UNIVERSITÀ DI PADOVA FACOLTÀ DI INGEGNERIA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

SCUOLA DI DOTTORATO IN INGEGNERIA DELL'INFORMAZIONE

INDIRIZZO IN SCIENZA E TECNOLOGIA DELL'INFORMAZIONE

XXV Ciclo

Coping with spectrum and energy scarcity in Wireless Networks: a Stochastic Optimization approach to Cognitive Radio and Energy Harvesting

Dottorando

NICOLÒ MICHELUSI

Supervisore:

Chiar.mo Prof. Michele Zorzi

Direttore della Scuola:

Chiar.mo Prof. Matteo Bertocco

Coordinatore di Indirizzo:

Chiar.mo Prof. Carlo Ferrari

Anno Accademico 2012/2013

To my family:
my parents, Adriana and Michele
and my siblings Mattia, Angelica, and Giulio
for supporting me with love
and for being part of myself, wherever I am

To Beatrice
for being my best half

Alla mia famiglia:
i miei genitori, Adriana e Michele
e i miei fratelli Mattia, Angelica e Giulio
per sostenermi sempre con amore
e per essere parte di me ovunque io sia

A Beatrice
per essere la mia miglior metà

Contents

Abstract

In the last decades, we have witnessed an explosion of wireless communications and networking, spurring a great interest in the research community. The design of wireless networks is challenged by the scarcity of resources, especially spectrum and energy. In this thesis, we explore the potential offered by two novel technologies to cope with spectrum and energy scarcity: Cognitive Radio (CR) and Energy Harvesting (EH). CR is a novel paradigm for improving the spectral efficiency in wireless networks, by enabling the coexistence of an incumbent legacy system and an opportunistic system with CR capability. We investigate a technique where the CR system exploits the temporal redundancy introduced by the Hybrid Automatic Retransmission reQuest (HARQ) protocol implemented by the legacy system to perform interference cancellation, thus enhancing its own throughput.

Recently, EH has been proposed to cope with energy scarcity in Wireless Sensor Networks (WSNs). Devices with EH capability harvest energy from the environment, *e.g.*, solar, wind, heat or piezo-electric, to power their circuitry and to perform data sensing, processing and communication tasks. Due to the random energy supply, how to best manage the available energy is an open research issue. In the second part of this thesis, we design control policies for EH devices, and investigate the impact of factors such as the finite battery storage, time-correlation in the EH process and battery degradation phenomena on the performance of such systems.

We cast both paradigms in a stochastic optimization framework, and investigate techniques to cope with spectrum and energy scarcity by opportunistically leveraging interference and ambient energy, respectively, whose benefits are demonstrated both by theoretical analysis and numerically.

As an additional topic, we investigate the issue of channel estimation in Ultra Wide-Band (UWB) systems. Due to the large transmission bandwidth, the channel has been typically modeled as sparse. However, some propagation phenomena, *e.g.*, scattering from rough surfaces and frequency distortion, are better modeled by a diffuse channel. We propose a novel Hybrid Sparse/Diffuse (HSD) channel model which captures both components, and design channel estimators based on it.

Sommario

Negli ultimi decenni, abbiamo assistito alla diffusione delle comunicazioni e reti wireless, suscitando un crescente interesse nella comunità scientifica. Tuttavia, la progettazione delle reti wireless è resa difficile dalla scarsità di risorse, in particolare, spettro ed energia. In questa tesi, si esplora il potenziale offerto da due nuove tecnologie nell'affrontare il problema della scarsità di spettro e di energia nelle future reti wireless: "Cognitive Radio" (CR) ed "Energy Harvesting" (EH). CR è un nuovo paradigma che consente di migliorare l'efficienza di utilizzo dello spettro nelle reti wireless, abilitando la coesistenza di un preesistente sistema titolare dello spettro, comunemente denominato *Utente Primario*, e un sistema opportunistico "intelligente", noto come *Utente Secondario*. In questa tesi, si sviluppa una tecnica per sfruttare, da parte di un utente secondario, la ridondanza temporale introdotta dal protocollo "Hybrid Automatic Retransmission reQuest" (HARQ) utilizzato da un Utente Primario, per eseguire tecniche di cancellazione di interferenza, consentendo così di migliorare il throuhgput secondario.

Recentemente, EH è stato proposto per superare il problema della scarsità di energia nelle "Wireless Sensor Networks" (WSNs). I dispositivi con capacità di EH accumulano energia resa disponibile nell'ambiente circostante, come, per esempio, energia solare, eolica, termica o piezo-elettrica, per alimentare il dispositivo e per eseguire compiti di "data sensing", processamento e comunicazione. Dato che la disponibilità di energia è aleatoria e intermittente, il problema di come utilizzare al meglio l'energia disponibile è di grande interesse nella comunità scientifica. Nella seconda parte di questa tesi, si propongono politiche di controllo per dispositivi con capacità di EH, e si analizza l'impatto di vari fattori quali la capacità finita della batteria, la correlazione temporale nel processo di EH, la conoscenza imperfetta dello stato di carica della batteria e i fenomeni di degrado della batteria.

Si studiano entrambi i paradigmi in un framework di ottimizzazione stocastica, e vengono proposte tecniche per far fronte alla scarsità di spettro ed energia sfruttando in modo opportunistico, rispettivamente, l'interferenza e l'energia ambientale. Si dimostrano i benefici delle tecniche pro-

poste per mezzo sia di un'analisi teorica che per via numerica.

Come argomento di ricerca aggiuntivo, nell'ultima parte di questa tesi, si studia il problema della stima di canale nei sistemi Ultra Wide-Band (UWB). Data la larga banda di trasmissione utilizzata in questi sistemi, il canale è stato tipicamente modellato come sparso. Tuttavia, alcuni fenomeni di propagazione come, per esempio, la dispersione dovuta a superfici scabrose e la distorsione in frequenza, sono modellabili in modo più accurato da un canale diffuso. Si propone un nuovo modello di canale denominato "Hybrid Sparse/Diffuse" (HSD) che cattura entrambe le componenti di canale, e si propongono stimatori di canale basati sul modello proposto.

List of Acronyms

ACK Acknowledgment

AWGN Additive White Gaussian Noise

ARQ Automatic Repeat reQuest

BER Bit Error Rate

CDF Cumulative Distribution Function

CIR Channel Impulse Response

CR Cognitive Radio

CSI Channel State Information

EH Energy Harvesting

EHS Energy Harvesting Sensor

EH-WSN Energy Harvesting Wireless Sensor Network

FC Fusion Center

FCC Federal Communications Commission

HARQ Hybrid Automatic Repeat reQuest

HSD Hybrid Sparse Diffuse

i.i.d. Independent and Identically Distributed

LS Least Squares

MAP Maximum A Posteriori

ML Maximum Likelihood

MMSE Minimum Mean Square Error

MSE Mean Square Error

NACK Not Acknowledgment

NE Nash Equilibrium

p.d.f. Probability Density Function

PDP Power Delay Profile

PU Primary User

QoS Quality of Service

SINR Signal to Interference and Noise Ratio

SNR Signal-to-Noise Ratio

SU Secondary User

UWB Ultra-WideBand

WSN Wireless Sensor Network

Chapter 1

Introduction

The development of wireless communications and networking in the last decades has enabled applications such as ubiquitous and mobile access to the internet, wireless sensor and cellular networks. However, the widespread and pervasive diffusion of these technologies is challenged by the scarcity of resources, most importantly, spectrum and energy. The *spectrum licensing* approach, commonly employed to reserve spectrum usage to specific classes of wireless users, has lead to a spectrum scarcity problem. On the other hand, the design of wireless systems has typically relied on the use of batteries to sustain the operation of the wireless terminals, posing an energy scarcity problem in those systems, *e.g.*, Wireless Sensor Networks (WSNs), where long-term and autonomous operation is required, and factors such as the sheer number of nodes or inaccessibility render battery replacement unrealistic and cost-prohibitive.

