available. The performance of a policy π is measured by the throughput $V_1^\pi(\omega(1))$ over the horizon $t \in \{1, ..., T\}$, where

$$V_1^\pi(\omega(1)) = \sum_{t=1}^{T} \beta^{t-1} \mathrm{E}^\pi\left[R(\mathbf{B}(t), \mathcal{U}^\pi(t))|\boldsymbol{\omega}(1)\right], \tag{6.3}$$

where $0 \leq \beta \leq 1$, and the expected value $\mathrm{E}^\pi[\cdot|\boldsymbol{\omega}(1)]$ is with respect to the probability distribution of the random process $\mathbf{B}(t)$, determined by the Markov chains in Figure 6.2 for fixed policy π and initial distribution $\boldsymbol{\omega}(1)$. Note that $V_1^\pi(\omega(1)) \leq K\frac{1-\beta^T}{1-\beta}$ for any π. The optimization goal is to find a policy π^*, with optimal actions $\mathcal{U}^*(t)$, for $t \in \{1, ..., T\}$, which maximizes the throughput[1] (6.3) so that

$$\pi^* = \arg\max_\pi V_1^\pi(\boldsymbol{\omega}(1)), \text{ and} \tag{6.4}$$

$$V_1^*(\boldsymbol{\omega}(1)) = V_1^{\pi^*}(\boldsymbol{\omega}(1)) = \max_\pi V_1^\pi(\boldsymbol{\omega}(1)). \tag{6.5}$$

6.2.2 Formulation as Belief MDP and RMAB

Problem (6.4)-(6.5) is a POMDP, since the FC has only partial information about the instantaneous state $\mathbf{B}(t)$ of the system through the observations $\mathcal{O}(t)$. The problem can be reformulated (6.4)-(6.5) as an equivalent MDP with full state knowledge, referred to as *belief* MDP [67]. To this end, it is worth noticing that, while decision $\mathcal{U}^\pi(t)$ at time t depends in general on the entire past history $\mathcal{H}(t)$, it is well-known

---

[1]An optimal policy exists given the discrete nature of the set of all possible policies.

that a sufficient statistics for the optimization problem (6.4)-(6.5) is given by the probability distribution of $\mathbf{B}(t)$ conditioned on the history $\mathcal{H}(t)$ [67]. This conditional probability is referred to as *belief* and it is given by vector $\boldsymbol{\omega}(t) = [\omega_1(t), ..., \omega_M(t)]$, with ith entry being

$$\omega_i(t) = \Pr\left[B_i(t) = 1 | \mathcal{H}(t)\right]. \tag{6.6}$$

The belief $\boldsymbol{\omega}(t)$ fully summarizes, without loss of optimality for problem (6.4)-(6.5), the entire history $\mathcal{H}(t)$ of past actions and observations. An optimal decision $\mathcal{U}^*(t)$ in each tth slot can thus be found as a function of the belief $\boldsymbol{\omega}(t)$ only, which is known by the FC. Therefore, a policy π can be equivalently defined by a collection of functions $\mathcal{U}^\pi(t)$ that map the current state $\boldsymbol{\omega}(t)$ (instead of the whole history $\mathcal{H}(t)$) into the set of the K scheduled nodes.

To define the belief MDP it is necessary to: i) verify that the belief $\boldsymbol{\omega}(t)$ evolves as a controlled Markov process, with control given by the scheduling decisions, and obtain the corresponding transition probabilities; ii) write the throughput (6.3) in terms of an immediate reward function that depends only on the belief $\boldsymbol{\omega}(t)$ and on the scheduling decision.

Transition probabilities: Since the batteries evolve independently at each node for a given scheduling decision, the same holds for the beliefs $\omega_i(\cdot)$. The transition probabilities over the beliefs, given decision $\mathcal{U}(t) = \mathcal{U}$ and belief $\boldsymbol{\omega}(t) = \boldsymbol{\omega} = [\omega_1, ..., \omega_M]$, are thus obtained as

$$\begin{aligned} p_{\boldsymbol{\omega\omega'}}^{(\mathcal{U})} &= \Pr\left[\boldsymbol{\omega}(t+1) = \boldsymbol{\omega}' | \boldsymbol{\omega}(t) = \boldsymbol{\omega}, \mathcal{U}(t) = \mathcal{U}\right] \\ &= \prod_{i=1}^{M} \Pr[\omega_i(t+1) = \omega_i' | \omega_i(t) = \omega_i, \mathcal{U}(t) = \mathcal{U}], \end{aligned} \tag{6.7}$$

where $\boldsymbol{\omega}(t+1) = \boldsymbol{\omega}' = [\omega_1', ..., \omega_M']$ is the next slot's state, while the transition probabilities of the belief $\omega_i(t)$ of node U_i, namely $\Pr[\omega_i(t+1) = \omega_i' | \omega_i(t) = \omega_i, \mathcal{U}(t) =$

$\mathcal{U}]$ are given by

$$
\omega_i(t+1) = \begin{cases} p_{11}^{(1)} & \text{w.p. } \omega_i(t) & \text{if } U_i \in \mathcal{U}(t) \\ p_{01}^{(1)} & \text{w.p. } (1-\omega_i(t)) & \text{if } U_i \in \mathcal{U}(t) \\ \tau_0^{(1)}(\omega_i(t)) & \text{w.p. } 1 & \text{if } U_i \notin \mathcal{U}(t) \end{cases} \cdot \tag{6.8}
$$

In (6.8), the first two lines reflect the fact that, when node U_i is scheduled ($U_i \in \mathcal{U}(t)$) it has an energy unit in its battery w.p. $\omega_i(t)$, and thus, from Figure 6.2-b), the probability that it will have an energy unit in the next slot, i.e., the belief $\omega_i(t+1)$, is $p_{11}^{(1)}$. Similarly, w.p. $(1-\omega_i(t))$ the scheduled node U_i does not have energy in its battery and hence, from Figure 6.2-a), the new belief is $p_{01}^{(1)}$. Finally, the last line in (6.8) states that, if node U_i is not scheduled (i.e., $U_i \notin \mathcal{U}(t)$), then its belief in the next slot can be calculated through the function

$$
\begin{aligned}
\tau_0^{(1)}(\omega) &= \Pr[B_i(t+1)=1|\omega_i(t)=\omega, U_i \notin \mathcal{U}(t)] \\
&= \omega p_{11}^{(0)} + (1-\omega)p_{01}^{(0)} = \omega\delta_0 + p_{01}^{(0)},
\end{aligned} \tag{6.9}
$$

where $\delta_0 = p_{11}^{(0)} - p_{01}^{(0)} \geq 0$ due to inequalities (6.1b). Eq. (6.9) follows from Figure 6.2-a), since the next slot's belief is either $p_{11}^{(0)}$ if $B_i(t)=1$ (w.p. ω) or $p_{01}^{(0)}$ if $B_i(t)=0$ (w.p. $(1-\omega)$). For convenience of notation, it is also useful to define the vector

$$
\boldsymbol{\tau}_0^{(1)}(\omega_1, ..., \omega_K) = [\tau_0^{(1)}(\omega_1), ..., \tau_0^{(1)}(\omega_K)]. \tag{6.10}
$$

A generalization of function $\tau_0^{(1)}(\omega)$ that computes the belief $\omega_i(t+k)$ of node U_i when it is not scheduled for k successive slots, e.g., slots $\{t, ..., t+k-1\}$, and $\omega_i(t) = \omega$, can be obtained as

$$
\begin{aligned}
\tau_0^{(k)}(\omega) &= \Pr[B_i(t+k)=1|\omega_i(t)=\omega, U_i \notin \mathcal{U}(t), ..., U_i \notin \mathcal{U}(t+k-1)]. \\
&= \omega\delta_0^k + p_{01}^{(0)}\frac{1-\delta_0^k}{1-\delta_0}.
\end{aligned} \tag{6.11}
$$

Eq. (6.11) can be obtained recursively from (6.9) as $\tau_0^{(k)}(\omega) = \tau_0^{(1)}(\tau_0^{(k-1)}(\omega))$, for all $k \geq 1$, with $\tau_0^{(0)}(\omega) = \omega$. Some fundamental properties of function (6.11) are summarized in the lemma below.

Lemma 10. If conditions (6.1b) hold, function (6.11) satisfies the inequalities

$$p_{11}^{(1)} \leq p_{01}^{(1)} \leq \tau_0^{(1)}(\omega), \text{ for all } \omega \in [0,1], \text{ and} \tag{6.12}$$

$$\tau_0^{(k)}(\omega) \leq \tau_0^{(k)}(\omega'), \text{ for all } \omega \leq \omega' \text{ with } \omega, \omega' \in [0,1]. \tag{6.13}$$

Proof. From (6.9), it results that $\tau_0^{(1)}(\omega) = \omega\delta_0 + p_{01}^{(0)} \geq p_{01}^{(0)} \geq p_{01}^{(1)} \geq p_{11}^{(1)}$, since $\delta_0 = p_{11}^{(0)} - p_{01}^{(0)} \geq 0$ given the conditions (6.1b), and thus (6.12) is proved. Inequality (6.13) instead follows since it results $\tau_0^{(k)}(\omega) - \tau_0^{(k)}(\omega') = (\omega' - \omega)\,\delta_0^k \geq 0$ for all $\omega \leq \omega'$ as $\delta_0 \geq 0$. $\qquad\square$

Inequalities (6.12) guarantee that the belief of a non-scheduled node is always larger than that of a scheduled one. Inequality (6.13), instead, says that the belief ordering of two non-scheduled nodes is maintained across a slot. Inequalities (6.12)-(6.13) play a crucial role in the analysis below.

Throughput: Similarly to (6.2) it is possible to define an average immediate reward $R(\boldsymbol{\omega},\mathcal{U})$, which depends only on the belief $\boldsymbol{\omega}$ and the scheduling decision $\mathcal{U}$, as

$$R(\boldsymbol{\omega},\mathcal{U}) = \sum_{i=1}^{M} \omega_i 1\,[U_i \in \mathcal{U}]\,. \tag{6.14}$$

In (6.14) the average reward accrued by the FC from a scheduled node $U_i \in \mathcal{U}$ corresponds to the probability ω_i that U_i has energy in its battery. Exploiting (6.14), the throughput (6.3) becomes

$$V_1^{\pi}(\boldsymbol{\omega}(1)) = \sum_{t=1}^{T} \beta^{t-1}\mathrm{E}^{\pi}\,[R\,(\boldsymbol{\omega}(t),\mathcal{U}^{\pi}(t))\,|\boldsymbol{\omega}(1)]\,. \tag{6.15}$$

In (6.15), the expectation $\mathrm{E}^\pi[\cdot|\boldsymbol{\omega}(1)]$ is intended with respect to the distribution of the Markov process $\boldsymbol{\omega}(t)$, as obtained from the transition probabilities (6.8), for fixed policy π and initial belief $\boldsymbol{\omega}(1)$. The optimal policy and the optimal throughput are defined as in (6.4)-(6.5).

Overall, the problem (6.4)-(6.5) has been converted from a POMDP with immediate reward (6.2) and partially observable state $\mathbf{B}(t)$ to an equivalent belief MDP, with immediate reward (6.14) and fully observable state given by the conditional probability of $\mathbf{B}(t)$ (i.e., the belief vector $\boldsymbol{\omega}(t)$). Such a belief MDP constitutes a RMAB with M arms given by the M nodes [70].

6.2.3 Optimality Equations

This section introduces the standard dynamic programming (DP) optimality conditions that characterize an optimal policy π^* in (6.4). To start with, it is possible to define the throughput $V_t^\pi(\boldsymbol{\omega})$ over the horizon $\{t, ..., T\}$ for policy π and initial belief $\boldsymbol{\omega}(t) = \boldsymbol{\omega}$ as

$$V_t^\pi(\boldsymbol{\omega}) = \sum_{j=t}^{T} \beta^{j-t}\mathrm{E}^\pi\left[R\left(\boldsymbol{\omega}(j), \mathcal{U}^\pi(j)\right)|\boldsymbol{\omega}(t) = \boldsymbol{\omega}\right], \tag{6.16}$$

consistently with (6.15). The DP optimality conditions are expressed in terms of *value functions* $V_t^*(\boldsymbol{\omega}) = \max_\pi V_t^\pi(\boldsymbol{\omega})$ that represent the optimal throughput in the interval $\{t, ..., T\}$. The probability that the K scheduled nodes have energies $b_1, ..., b_K \in \{0, 1\}$ for a given belief $\boldsymbol{\omega}$ can be calculated as

$$b(b_1, ..., b_K, \omega_1, .., \omega_K) = \prod_{i=1}^{K} \omega_i^{b_i}(1 - \omega_i)^{1-b_i}. \tag{6.17}$$

Lemma 11. *DP optimality conditions:* The throughput $V_t^\pi(\boldsymbol{\omega})$ in (6.16), given belief $\boldsymbol{\omega} = [\boldsymbol{\omega}_{\mathcal{U}^\pi(t)}, \boldsymbol{\omega}_{(\mathcal{U}^\pi(t))^c}]$, with $(\mathcal{U}^\pi(t))^c = \{U_1, ..., U_M\} \setminus \mathcal{U}^\pi(t)$, satisfies the recursive

conditions

$$V_T^\pi(\boldsymbol{\omega}) = R(\boldsymbol{\omega}, \mathcal{U}^\pi(T)) = \sum_{U_i \in \mathcal{U}^\pi(T)} \omega_i \tag{6.18}$$

$$V_t^\pi(\boldsymbol{\omega}) = R(\boldsymbol{\omega}, \mathcal{U}^\pi(t)) + \beta \sum_{\boldsymbol{\omega}'} V_{t+1}^\pi(\boldsymbol{\omega}') p_{\boldsymbol{\omega\omega}'}^{(\mathcal{U}^\pi(t))}$$

$$= \sum_{U_i \in \mathcal{U}^\pi(T)} \omega_i + \beta \sum_{b_1,\ldots,b_K \in \{0,1\}} b(b_1, \ldots, b_K, \boldsymbol{\omega}_{\mathcal{U}^\pi(t)}) \cdot \tag{6.19}$$

$$V_{t+1}^\pi \left(\gamma(b_1), \ldots, \gamma(b_K), \boldsymbol{\tau}_0^{(1)}(\boldsymbol{\omega}_{(\mathcal{U}^\pi(t))^c}) \right), \quad for\ t \in \{1, \ldots, T-1\},$$

where $\gamma(b) = p_{01}^{(1)}(1-b) + p_{11}^{(1)}b$.

Moreover, the value function $V_t^*(\boldsymbol{\omega}) = \max_\pi V_t^\pi(\boldsymbol{\omega})$ satisfies the DP optimality conditions

$$V_T^*(\boldsymbol{\omega}) = \max_{\mathcal{U}(T) \subseteq \{U_1,\ldots,U_M\}} \sum_{U_i \in \mathcal{U}(T)} \omega_i \tag{6.20}$$

$$V_t^*(\boldsymbol{\omega}) = \max_{\mathcal{U}(t) \subseteq \{U_1,\ldots,U_M\}} \left\{ \sum_{U_i \in \mathcal{U}(t)} \omega_i + \beta \sum_{b_1,\ldots,b_K \in \{0,1\}} b(b_1, \ldots, b_K, \boldsymbol{\omega}_{\mathcal{U}(t)}) \cdot \right. \tag{6.21}$$

$$\left. V_{t+1}^* \left(\boldsymbol{\tau}_0^{(1)}(\boldsymbol{\omega}_{(\mathcal{U}(t))^c}), \ p_{01}^{(1)}\mathbf{1}_{K-\sum_{i=1}^K b_i}, \ p_{11}^{(1)}\mathbf{1}_{\sum_{i=1}^K b_i} \right) \right\}, \quad for\ t \in \{1, \ldots, T-1\}.$$

Finally, an optimal policy π^* (6.4) is such that $\mathcal{U}^*(t)$ attains the maximum in the conditions (6.20)-(6.21) for $t = 1, 2, \ldots, T$.