In this thesis, we investigate the potential offered by two approaches to cope with spectrum and energy scarcity in wireless networks: Cognitive Radio (CR) and Energy Harvesting (EH). CR is a novel paradigm for improving the efficiency of spectrum usage in wireless networks, by enabling the coexistence of an incumbent legacy system, commonly referred to as *Primary Users* (PU), and an opportunistic system with CR capability, known as *Secondary Users* (SU). The latter adapt their operation by opportunistically leveraging the information collected about the PUs, *e.g.*, primary message, channel state, idle/busy state, protocols, so as to earn a performance gain, *e.g.*, in terms of secondary throughput. In a widely used model for cognitive radio, the legacy system is oblivious to the presence of the SUs, which need to satisfy given constraints on the performance loss caused to the PUs.

Within this framework, in Chapter 2, we investigate a technique to exploit the Type-I Hybrid Automatic Retransmission reQuest (Type-I HARQ) protocol implemented by the PU. In fact, HARQ

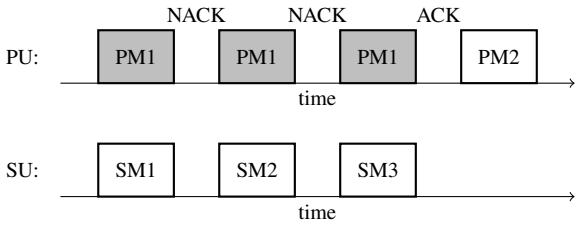
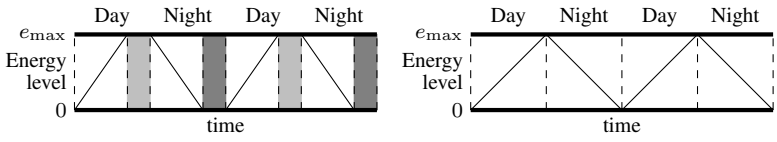


Figure 1.1. *PU with HARQ scheme*

introduces temporal redundancy in the wireless channel, in that copies of the same primary data packet are retransmitted over subsequent time-slots. Opportunities thus arise for the SU to improve its throughput, as explained in the following example. Consider the scenario depicted in Fig. 1.1, where a PU subsequently retransmits the same packet PM1, in response to retransmission requests by its intended receiver. Different options are available for the SU, depending on the side information about the PU: if the SU does not know the codebook employed by the PU, then the secondary receiver treats the signal coming from the primary transmitter as noise, which degrades the secondary outage performance. If the primary codebook is known at the secondary receiver, such knowledge can be leveraged for interference cancellation. For instance, if the signal from the primary transmitter is strong, the secondary receiver can, in sequence, decode the primary message, remove its interference from the received signal, and then take advantage of a "clean" channel to decode its intended message. If, in addition, the secondary receiver is able to track the retransmission process of the PU then, after decoding the primary message in the first time-slot, it can leverage this knowledge to perform interference cancellation in the following PU retransmissions of PM1, not only in the first time-slot where the actual decoding of PM1 takes place. It is thus clear that the use of HARQ by the PU opens up intriguing opportunities for a more efficient utilization of the spectrum. We employ a stochastic optimization approach to optimize the control policy of the SU, which determines its access pattern, based on the state of the system, so as to maximize its own throughput, while bounding the performance degradation incurred to the PU.

Recently, EH has been proposed to cope with energy scarcity in wireless systems. Devices with EH capability harvest energy from the environment, *e.g.*, solar, wind, heat or piezo-electric, to power their circuitry and to perform data sensing, processing and communication tasks. By relying on a potentially unlimited energy reservoir, the ambient energy, the EH technology is particularly appealing in the deployment of WSNs, where battery replacement is typically prohibitive. In contrast to battery-operated sensors, where energy efficiency and conservation are crucial to prolong lifetime, in EH powered systems the energy supply is unlimited, but its availability is random and intermittent



(a) $p(t)$ from battery-powered systems (b) $p(t)$ optimized for the EH setting

Figure 1.2. *The battery is recharged during daylight, and discharged during night. Light gray boxes denote time intervals during which part of the harvested energy is lost due to overflow. Gray boxes denote time intervals during which the battery is depleted, hence the transmit power is forced to zero ($p(t) = 0$)*

over time. The objective thus shifts from energy efficiency and conservation to the management of the harvested energy, so as to provide a stable energy supply to the sensor node by minimizing the deleterious impact of energy depletion. We remark that the random and intermittent nature of the EH supply gives rise to new dynamics and trade-offs with respect to traditional battery powered systems. For example, one aspect which plays a crucial role in determining the performance is the interplay between the finite battery capacity and the intermittent EH process. Consider, for instance, a device which aims at maximizing a time-average of a concave function $g(p(t))$ of the transmit power $p(t)$. In traditional battery powered systems, where energy conservation is typically handled as a time-average power constraint β, the device should transmit with constant power β, owing to the concavity of $g(p(t))$. In contrast, such solution may not be optimal for an EH powered device as can be seen with the help of Fig. 1.2: assuming the device is powered by solar energy with average EH rate β (*i.e.*, the power supplied by the environment is, on average, β), by transmitting with constant power $p(t) = \beta$, the device may quickly run out of energy during night (gray boxes in the figure), when the power is solely supplied by the rechargeable battery, thus forcing the transmit power to zero; on the other hand, the battery may be quickly recharged during daylight and, upon fully charging it, part of the harvested energy may be lost due to overflow (light gray boxes in the figure). A better approach would be, instead, to adapt the transmit power to the state of the EH process (day,night), hence to transmit with a smaller power $p(t) < \beta$ during night, so as to avoid energy depletion, and with a larger power $p(t) > \beta$ during daylight, in such a way as to avoid energy overflow.

In Chapter 3 of this thesis, we present a general model for EH-WSN where an EH Sensor (EHS) needs to report data of varying importance to a Fusion Center (FC). The importance models, for example, the priority of data packets, the importance of the sensed events, *e.g.*, temperature or humidity, the channel fading state, or the achievable rate in a Rayleigh fading channel. Using a stochastic optimization approach, we design control policies for EH devices, which determine, based on the state of the system (energy level in the battery, state of the EH process and importance of the current data

packet), whether to report the data packet to the FC or to drop it. In particular, due to the limited processing capability typically found in practical WSN deployments, we focus on the design of low-complexity control policies, which are shown to achieve close-to-optimal performance with respect to the globally optimal policy. We investigate the impact of factors such as the finite battery storage and time-correlation in the EH process.

While in Chapter 3 it is ideally assumed that the battery used by the EH device to store the incoming ambient energy can perpetually operate without incurring a performance degradation, in Chapter 4 we investigate the impact of degradation phenomena, which cause the storage capability of a battery to diminish over time. This poses a problem to the operation of the EH device, hence of the WSN as a whole, since, the smaller the battery capacity, the faster the battery depletion during periods of limited ambient energy supply, hence, in turn, the worse the performance. We propose a stochastic framework, suitable for policy optimization, which captures the trade-off between QoS and battery degradation, and its interplay with the control policy implemented by the EHS controller. We believe that acknowledging the degradation of the battery capacity represents an important step towards the realistic characterization of rechargeable batteries and, by extension, of WSNs and their management strategies.

Despite the different objectives and application scenarios which CR and EH have been envisioned to, in this thesis we employ similar methodologies and techniques based on *stochastic optimization* to address the problem of spectrum and energy scarcity in wireless networks. In particular, we will resort to the theory of Markov Decision Processes [1]. Stochastic optimization is of crucial importance to optimize the operation of the wireless terminals and achieve the best performance in resource limited settings, as the ones considered in this thesis. In fact, the common feature of CR and EH is *resource limitation*. In CR, the SU is required to communicate over a shared wireless channel, posing the problem of how to best manage the knowledge about the incumbent PU (*e.g.*, the primary HARQ process), and the interference to the PU, in order to maximize its own performance, while bounding the performance loss to the PU. On the other hand, EH devices are required to operate under a stochastic and intermittent energy supply, which poses the problem of how to best utilize the available energy (as seen in the previous example, depicted in Fig. 1.2), in order to minimize the deleterious impact of energy depletion and overflow.

As an additional topic, in the last part of this thesis, we investigate the issue of channel estimation in Ultra Wide-Band (UWB) systems. This work is the result of my visit at the University

of Southern California, Los Angeles, USA, from January to July 2011, under the supervision of Prof. Urbashi Mitra. Due to the large transmission bandwidth, the channel has been typically modeled as sparse. However, some propagation phenomena, *e.g.*, scattering from rough surfaces and frequency distortion, are better modeled by a diffuse channel. In this context, we propose a novel Hybrid Sparse/Diffuse (HSD) channel model, and design channel estimators based on it. Moreover, we provide a Mean Square Error (MSE) analysis of the proposed estimators, and demonstrate, based on a realistic channel emulator, the benefits in terms of MSE and Bit-Error-Rate performance, with respect to unstructured and purely sparse estimators.

1.1 Organization of the Thesis

The rest of the thesis is subdivided into four chapters, each addressing a specific topic and the corresponding results. Each chapter can be read separately.

In Chapter 2, we study the problem of designing optimal secondary access strategies in cognitive radio networks, which leverage the HARQ protocol implemented by the primary user. This work is based on the journal paper [J1] and on the conference papers [C1], [C2] (see page 193 for a list of my publications).

In Chapter 3, we focus on the design of energy management polices for EH devices, and we evaluate, both theoretically and numerically, the impact of factors such as the finite battery capacity and time-correlation in the EH process. This work is based on the journal paper [J2] and on the conference papers [C3], [C4] and [C5].

In Chapter 4, we investigate the impact of battery degradation on the lifetime of EH devices. This work is based on the journal paper [J3] and on the conference paper [C6].

Chapter 5 concludes this thesis.

In Appendix A, we investigate the issue of channel estimation in UWB systems, which is based on the journal papers [J4], [J5].

Chapter 2

Optimal Secondary Access in Cognitive Radio Networks

2.1 Introduction

Spectrum licensing has been traditionally employed to protect wireless systems against mutual interference. While effective in avoiding multi user interference, this approach has led to an inefficient utilization of the available resources, hence to spectrum scarcity [2–4], as can be seen from the 2003 FCC spectrum allocation chart, depicted in Fig. 2.1. Cognitive radio networks, a concept first proposed by Mitola in his seminal work [5], hold the promise to improve the spectral efficiency of wireless networks with respect to conventional licensing, by allowing the coexistence of Primary (licensed) and Secondary (unlicensed) Users (PUs and SUs, respectively) on the same radio band. In order to achieve such objective, SUs are equipped with smart, cognitive radios through which they can sense the radio environment and collect side information about the presence and the operation of active primary transmitters. This information is then used by the cognitive radios to make decisions and dynamically adapt their operation, so as to optimize a given performance metric, while limiting their interference to the incumbent licensed system. For a survey on cognitive radio, dynamic spectrum access and the related research challenges, we refer the interested reader to [4, 6–8].

Most prior works on cognitive radio networks are based on the assumption that the SUs are allowed to operate only in time-frequency slots left unused by the licensed system (*interweave* cognitive radio paradigm [7]). A crucial aspect in these systems is the ability of SUs to detect, as accurately and quickly as possible, the activity of licensed users in a given time-frequency slot [9], so that lit-

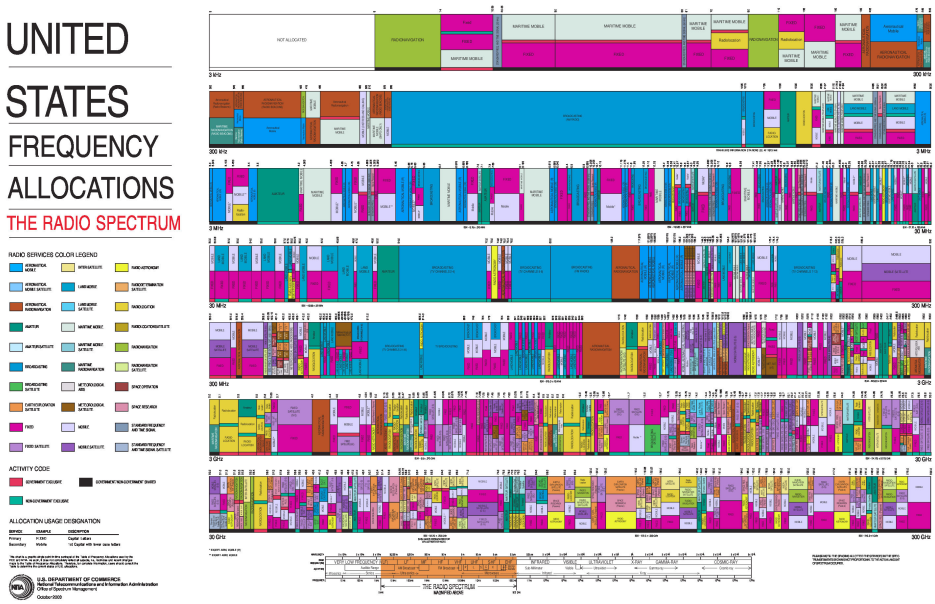


Figure 2.1. *2003 FCC spectrum allocation chart, from http://www.ntia.doc.gov/files/ntia/publications/2003-allochrt.pdf*

tle or no harm is caused to the licensed radios. In *overlay* systems, on the other hand, the SUs use sophisticated signal processing and coding to maintain or even improve the performance of the PUs, while also obtaining some additional bandwidth for their own communication. A more general and advanced paradigm than interweave cognitive radio is *underlay* cognitive radio [7], where the SUs are allowed to operate also in time-frequency slots used by PUs, but need to satisfy given constraints on the performance loss caused to the PU, *e.g.*, the interference to each PU should be kept within a tolerable limit [4, 10]. Within this framework, the problem of how the SUs should best utilize the side information about the primary system, *e.g.*, codebook, protocol, retransmission schemes, channel state information, is still an open research issue.