Proof. The equalities (6.18)-(6.19) follow from DP backward induction from definition (6.16) (see [69]), and are a consequence of (6.7), (6.8), (6.10) and (6.14). The DP optimality conditions (6.20)-(6.21) follow from standard DP theory [69], where it has been exploited the fact that the nodes are stochastically equivalent, and thus $V_t^*(\boldsymbol{\omega})$ only depends on the numerical values of the entries of the belief vector $\boldsymbol{\omega}$ regardless of the way it is ordered. $\qquad\square$

Some comments on (6.19) and (6.21) are now in order. In the second term of the right hand side (RHS) of (6.19) and (6.21), one averages over the distribution

$p_{\boldsymbol{\omega}\boldsymbol{\omega}'}^{(\mathcal{U})}$ (6.7) of the next-slot belief given the current belief and scheduling decision $\mathcal{U}$. From (6.8), the beliefs of all the unscheduled nodes in $\mathcal{U}^c$ evolve deterministically as $\boldsymbol{\tau}_0^{(1)}(\boldsymbol{\omega}_{\mathcal{U}^c})$. Instead, the beliefs of each scheduled node $U_i \in \mathcal{U}$ can be either equal to $p_{11}^{(1)}$ or $p_{01}^{(1)}$ w.p. ω_i and $(1 - \omega_i)$, respectively. This is accounted for by the sum in the last equality in (6.19) and (6.21).

6.3   Optimality of the Myopic Policy

This section first defines the myopic policy (MP) and then shows that, under conditions (6.1), the MP is a round-robin (RR) strategy that schedules nodes periodically. It is then proved that the MP is optimal for problem (6.4), and its throughput (6.5) is computed in closed form.

6.3.1   The Myopic Policy is Round-Robin

The MP $\pi^{MP} = \{\mathcal{U}^{MP}(1), ..., \mathcal{U}^{MP}(T)\}$ is a greedy policy that in each tth slot schedules the K nodes with the largest beliefs so as to maximize the immediate reward (6.14) as

$$\mathcal{U}^{MP}(t) = \arg\max_{\mathcal{U}} R(\boldsymbol{\omega}(t), \mathcal{U}) = \arg\max_{\mathcal{U}} \sum_{U_i \in \mathcal{U}} \omega_i(t). \qquad (6.22)$$

Note that the MP is a *stationary* policy in the sense that the scheduling decision $\mathcal{U}^{MP}(t)$ depends only on the value of the belief $\boldsymbol{\omega}(t)$ regardless the slot t.

Proposition 12. If conditions (6.1) hold, the MP π^{MP} (6.22), given an initial belief $\boldsymbol{\omega}'(1)$, is a RR policy that operates as follows: **1)** Sort vector $\boldsymbol{\omega}'(1)$ in a decreasing order to obtain $\boldsymbol{\omega}(1) = [\omega_1(1), ..., \omega_M(1)]$ such that $\omega_1(1) \geq ... \geq \omega_M(1)$. Renumber the nodes so that U_i has belief $\omega_i(1)$; **2)** Divide the nodes into m groups of K nodes each, so that the gth group $\mathcal{G}_g$, $g \in \{1, ..., m\}$, contains all nodes U_i such that $g = \lfloor \frac{i-1}{K} \rfloor + 1$, namely: $\mathcal{G}_1 = \{U_1, ..., U_K\}, \mathcal{G}_2 = \{U_{K+1}, ..., U_{2K}\}$, and so on; **3)** Schedule the

groups in a RR (periodic) fashion with period m slots, so that groups $\mathcal{G}_1, ..., \mathcal{G}_m, \mathcal{G}_1, ...$ are sequentially scheduled at slot $t = 1, ..., m, m + 1, ...$ and so on.

Proof. According to (6.22), the first scheduled set is $\mathcal{U}^{MP}(1) = \mathcal{G}_1 = \{U_1, U_2, ..., U_K\}$. The beliefs are then updated through (6.8). Recalling (6.12), the scheduled nodes, in $\mathcal{G}_1$, have their belief updated to either $p_{11}^{(1)}$ or $p_{01}^{(1)}$, which are both smaller than the belief of any non-scheduled node in $\{U_1, ..., U_M\} \setminus \mathcal{G}_1$. Moreover, the ordering of the non-scheduled nodes' beliefs is preserved due to (6.13). Hence, the second scheduled group is $\mathcal{U}^{MP}(2) = \mathcal{G}_2$, the third is $\mathcal{U}^{MP}(3) = \mathcal{G}_3$, and so on. This proves that the MP, upon an initial ordering of the beliefs, is a RR policy. $\qquad\square$

It now possible to make an useful observation for proving the optimality results of this section. Consider a RR policy π^{RR} that operates according to steps **2)** and **3)** of Proposition 12 (i.e., without re-ordering the initial belief). The throughput (6.16) of π^{RR} can be expressed recursively through functions $\tilde{V}_t(\boldsymbol{\omega})$ as

$$\tilde{V}_T(\boldsymbol{\omega}) \;=\; \sum_{i=1}^{K} \omega_i \tag{6.23}$$

$$\tilde{V}_t(\boldsymbol{\omega}) \;=\; \sum_{i=1}^{K} \omega_i + \beta \sum_{b_1, ..., b_K \in \{0,1\}} b(b_1, ..., b_K, \omega_1, ..., \omega_K) \cdot \tag{6.24}$$

$$\tilde{V}_{t+1}\left(\boldsymbol{\tau}_0^{(1)}(\omega_{K+1}, ..., \omega_M), \; p_{01}^{(1)}\mathbf{1}_{K - \sum_{i=1}^{K} b_i}, \; p_{11}^{(1)}\mathbf{1}_{K - \sum_{i=1}^{K} b_i}\right),$$

$$\text{for } t = 1, ..., T - 1. \tag{6.25}$$

The policy π^{RR} in each slot: $i)$ schedules the K nodes whose beliefs are in the first K positions of the argument $\boldsymbol{\omega}$ of $\tilde{V}_T(\boldsymbol{\omega})$; $ii)$ the argument $\boldsymbol{\omega}'$ for the next slot is updated (through (6.8)) so that the beliefs of the scheduled nodes are decreasingly ordered and put at the K rightmost positions of $\boldsymbol{\omega}'$ so that $\boldsymbol{\omega}' = [\boldsymbol{\tau}_0^{(1)}(\omega_{K+1}, ..., \omega_M), \; p_{01}^{(1)}\mathbf{1}_{K - \sum_{i=1}^{K} b_i}, \; p_{11}^{(1)}\mathbf{1}_{K - \sum_{i=1}^{K} b_i}]$. Note that, when the initial belief $\boldsymbol{\omega}$ is ordered so that $\omega_1 \geq ... \geq \omega_M$, then $\tilde{V}_t(\boldsymbol{\omega}) = V_t^{MP}(\boldsymbol{\omega})$.

6.3.2   Optimality of the Myopic Policy

This section proves the optimality of the MP described above by showing that it satisfies the DP optimality conditions of Lemma 11. The proof is based on backward induction arguments similarly to [39, 73]. The following lemma establishes a sufficient condition for the optimality of the MP.

Lemma 13.   Assume that the MP is optimal at slot $t + 1, ..., T$, in the sense that $\mathcal{U}^{MP}(t + 1), ..., \mathcal{U}^{MP}(T)$ attain the maximum in the corresponding DP optimality conditions (6.20)-(6.21). To show that the MP is optimal also at slot t it is sufficient to show that

$$\tilde{V}_t(\boldsymbol{\omega}_{\mathcal{S}}, \boldsymbol{\omega}_{\mathcal{S}^c}) \leq V_t^{MP}(\boldsymbol{\omega}_{\mathcal{S}}, \boldsymbol{\omega}_{\mathcal{S}^c}) = \tilde{V}_t(\omega_1, \omega_2, ..., \omega_M), \quad \text{for all } \omega_1 \geq \omega_2 \geq ... \geq \omega_M,$$

$$(6.26)$$

and all sets $\mathcal{S} \subseteq \{1, ..., M\}$ of K elements, with the elements in $\boldsymbol{\omega}_{\mathcal{S}^c}$ decreasingly ordered.

Proof. Since by assumption the MP is optimal from $t + 1$ onward, it is sufficient to show that scheduling K nodes with arbitrary beliefs at slot t and then following the MP from slot $t + 1$ on, is no better than following the MP immediately at slot t. The performance of the former policy is given by the left-hand side (LHS) of (6.26). In fact $\tilde{V}_t(\boldsymbol{\omega}_{\mathcal{S}}, \boldsymbol{\omega}_{\mathcal{S}^c})$, for any set $\mathcal{S}$, represents the throughput of a policy that schedules the K nodes with beliefs $\boldsymbol{\omega}_{\mathcal{S}}$ at slot t, and then operates as the MP from $t + 1$ onward, since beliefs $\boldsymbol{\omega}_{\mathcal{S}^c}$ are in decreasing order (see (6.23)-(6.24)). The MP's performance is instead given by the RHS of (6.26). This concludes the proof. □

The following lemma demonstrates that inequality (6.26) holds.

Lemma 14. If conditions (6.1) hold, then: **a)** inequality (6.27) holds for all $x \geq y$ and $0 \leq j \leq M - 2$

$$\tilde{V}_t(\omega_1, ..., \omega_j, y, x, ..., \omega_M) \quad \leq \quad \tilde{V}_t(\omega_1, ..., \omega_j, x, y, ..., \omega_M), \qquad (6.27)$$

where for $j = 0$ inequality (6.27) is intended as $\tilde{V}_t(y, x, ..., \omega_M) \leq \tilde{V}_t(x, y, ..., \omega_M)$; and

b) inequality (6.26) is satisfied for all $\omega_1 \geq ... \geq \omega_M$ and all subsets $\mathcal{S} \subseteq \{1, ..., M\}$ of K elements.

Proof. Part **a**) see Appendix F. Part **b**). By part **a**) condition (6.27) holds. This implies that, for $\omega_1 \geq ... \geq \omega_M$, switching positions of neighboring elements ω_i in the RHS of (6.27) does not increase function $\tilde{V}_t(\cdot)$. But, through a sequence of switching operations between neighboring elements of $\boldsymbol{\omega}$, it is possible to obtain an arbitrary vector $(\boldsymbol{\omega}_\mathcal{S}, \boldsymbol{\omega}_{\mathcal{S}^c})$, which proves that (6.26) holds. $\qquad\square$

It is now possible to establish the optimality of the MP.

Theorem 15. If conditions (6.1) hold then the MP is optimal for problem (6.4)-(6.5) so that $\pi^{MP} = \pi^*$ and its throughput $V_1^{MP}(\boldsymbol{\omega}(1)) = V_1^*(\boldsymbol{\omega}(1))$ is calculated in closed form in Appendix E.

Proof. Using Lemma 14, the proof is concluded immediately by Lemma 13. $\qquad\square$

6.4 Extension to the Infinite-Horizon Case

This section briefly describes the extension of the problem (6.15) (and thus (6.4)-(6.5)) to the infinite-horizon case. Beside its independent interest, this will be useful in the next section where the optimality of the Whittle index policy will be discussed. The throughput in the infinite-horizon case under policy π and discount factor $0 \leq \beta < 1$, and its optimal value, are given by [39]

$$V^\pi(\boldsymbol{\omega}(1)) = \sum_{t=1}^\infty \beta^{t-1} \mathrm{E}^\pi\left[R\left(\boldsymbol{\omega}(t), \mathcal{U}^\pi(t)\right) | \boldsymbol{\omega}(1)\right], \text{ and} \qquad (6.28)$$

$$V^*(\boldsymbol{\omega}(1)) = \max_\pi V^\pi(\boldsymbol{\omega}(1)), \qquad (6.29)$$

where the optimal policy is $\pi^* = \arg\max_\pi V^\pi(\boldsymbol{\omega}(1))$. From standard DP theory, the optimal policy π^* is *stationary*, so that π^* is such that the optimal scheduling decision

$\mathcal{U}^*(t)$ is a function of the current state $\omega(t)$ only independently of slot t [69]. Following the same reasoning as in [39, Theorem 3], it is easy to show that the optimality of the MP for the finite-horizon setting implies the optimality also for the infinite-horizon scenario.

6.5   Optimality of the Whittle Index Policy

This section briefly reviews the Whittle index policy for RMAB problems [70], and then focuses on the infinite-horizon scenario of Section 6.4, when conditions (6.1b) are specialized to

$$0 = p_{11}^{(1)} \leq p_{01}^{(1)} = p_{01}^{(0)} = p_{01} \leq p_{11}^{(0)} = 1. \tag{6.30}$$

It will be shown that under the assumption (6.30) (see Section 6.1.2 for a discussion on these conditions), the RMAB at hand is indexable and it is possible to calculate its Whittle index in closed-form. It will be then shown that the Whittle index policy is equivalent to the MP, and thus optimal for the problem (6.29).

It is emphasized that, the results in this section provide a rare example [70] in which, as in [40], not only indexability is established, but also the Whittle index is obtained in closed form and the Whittle policy proved to be optimal. It is finally remarked that the proof technique of this section is inspired by [40], but the different system model poses new challenges that require significant work.

6.5.1   Whittle Index

The Whittle index policy assigns a numerical value $W(\omega_i)$ to each state ω_i of node U_i, referred to as *index*, to measure how rewarding it is to schedule node U_i in the current slot. The K nodes with the largest index are then scheduled in each slot. As detailed below, the Whittle index is calculated independently for each node, and thus the Whittle index policy is not generally optimal for RMAB problems. Moreover,

even the existence of a well-defined Whittle index is not guaranteed [70]. To study the indexability and the Whittle index for the RMAB at hand, it is possible to focus on a restless single-armed bandit (RSAB) model, as defined below [70].