In the information theoretic community, cognitive radio network models have often been proposed by assuming a genie-aided SU with non-causal access to the whole or part of the active primary message (side information about the primary message) [7, 11, 12]. While this assumption allows for analysis of information-theoretic optimal transmission strategies and codebook design, it is not able to capture critical aspects of a cognitive radio network, related to the imperfect sensing and the dynamic acquisition of the knowledge about the primary message. Another line of inquiry is resource management, which employs various tools from stochastic optimization or machine learning to design optimal secondary strategies which best utilize the available resources and the side information, *e.g.*,

see [13] and references therein. This approach allows to consider network constraints, such as delay or other QoS guarantees, as well as to model the dynamic acquisition of the side information by the SUs, *e.g.*, by a proper Markov chain representation of the system.

Based on the interweave cognitive radio paradigm, we propose to exploit the Hybrid Automatic Retransmission reQuest (HARQ [14]) protocol implemented by the PU. The use of such protocol introduces temporal redundancy in the wireless channel, in the form of copies of primary packets transmitted in subsequent time-slots in response to retransmission requests by the primary receiver. Opportunities for secondary access thus arise: by tracking the retransmission process of the PU and by decoding the current primary message, the secondary receiver can remove its interference by employing Interference Cancellation (IC) techniques over the entire interval over which retransmissions of the same primary message take place, thus enhancing the secondary outage performance and improving the spectral efficiency of the system. We believe that the ability of the SU to best manage the interference from nearby terminals is crucial to achieve high spectral efficiency in cognitive radio networks, since interference is a limiting factor in wireless networks. For this reason, the strategy of the SU, which prescribes whether to access the channel or remain idle, based on the HARQ state of the PU and on the state of the SU, is optimized by using stochastic optimization tools. However, interference cancellation may not be successfully employed by the PUs, which are typically assumed to be oblivious to the presence of SUs in the network. Hence, the interference produced by the SUs to the PUs should be kept within tolerable limits.

We consider a simple network topology consisting of a pair of PUs and a pair of SUs (transmitter and receiver), as depicted in Fig. 2.2. Despite the simplicity of such network topology, understanding its fundamental limits is still an open research issue which requires in-depth investigation. Moreover, we believe that this topology represents a building block of more general network settings, consisting of multiple primary and SU pairs.

The idea of exploiting the primary HARQ process to perform IC on future packets was put forth by [15], which devises several cognitive radio protocols exploiting the HARQ protocol of the PU. Therein, the PU employs HARQ with incremental redundancy and the ARQ mechanism is limited to at most one retransmission. The SU receiver attempts to decode the PU message in the first time-slot. If successful, the SU transmitter sends its packet and the SU receiver decodes it by using IC on the received signal. In contrast, in this chapter, we address the more general case of an arbitrary number of primary ARQ retransmissions, and we allow a more general access pattern for the SU pair over the

entire primary ARQ window, as detailed in the next section.

Other related works include [16], which devises an opportunistic sharing scheme with channel probing based on the ARQ feedback from the PU receiver. An information theoretic framework for cognitive radio is investigated in [12], where the SU transmitter has non-casual knowledge of the PU's codeword. In [17], the data transmitted by the PU is obtained causally at the SU receiver. However, this model requires a joint design of the PU and SU signaling and channel state information at the transmitters. In contrast, we explicitly model the dynamic acquisition of the PU message at the SU receiver, which enables IC. Moreover, the PU is oblivious to the presence of the SU.

2.1.1 Contributions

Within this framework, we propose to exploit the primary HARQ process and introduce two IC schemes that work in concert, both enabled by the underlying retransmission process of the PU. With *Forward IC* (FIC), SUrx, after decoding the PU message, performs IC in the next PU retransmission attempts, if these occur. While FIC provides IC on SU transmissions performed in future time-slots, *Backward IC* (BIC) provides IC on SU transmissions performed in previous time-slots within the same primary ARQ retransmission window, whose decoding failed due to severe interference from the PU. BIC relies on buffering of the received signals at the SU receiver. Based on these IC schemes, we model the state evolution of the PU-SU network as a Markov Decision Process [1, 18], induced by the specific access policy used by the SU, which determines its access probability in each state of the network.

As an application of this framework, we study the problem of designing optimal secondary access policies that maximize the average long-term SU throughput by opportunistically leveraging FIC and BIC, while causing a bounded average long-term throughput loss to the PU and a bounded average long-term SU power expenditure. A similar problem has been studied in [19]. However, therein the secondary receiver is not allowed to perform interference cancellation based on decoding of the PU's message. This aspect plays instead a central role in our work. We show that the optimal strategy dictates that the SU prioritizes its channel access in the states where SUrx knows the PU message, thus enabling IC; moreover, we provide an algorithm to optimally allocate additional secondary access opportunities in the states where the PU message is unknown. In order to derive further insights in the interaction between the PU and SU in the network, we consider a *degenerate cognitive radio* network

Table 2.1. *List of symbols.*

D	Primary HARQ deadline
$t \in \mathbb{N}(1, D)$	primary ARQ state (retransmission index)
$b \in \mathbb{N}(0, B)$	SU buffer state (number of received signals currently buffered at SUrx)
$\Phi \in \{\mathrm{K}, \mathrm{U}\}$	PU message knowledge state
	($\Phi = \mathrm{K}$, if the current PU message is known to SUrx; otherwise, $\Phi = \mathrm{U}$)
R_p	PU transmission rate
$R_{s\mathrm{U}}$	SU transmission rate when PU message is unknown at SUrx
$R_{s\mathrm{K}}$	SU transmission rate when PU message is known at SUrx
$T_p^{(\mathrm{I})}$	PU throughput when SU is idle
$T_p^{(\mathrm{A})}$	PU throughput when SU is active
$T_{s\mathrm{U}}$	SU throughput when $\Phi = \mathrm{U}$
$T_{s\mathrm{K}}$	PU throughput when $\Phi = \mathrm{K}$
μ	SU access policy
$\bar{T}_s(\mu)$	average long-term SU throughput under policy μ
$\bar{W}_s(\mu)$	average long-term SU access rate under policy μ
$\bar{T}_p(\mu)$	average long-term PU throughput under policy μ
$q_{pp}^{(\mathrm{I})}$	outage prob. at PUrx, when SU is idle
$q_{pp}^{(\mathrm{A})}$	outage prob. at PUrx, when SU is active
$q_{ps}^{(\mathrm{I})}$	prob. that current PU message is in outage at SUrx, given that SU is idle
$q_{ps}^{(\mathrm{A})}$	prob. that current PU message is in outage at SUrx, given that SU is active
$p_{s,\mathrm{buf}}$	prob. that current SU message is buffered (it can be decoded via BIC)

scenario, where the SU transmitter is *far away* from the PU receiver and thus generates negligible interference to the PU.