6.5.2 RSAB with Subsidy for Passivity

The Whittle index is based on the concept of *subsidy for passivity*, whereby the FC is given a subsidy $m \in \mathbb{R}$ when the arm is not scheduled. At each slot t, the CC, based on the state $\omega(t)$ of the arm, can decide to activate (or schedule) it, i.e., to set $u(t) = 1$, obtaining an immediate reward $R_m(\omega(t), 1) = \omega(t)$. If, instead, the arm is kept passive, i.e., $u(t) = 0$, a reward $R_m(\omega(t), 0) = m$ equal to the subsidy is accrued. The state $\omega(t)$ evolves through (6.8), which under (6.30) and adapted to the simplified notation used here becomes

$$\omega(t+1) = \begin{cases} 0 & \text{w.p. } \omega(t) & \text{if } u(t) = 1 \\ p_{01} & \text{w.p. } (1 - \omega(t)) & \text{if } u(t) = 1 \\ \tau_0^{(1)}(\omega(t)) & \text{w.p. } 1 & \text{if } u(t) = 0 \end{cases} . \tag{6.31}$$

The throughput, given policy $\pi = \{u^\pi(1), u^\pi(2), ...\}$ and initial belief $\omega(1)$, is

$$V_m^\pi(\omega(1)) = \sum_{t=1}^{\infty} \beta^{t-1} \mathrm{E}^\pi \left[R_m(\omega(t), u^\pi(t)) | \omega(1) \right]. \tag{6.32}$$

The optimal throughput is $V_m^*(\omega(1)) = \max_\pi V_m^\pi(\omega(1))$, while the optimal policy $\pi^* = \arg\max_\pi V_m^\pi(\omega(1))$ is stationary in the sense that the optimal decisions $u_m^*(\omega) \in \{0, 1\}$ are functions of the belief ω only, independently of slot t [40]. Removing the slot index from the initial belief, the optimal throughput $V_m^*(\omega)$ and the optimal decision $u_m^*(\omega)$ satisfy the following DP optimality equations for the infinite-horizon scenario (see [40])

$$V_m^*(\omega) = \max_{u \in \{0,1\}} \{V_m(\omega | u)\}, \tag{6.33}$$

$$\text{and } u_m^*(\omega) = \arg\max_{u \in \{0,1\}} \{V_m(\omega | u)\}. \tag{6.34}$$

In (6.33)-(6.34) it has been defined $V_m(\omega|u)$, $u \in \{0, 1\}$, as the throughput (6.32) of a policy that takes action u at the current slot and then uses the optimal policy $u_m^*(\omega)$ onward, it results

$$V_m(\omega|0) \;=\; m + \beta V_m^*(\tau_0^{(1)}(\omega)), \text{ and} \tag{6.35}$$

$$V_m(\omega|1) \;=\; \omega + \beta \left[\omega V_m^*(0) + (1-\omega)V_m^*(p_{01})\right]. \tag{6.36}$$

6.5.3 Indexability and Whittle Index

This section adopts the notation of [40] to define indexability and Whittle index for the RSAB at hand. The first definition is the so called *passive set*

$$\mathcal{P}(m) = \{\omega\colon\, 0 \leq \omega \leq 1 \text{ and } u_m^*(\omega) = 0\} \tag{6.37}$$

that contains all the beliefs ω for which the passive action is optimal (i.e., all $0 \leq \omega \leq 1$ such that $V_m(\omega|0) \geq V_m(\omega|1)$, see (6.35)-(6.36)) under the given subsidy for passivity $m \in \mathbb{R}$. The RMAB at hand is said to be *indexable* if the passive set $\mathcal{P}(m)$, for the associated RSAB problem[2], is monotonically increasing as m increases within the interval $(-\infty, +\infty)$, in the sense that $\mathcal{P}(m') \subseteq \mathcal{P}(m)$ if $m' \leq m$ and $\mathcal{P}(-\infty) = \emptyset$ and $\mathcal{P}(+\infty) = [0, 1]$.

If the RMAB is indexable, the Whittle index $W(\omega)$ for each arm with state ω is the infimum subsidy m such that it is optimal to make the arm passive. Equivalently, the Whittle index $W(\omega)$ is the infimum subsidy m that makes passive and active actions equally rewarding, i.e.,

$$W(\omega) \;=\; \inf\{m\colon\, u_m^*(\omega) = 0\} = \inf\{m\colon\, V_m(\omega|0) = V_m(\omega|1)\}. \tag{6.38}$$

---

[2]Note that in a RMAB with arms characterized by different statistics this condition must be checked for all arms.

6.5.4 Optimality of the Threshold Policy

In this section it is shown that the optimal policy $u_m^*(\omega)$ for the RSAB of Section 6.5.2 is a threshold policy over the belief ω. This is crucial in the proof of indexability of the RMAB at hand given in Section 6.5.6. To this end, it is possible to observe that: i) function $V_m(\omega|1)$ in (6.36) is linear over the belief ω; ii) function $V_m(\omega|0) = m + \beta V_m^*(\tau_0^{(1)}(\omega))$ in (6.35) is convex over ω, since the convexity of $V_m^*(\omega)$ is a general property of POMDPs (see [40, 67]). The the following lemma establishes useful results.

Lemma 16. The following inequalities hold:

a) For $0 \le m < 1$:   a.1) $V_m(0|1) \le V_m(0|0) \le V_m(1|1)$;   a.2) $V_m(1|0) \le V_m(1|1)$;

$$(6.39a)$$

b) For $m < 0$:   b.1) $V_m(0|0) \le V_m(0|1) \le V_m(1|1)$;   b.2) $V_m(1|0) \le V_m(1|1)$;

$$(6.39b)$$

c) For $m \ge 1$:   c.1) $V_m(0|0) \le V_m(1|1) \le V_m(0|1)$;   c.2) $V_m(1|1) \le V_m(1|0)$.

$$(6.39c)$$

Proof. See Appendix G.   □

Leveraging Lemma 16, it is now possible to establish the optimality of a threshold policy $u_m^*(\omega)$.

Proposition 17. The optimal policy $u_m^*(\omega)$ in (6.34) for subsidy $m \in \mathbb{R}$ is given by

$$u_m^*(\omega) = \begin{cases} 1, & if \ \omega > \omega^*(m) \\ 0, & if \ \omega \le \omega^*(m) \end{cases}, \qquad (6.40)$$

where $\omega^*(m) \in \mathbb{R}$ is the optimal threshold for a given subsidy m. The optimal threshold $\omega^*(m)$ is $0 \le \omega^*(m) \le 1$ if $0 \le m < 1$, while it is arbitrary negative for

$m < 0$ and arbitrary greater than unity for $m \geq 1$. In other words it results $u_m^*(\omega) = 1$ if $m < 0$ and $u_m^*(\omega) = 0$ if $m \geq 1$.

Proof. The proof starts by showing that (6.40), for $0 \leq m < 1$, satisfies (6.34) and is thus an optimal policy. To see this, it is possible to refer to Figure 6.3, where a sketch of the functions $V_m(\omega|1)$ and $V_m(\omega|0)$ for different values of the subsidy m is provided. From (6.34), it results that $u_m^*(\omega) = 1$ for all ω such that $V_m(\omega|1) > V_m(\omega|0)$ and $u_m^*(\omega) = 0$ otherwise. For $0 \leq m < 1$, from the inequalities of Lemma 16-a), the linearity of $V_m(\omega|1)$ and the convexity of $V_m(\omega|0)$, it follows that there is only one intersection $\omega^*(m)$ between $V_m(\omega|1)$ and $V_m(\omega|0)$ with $0 \leq \omega^*(m) \leq 1$, as shown in Figure 6.3-a). Instead, when $m < 0$, by Lemma 16-b), arm activation is always optimal, that is, $u_m^*(\omega) = 1$, since $V_m(\omega|1) > V_m(\omega|0)$ for any $0 \leq \omega \leq 1$ as shown in Figure 6.3-b). Conversely, when $m \geq 1$, by Lemma 16-c), it follows that passivity is always optimal, that is, $u_m^*(\omega) = 0$, since $V_m(\omega|0) \geq V_m(\omega|1)$ for any $0 \leq \omega \leq 1$ as shown in Figure 6.3-c). □
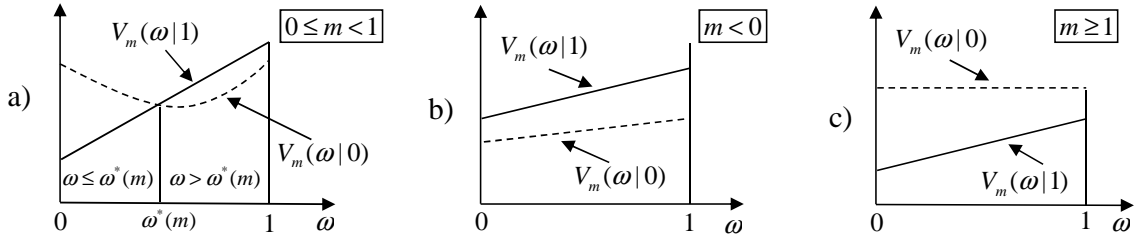
Figure 6.3 Illustration of the optimality of a threshold policy for different values of the subsidy for passivity m: a) $0 \leq m < 1$; b) $m < 0$; c) $m \geq 1$.

6.5.5 Closed-Form Expression of the Value Function

By leveraging the optimality of the threshold policy (6.40) this section derives a closed-form expression of $V_m^*(\omega)$ in (6.33), being a key step in establishing the RMAB's indexability in Section 6.5.6.

Notice that function $\tau_0^{(k)}(\omega)$ in (6.11), when specialized to conditions (6.30), becomes

$$\tau_0^{(k)}(\omega) = 1 - (1 - p_{01})^k(1 - \omega), \qquad (6.41)$$

which is a monotonically increasing function of k, so that $\tau_0^{(k)}(\omega) \geq \tau_0^{(i)}(\omega)$ for any $k \geq i$. Based on such monotonicity, it is possible to define the average number $L(\omega, \omega')$ of slots it takes for the belief to become larger than ω' when starting from ω while the arm is kept passive, as

$$L(\omega, \omega') = \min\left\{k\colon \tau_0^{(k)}(\omega) > \omega'\right\} = \begin{cases} 0 & \omega > \omega' \\ \left\lfloor \frac{\ln\left(\frac{1-\omega'}{1-\omega}\right)}{\ln(1-p_{01})} \right\rfloor + 1 & \omega \leq \omega' \\ \infty & \omega \leq 1 \leq \omega' \end{cases} . (6.42)$$

From (6.42) it results that $L(\omega, \omega') = 1$ for $\omega = \omega'$ since, without loss of optimality, it is possible to assume that the passive action is optimal (i.e., $u_m^*(\omega) = 0$) when $V_m(\omega|0) = V_m(\omega|1)$. For $\omega' \geq 1$ instead (according to Proposition 17), the arm is always kept passive and thus $L(\omega, \omega') = \infty$.

Lemma 18. The optimal throughput $V_m^*(\omega)$ in (6.33) can be written as

$$V_m^*(\omega) = \frac{1 - \beta^{L(\omega, \omega^*(m))}}{1 - \beta} m + \beta^{L(\omega, \omega^*(m))} V_m(\tau_0^{(L(\omega, \omega^*(m)))}(\omega)|1), \qquad (6.43)$$

where $\omega^*(m)$ is the optimal threshold obtained from Proposition 17.

Proof. According to Proposition 17, the optimal policy $u_m^*(\omega)$ keeps the arm passive as long as the current belief is $\omega \leq \omega^*(m)$. Therefore, the arm is kept passive for $L(\omega, \omega^*(m))$ slots, during which a reward $R_m(\omega, 0) = m$ is accrued in each slot. This leads to a total reward within the passivity time given by the following geometric series $\sum_{k=0}^{L(\omega, \omega^*(m))-1} \beta^k m = \frac{1 - \beta^{L(\omega, \omega^*(m))}}{1 - \beta} m$, which corresponds to the first term in the RHS of (6.43). After $L(\omega, \omega^*(m))$ slots of passivity, the belief becomes larger than

the threshold $\omega^*(m)$ and the arm is activated. The contribution to the value function $V(\omega)$ thus becomes $\beta^{L(\omega,\omega^*(m))} V_m(\tau_0^{(L(\omega,\omega^*(m)))}(\omega)|1)$, which is the second term in the RHS of (6.43). Note that, when $\omega > \omega^*(m)$, activation is optimal, and $V^*(\omega) = V(\omega|1)$. □

To evaluate $V_m^*(\omega)$ from (6.43), it is only necessary to calculate $V_m(\omega|1)$ since the other terms, thanks to (6.42) are explicitly given once $\omega^*(m)$ is obtained from Proposition 17. However, from (6.36), evaluating $V_m(\omega|1)$ only requires $V_m^*(0)$ and $V_m^*(p_{01})$, which are calculated in the lemma below.

Lemma 19. It follows

$$V_m^*(0) = \frac{\left(m - 2m\beta^{L_m^*} + \beta^{L_m^*}v_m^* - \beta^{L_m^*+1}v_m^* + m\beta^{L_m^*+1} + m\beta^{L_m^*}v_m^* - m\beta^{L_m^*+1}v_m^*\right)}{(\beta - 1)\left(\beta^{L_m^*} - \beta^{L_m^*}v_m^* + \beta^{L_m^*+1}v_m^* - 1\right)}$$

(6.44a)

$$V_m^*(p_{01}) = \frac{\left(m\beta - m\beta^{L_m^*} + \beta^{L_m^*}v_m^* - \beta^{L_m^*+1}v_m^* + m\beta^{L_m^*+1}v_m^* - m\beta^{L_m^*+2}v_m^*\right)}{\beta\,(\beta - 1)\left(\beta^{L_m^*} - \beta^{L_m^*}v_m^* + \beta^{L_m^*+1}v_m^* - 1\right)}.$$

(6.44b)

where $L_m^* = L(0, \omega^*(m))$ and $v_m^* = \tau_0^{(L(0,\omega^*(m)))}(0)$.

Proof. By plugging (6.36) into (6.43), and evaluating (6.43) for $\omega = 0$ and $\omega = p_{01}$, a linear system in the two unknowns $V_m^*(0)$ and $V_m^*(p_{01})$ is obtained, and it can be solved leading to (6.44). □

6.5.6 Indexability and Whittle Index

This section proves that the RMAB at hand is indexable, and derives the Whittle index in closed form and shows that it is equivalent to the MP and thus optimal for the RMAB problem (6.29).

Theorem 20. a) The RMAB at hand is indexable and **b)** its Whittle index is

$$W(\omega) = \frac{\left(1 - \beta^{L(0,\omega)}\left(1 - \beta\tau_0^{L(0,\omega)}(0)\bar{\beta}\left(1-h\right)\right)\right)\omega + \beta^{L(0,\omega)}\tau_0^{L(0,\omega)}(0)\bar{\beta}\left(h\beta + 1\right)}{-\bar{\beta}\left(\beta^{L(0,\omega)}\left(1 - \beta(1-h)\right)\omega - \left(1 + \beta^{L(0,\omega)}\left(\tau_0^{L(0,\omega)}(0)\bar{\beta} + h\beta\right)\right)\right)},$$

(6.45)

where $\bar{\beta} = 1 - \beta$.

Proof. Part **a)**. See Appendix H. Part **b)**. By (6.38), the Whittle index $W(\omega)$ of state ω is the value of the subsidy m for which activating or not the arm is equally rewarding so that $V_m(\omega|0) = V_m(\omega|1)$. By using (6.35)-(6.36) this becomes $\omega + \beta\left[\omega V_m^*(0) + (1-\omega)V_m^*(p_{01})\right] = m + \beta V_m^*(\tau_0^{(1)}(\omega))$. Moreover, since the threshold policy is optimal and $\tau_0^{(1)}(\omega) > \omega$, it follows that, when the belief becomes $\tau_0^{(1)}(\omega)$, it is optimal to activate the arm and thus $V_m^*(\tau_0^{(1)}(\omega)) = V_m(\tau_0^{(1)}(\omega)|1) = \beta\tau_0^{(1)}(\omega)V_m^*(0) + \beta(1 - \tau_0^{(1)}(\omega))V_m^*(p_{01})$. Plugging this result into $V_m(\omega|0) = V_m(\omega|1)$, along with (6.44a) and (6.44b), leads to (6.45), which concludes the proof. $\square$

It can be show that the Whittle index $W(\omega)$ in (6.45) is an increasing function of ω. Therefore, since the Whittle policy selects the K arms with the largest index at each slot, it follows that

Corollary 21. The Whittle index policy is equivalent to the MP and is thus optimal.

6.6 Extension to Batteries of Arbitrary Capacity $C > 1$

The problem of characterizing the optimal policies when $C > 1$ is significantly more complicated than for $C = 1$ and is left open by this work. Moreover, since the dimension of the state space of the belief MDP grows with C, even the numerical computation of the optimal policies is quite cumbersome. Due to these difficulties, this section compares the performance of the MP, inspired by its optimality for $C = 1$, with a performance upper bound obtained following the relaxation approach of [38].
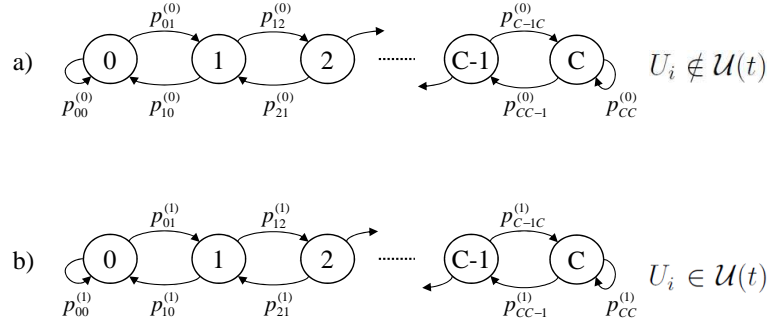
Figure 6.4 Markov model for the evolution of the batteries $B_i(t)$, of arbitrary capacity C, when the node U_i: a) is not scheduled in slot t (i.e., $U_i \notin \mathcal{U}(t)$); b) is scheduled in slot t (i.e., $U_i \in \mathcal{U}(t)$).

6.6.1   System Model and Myopic Policy

Each node U_i has a battery $B_i(t) \in \{0, 1, ..., C\}$ of capacity C. In this section the EH and leakage processes are modeled as the controlled Markov processes drawn in Figure 6.4 (cf. Section 6.1.1). The transition probabilities between battery states when node U_i is not scheduled are $p_{xy}^{(0)} = \Pr[B_i(t+1) = y | B_i(t) = x, U_i \notin \mathcal{U}(t)]$, whereas when U_i is scheduled one has $p_{xy}^{(1)} = \Pr[B_i(t+1) = y | B_i(t) = x, U_i \in \mathcal{U}(t)]$, for $x, y \in \{0, 1, ..., C\}$. When node U_i is scheduled at slot t, and $B_i(t) \geq 1$, an energy unit is drawn from its battery, and the node also informs the FC about the remaining energy in the battery (observation). It is assumed that at most one energy unit can be harvested (or lost) in a slot, so that $p_{xy}^{(u)} = 0$ for $y < x - 1$ and $y > x + 1$, with $u \in \{0, 1\}$ as shown in Figure 6.4.

The belief of each ith node is represented by a $(C \times 1)$ vector $\boldsymbol{\omega}_i = [\omega_{i,0}, ..., \omega_{i,C-1}]$ whose kth entry $\omega_{i,k}$, for $k \in \{0, 1, ..., C-1\}$, is given by (cf. (6.6)) $\omega_{i,k} = \Pr[B_i(t) = k | \mathcal{H}(t)]$. The immediate reward (6.14), given the initial belief vectors $\boldsymbol{\omega}_1(t), ..., \boldsymbol{\omega}_M(t)$ and action $\mathcal{U}$, becomes

$$
\begin{aligned}
R(\boldsymbol{\omega}_1(t), ..., \boldsymbol{\omega}_M(t), \mathcal{U}) &= \sum_{i=1}^{M} \Pr[B_i(t) > 0 | \mathcal{H}(t)] \, \mathbb{1}(U_i \in \mathcal{U}) \\
&= K - \sum_{i \in \mathcal{U}} \omega_{i,0}(t).
\end{aligned}
\tag{6.46}
$$

The performance of interest is the infinite-horizon throughput (6.28).

The MP (6.22), specialized to the immediate reward (6.46), becomes

$$\mathcal{U}^{MP}(t) = \underset{\mathcal{U}}{\operatorname{argmax}} R(\boldsymbol{\omega}_1(t), ..., \boldsymbol{\omega}_M(t), \mathcal{U}) = \underset{\mathcal{U}}{\operatorname{argmin}} \sum_{i \in \mathcal{U}} \omega_{i,0}(t). \qquad (6.47)$$

Note that, unlike Section 6.3.1, when $C > 1$ the MP does not generally have a RR structure.

6.6.2 Upper Bound

This section presents an upper bound to the throughput (6.28) by following the approach for general RMAB problems proposed in [38]. The upper bound relaxes the constraint that exactly K nodes must be scheduled in each slot. Specifically, it allows a variable number $K^\pi(t)$ of scheduled nodes in each tth slot under policy π, with the only constraint that its discounted average satisfies

$$E^\pi \left[\sum_{t=1}^{\infty} \beta^{t-1} K^\pi(t) \right] = \frac{K}{1 - \beta}. \qquad (6.48)$$

The advantage of this relaxed version of the scheduling problem is that it can be tackled by focusing on each single arm independently from the others [38, 74]. This is because, by the symmetry of the nodes, the constraint (6.48) can be equivalently handled by imposing that each node is active on average for a discounted time $E^\pi[\sum_{t=1}^{\infty} \beta^{t-1} 1(U_i \in \mathcal{U}^\pi(t))] = \frac{K}{M(1-\beta)}$. It is thus possible to calculate the optimal solution of the relaxed problem by solving a single RSAB problem.

The RSAB model is now elaborated by dropping the node index. Here, the immediate reward when the arm is in state $\boldsymbol{\omega}$ (a vector since $C > 1$, see Section 6.6.1), and action $u \in \{0, 1\}$ is chosen, is $R(\boldsymbol{\omega}, u) = 1 - \omega_0$ if $u = 1$ and $R(\boldsymbol{\omega}, u) = 0$ if $u = 0$, while the Markov evolution of the belief follows from Figure 6.4 and similarly to Section 6.2.2. The problem consists in optimizing the throughput under the constraint $E^\pi[\sum_{t=1}^{\infty} \beta^{t-1} 1(U_i \in \mathcal{U}^\pi(t))] = \sum_{t=1}^{\infty} \beta^{t-1} E^\pi[u^\pi(t)] = K/(M(1 - \beta))$, as introduced

above. Under the assumption that the state $\boldsymbol{\omega}$ belongs to a finite state space $\mathcal{W}$ (to be discussed below), this optimization can be done by resorting to a linear programming (LP) formulation [74]. Specifically, let $z_{\boldsymbol{\omega}}^{(u)}$ be the probability of being in state $\boldsymbol{\omega}$ and selecting action $u \in \{0, 1\}$ under a given policy. The optimization at hand leads to the following LP

$$\text{maximize} \quad \sum_{\boldsymbol{\omega}, u} R(\boldsymbol{\omega}, u) z_{\boldsymbol{\omega}}^{(u)}, \tag{6.49a}$$

$$\text{subject to :} \sum_{\boldsymbol{\omega}, u} z_{\boldsymbol{\omega}}^{(u)} \quad = \quad 1, \tag{6.49b}$$

$$\sum_{\boldsymbol{\omega}} z_{\boldsymbol{\omega}}^{(1)} \quad = \quad \frac{K}{M(1 - \beta)}, \tag{6.49c}$$

$$z_{\boldsymbol{\omega}}^{(0)} + z_{\boldsymbol{\omega}}^{(1)} \quad = \quad \delta\left(\boldsymbol{\omega} - \boldsymbol{\omega}(1)\right) + \beta \sum_{\boldsymbol{\omega}', u} z_{\boldsymbol{\omega}'}^{(u)} p_{\boldsymbol{\omega}\boldsymbol{\omega}'}^{(u)}, \tag{6.49d}$$

$$\text{for all} \ \boldsymbol{\omega} \in \mathcal{W},$$

where (6.49c) is the constraint on the average time in which the node is scheduled, while (6.49d) guarantees that $z_{\boldsymbol{\omega}}^{(u)}$ is the stationary distribution [74], in which $\delta\left(\boldsymbol{\omega} - \boldsymbol{\omega}(1)\right) = 1$ if $\boldsymbol{\omega} = \boldsymbol{\omega}(1)$ and $\delta\left(\boldsymbol{\omega} - \boldsymbol{\omega}(1)\right) = 0$ if $\boldsymbol{\omega} \neq \boldsymbol{\omega}(1)$. Note that, as discussed in Section 6.2.2, the term $p_{\boldsymbol{\omega}\boldsymbol{\omega}'}^{(u)}$ is the probability that the next state is $\boldsymbol{\omega}'$ given that action u is taken in state $\boldsymbol{\omega}$.

It is now left to discuss the cardinality of the set $\mathcal{W}$. While the belief $\boldsymbol{\omega}$ can generally assume any value in the C-dimensional probability simplex, the number of states actually assumed by ω during any *limited* horizon of time is finite due to the finiteness of the action space [67]. In the problem of this section, since the time horizon is unlimited, this fact alone is not sufficient to conclude that the set $\mathcal{W}$ is finite. However, after each tth slot in which the arm is activated, the belief at the $(t+1)$th slot can only takes C values given that the battery state is learned by the FC. Therefore, the evolution of the belief is reset after each activation, and in practice, the time between two activations is finite since the node must be kept active for a discounted fraction of time $K/\left(M(1 - \beta)\right)$. Hence, by constraining the maximum

time interval between two activations to a sufficiently large value, the state space $\mathcal{W}$ remains finite and the optimal performance is not affected. The latter approach has been used for the numerical evaluation of the upper bound in Section 6.6.3.

6.6.3 Numerical Results

This section presents some numerical results to compare the performance of the MP with the upper bound of Section 6.6.2. The performance is the throughput (6.28) normalized by its ideal value $K/(1-\beta)$ that is obtained if the nodes always have energy in their batteries when scheduled.

In Figure 6.5 it is shown the normalized throughput versus the battery capacity C for different ratio M/K between the number M of nodes and the number K of nodes scheduled in each slot. The value $K = 3$ is kept fixed while M varies. It is assumed a uniform distribution for the initial energy in the batteries $B_i(1)$ for all the nodes, so that $\omega_{i,k}(1) = 1/(C+1)$ for all i, k. The probabilities that an energy unit is harvested when the arm is kept passive are $p_{01}^{(0)} = 0.15$ and $p_{kk+1}^{(0)} = 0.1$, for $k \in \{1, C-1\}$, while under activation they are $p_{01}^{(1)} = 0.05$ and $p_{kk+1}^{(1)} = 0$. The probability that an energy unit is lost when the arm is kept passive and activated are $p_{kk-1}^{(0)} = 0.05$ and $p_{kk-1}^{(1)} = 0.95$, respectively. The remaining transitions probabilities are $p_{CC}^{(0)} = 0.9$, $p_{CC}^{(1)} = 0.05$, while $\beta = 0.95$.

From Figure 6.5 it can be seen that when C and/or M/K are small the MP's performance is close to the upper bound. In fact, for small M/K, most of the nodes are scheduled in each slot and the relaxed system in Section 6.6.2 approaches the original one, while for small C the scenario gets closer to the optimality of the MP for $C = 1$. For moderate to large values of M/K and/or C instead, the more flexibility in the relaxed system enables larger gains over the MP.
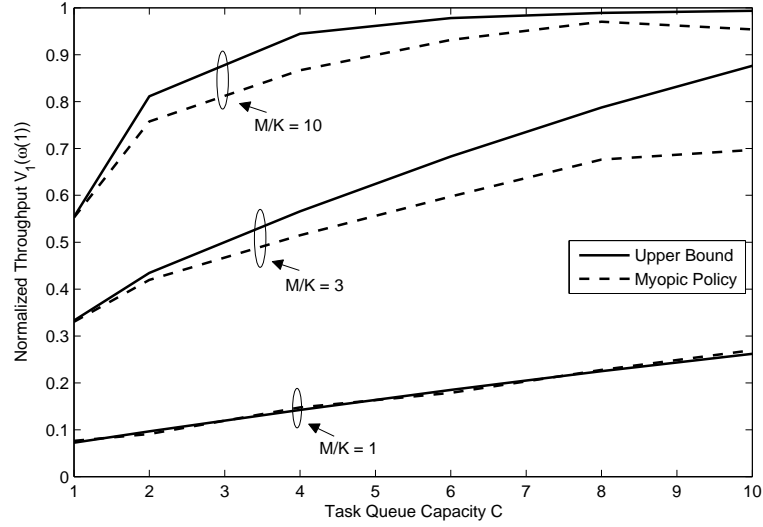
Figure 6.5 Normalized optimal throughput of the MP in (6.47) as compared to the upper bound versus the battery capacity C for different ratios $M/K \in \{1, 3, 10\}$ (system parameters are $K = 3$, $\beta = 0.95$, $\omega_{i,k}(1) = 1/(C+1)$ for all i, k, $p_{01}^{(0)} = 0.15$, $p_{01}^{(1)} = 0.05$, $p_{CC}^{(0)} = 0.9$, $p_{CC}^{(1)} = 0.05$, $p_{kk-1}^{(0)} = 0.05$, $p_{kk-1}^{(1)} = 0.95$, $p_{kk+1}^{(0)} = 0.1$, $p_{kk+1}^{(1)} = 0$, for $k \in \{1, C-1\}$).

6.7 Conclusions

This chapter considered a scheduling problem with applications to energy harvesting (EH) networks, where a fusion center (FC) schedules a set of nodes uncertainties on the energy available at each node. EH and battery leakage are accounted for via simple Markov models. The problem is formulated as a partially observable Markov decision process (POMDP), and converted into a restless multi-armed bandit (RMAB) problem. Under the assumption that the battery capacity is unitary, a stationary myopic policy (MP) that operates in the space of the a posteriori probabilities (beliefs) of the battery levels is proved to be optimal for both finite horizon and (discounted) infinite-horizon throughput criteria. The MP selects at each time-slot the nodes with the largest probability of having enough energy to transmit. It is shown that such policy is round-robin in the sense that it schedules nodes periodically. Closed-form expressions for the optimal throughput performance metrics are also derived. Finally, it has been established that the considered RMAB

problem is indexable and the Whittle index has been derived in closed form. From the expression of the Whittle index, it is concluded that the Whittle index policy is equivalent to the MP and thus is optimal. Arbitrary battery capacities have been investigated as well by comparing the performance of a generally suboptimal myopic policy with an upper bound based on a relaxation approach of the scheduling constraint.

CHAPTER 7

CONCLUSIONS

Energy harvesting (EH) technologies represent nowadays a promising solution to mitigate the energy footprint of wireless communications. This is especially true for mobile applications, where the nodes rely on batteries for their operations, such as cell phones or wireless sensors. EH technologies thus provides not only a support for batteries to reduce their maintenance requirements, but they also enable the deployment of electronic devices that are exclusively powered via EH, for which the maintenance is virtually unnecessary. However, unlike battery-powered nodes, the main drawback of EH-devices is due to the fact that they generally depend on unpredictable energy sources, and hence they call for the development of energy management strategies that need to be designed so as to be robust to uncertainties in energy availability.

While most previous work on EH-capable systems has focused on energy management for single device, the main contributions of this dissertation has been the analysis and design of medium access control (MAC) protocols for EH networks. In particular, two main categories of MAC protocols have been considered: Random access and centralized scheduling-based schemes. Within this framework, the new trade-offs enabled by EH have been investigated for random MAC protocols conventionally used in wireless networks, such as framed-ALOHA and dynamic framed-ALOHA. Furthermore, a novel random MAC protocol, tailored to EH networks, has been proposed and shown to outperform conventional solutions. Moreover, centralized scheduling-based MAC protocols have been investigated under several system settings, including scenarios in which the networks is operated either by nodes exclusively powered via EH or by nodes powered by a hybrid energy storage

system, composed by a non-rechargeable battery and a capacitor recharged via EH. Optimal scheduling policies have been derived under several scenarios and for different level of information available at the central scheduler.

This dissertation also considered the design of energy management techniques for single-device networks within the framework of EH-enhanced RFID systems. Here, it has been shown that a careful utilization of the energy harvested from the environment can lead to remarkable performance gain with respect to conventional systems.