2.1.2 Structure of the chapter

This chapter is organized as follows. Sec. 2.2 presents the system model. Sec. 2.3 introduces the secondary access policy, the performance metrics and the optimization problem, which is addressed in Sec. 2.4. Sec. 2.5 discusses and analyzes the degenerate cognitive radio network scenario. Sec. 2.6 presents and discusses the numerical results. Finally, Sec. 2.7 concludes the chapter. The proofs of the theorems and lemmas are provided in the appendices at the end of the chapter.

The main symbols used in this chapter are listed in Table 2.1. The notation $\mathbb{N}(x, y)$ for integers x, y denotes the set $\mathbb{N}(x, y) \equiv \{x, x+1, \ldots, y\}$.

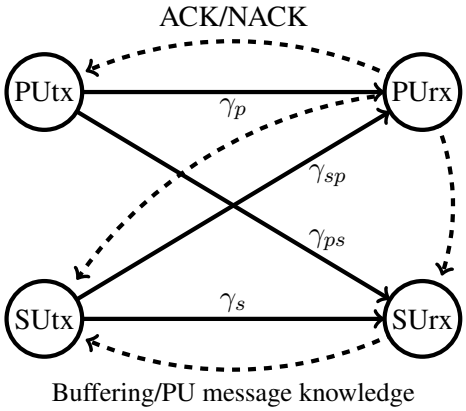


Figure 2.2. *System model*

2.2 System Model

We consider a two-user interference network, depicted in Fig. 2.2, where a primary transmitter and a secondary transmitter, denoted by PUtx and SUtx, respectively, transmit to their respective receivers, PUrx and SUrx, over the direct links PUtx→PUrx and SUtx→SUrx. Their transmissions generate mutual interference over the links PUtx→SUrx and SUtx→PUrx.

Time is divided into time-slots of fixed duration. Each time-slot matches the length of the PU and SU packets, and the transmissions of the PU and SU are assumed to be perfectly synchronized. We adopt the block-fading channel model, *i.e.*, the channel gains are constant within the time-slot duration, and change from time-slot to time-slot. Assuming that the SU and the PU transmit with constant power P_s and P_p, respectively, and that noise at the receivers is zero mean Gaussian with variance σ_w^2, we define the instantaneous Signal to Noise Ratios (SNR) of the links SUtx→SUrx, PUtx→PUrx, SUtx→PUrx and PUtx→SUrx, during the nth time-slot, as $\gamma_s(n)$, $\gamma_p(n)$, $\gamma_{sp}(n)$ and $\gamma_{ps}(n)$, respectively. We model the SNR process $\{\gamma_x(n), n = 0, 1, \dots\}$, where $x \in \{s, p, sp, ps\}$, as i.i.d. over time-slots and independent over the different links, and we denote the average SNR as $\bar{\gamma}_x = \mathbb{E}[\gamma_x]$.

We assume that no Channel State Information (CSI) is available at the transmitters, so that the latter cannot allocate their rate based on the instantaneous link quality, to ensure correct delivery of the packets to their respective receivers. Transmissions may thus undergo outage, when the selected rate is not supported by the current channel quality.

In order to improve reliability, the PU employs Type-I HARQ [14] with deadline $D \geq 1$, *i.e.*, at most D transmissions of the same PU message can be performed, after which the packet is discarded

and a new transmission is performed (the PU is assumed to be backlogged). We define the *primary ARQ state* $t \in \mathbb{N}(1, D)$[1] as the number of ARQ transmission attempts already performed on the current PU message, plus the current one. Namely, $t = 1$ indicates a new PU transmission, and the counter t is increased at each ARQ retransmission, until the deadline D is reached. We assume that the ARQ feedback is received at the PU transmitter by the end of the time-slot, so that, if requested, a retransmission can be performed in the next time-slot.

On the other hand, the SU, in each time-slot, either accesses the channel by transmitting its own message, or stays idle. This decision is based on the access policy μ, defined in Sec. 2.3. The activity of the SU, which is governed by μ, affects the outage performance of the PU, by creating interference to the PU over the link SUtx→PUrx. We denote the primary outage probability when the SU is idle and accesses the channel, respectively, as[2]

$$q_{pp}^{(\mathrm{I})}(R_p) \triangleq \Pr\left(R_p > C\left(\gamma_p\right)\right),\ q_{pp}^{(\mathrm{A})}(R_p) \triangleq \Pr\left(R_p > C\left(\frac{\gamma_p}{1+\gamma_{sp}}\right)\right), \tag{2.1}$$

where R_p denotes the PU transmission rate, measured in bits/s/Hz, $C(x) \triangleq \log_2(1+x)$ is the (normalized) capacity of the Gaussian channel with SNR x at the receiver [20]. This outage definition, as well as the ones introduced later on, assume the use of Gaussian signaling and capacity-achieving coding with sufficiently long codewords. However, our analysis can be extended to include practical codes by computing the outage probabilities for the specific code considered. In (2.1), it is assumed that SU transmissions are treated as background Gaussian noise by the PU. This is a reasonable assumption in CRs in which the PU is oblivious to the presence of SUs. In general, we have $q_{pp}^{(\mathrm{A})}(R_p) \geq q_{pp}^{(\mathrm{I})}(R_p)$, where equality holds if and only if $\gamma_{sp} \equiv 0$ deterministically. We denote the expected PU throughput accrued in each time-slot, when the SU is idle and accesses the channel, as $T_p^{(\mathrm{I})}(R_p) = R_p[1 - q_{pp}^{(\mathrm{I})}(R_p)]$ and $T_p^{(\mathrm{A})}(R_p) = R_p[1 - q_{pp}^{(\mathrm{A})}(R_p)]$, respectively.

2.2.1 Operation of the SU

Unlike the PU that uses a simple Type-I Hybrid ARQ mechanism, it is assumed that the SU uses "best effort" transmission. Moreover, the SU is provided with side-information about the PU, *e.g.*,

[1] We define $\mathbb{N}(n_0, n_1) = \{t \in \mathbb{N}, n_0 \leq t \leq n_1\}$ for $n_0 \leq n_1 \in \mathbb{N}$

[2] Herein, we denote the outage probability as $q_{xy}^{(Z)}$, where x and y are the source and the recipient of the message, respectively (PU if $x, y = p$, SU if $x, y = s$), and $Z \in \{\mathrm{A}, \mathrm{I}\}$ denotes the action of the SU (A if the SU is active and it accesses the channel, I if the SU remains idle). For example, $q_{ps}^{(\mathrm{A})}$ is the probability that the PU message is in outage at SUrx, when SUtx transmits.