Overall, this dissertation shown that wireless networks can greatly benefit from the adoption of EH technologies. However, to fully exploit their potentialities, the novel challenges that they introduce should be tackled by designing smart energy management algorithms that are tailored to the EH dynamics.

This section calculates the probability $p_{c,k}$ that a node transmits successfully within the kth frame after it has collided in all the previous $(k-1)$ frames. Since all the M nodes are stochastically symmetric, it is possible to focus on the mth node U_m without loss of generality. In the remaining of this section the IR index is dropped as the derivations below are valid for any nth IR.

Let h_m be the channel gain of node U_m during the current IR. According to the assumptions in Section 3.2, channel gains are independent and identically distributed (i.i.d.) across nodes, and they are constant within the whole IR (i.e., they are "fixed" at the beginning of the IR and they remain constant across all the frames in which a node transmits).

The (unconditional) probability density function (pdf) of the channel gain h_m is $f_h(\cdot)$. Let $\mathcal{B}_k \subseteq \{1,...,M\}$ be the set of nodes that are active at frame k (i.e., the *backlog* for frame k). $\mathcal{B}_k$ contains all the nodes that: i) have a new measure to transmit within the current IR; ii) collided in all the first $(k-1)$ frames; iii) have enough residual energy to transmit in frame k. Let $\mathcal{C}_k$ be the set of node U_i that collided in all the first $(k-1)$ frames and let $\mathcal{N}$ be the set of all nodes that have a measure to report to the FC within the considered IR. Then sets $\mathcal{B}_k$, for $k = 1,...,F_\varepsilon$, are defined as

$$
\begin{aligned}
\mathcal{B}_1 &= \left\{ U_m : U_m \in \mathcal{N}, E_m \geq \varepsilon \right\}, \\
&\vdots \\
\mathcal{B}_k &= \left\{ U_m : U_m \in \mathcal{N} \cap C_{k-1}, E_m \geq k\varepsilon \right\}, \\
&\vdots \\
\mathcal{B}_{F_\varepsilon} &= \left\{ U_m : U_m \in \mathcal{N} \cap C_{F_\varepsilon-1}, E_m \geq F_\varepsilon \varepsilon \right\},
\end{aligned}
$$

with $m \in \{1, ..., M\}$. Note that $\mathcal{B}_k \subseteq \mathcal{B}_l$ for any $l \leq k \leq F_\varepsilon$, where F_ε is the maximum number of frames to which a node can participate (i.e., the normalized capacity of the ESD as defined in Section 3.2.2).

Now assume that $U_m \in \mathcal{B}_k$, then let $\mathcal{I}_{m,k} \subseteq \mathcal{B}_k \backslash \{U_m\}$ be the set of other active nodes that select the same slot of node U_m within the kth frame. The probability that U_m transmits successfully within the kth frame in the selected slot, when there are $|\mathcal{I}_{m,k}| = j$ interfering nodes, is given by

$$p_{c,k}(j) = \Pr[\gamma_m \geq \gamma_{th}|(\mathcal{I}_{m,k}, \mathcal{B}_k) \text{ such that } U_m \in \mathcal{B}_k, |\mathcal{I}_{m,k}| = j, \mathcal{I}_{m,k} \subseteq \mathcal{B}_k \backslash \{U_m\}],$$
(A.1)

where the signal-to-interference ration (SIR) γ_m is

$$\gamma_m = \frac{h_m}{\sum_{i \in \mathcal{I}_{m,k}} h_i},$$
(A.2)

and

$$p_{c,k} = \sum_j p_{c,k}(j) \Pr\left[(\mathcal{I}_{m,k}, \mathcal{B}_k) \text{ such that } U_m \in \mathcal{B}_k, |\mathcal{I}_{m,k}| = j, \mathcal{I}_{m,k} \subseteq \mathcal{B}_k \backslash \{U_m\}\right].$$
(A.3)

Note that the conditioning on the fact that $\mathcal{I}_{m,k} \subseteq \mathcal{B}_k$ in (A.1) indicates that all the interfering nodes for U_m, in frame $k \geq 2$, have also collided in all the previous $(k-1)$ frames, and still have residual energy for transmission. This implies that any node within $\mathcal{B}_k$ might have already collided with some of the other nodes within $\mathcal{B}_k$, and hence the channel gains h_i, for any $U_i \in \mathcal{B}_k$, become generally correlated with each other for any $k \geq 2$. Instead, this is not the case for $k = 1$, since there has not been any interaction among nodes.

Due to the reasons explained above, computing probabilities (A.3) for any $k \geq 2$ is quite cumbersome, while it is possible to obtain closed form solutions for $k = 1$ (see, e.g., [58] when $f_h(\cdot)$ is exponential). However, under the large backlog assumption $\mathcal{A}.2$ considered in Chapter 3, evaluating (A.1) simplifies to a manageable

level. In fact, when the number of nodes in the system is large, the channel gains remain independent for any $k \geq 1$ as it is unlikely that any two nodes collide with each other in more than one frame within the same IR. Nevertheless, while the channel gains remain independent, the pdf of each channel gain changes, with respect to the initial one $f_h(\cdot)$, when conditioned as in (A.1).

To summarize, under assumption $\mathcal{A}.2$ the conditional probabilities (A.1) only depend on: i) the number of interfering nodes; ii) the approximated pdf $f_h^{(k)}(\cdot)$ of the channel gain of any node U_m conditioned on having collided in all the first $(k-1)$ frames (i.e., for any $U_m \in \mathcal{B}_k$). Note that, pdf $f_h^{(k)}(\cdot)$ is calculated for any k under assumption $\mathcal{A}.2$ as it will be clarified below, while $f_h^{(1)}(\cdot) = f_h(\cdot)$ as no approximations are required for the first frame as described above.

It is now possible to define random variables $\tilde{h}_1^{(k)}, ..., \tilde{h}_M^{(k)}$ that are i.i.d. with pdf $f_h^{(k)}(\cdot)$ for any $k \in \{1, ..., F_\varepsilon\}$. Roughly speaking, under the approximation described above, random variable $\tilde{h}_i^{(k)}$ is used to represent the channel gain of node U_i at frame k when U_i collided in all of the first $(k-1)$ frames. It is possible to approximate (A.1) as

$$p_{c,k}(j) \simeq \Pr[\tilde{h}_m^{(k)} \geq \gamma_{th} \sum_{\substack{i=1 \\ m \notin \{1,...,j\}}}^{j} \tilde{h}_i^{(k)}], \tag{A.4}$$

Note that, (A.4) is exact for $k = 1$ as $f_h^{(1)}(\cdot) = f_h(\cdot)$.

Moreover, under assumption $\mathcal{A}.2$, (A.3) becomes

$$p_{c,k} \simeq e^{-\frac{1}{\rho}} \sum_{j=0}^{\infty} \frac{1}{\rho^j j!} p_{c,k}(j) \tag{A.5}$$

which is due to the Poisson approximation of the binomial distribution (as described in Section 3.5), which implies that

$$\Pr\left[(\mathcal{I}_{m,k}, \mathcal{B}_k) \text{ such that } U_m \in \mathcal{B}_k, |\mathcal{I}_{m,k}| = j, \mathcal{I}_{m,k} \subseteq \mathcal{B}_k \backslash \{U_m\}\right] \simeq \frac{e^{-\frac{1}{\rho}}}{\rho^j j!}. \tag{A.6}$$

In fact, under assumption $\mathcal{A}.2$, set $\mathcal{B}_k \backslash \{U_m\}$ contains a large number of elements, and thus the probability that j nodes within $\mathcal{B}_k \backslash \{U_m\}$ select the same slot of U_m in a frame of size $L = \lceil \rho |\mathcal{B}_k| \rceil$ can be approximate as $\frac{e^{-\frac{1}{\rho}}}{\rho^j j!}$, which is independent of the actual backlog size $|\mathcal{B}_k|$.

In order to evaluate (A.4)-(A.5) one thus needs to calculate

$$\Pr\left[\tilde{h}_m^{(k)} \geq \gamma_{th} x^{(k)}(j)\right] = \Pr[\tilde{h}_m^{(k)} \geq \gamma_{th} \sum_{\substack{i=1 \\ m \notin \{1,\ldots,j\}}}^{j} \tilde{h}_i^{(k)}], \tag{A.7}$$

where

$$x^{(k)}(j) = \sum_{\substack{i=1 \\ m \notin \{1,\ldots,j\}}}^{j} \tilde{h}_i^{(k)}. \tag{A.8}$$

Recalling that gains $\tilde{h}_i^{(k)}$ are independent under $\mathcal{A}.2$, the pdf $f_{x^{(k)}(j)}(\cdot)$ of the *cumulative interference* $x^{(k)}(j)$ can be calculated as the convolution of $f_h^{(k)}(\cdot)$ with itself for j times. Note also that $x^{(k)}(j)$ and $\tilde{h}_i^{(k)}$ are independent each other.

A.1 Computation of the pdf $f_h^{(k)}(\cdot)$ of Random Variables $\tilde{h}_i^{(k)}$

It is now left to calculate the pdf $f_h^{(k)}(\cdot)$ of the channel gains, or more precisely of random variables $\tilde{h}_i^{(k)}$, for each kth frame. It is possible to proceed as follows. Let p be a random variable with Poisson distribution of parameter $1/\rho$ so that

$$\Pr\left[p = j\right] = \frac{e^{-\frac{1}{\rho}}}{\rho^j j!}. \tag{A.9}$$

Random variable p can be thought of representing the number of nodes interfering node U_m, which has already been shown as being Poisson-distributed under assumption $\mathcal{A}.2$ (see (A.6)). Now, the cumulative density function (cdf) $F_h^{(k)}(z)$

of gain $\tilde{h}_m^{(k)}$ can be calculated as

$$
\begin{aligned}
F_h^{(k)}(z) &= \Pr\left[\tilde{h}_m^{(k)} < z\right] \\
&= \sum_{j=0}^{\infty} \Pr\left[\tilde{h}_m^{(k-1)} < z | \tilde{h}_m^{(k-1)} < \gamma_{th} x^{(k-1)}(j)\right] \Pr\left[p = j | p \geq 1\right] \quad \text{(A.10a)} \\
&= \frac{1}{e^{\frac{1}{\rho}} - 1} \sum_{j=1}^{\infty} \frac{1}{\rho^j j!} \Pr\left[\tilde{h}_m^{(k-1)} < z | \tilde{h}_m^{(k-1)} < \gamma_{th} x^{(k-1)}(j)\right], \quad \text{(A.10b)}
\end{aligned}
$$

where in (A.10a) one has conditioned on having at least one interfering node, otherwise there would not be collision. This is accounted for through the term $\Pr\left[p = j | p \geq 1\right]$, which can be calculated through the Bayes rule as

$$
\Pr\left[p = j | p \geq 1\right] = \Pr\left[p \geq 1 | p = j\right] \frac{\Pr[p = j]}{\Pr[p \geq 1]} = \begin{cases} \frac{\Pr[p=j]}{\Pr[p\geq 1]} = \frac{e^{-\frac{1}{\rho}}/\rho^j j!}{1 - e^{-\frac{1}{\rho}}} & j \geq 1 \\ 0 & j = 0 \end{cases},
$$
(A.11)

since $\Pr\left[p \geq 1\right] = 1 - \Pr[p = 0] = 1 - e^{-\frac{1}{\rho}}$. Eq. (A.10b) then follows immediately from (A.10a) and (A.11) by considering that $e^{-\frac{1}{\rho}} / \left(1 - e^{-\frac{1}{\rho}}\right) = 1/\left(e^{\frac{1}{\rho}} - 1\right)$.

To calculate the conditional probability $\Pr\left[\tilde{h}_m^{(k-1)} < z | \tilde{h}_m^{(k-1)} < \gamma_{th} x^{(k-1)}(j)\right]$ it is possible to proceed as follows

$$
\Pr\left[\tilde{h}_m^{(k-1)} < z | \tilde{h}_m^{(k-1)} < \gamma_{th} x^{(k-1)}(j)\right] = \frac{\Pr\left[\tilde{h}_m^{(k-1)} < z, \tilde{h}_m^{(k-1)} < \gamma_{th} x^{(k-1)}(j)\right]}{\Pr\left[\tilde{h}_m^{(k-1)} < \gamma_{th} x^{(k-1)}(j)\right]},
$$
(A.12)

which can be obtained by integrating the joint pdf of $\tilde{h}_m^{(k-1)}$ and $x^{(k-1)}(j)$ over a convenient 2-D region with standard techniques. Note that the joint pdf between $\tilde{h}_m^{(k-1)}$ and $x^{(k-1)}(j)$ is simply the product between $f_h^{(k)}(\cdot)$ and $f_{x^{(k)}(j)}(\cdot)$ as $\tilde{h}_m^{(k-1)}$ and $x^{(k-1)}(j)$ are independent each other.

Finally, once $F_h^{(k)}(\cdot)$ is obtained, one can easily calculate $f_h^{(k)}(\cdot)$ and then probabilities (A.5) still by integration over convenient regions.

APPENDIX B

NETWORK LIFETIME CALCULATION FOR $K = 1$

This section proposes a procedure to calculate the network lifetime with reduced complexity for the full state information scenario when the following two conditions apply: i) the number of scheduled node in each slot is $K = 1$; ii) the adopted policy π is stationary and always schedules in each slot a node with a non-empty capacitor if at least one is available, otherwise it operates according to steps 2) and 3) of the MCF policy in Definition 6. In other words, under the conditions i) and ii) above, the FC keeps scheduling in each slot a node that draws energy from its capacitor for transmission, until all the nodes have their capacitors simultaneously empty. When this occurs a node that draws energy from its battery is scheduled. Accordingly, the network lifetime for such policies can be calculated by accounting for the contribution due to capacitors (and thus due to EH) and due to batteries separately. Specifically, for any initial state of the energy stored in the capacitors, it is possible to calculate the average time before that all the capacitors become empty, i.e., $\mathbf{c}(\cdot) = \mathbf{0} = [0, ..., 0]$, by resorting to an absorbing Markov chain model, in which the absorbing state is $\mathbf{c}(\cdot) = \mathbf{0}$. When the absorbing state is reached, then a node that draws energy from its battery is scheduled, while in the meantime the capacitors are possibly recharged via EH.

To elaborate on this point, it is possible to define the *first hitting time* $\tau^\pi(\mathbf{c}(1))$ for the absorbing state $\mathbf{0}$, when the starting state is $\mathbf{c}(1)$ and policy π is followed, as

$$\tau^\pi(\mathbf{c}(1)) = inf \left\{ t \geq 1 : \mathbf{c}(t) = \mathbf{0} | \mathbf{c}(1) \right\}, \tag{B.1}$$

while its average value can be calculated as

$$T_{\mathbf{0}}(\mathbf{c}(1)) = E^\pi \left[\tau^\pi(\mathbf{c}) \right]. \tag{B.2}$$

Note that, the average value in (B.2) is calculated with respect to the distribution (5.8), where only the marginal distribution with respect to the energy in the capacitors is considered. Moreover, the average time (B.2) can be calculated by resorting to standard techniques utilized for absorbing Markov chain [47].

When all the capacitors become simultaneously empty and the system state is not terminal (i.e., $(\mathbf{c} = \mathbf{0}, \mathbf{b}) \notin \mathcal{T}$), the node that is scheduled by the FC transmits by using its battery. Therefore, the energy in the battery of the scheduled node is decremented by one, while the capacitors of all the nodes possibly harvest energy. Now, recalling that the probability distribution of the next slot's state is

$$\Pr\left[\mathbf{c}', \mathbf{b}' | \left(\mathbf{0}, \mathbf{b}\right), \mathcal{U}^\pi\left(\cdot\right)\right] \tag{B.3}$$

(see eq. (5.8)), then the next slot in which all the capacitors become simultaneously empty again, given initial state $(\mathbf{0}, \mathbf{b})$, occurs on average after

$$\sum_{\mathbf{c}'} T_{\mathbf{0}}(\mathbf{c}') \Pr\left[\mathbf{c}', \mathbf{b}' | \left(\mathbf{0}, \mathbf{b}\right), \mathcal{U}^\pi\left(\cdot\right)\right] \tag{B.4}$$

slots, where $T_{\mathbf{0}}(\mathbf{c}')$ is (B.2) with $\mathbf{c}(1) = \mathbf{c}'$. In other words, after the slot in which the scheduled node uses its battery, the capacitors harvest energy and their overall state at the next slot is $\mathbf{c}'$ with probability $\Pr\left[\mathbf{c}', \mathbf{b}' | \left(\mathbf{0}, \mathbf{b}\right), \mathcal{U}^\pi\left(\cdot\right)\right]$. Therefore, by averaging over the possible next states $\mathbf{c}'$ it is possible to calculate the average time after which the capacitors will be simultaneously empty again. Note that, the marginal distribution of $\mathbf{c}'$ obtained through $\Pr\left[\mathbf{c}', \mathbf{b}' | \left(\mathbf{0}, \mathbf{b}\right), \mathcal{U}^\pi\left(\cdot\right)\right]$ is always the same for any non terminal state $(\mathbf{0}, \mathbf{b}) \notin \mathcal{T}$, and thus the same distribution is repeated after each slot in which a battery is used. Now, let $\mathbf{b}(1)$ be the initial state of the energy in the batteries, and let $\tilde{L}\left(\mathbf{b}(1)\right) = \sum_{i=1}^{M} 1[b_i \geq 1]\left(b_i - 1\right) + 1$ be the (optimal) network lifetime if there were no capacitors in the system, and thus no harvesting and no leakage. The term $1[b_i \geq 1]\left(b_i - 1\right)$ is a consequence of the definition of terminal set in eq. (5.3), since a node with battery $b_i = 0$ implies that the network is in the

terminal set. Finally, the network lifetime under any policy π (defined as in point ii) above) can be calculated as

$$
\begin{aligned}
L^\pi(\mathbf{c}(1), \mathbf{b}(1)) \;=\;\; & \tilde{L}\left(\mathbf{b}(1)\right) + \\
& T_{\mathbf{0}}(\mathbf{c}(1)) + \left(\tilde{L}\left(\mathbf{b}(1)\right) - 1\right) \sum_{\mathbf{c}'} T_{\mathbf{0}}(\mathbf{c}') \Pr\left[\mathbf{c}', \mathbf{b}' \middle| (\mathbf{0}, \mathbf{b}), \mathcal{U}^\pi\left(\cdot\right)\right].
\end{aligned}
\tag{B.5}
$$

APPENDIX C

UPPER BOUND OF THE NETWORK LIFETIME

This section proposes an upper bound for the network lifetime in eq. (5.5) that can be useful when the system size is large and the evaluation of the optimal lifetime via dynamic programming tools becomes prohibitive (see Section 5.2.1). The proposed upper bound is obtained by considering a system with a *super-node U_{UB}* (i.e., $M = 1$) that is equipped with one capacitor C_{UB} and one battery B_{UB}. The size E_c^{UB} of C_{UB} and the size E_b^{UB} of B_{UB} are equal to the sum of the size of the capacitors and of the batteries of the M original nodes in the system, i.e., $E_c^{UB} = ME_c$ and $E_b^{UB} = ME_b$, respectively. The energy stored in the capacitor and in the battery at slot t are denoted as $c_{UB}(t)$ and $b_{UB}(t)$, while the state of super-node U_{UB} is $(c_{UB}(t), b_{UB}(t)) \in \mathcal{S}^{UB}$, with $\mathcal{S}^{UB} = \{0, ..., ME_c\} \times \{0, ..., ME_b\}$. The super-node U_{UB} consumes in each slot K energy units for transmission by drawing energy from its capacitor first (see Lemma 5-IV). Note that the size of the space $\mathcal{S}^{UB}$ is quadratic in the number of nodes, namely, $\left|\mathcal{S}^{UB}\right| \propto M^2 E_c E_b$, which is thus numerically tractable even for moderately large systems, differently from the original system that it can be easily shown to be exponentially large in the number of nodes (see Section 5.2.1).

The transition probabilities $\Pr\left[(c', b') | (c, b)\right]$ for the state of the super-node can be identified in such a way that the lifetime of the induced single-node system provides an upper bound on the lifetime of the original system. The main idea is to guarantee that, for any realization of the harvesting and leakage processes, the energy gained by the super-node via harvesting is at least as large as the total energy harvested in the original system, and that the energy lost due to leakage at the super-node is no larger than that in the original system. In other words, the super-node system is stochastically dominant with respect to the original system.

To derive the transition probabilities for the super-node system, let $c_{UB}(t) = c$ for any $c \in \{0, ..., E_c^{UB}\}$ and consider the evolution of the system to slot $t + 1$. Recall that $c_{UB}(t)$ represents the total energy available in the capacitors of the M nodes. In the original system, after that K nodes transmit, there must at least $N_{ne}(c) = \lceil (c - K)^+ / E_c \rceil$ nodes with a non-empty capacitor and at least $N_f(c) = (c - K - M(E_c - 1))^+$ nodes with a full capacitor, where $\lceil \cdot \rceil$ is the upper nearest integer operator. Therefore, the overall increment in the energy available in the original system is upper bounded by

$$e(t|c) = \sum_{j=1}^{M - N_f(c)} h_j(t) - \sum_{j=1}^{N_{ne}(c) - N_f(c)} d_j(t) - \sum_{i=1}^{N_f(c)} \tilde{d}_i(t), \tag{C.1}$$

where $h_j(t)$ and $d_j(t)$ are defined as in Section 5.2.2 and $\tilde{d}_i(t)$ for $i \in \{1, ..., N_f(c)\}$ are independent binary random variables with pmf $\Pr[\tilde{d}_i(t) = -1] = \mu$ and $\Pr[\tilde{d}_i(t) = 0] = 1 - \mu$, where μ is given in eq. (5.7). The first and the third terms in (C.1) account for the fact that nodes with full capacitors cannot harvest energy but can lose energy with probability μ, while the second reflects the fact that nodes with non-empty but non-full capacitors lose an energy unit with probability p_d. The upper bound (C.1) will be used as the increment of energy in the super-node. Note that the energy (C.1) can be negative as well.

The transition probabilities for state $(c_{UB}(t), b_{UB}(t))$ can then be calculated as follows. Let $(c_{UB}(t + 1), b_{UB}(t + 1)) = (c', b')$ be state at slot $t + 1$ and let $(c_{UB}(t), b_{UB}(t)) = (c, b)$ be the state at slot t, then the probability $\Pr[(c', b') | (c, b)]$ is given by

$$\Pr[(c', b') | (c, b)] = \begin{cases} \Pr[e(t|c) = c' - (c - K)^+] & \text{if } b' = b - (K - c)^+ \\ 0 & \text{otherwise} \end{cases} \tag{C.2}$$

where the pmf $\Pr[e(t|c) = k]$ can be easily calculated as the pmf of the sum of the independent random variables that are involved in (C.1). The set of terminal states

for the Markov chain $(c_{UB}(t), b_{UB}(t))$ is defined as the set of all the states (c_{UB}, b_{UB}) for which the energy b_{UB} in the battery B_{UB} is strictly smaller than M, namely, $\mathcal{T}_{UB} = \{(c_{UB}, b_{UB}) : b_{UB}(t) < M\}$. Note that the condition $b_{UB}(t) < M$ simply states that the total energy in the M original batteries is smaller than M, which is a weaker condition than the one considered in Definition 1, where an energy unit is required to be stored in each battery. This condition guarantees that the average time to absorption $L^{UB}(\mathbf{c}, \mathbf{b})$ of the Markov chain $(c_{UB}(t), b_{UB}(t))$ with respect to the terminal set $\mathcal{T}_{UB}$ provides an upper bound for the network lifetime of the original system, where $L^{UB}(\mathbf{c}, \mathbf{b})$ is given by

$$L^{UB}(\mathbf{c}, \mathbf{b}) = \lim_{T \to \infty} E\left[\sum_{t=1}^{T} 1\left[(c_{UB}(t), b_{UB}(t)) \notin \mathcal{T}_{UB}\right] \bigg| (c_{UB}(1), b_{UB}(1))\right], \qquad \text{(C.3)}$$

where $c_{UB}(1) = \sum_{i=1}^{M} c_i$ and $b_{UB}(1) = \sum_{i=1}^{M} b_i$. The limit in (C.3) always exists and it is finite for any $p_h < 1$ for the same reason explained in the proof of Lemma 4. The upper bound $L^{UB}(\mathbf{c}, \mathbf{b})$ can be calculated by resorting to the standard techniques for absorbing Markov chains [47]. It is also emphasized that the upper bound is actually exact when the capacitors are of size one, i.e., $E_c = 1$. In fact, in such a case the number of non-empty and full capacitors coincide $N_{ne}(c) = N_f(c) = (c - K)^+$ for any total energy c, and thus the upper bound is achievable.

APPENDIX D

PROOF OF PROPOSITION 8

This section provides the remaining details for the proof of Proposition 8, whose main ideas were sketched in Section 5.6. To start with, note that the MCF policy selects the nodes with the largest energy in the capacitors, and thus it follows that

$$c_i(1) > c_j(1) \text{ for all } U_i \in \Omega^\gamma(1) \text{ and } U_j \in \Omega^\pi(1). \tag{D.1}$$

Moreover, for a given sample path $h_i(t)$, for $i \in \{1, ..., M\}$, of the EH processes the energies in the capacitors under γ and π at any slot $t \geq 2$ are given by

$$c_i^\gamma(t) = \min\left(c_i(t-1) + h_i(t-1) - 1\left[U_i \in \mathcal{U}^\gamma(t-1)\right], \ E_c\right) \tag{D.2}$$

$$c_i^\pi(t) = \min\left(c_i(t-1) + h_i(t-1) - 1\left[U_i \in \mathcal{U}^\pi(t-1)\right], \ E_c\right). \tag{D.3}$$

Moreover, the relations between the energies in the capacitors under the two policies at slot $t = 2$ are given, from (D.2) and (D.3), as

$$c_i^\gamma(2) = \begin{cases} c_i^\pi(2) + 1 & \text{for all } U_i \in \Omega^\pi(1) \\ c_i^\pi(2) & \text{for all } U_i \in \Omega^\gamma(1), \text{ such that } c_i(1) + h_i(1) > E_c \\ c_i^\pi(2) - 1 & \text{for all } U_i \in \Omega^\gamma(1), \text{ such that } c_i(1) + h_i(1) \leq E_c \end{cases} \tag{D.4}$$

$$c_i^\pi(2) \geq c_j^\gamma(2) \text{ for all } U_i \in \Omega^\gamma(1) \text{ and } U_j \in \Omega^\pi(1). \tag{D.5}$$

The first line of (D.4) accounts for the fact that there cannot be energy overflows under policy γ at the end of slot $t = 1$ for nodes $U_i \in \Omega^\pi(1)$ since $c_i(1) < E_c$ due to (D.1); while the second and third lines of (D.4) account for possible energy overflows under policy π for nodes $U_i \in \Omega^\gamma(1)$. Note that, in case an energy overflow occurs under policy π at any node $U_i \in \Omega^\pi(1)$ between slots t and $t+1$, then this node will

have the same stored energy under the two policies at slot $t = 2$, i.e., $c_i^\gamma(2) = c_i^\pi(2)$, and it will be put into the common set $\Omega^c(2)$. Finally, inequality (D.5) is an updated version of (D.1).

Since equivalent nodes, due to either energy overflows or index permutation, are put into the common set, and since the sample path of the EH processes is fixed and common under the two policies, then it is possible to rewrite the relations (D.4), for any arbitrary slot t, as a function of the sets $\Omega^\gamma(t)$ and $\Omega^\pi(t)$ only. It results

$$c_i^\gamma(t) = \begin{cases} c_i^\pi(t) + 1 & \text{for all } U_i \in \Omega^\pi(t) \\ c_i^\pi(t) - 1 & \text{for all } U_i \in \Omega^\gamma(t) \end{cases}. \tag{D.6}$$

It is now evident from (D.6) that an energy overflow at some node $U_i \in \Omega^\gamma(t)$ or $U_i \in \Omega^\pi(t)$, between slots t and $t + 1$, implies that $c_i^\gamma(t + 1) = c_i^\pi(t + 1)$ (due to (D.2)-(D.3)), and thus that node U_i becomes equivalent under the two policies γ and π starting from slot $t + 1$. Furthermore, when a switch event between policy γ and π occurs, then the nodes involved in the switch become equivalent under the two policies (for nodes in both sets $\Omega^\gamma(t)$ and $\Omega^\pi(t)$, see (D.6)). Equivalent nodes are then put into the common set.

All the elements are now available to show that the case iii) of Section 5.6 (i.e., that $|\Omega^\gamma(t)| > |\Omega^\pi(t)| = 0$) can never occur, and thus that policy γ has a lifetime never smaller than policy π for any sample path of the EH processes. The key to prove this result is to observe that no energy overflow can occur at any node $U_i \in \Omega^\pi(t)$ under γ before that such node becomes equivalent (possibly upon an index permutation) to another node $U_j \in \Omega^\gamma(t)$. This can be recognized by looking at the inequality (D.5) and the capacitor updating rules (D.2)-(D.3). To do so, assume for simplicity that policy γ always schedules the same nodes of policy π with no switches (i.e., policy π does not schedule any node in set $\Omega^\gamma(t)$) and assume that no energy overflows occur at any node in the set $\Omega^\gamma(t)$ under policy π. This is done without

loss of generality since both switches and energy overflows under policy π generate an equivalence between nodes as described above. Now, due to inequality (D.5) at slot $t = 2$, for any realization of the EH at the nodes, to have an overflow at any node $U_i \in \Omega^\pi(t)$ under policy γ, there must be a slot $u \leq t$ in which $c_i^\gamma(u) = c_j^\pi(u)$, for $U_i \in \Omega^\pi(u)$ and some $U_j \in \Omega^\gamma(u)$, which implies that $c_i^\pi(u) = c_i^\gamma(u) - 1$ and $c_j^\gamma(u) = c_j^\pi(u) - 1 = c_i^\gamma(u) - 1$ (due to (D.6)). But if this occurs, then nodes U_i and U_j are equivalent under the two policies γ and π upon an index permutation, so that they are put into the common set $\Omega^c(u)$. Therefore, since there cannot be any energy overflow under policy γ at any node $U_i \in \Omega^\pi(t)$ for all t, before that $U_i \in \Omega^\pi(t)$ becomes equivalent to another node, then it must hold that $|\Omega^\pi(t)| \geq |\Omega^\gamma(t)|$ for any t, and consequently, the event $iii)$ above can never occur, and thus policy γ has a lifetime never smaller than policy π for any realization of the EH processes. This proves the inequalities (5.17) and (5.18) and thus that the MCF is optimal.