ARQ deadline D, PU codebook and feedback information from PUrx (ACK/NACK messages). This is consistent with the common characterization of the PU as a legacy system, and of the SU as an opportunistic and cognitive system, which exploits the primary ARQ feedback to create a best-effort link with maximized throughput, while the flow control mechanisms are left to the upper layers. By overhearing the feedback information from PUrx, the SU can thus track the primary ARQ state t. Moreover, by leveraging the PU codebook, SUrx attempts, in any time-slot, to decode the PU message, which enables the following IC techniques at SUrx:

- *Forward IC* (FIC): by decoding the PU message, SUrx can perform IC in the current as well as in the following ARQ retransmissions, if these occur, to achieve a larger SU throughput;

- *Backward IC* (BIC): SUrx buffers the received signals corresponding to SU transmissions which undergo outage due to severe interference from the PU. These transmissions can later be recovered using IC on the buffered received signals, if the interfering PU message is successfully decoded by SUrx in a subsequent primary ARQ retransmission attempt.

We define the *SU buffer state* $b \in \mathbb{N}(0, B)$ as the number of received signals currently buffered at SUrx, where $B \in \mathbb{N}(0, D-1)$[3] denotes the buffer size. Moreover, we define the *PU message knowledge state* $\Phi \in \{\mathrm{K}, \mathrm{U}\}$, which denotes the knowledge at SUrx about the PU message currently handled by the PU. Namely, if $\Phi = \mathrm{K}$, then SUrx knows the PU message, thus enabling FIC/BIC; conversely ($\Phi = \mathrm{U}$), the PU message is unknown to SUrx.

Remark 2.2.1 (Feedback Information). Note that PUrx needs to report one feedback bit to inform PUtx (and the SU, which overhears the feedback) on the transmission outcome (ACK/NACK). On the other hand, two feedback bits need to be reported by SUrx to SUtx: one bit to inform SUtx as to whether the PU message has been successfully decoded, so that SUtx can track the PU message knowledge state Φ; and one bit to inform SUtx as to whether the received signal has been buffered, so that SUtx can track the SU buffer state b. Herein, we assume ideal (error-free) feedback channels, so that the SU can track (t, b, Φ), and the PU can track the ARQ state t. However, optimization is possible with imperfect observations as well [21].

We now further detail the operation of the SU for $\Phi \in \{\mathrm{K}, \mathrm{U}\}$.

[3]Note that $B \leq D-1$, since the same PU message is transmitted at most D times by PUtx. Once the ARQ deadline D is reached, a new PU transmission occurs, and the buffer is emptied.

2.2.1.1 PU message unknown to SUrx ($\Phi = \mathrm{U}$)

When $\Phi = \mathrm{U}$ and the SU is idle, SUrx attempts to decode the PU message, so as to enable FIC/BIC. A decoding failure occurs if the rate of the PU message, R_p, exceeds the capacity of the channel PUtx→SUrx, with SNR γ_{ps}. We denote the corresponding outage probability as $q_{ps}^{(\mathrm{I})}(R_p) = \Pr(R_p > C(\gamma_{ps}))$.

If the SU accesses the channel, SU transmissions are performed with rate $R_{s\mathrm{U}}$ (bits/s/Hz) and are interfered by the PU. SUrx thus attempts to decode both the SU and PU messages; moreover, if the decoding of the SU message fails due to severe interference from the PU, the received signal is buffered for future BIC recovery. Using standard information-theoretic results [20], with the help of Fig. 2.3, we define the following SNR regions associated with the decodability of the SU and PU messages at SUrx, where $\mathcal{A}^c$ denotes the complementary set of $\mathcal{A}$:[4]

$$\Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p) \triangleq \left\{ (\gamma_s, \gamma_{ps}) : R_{s\mathrm{U}} \leq C(\gamma_s), R_p \leq C(\gamma_{ps}), R_{s\mathrm{U}} + R_p \leq C(\gamma_s + \gamma_{ps}) \right\}, \quad (2.2)$$

$$\bigcup \left\{ (\gamma_s, \gamma_{ps}) : R_{s\mathrm{U}} > C(\gamma_s), R_p \leq C\left(\frac{\gamma_{ps}}{1+\gamma_s}\right) \right\}, \quad (2.3)$$

$$\Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p) \triangleq \left\{ (\gamma_s, \gamma_{ps}) : R_{s\mathrm{U}} \leq C(\gamma_s), R_p \leq C(\gamma_{ps}), R_{s\mathrm{U}} + R_p \leq C(\gamma_s + \gamma_{ps}) \right\} \quad (2.4)$$

$$\bigcup \left\{ (\gamma_s, \gamma_{ps}) : R_p > C(\gamma_{ps}), R_{s\mathrm{U}} \leq C\left(\frac{\gamma_{ps}}{1+\gamma_s}\right) \right\}, \quad (2.5)$$

$$\Gamma_{\mathrm{buf}}(R_{s\mathrm{U}}, R_p) \triangleq \left\{ \Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p) \cup \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p) \right\}^c \bigcap \left\{ (\gamma_s, \gamma_{ps}) : R_{s\mathrm{U}} \leq C(\gamma_s) \right\}. \quad (2.6)$$

The SNR regions (2.2) and (2.4) guarantee that the two rates R_p and $R_{s\mathrm{U}}$ are within the multiple access channel region formed by the two transmitters (PUtx and SUtx) and SUrx [20], so that both the SU and PU messages are correctly decoded via joint decoding techniques. On the other hand, in the SNR region (2.5) (respectively, (2.3)), only the SU (PU) message is successfully decoded at SUrx by treating the interference from the PU (SU) as background noise. If the SNR pair falls outside the two regions (2.4) and (2.5) (respectively, (2.2) and (2.3)), then SUrx incurs a failure in decoding the SU (PU) message. Therefore, when $(\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p)$, SUrx successfully decodes the SU message. The corresponding expected SU throughput is thus given by

$$T_{s\mathrm{U}}(R_{s\mathrm{U}}, R_p) \triangleq R_{s\mathrm{U}} \Pr\left((\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p)\right). \quad (2.7)$$

[4]Herein, we assume optimal joint decoding techniques of the SU and PU messages. Using other techniques, *e.g.*, successive IC, the SNR regions may change accordingly, without providing any further insights in the following analysis.

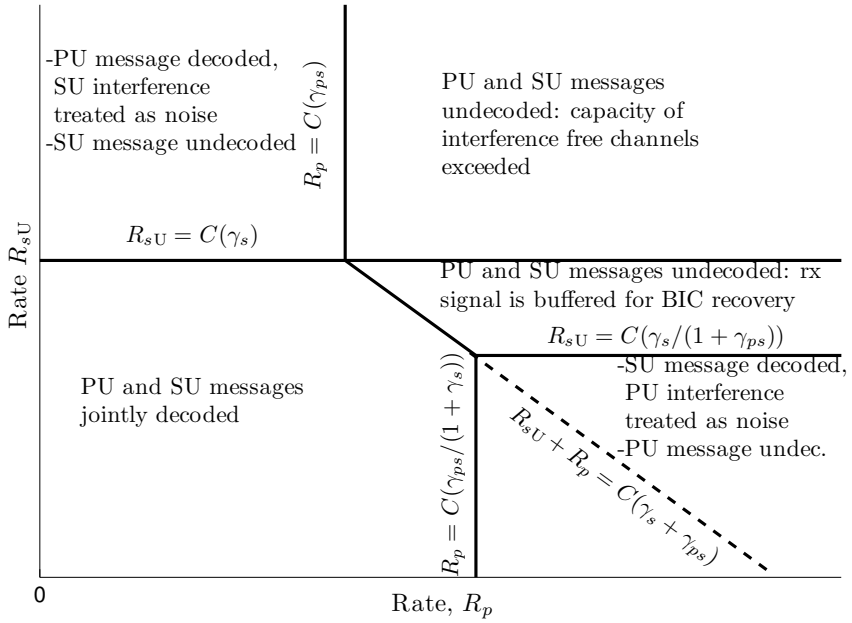


Figure 2.3. *Decodability regions for PU message (rate R_p) and SU message (rate $R_{s\mathrm{U}}$) at SUrx, for a fixed SNR pair (γ_s, γ_{ps}); these regions change according to the fading state (γ_s, γ_{ps})*