Optimality of the LCF policy in the leakage-only scenario. The proof leverages the same technique used above for Part a). Hence, this section only sketches the main difference between the two proofs. Specifically, in the proof of the optimality of the MCF policy above, the key was to show that there cannot be energy overflows under policy γ, for any node in the set $U_i \in \Omega^\pi(t)$, before that U_i becomes equivalent to another node in the set $\Omega^\gamma(t)$. The complementary event of the energy overflows in the leakage-only scenario is the case in which $c_i^\gamma(t) = 1$ and $c_i^\pi(t) = 0$ for $U_i \in \Omega^\pi(t)$, and the common leakage is $d_i(t) = 1$. In such a case, node U_i would lose an energy unit under policy γ but not under π. However, by following steps similar to the ones taken in Part a) it can be easily shown that this can never occur before that an equivalence between two nodes occur, and thus it can be concluded that the LCF policy is optimal in the leakage-only scenario.

APPENDIX E

THROUGHPUT OF THE MYOPIC POLICY

By exploiting the RR structure of the MP (Proposition 12), it is possible to derive the throughput $V_1^{MP}(\omega(1))$ of the MP as the sum of the contributions of each node separately. To elaborate, focus on node U_i, with initial belief $\omega_i(1)$, and assume that $U_i \in \mathcal{G}_1$. Nodes in group $\mathcal{G}_1$ are scheduled at slots $t = 1 + (j-1)m$, for $j \in \{1, 2, ...\}$. Let $r_i(j) = \mathrm{E}^{MP}[\omega_i(1+(j-1)m)|\omega_i(1)]$ be the average reward accrued by the FC from node U_i only, when scheduling it for the jth time at slot $t = 1+(j-1)m$ (see the RHS of eq. (6.14)). At slot $t = 1$ it results $r_i(1) = \omega_i(1)$. To calculate $r_i(2)$ it is possible to first derive the average value of the belief (see (6.8)) after the slot of activity in $t = 1$ as $\mathrm{E}^{MP}[\omega_i(2)|\omega_i(1)] = \tau_1^{(1)}(\omega_i(1))$, where $\tau_1^{(1)} = \omega\delta_1 + p_{01}^{(1)}$ with $\delta_u = \left(p_{11}^{(u)} - p_{01}^{(u)} \right)$ (cf. eq. (6.9)), then it is possible to account for the $(m-1)$ slots of passivity by exploiting eq. (6.11), so that $r_i(2) = \mathrm{E}^{MP}[\omega_i(1+m)|\omega_i(1)] = \phi^{(1)}(\omega_i(t))$, where the following quantities have been defined $\phi^{(1)}(\omega) = \tau_0^{(m-1)}(\tau_1^{(1)}(\omega)) = \omega\alpha_m + \psi_m$ with $\alpha_m = \delta_1\delta_0^{m-1}$ and $\psi_m = p_{01}^{(1)}\delta_0^{m-1} + p_{01}^{(0)}\frac{1-\delta_0^{m-1}}{1-\delta_0}$. In general, it is possible to obtain $r_i(j) = \mathrm{E}^{MP}[\omega_i(1+(j-1)m)|\omega_i(1)]$, for $j \geq 2$, by iterating the procedure above by applying $\phi^{(1)}(\omega)$ to itself $(j-1)$ times. After a little algebra it follows $\phi^{(j-1)}(\omega) = \phi^{(1)}(\phi^{(j-2)}(\omega)) = \omega\alpha_m^{j-1} + \psi_m\frac{1-\alpha_m^{j-1}}{1-\alpha_m}$, so that $r_i(j) = \phi^{(j-1)}(\omega_i(1))$, where $\phi^{(0)}(\omega) = \omega$. By recalling that a node $U_i \in \mathcal{G}_g$, for $g \geq 1$, is scheduled the first time at slot $t = g$ while its belief is $\omega_i(t = g) = \tau_0^{(g-1)}(\omega_i(1))$ (see eq. (6.8)), and following the same reasoning as above it follows $r_i(j) = \phi^{(j-1)}\left(\tau_0^{(g-1)}(\omega_i(1))\right)$, for any $U_i \in \mathcal{G}_g$.

Let $H_g = \left\lfloor \frac{T-g}{m} \right\rfloor + 1$ be the number of times any node $U_i \in \mathcal{G}_g$ is scheduled in the horizon $t \in \{1, ..., T\}$. The contribution to $V_1^{MP}(\omega(1))$ from $U_i \in \mathcal{G}_g$ is $\beta^{g-1}\sum_{j=1}^{H_g} r_i(j)$. By summing up the contribution from all the groups, after a little

algebra and by substituting terms it follows

$$V_1^{MP}(\omega(1)) = \sum_{g=1}^{m} \beta^{g-1} \theta^{(H_g,K)} \left(\sum_{i=1+(g-1)K}^{gK} \tau_0^{(g-1)}(\omega_i(1)) \right), \qquad (E.1)$$

where $\theta^{(H,K)}(x)$ is the contribution to the whole throughput from K nodes that belong to the same group and whose initial beliefs sum up to σ, with

$$\theta^{(H,K)}(\sigma) = K \frac{\psi_m}{1-\alpha_m} \left(\frac{1-\beta^{mH}}{1-\beta^m} - \frac{1-(\beta^m \alpha_m)^H}{1-\beta^m \alpha_m} \right) + \frac{1-(\beta^m \alpha_m)^H}{1-\beta^m \alpha_m} \sigma. \qquad (E.2)$$

APPENDIX F

PROOF OF LEMMA 14

The proof is by backward induction on t. At $t = T$, inequality (6.27) is easily seen to hold from eq. (6.23). Suppose now that (6.27) holds at slots $t + 1, ..., T$. It is then necessary to show that it also holds at t. To do so, denote as $\mathcal{L}$ and $\mathcal{R}$ the RR policies whose throughputs are given by the LHS and RHS of (6.27), respectively. The differences between $\mathcal{L}$ and $\mathcal{R}$ are the positions of the nodes with belief x and y in the initial belief vectors, being $[\omega_1, ..., \omega_j, y, x, ..., \omega_M]$ and $[\omega_1, ..., \omega_j, x, y, ..., \omega_M]$, respectively. Beliefs y and x occupy the positions $(j + 1)$ and $(j + 2)$ under policy $\mathcal{L}$ and viceversa for policy $\mathcal{R}$. Therefore, some of the m groups created by the two policies might have different nodes (see the RR operations in Proposition 12). Let $\mathcal{G}_g^{\mathcal{L}}$ and $\mathcal{G}_g^{\mathcal{R}}$, for $g \in \{1, ..., m\}$, be the gth group of nodes created by policies $\mathcal{L}$ and $\mathcal{R}$, respectively. To simplify, the node with belief x (y) is referred to as node x (y). It is possible to distinguish three cases. **Case a)** If $j \leq K - 2$, the groups $\mathcal{G}_g^{\mathcal{L}}$ and $\mathcal{G}_g^{\mathcal{R}}$ under policies $\mathcal{L}$ and $\mathcal{R}$ coincide for any $1 \leq g \leq m$, and thus (6.27) holds with equality. **Case b)** If $j \geq K$, neither node x nor node y belong to the first group in both policies $\mathcal{L}$ and $\mathcal{R}$, i.e., $x, y \notin \mathcal{G}_1^{\mathcal{L}}$ and $x, y \notin \mathcal{G}_1^{\mathcal{R}}$. Here, by using eq. (6.24) one can expand (6.27) and promptly verify that the inequality holds due to the induction hypothesis since $\tau_0^{(1)}(x) \geq \tau_0^{(1)}(y)$ due to inequality (6.13). **Case c)** If $j = K - 1$, nodes x and y belong to different groups under policies $\mathcal{L}$ and $\mathcal{R}$, specifically $y \in \mathcal{G}_1^{\mathcal{L}}$, $x \in \mathcal{G}_2^{\mathcal{L}}$ and $y \in \mathcal{G}_2^{\mathcal{R}}$, $x \in \mathcal{G}_1^{\mathcal{R}}$, while the other groups $\mathcal{G}_g^{\mathcal{L}}$ and $\mathcal{G}_g^{\mathcal{R}}$ coincide for $g \in \{3, ..., m\}$. Hence, the only difference between policies $\mathcal{L}$ and $\mathcal{R}$ is the scheduling order of nodes x and y. To verify that inequality (6.27) holds, it is necessary to prove that scheduling node y in the first group and node x in second one is not better than doing the opposite for any $x \geq y$. To start with, let $H_x^{\mathcal{R}}(t) = H_y^{\mathcal{L}}(t)$ and $H_y^{\mathcal{R}}(t) = H_x^{\mathcal{L}}(t)$

be the number of times that node x (or y) is scheduled under policy $\mathcal{R}$ (or $\mathcal{L}$) and node y (or x) is scheduled under policy $\mathcal{L}$ (or $\mathcal{R}$), respectively. Now, recall that the contribution to the whole throughput from a single node that is scheduled H times with initial belief ω can be accounted for separately through (E.2) by setting $K = 1$ and $\sigma = \omega$. Then, the contribution generated by node x and y under policy $\mathcal{R}$ is $\theta^{\left(H_x^{\mathcal{R}}(t),1\right)}(x)$ and $\beta\theta^{\left(H_y^{\mathcal{R}}(t),1\right)}(\tau_0^{(1)}(y))$, respectively, and similarly under policy $\mathcal{L}$ it results $\beta\theta^{\left(H_x^{\mathcal{L}}(t),1\right)}(\tau_0^{(1)}(x))$ and $\theta^{\left(H_y^{\mathcal{L}}(t),1\right)}(y)$. Inequality (6.27) can thus be rewritten as $\theta^{\left(H_x^{\mathcal{R}}(t),1\right)}(x) + \beta\theta^{\left(H_y^{\mathcal{R}}(t),1\right)}(\tau_0^{(1)}(y)) - \beta\theta^{\left(H_x^{\mathcal{L}}(t),1\right)}(\tau_0^{(1)}(x)) - \theta^{\left(H_y^{\mathcal{L}}(t),1\right)}(y) \geq 0$, which must hold for all admissible $H_x^{\mathcal{R}}(t) = H_y^{\mathcal{L}}(t)$ and $H_y^{\mathcal{R}}(t) = H_x^{\mathcal{L}}(t)$. There are two subcases: **c.1)** $H_x^{\mathcal{R}}(t) = H_y^{\mathcal{L}}(t) = H_y^{\mathcal{R}}(t) = H_x^{\mathcal{L}}(t) = H \geq 1$; **c.2)** $H_x^{\mathcal{R}}(t) = H_y^{\mathcal{L}}(t) = H$, and $H_y^{\mathcal{R}}(t) = H_x^{\mathcal{L}}(t) = H - 1$, for $H \geq 1$. By exploiting the RHS of (E.2), after a little algebra one can promptly verify that the inequality above holds in both subcases, which concludes the proof.

APPENDIX G

PROOF OF LEMMA 16

Proof of case a). From equations (6.35)-(6.36), and recalling that $\tau_0^{(1)}(0) = p_{01}$ from eq. (6.41), the leftmost inequality in (6.39a.1) follows immediately as it becomes $V_m(0|1) = \beta V_m^*(p_{01}) \leq m + \beta V_m^*(p_{01}) = V_m(0|0)$. For the rightmost inequality in (6.39a.1), it follows that $V_m(1|1) = 1 + \beta V_m^*(0)$, while from eq. (6.33) and the fact that $V_m(0|1) \leq V_m(0|0)$ it follows $V_m^*(0) = \max\{V_m(0|0), V_m(0|1)\} = V_m(0|0)$. Therefore, it results $V_m(1|1) = 1 + \beta V_m^*(0)1 + \beta V_m(0|0) \geq V_m(0|0)$, which holds since $1 + \beta V_m(0|0) \geq V_m(0|0)$ implies $V_m(0|0) \leq \frac{1}{1-\beta}$. The latter bound always holds, since for $m < 1$ the infinite horizon throughput is upper bounded as $V_m^*(\omega) \leq \sum_{t=0}^{\infty} \beta = \frac{1}{1-\beta}$ given that the FC can get at most a reward of $R_m(\omega, u) \leq 1$ in each slot. Hence, inequalities (6.39a.1) are proved. Inequality (6.39a.2) can be proved by contradiction. Specifically, assume that: *hp.1*) $V_m(1|0) \geq V_m(1|1)$. From (6.33) one would have $V_m^*(1) = \max\{V_m(1|0), V_m(1|1)\} = V_m(1|0)$, i.e., the passive action would be optimal when $\omega = 1$. Moreover, from (6.35) one would have $V_m(1|0) = m + \beta V_m^*(1) = m + \beta V_m(1|0)$, which can be solved with respect to $V_m(1|0)$ to get $V_m(1|0) = \frac{m}{1-\beta} = V_m^*(1)$. Therefore, if hypothesis *hp.1*) holds, one also has that $V_m(1|1) = 1 + \beta V_m^*(0) \leq V_m(1|0) = V_m^*(1) = \frac{m}{1-\beta}$. However, the value function $V_m^*(\omega)$ is bounded $\frac{m}{1-\beta} \leq V_m^*(\omega) \leq \frac{1}{1-\beta}$, where the lower bound is obtained considering a policy that always chooses the passive action for any belief ω. The boundedness of the value function, thus implies that if *hp.1*) holds then $1 + \beta \frac{m}{1-\beta} \leq 1 + \beta V_m(0) = V_m(1|1) \leq V_m(1|0) = \frac{m}{1-\beta}$, which yields $1 + \beta \frac{m}{1-\beta} \leq \frac{m}{1-\beta}$ and thus $(1-\beta)(1-m) \leq 0$. But this is clearly impossible as $m, \beta < 1$. Consequently, it has been proved that $V_m(1|1) \geq V_m(1|0)$.

Proof of case b) Inequality $V_m(0|0) \leq V_m(0|1)$ follows immediately since $m + \beta V_m^*(p_{01}) \leq \beta V_m^*(p_{01})$ holds for $m < 0$. The second inequality $V_m(0|1) \leq V_m(1|1)$

becomes $V_m(0|1) \leq = 1 + \beta V_m^*(0)1 + \beta V_m(0|1)$, which leads to $V_m(0|1) \leq \frac{1}{1-\beta}$, which always holds as discussed above. Inequality $V_m(1|0) \leq V_m(1|1)$ holds since an active action is always optimal when $m < 0$.

Proof of case c) The inequality holds since a passive action is always optimal for any $m \geq 1$.

APPENDIX H

PROOF OF THEOREM 20

Following the discussion in Section 6.5.3, to prove indexability it is sufficient to show that the threshold $\omega^*(m)$ is monotonically increasing with the subsidy m, for $0 \leq m < 1$. In fact, from Proposition 17 the passive set (6.37) for $m < 0$ is $\mathcal{P}(m) = \emptyset$, while for $m \geq 1$, it results $\mathcal{P}(m) = [0,1]$. It is thus only necessary to prove the monotonicity of $\omega^*(m)$ for $0 \leq m < 1$, which has been shown to hold in [40, Lemma 9] if

$$\left.\frac{dV_m(\omega|1)}{dm}\right|_{\omega=\omega^*(m)} < \left.\frac{dV_m(\omega|0)}{dm}\right|_{\omega=\omega^*(m)}. \tag{H.1}$$

To check if (H.1) holds, it is possible to differentiate (6.35)-(6.36) at the optimal threshold $\omega = \omega^*(m)$ as

$$V_m(\omega^*(m)|1) = \omega^*(m) + \beta\omega^*(m)V_m^*(0) + \beta(1-\omega^*(m))V_m^*(p_{01}), \text{ and} \tag{H.2}$$

$$V_m(\omega^*(m)|0) = m + \beta\tau_0^{(1)}(\omega^*(m))\left(1 + \beta V_m^*(0)\right) + $$
$$\beta^2(1-\tau_0^{(1)}(\omega^*(m)))V_m^*(p_{01}), \tag{H.3}$$

where (H.3) follows from (6.36) and from the fact that $\tau_0^{(1)}(\omega) \geq \omega$, for any ω (see eq. (6.41)), and hence $V_m^*(\tau_0^{(1)}(\omega^*(m))) = V_m(\tau_0^{(1)}(\omega^*(m))|1)$, since arm activation is optimal for any $\omega > \omega^*(m)$.

By letting $D_m(\omega) = \frac{dV_m^*(\omega)}{dm}$, then from (H.2) it follows

$$\left.\frac{dV_m(\omega|1)}{dm}\right|_{\omega=\omega^*(m)} = \beta\omega^*(m)D_m(0) + \beta(1-\omega^*(m))D_m(p_{01}), \tag{H.4}$$

while from (H.3) it results

$$\left.\frac{dV_m(\omega|0)}{dm}\right|_{\omega=\omega^*(m)} = 1 + \beta^2\tau_0^{(1)}(\omega^*)D_m(0) + \beta^2(1-\tau_0^{(1)}(\omega^*))D_m(p_{01}) \tag{H.5}$$

Finally, after some algebraic manipulations, and recalling that $D_m(0) = \frac{dV_m^*(0)}{dm} = \frac{d(m+\beta V^*(p_{01}))}{dm} = 1 + \beta D_m(p_{01})$, one can rewrite (H.1) as

$$D_m(p_{01})\beta\,(1-\beta)\,[1 - \omega\,(1 - \beta(1-p_{01}))] + \beta\,[\omega\,(1 - \beta(1-p_{01})) - \beta p_{01}] < 1. \quad \text{(H.6)}$$

To show that the inequality (H.6) holds when $0 \leq m < 1$, it is first introduced an upper bound to the derivative of the value function as $D_m(\omega) \leq \frac{1}{1-\beta}$, since $\frac{d}{dm} R_m(\omega) \leq 1$. Finally, by using this upper bound as $D_m(p_{01}) \leq \frac{1}{1-\beta}$ and after a little algebra then (H.1) reduces to $\beta(1 - \beta p_{01}) < 1$, which clearly holds for any $\beta \in [0, 1)$ as $0 \leq p_{01} \leq 1$. This concludes the proof of Theorem 20.

REFERENCES

[1] J. A. Paradiso, T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Perv. Comput. Mag.*, vol. 4, no. 1, pp. 18-27, Jan.-Mar. 2005.

[2] The Economist. (2010, Nov. 4). *It's a smart world* [Online]. Available: http://www.economist.com/node/17388368 [Feb. 7, 2012].

[3] T. Esram and P. L. Chapman, "Comparison of photovoltaic array maximum power point tracking techniques," *IEEE Trans. Energy Convers.*, vol. 22, no. 2, pp. 439-449, Jun. 2007.

[4] A. P. Sample, D.J. Yeager, P.S. Powledgde, A.V. Mamishev, and J.R. Smith, "Design of an RFID-based battery-free programmable sensing platform," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 11, pp. 2608-2615, Nov. 2008.

[5] I. F. Akyildiz, S. Weilian, Y. Sankarasubramaniam, E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102-114, Aug. 2002.

[6] S. Roundy, P.K. Wright and J. Rabaey, *Energy Scavenging for Wireless Sensor Networks with Special Focus on Vibrations*. Norwell, MA: Kluwer Academic Publisher, 2003.

[7] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292-2330, Aug. 2008.

[8] D. Bertsekas, R. G. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice Hall, 1992.

[9] A. Chandra, V. Gummalla, and J. O. Limb, "Wireless medium access control protocols," *IEEE Commun. Surveys Tutorials*, vol. 3, no. 2, pp. 2-15, Apr. 2000.

[10] S. Kumar, V. S. Raghavan, and J. Deng, "Medium access control protocols for ad hoc wireless networks: a survey," *Ad Hoc Networks*, vol. 4, no. 3, pp. 326-358, May 2006.

[11] N. Abramson, "The ALOHA system - another alternative for computer communications," in *Proc. Fall Joint Comp. Conf.*, vol. 37, pp. 281-285, Montvale, NJ, Nov. 1970.

[12] L. Roberts, "ALOHA packet system with and without slots and capture," in *ACM SIGCOMM Comp. Commun. Review*, vol. 5, no. 2, pp. 28-42, Apr. 1975.

[13] F. C. Schoute, "Dynamic frame length ALOHA," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 565-568, Apr. 1983.

[14] J. E. Wieselthier, A. Ephremides, and L. A. Michaels, "An exact analysis and performance evaluation of framed ALOHA with capture," *IEEE Trans. Commun.*, vol. 37, no. 2, pp. 125-137, Feb. 1989.

[15] L. Kleinrock and F. Tobagi, "Packet switching in radio channels: Part I - carrier sense multiple-access modes and their throughput-delay characteristics," *IEEE Trans. Commun.*, vol.23, no.12, pp. 1400-1416, Dec. 1975.

[16] X. Liu, E. K. P. Chong, and N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 10, pp. 2053-2064, Oct. 2001.

[17] J. W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proc. IEEE INFOCOM* 2002, New York, NY, Jun. 2002.

[18] L. Benini, A. Bogliolo, G. Paleologo, G. De Micheli, "Policy optimization for dynamic power management," *IEEE Trans. Computer-Aided Des. Integr. Circuits Syst.*, vol. 18, no. 6, pp. 813-833, Jun. 1999.

[19] S. Priya and D. J. Inman, *Energy Harvesting Technologies*. New York, NY: Springer, 2009.

[20] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. on Embedded Comput. Syst.*, vol. 6, no. 4, art. 32, Sep. 2007.

[21] J. Lei, R. Yates and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replinishable sensors," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 547-551, Feb. 2009.

[22] V. Sharma, U. Mukherji, V. Joseph and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1326-1336, Apr. 2010.

[23] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1180-1189, Mar. 2012.

[24] J. Yang and S. Ulukus, "Transmission completion time minimization in an energy harvesting system," in *Proc. 44th Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, Mar. 2010.

[25] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: optimal policies," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1732-1743, Sep. 2011.

[26] K. Tutuncuoglu and A. Yener, "Optimal power policy for energy harvesting transmitters with inefficient energy storage," in *Proc. 46th Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, Mar. 2012.

[27] V. Sharma, U. Mukherji, V. Joseph, "Efficient energy management policies for networks with energy harvesting sensor nodes," in *Proc. Allerton Conf. Commun., Control Comput.*, Monticello, IL, pp. 375-383, Sep. 2008.

[28] L. Huang and M. J. Neely, "Utility optimal scheduling in energy harvesting networks," in *Proc. 12th ACM Int. Sym. Mobile Ad Hoc Networking Comput. (MobiHoc)*, Paris, France, May 2011.

[29] Y. Jing, S. Ulukus, "Optimal packet scheduling in a multiple access channel with rechargeable nodes," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kyoto, Japan, Jun. 2011.

[30] J. Yang, O. Ozel and S. Ulukus, "Optimal packet scheduling in a broadcast channel with an energy harvesting transmitter," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kyoto, Japan, Jun. 2011.

[31] 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.

[32] L. Huijiang, N. Jaggi, B. Sikdar, "Relay scheduling for cooperative communications in sensor networks with energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 2918-2928, Sep. 2011.

[33] V. Joseph, V. Sharma, U. Mukherji, M. Kashyap, "Joint power control, scheduling and routing for multicast in multihop energy harvesting sensor networks," in *Proc. Int. Conf. Ultra Modern Telecommun. Workshops (ICUMT)*, St. Petersburg, Russia, Oct. 2009.

[34] P. Nikitin and K. Rao, "Antennas and propagation in UHF RFID systems," in *Proc. IEEE RFID conf.*, Las Vegas, NV, Apr. 2008.

[35] Mays et al., "RFID transponder having active backscatter amplifier for retransmitting a received signal," U. S. Patent 6 838 989 B1, Jan. 2005.

[36] P. Popovski, F. H. P. Fitzek, and R. Prasad, "A class of algorithms for batch conflict resolution algorithms with multiplicity estimation," *Algorithmica*, Special issue on Principles of Mobile Computing (Springer-Verlag), Dec. 2007.

[37] A. Kailas and M. A. Ingram, "A novel routing metric for environmentally-powered sensors with hybrid energy storage systems," in *Proc. IEEE Wireless VITAE Conf.*, Aalborg, Denmark, May 2009.

[38] P. Whittle, "Restless bandits: Activity allocation in a changing world," *J. Appl. Probab.*, vol. 25, pp. 287-298, 1988.

[39] S. H. A. Ahmad, M. Liu, T. Javidi, Q. Zhao and B. Krishnamachari, "Optimality of myopic sensing in multi-channel opportunistic access," *IEEE Trans. Inf. Theory*, vol. 55, No. 9, pp. 4040-4050, Sept. 2009.

[40] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5547-5567, Nov. 2010.

[41] N. Vaidya and S. R. Das, "RFID-based networks: exploiting diversity and redundancy," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 12, no. 1, pp. 2-14, Jan. 2008.

[42] K. Finkenzeller, *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, Second Edition.* Hoboken, NJ: John Wiley and sons, 2003.

[43] H.C. Liu, M.C. Hua, C.G. Peng, and J.P. Ciou, "A novel battery-assisted class-1 generation-2 RF identification tag design," *IEEE Trans. Microw. Theory Tech.*, vol. 57, no. 5, pp. 1388-1397, May 2009.

[44] A. Janek, C. Trummer, C. Steger, R. Weiss, J. Preishuber-Pfluegl, and M. Pistauer, "Lifecycle extension of long range UHF RFID tags based on energy harvesting," in *Proc. EURASIP Workshop RFID Tech.*, Vienna, Austria, Sep. 2007.

[45] A. Janek, C. Trummer, C. Steger, R. Weiss, J. Preishuber-Pfluegl, and M. Pistauer, "Simulation based verification of energy storage architectures for higher class tags supported by energy harvesting devices," in *Proc. 10th Euromicro Conf. Digital Syst. Des. Arch.,* Lubeck, Germany, Oct. 2007.

[46] D. Kim, M. A. Ingram, and W. W. Smith, "Measurements of small-scale fading and path loss for long range RF tags," *IEEE Trans. Antennas Propag.*, vol. 51, no. 8, pp. 1740-1749, Nov. 2003.

[47] R. Gallager, *Discrete Stochastic Processes.* Norwell, MA: Kluwer Academic Publisher, 1995.

[48] GS1 EPCglobal. *UHF Class 1 Gen 2 Standard v. 1.2.0* [Online]. Available: http://www.gs1.org/gsmp/kc/epcglobal/uhfc1g2 [Mar. 24, 2012].

[49] G. De Vita and G. Iannaccone, "Design criteria for the RF section of UHF and microwave passive RFID transponders," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 9, pp. 2978-2990, Sep. 2005.

[50] J.D. Griffin, G. Durgin, A. Haldi, and B. Kippelen, "Radio link budgets for 915MHz RFID antennas placed on various objects," in *Proc. WCNG Wireless Sym.*, Austin, TX, Oct. 2005.

[51] D. Panigrahi, S. Dey, R. Rao, K. Lahiri, C. Chiasserini, A. Raghunathan, "Battery life estimation of mobile embedded systems," in *Proc. 14th Int. Conf. VLSI Des.*, Bangalore, India, Jan. 2001.

[52] K.W. Ross, "Randomized and past-dependent policies for Markov decision processes with multiple constraints," *Operations Research*, vol. 37, no. 3, pp. 474-477, May-Jun. 1989.

[53] F. Iannello, O. Simeone and U. Spagnolini, "Dynamic framed-ALOHA for energy-constrained wireless sensor networks with energy harvesting," in *Proc. IEEE GLOBECOM*, Miami, FL, Dec. 2010.

[54] F. Iannello, O. Simeone and U. Spagnolini, "Medium access control protocols for wireless sensor networks with energy harvesting," *IEEE Trans. Commun.*, May 2012 (in press).

[55] M. Gatzianas, L. Georgiadis, L. Tassiulas, "Control of wireless networks with rechargeable batteries," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 581-593, Feb. 2010.

[56] EnOcean. *EnOcean Technology - Energy Harvesting Wireless* [Online]. Available: http://www.enocean.com/en/white-papers/ [Oct. 14, 2011].

[57] M. Kodialam and T. Nandagopal, "Fast and reliable estimation schemes in RFID systems," in *Proc. MOBICOM*, Los Angeles, CA, Sep. 2006.

[58] S. Kandukuri, S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. Commun.*, vol. 1, no. 1, pp. 46-55, Jan 2002.

[59] F. Iannello, O. Simeone and U. Spagnolini, "Medium access control protocols for wireless sensor networks with energy harvesting," [Online]. Available: http://arxiv.org/abs/1112.2409 [Feb. 9, 2012].

[60] B. Knerr, M. Holzer, C. Angerer, M. Rupp, "Slot-by-slot maximum likelihood estimation of tag populations in framed slotted aloha protocols," in *Proc. SPECTS*, Edinburgh, UK, pp. 303-308, Jun. 2008.

[61] C.Y. Wang, C.C. Lee, "A grouping-based dynamic framed slotted ALOHA anti-collision method with fine groups in RFID systems," in *Proc. FutureTech*, Busan, Korea, May 2010.

[62] A. Kailas and M. A. Ingram, "A novel routing metric for environmentally-powered sensors with hybrid energy storage systems," in *Proc. IEEE Wireless VITAE Conf.*, Aalborg, Denmark, May 2009.

[63] R. Vullers, R. van Schaijk, I. Doms, C. Van Hoof, R. Mertens, "Micropower energy harvesting," *Solid-State Electron.*, vol. 53, no. 7, pp. 684-693, 2009.

[64] D. P. Bertsekas, J. N. Tsitsiklis, "An analysis of stochastic shortest path problems," *Math. Oper. Res.*, vol. 16, no. 3, pp. 580-595, Aug. 1991.

[65] Y. Chen, Q. Zhao, V. Krishnamurthy, and D. Djonin, "Transmission scheduling for optimizing sensor network lifetime: A stochastic shortest path approach," *IEEE Trans. Signal Proc.*, vol. 55, no. 5, pp. 2294-2309, May 2007.

[66] K. Cohen and A. Leshem, "A time-varying opportunistic approach to lifetime maximization of wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5307-5319, Oct. 2010.

[67] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artif. Intell.*, vol. 101, pp. 99-134, May 1998.

[68] G. E. Monahan, "A survey of partially observable Markov decision processes: Theory, models, and algorithms," *Manag. Sci.*, vol. 28, no. 1, pp. 1-16, 1982.

[69] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Hoboken, NJ: Wiley, 2005.

[70] J. Gittins, K, Glazerbrook, R. Weber, *Multi-armed Bandit Allocation Indices.* West Sussex, UK: Wiley, 2011.

[71] C. H. Papadimitriou, J. N. Tsitsiklis, "The complexity of Markov decision processes," *Math. Oper. Res.*, vol. 12, no. 3, pp. 441-450, Aug. 1987.

[72] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of optimal queueing network control," *Math. Oper. Res.*, vol. 24, pp. 293-305, May 1999.

[73] S. H. A. Ahmad, L. Mingyan, "Multi-channel opportunistic access: A case of restless bandits with multiple plays," in *Proc. 47th Ann. Allerton Conf. Commun., Contr., Comput.*, Monticello, IL, pp. 1361-1368, Sept. 2009.

[74] D. Bertsimas and J. E. Nio-Mora, "Restless bandits, linear programming relaxations, and a primal-dual heuristic," *Oper. Res.*, vol. 48, no. 1, pp. 80-90, Jan. 2000.