Similarly, when $(\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p)$, SUrx successfully decodes the PU message. We denote the corresponding outage probability as $q_{ps}^{(\mathrm{A})}(R_{s\mathrm{U}}, R_p) \triangleq \Pr\left((\gamma_s, \gamma_{ps}) \notin \Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p)\right)$. Note that $q_{ps}^{(\mathrm{A})}(R_{s\mathrm{U}}, R_p) > q_{ps}^{(\mathrm{I})}(R_p)$, since SU transmissions interfere with the decoding of the PU message.

Finally, in (2.6), the decoding of both the SU and PU messages fails, since the SNR pair (γ_s, γ_{ps}) falls outside both regions $\Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p)$ and $\Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p)$. However, the rate $R_{s\mathrm{U}}$ is within the capacity region of the interference free channel ($R_{s\mathrm{U}} \leq C\left(\gamma_s\right)$), so that the SU message can be recovered via BIC, should the PU message become available in a future ARQ retransmission attempt. The received signal is thus buffered at SUrx. We denote the *buffering probability* as

$$\begin{aligned} p_{s,\mathrm{buf}}(R_{s\mathrm{U}}, R_p) &\triangleq \Pr\left((\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{buf}}(R_{s\mathrm{U}}, R_p)\right) \\ &= \Pr\left((\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, 0)\right) - \Pr\left((\gamma_s, \gamma_{ps}) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p)\right) > 0, \end{aligned} \tag{2.8}$$

where the second equality follows from inspection of Fig. 2.3.

2.2.1.2 PU message known to SUrx ($\Phi = \mathrm{K}$)

When $\Phi = \mathrm{K}$, SUrx performs FIC on the received signal, thus enabling interference free SU transmissions. The SU transmits with rate $R_{s\mathrm{K}}$, and the accrued throughput is given by

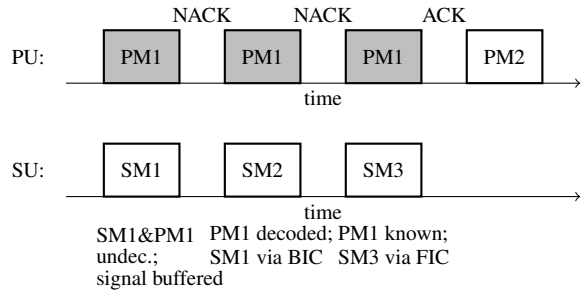


Figure 2.4. *Example of operation of FIC/BIC schemes*

$T_{s\mathrm{K}}(R_{s\mathrm{K}}) = R_{s\mathrm{K}}\Pr\left(R_{s\mathrm{K}} < C(\gamma_s)\right)$.

We now provide an example to illustrate the use of FIC/BIC at SUrx.

Example 1. *Consider a sequence of* 3 *primary retransmission attempts in which the SU always accesses the channel, as depicted in Fig. 2.4. Initially, the PU message is unknown to SUrx, hence the PU message knowledge state is set to* $\Phi = \mathrm{U}$ *in the first time-slot, and the SU transmits with rate* $R_{s\mathrm{U}}$*. Assume that the SNR pair* $(\gamma_s(1), \gamma_{ps}(1))$ *falls in* $\Gamma_{\mathrm{buf}}(R_{s\mathrm{U}}, R_p)$*. Then, neither the SU nor the PU messages are successfully decoded by SUrx, but the received signal is buffered for future BIC recovery. In the second time-slot,* $(\gamma_s(2), \gamma_{ps}(2)) \in \Gamma_{\mathrm{s}}(R_{s\mathrm{U}}, R_p) \cap \Gamma_{\mathrm{p}}(R_{s\mathrm{U}}, R_p)$*, hence both the SU and PU messages are correctly decoded by SUrx, and the PU message knowledge state switches to* $\Phi = \mathrm{K}$*. At this point, SUrx performs BIC on the previously buffered received signal to recover the corresponding SU message. In the third time-slot, SUtx transmits with rate* $R_{s\mathrm{K}}$*, and decoding at SUrx takes place after cancellation of the interference from the PU via FIC.*

We now briefly elaborate on the choice of the transmission rate $R_{s\mathrm{K}}$. Since its value does not affect the outage performance at PUrx (2.1) and the evolution of the ARQ process, $R_{s\mathrm{K}}$ is chosen so as to maximize $T_{s\mathrm{K}}(R_{s\mathrm{K}})$. Therefore, from (2.8) we obtain

$$T_{s\mathrm{K}}(R_{s\mathrm{K}}) \geq T_{s\mathrm{K}}(R_{s\mathrm{U}}) = T_{s\mathrm{U}}(R_{s\mathrm{U}}, R_p) + p_{s,\mathrm{buf}}(R_{s\mathrm{U}}, R_p)R_{s\mathrm{U}} > T_{s\mathrm{U}}(R_{s\mathrm{U}}, R_p). \tag{2.9}$$

Conversely, the choice of the rate $R_{s\mathrm{U}}$ is not as straightforward, since its value reflects a trade-off between the potentially larger throughput accrued with a larger rate $R_{s\mathrm{U}}$ and the corresponding diminished capabilities for IC caused by the more difficult decoding of the PU message by SUrx.

In the following treatment, the rates $R_{s\mathrm{K}}$, $R_{s\mathrm{U}}$ and R_p are assumed to be fixed parameters of the system, and they are not considered part of the optimization (see Sec. 2.6 for further elaboration in this regard). For the sake of notational convenience, we omit the dependence of the quantities defined

above on them. Moreover, for clarity, we consider the case $B = D - 1$ in which SUrx can buffer up to $D - 1$ received signals. However, the following analysis can be extended to a generic value of B.

2.3 Policy Definition and Optimization Problem

We model the evolution of the network as a Markov Decision Process [1, 18]. Namely, we denote the state of the PU-SU system by the tuple (t, b, Φ), where $t \in \mathbb{N}(1, D)$ is the primary ARQ state, $b \in \mathbb{N}(0, B)$ is the SU buffer state and $\Phi \in \{\mathrm{U}, \mathrm{K}\}$ is the PU message knowledge state. (t, b, Φ) takes values in the state space $\mathcal{S} \equiv \mathcal{S}_{\mathrm{U}} \cup \mathcal{S}_{\mathrm{K}}$, where $\mathcal{S}_{\mathrm{K}} \equiv \{(t, 0, \mathrm{K}) : t \in \mathbb{N}(2, D)\}$ and $\mathcal{S}_{\mathrm{U}} \equiv \{(t, b, \mathrm{U}) : t \in \mathbb{N}(1, D), b \in \mathbb{N}(0, t-1)\}$ are the sets of states where the PU message is known and unknown to SUrx, respectively.

The SU follows a *stationary randomized access policy* $\mu \in \mathcal{U} \equiv \{\mu : \mathcal{S} \mapsto [0, 1]\}$, which determines the secondary access probability for each state $\mathbf{s} \in \mathcal{S}$. Note that, from [22], this choice is without loss of optimality for the specific problem at hand. Namely, in state $(t, b, \Phi) \in \mathcal{S}$, the SU is "active", *i.e.*, it accesses the channel, with probability $\mu(t, b, \Phi)$ and stays "idle" with probability $1 - \mu(t, b, \Phi)$. We denote the "active" and "idle" actions as A and I, respectively.

With these definitions at hand, we define the following average long-term metrics under μ: the SU throughput $\bar{T}_s(\mu)$, the SU power expenditure $\bar{P}_s(\mu)$ and the PU throughput $\bar{T}_p(\mu)$, given by

$$\bar{T}_s(\mu) = \lim_{N\to+\infty} \frac{1}{N} \mathbb{E}\left[\sum_{n=0}^{N-1} R_{s\Phi_n} \mathbf{1}\left(\{Q_n = \mathrm{A}\} \cap O_{s,n}^c\right) + R_{s\mathrm{U}} B_n \mathbf{1}(O_{ps,n}^c) \middle| \mathbf{s}_0 \right], \tag{2.10}$$

$$\bar{P}_s(\mu) = P_s \lim_{N\to+\infty} \frac{1}{N} \mathbb{E}\left[\sum_{n=0}^{N-1} \mathbf{1}\left(\{Q_n = \mathrm{A}\}\right) \middle| \mathbf{s}_0 \right], \tag{2.11}$$

$$\bar{T}_p(\mu) = \lim_{N\to+\infty} \frac{1}{N} \mathbb{E}\left[\sum_{n=0}^{N-1} R_p \mathbf{1}\left(O_{p,n}^c\right) \middle| \mathbf{s}_0 \right], \tag{2.12}$$

where n is the time-slot index, $\mathbf{s}_0 \in \mathcal{S}$ is the initial state in time-slot 0; $\Phi_n \in \{\mathrm{K}, \mathrm{U}\}$ is the PU message knowledge state and B_n is the SU buffer state in time-slot n; $Q_n \in \{\mathrm{A}, \mathrm{I}\}$ is the action of the SU, drawn according to the access policy μ; $O_{s,n}$ and $O_{ps,n}$ denote the outage events at SUrx for the decoding of the SU and PU messages, so that $O_{s,n}^c$ and $O_{ps,n}^c$ denote successful decoding of the SU and PU messages by SUrx, respectively; $O_{p,n}$ denotes the outage event at PUrx, so that $O_{p,n}^c$ denotes successful decoding of the PU message by PUrx; and $\mathbf{1}(E)$ is the indicator function of the event E. Note that all the quantities defined above are independent of the initial state $\mathbf{s}_0$. In fact,

starting from any $\mathbf{s}_0 \in \mathcal{S}$, the system reaches with probability 1 the positive recurrent state $(1, 0, \mathrm{U})$ (new PU transmission) within a finite number of time-slots, due to the ARQ deadline. Due to the Markov property, from this state on, the evolution of the process is independent of the initial transient behavior, which has no effect on the time averages defined in (2.10), (2.11) and (2.12).

We study the problem of maximizing the average long-term SU throughput subject to constraints on the average long-term PU throughput loss and SU power. Specifically,

$$\mu^* = \arg\max_{\mu} \bar{T}_s(\mu) \text{ s.t. } \bar{T}_p(\mu) \geq T_p^{(\mathrm{I})}(1 - \epsilon_{\mathrm{PU}}),\ \bar{P}_s(\mu) \leq \mathcal{P}_s^{(\mathrm{th})}, \tag{2.13}$$

where $\epsilon_{\mathrm{PU}} \in [0, 1]$ and $\mathcal{P}_s^{(\mathrm{th})} \in [0, P_s]$ represent the (normalized) maximum tolerated PU throughput loss with respect to the case in which the SU is idle and the SU power constraint, respectively. This problem entails a trade-off in the operation of the SU. On the one hand, the SU is incentivized to transmit in order to increase its throughput and to optimize the buffer occupancy at SUrx (*i.e.*, failed SU transmissions which are potentially recovered via BIC). On the other hand, SU transmissions might jeopardize the correct decoding of the PU message at SUrx, thus impairing the use of FIC/BIC, and might violate the constraints in (2.13).

Under $\mu \in \mathcal{U}$, the state process is a stationary Markov chain, with steady state distribution π_μ [18, 23]. $\pi_\mu(\mathbf{s}), \mathbf{s} \in \mathcal{S}$, is the long-term fraction of the time-slots spent in state $\mathbf{s}$, *i.e.*, $\pi_\mu(\mathbf{s}) = \lim_{N\to+\infty} \frac{1}{N} \sum_{n=0}^{N-1} \Pr_\mu^{(n)}(\mathbf{s}|\mathbf{s}_0)$, where $\Pr_\mu^{(n)}(\mathbf{s}|\mathbf{s}_0)$ is the n-step transition probability of the chain from state $\mathbf{s}_0$.[5] In state (t, b, U), the SU accesses the channel with probability $\mu(t, b, \mathrm{U})$, thus accruing the throughput $\mu(t, b, \mathrm{U}) T_{s\mathrm{U}}$. Moreover, if SUrx successfully decodes the PU message (with probability $1 - q_{ps}^{(\mathrm{I})} - \mu(t, b, \mathrm{U})(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})})$), $bR_{s\mathrm{U}}$ bits are recovered by performing BIC on the buffered received signals, yielding an additional BIC throughput. Similarly, in state $(t, 0, \mathrm{K})$, the SU accrues the throughput $\mu(t, 0, \mathrm{K}) T_{s\mathrm{K}}$. Then, we can rewrite (2.10) and (2.11) in terms of the steady state distribution and of the cost/reward in each state as

$$\bar{T}_s(\mu) = T_{s\mathrm{U}} \bar{W}_s(\mu) + \bar{F}_s(\mu) + \bar{B}_s(\mu),\ \bar{P}_s(\mu) = P_s \bar{W}_s(\mu), \tag{2.14}$$

where the *SU access rate* $\bar{W}_s(\mu)$, *i.e.*, the average long-term number of secondary channel accesses

[5] Similarly to (2.10), (2.11) and (2.12), $\pi_\mu(\mathbf{s})$ is independent of the initial state $\mathbf{s}_0$, due to the recurrence of state $(1, 0, \mathrm{U})$.

per time-slot, the *FIC throughput* $\bar{F}_s(\mu)$ and the *BIC throughput* $\bar{B}_s(\mu)$ are defined as

$$\bar{W}_s(\mu) \triangleq \sum_{\mathbf{s}\in\mathcal{S}} \pi_\mu(\mathbf{s})\,\mu(\mathbf{s}), \tag{2.15}$$

$$\bar{F}_s(\mu) \triangleq \sum_{t=2}^{D} \pi_\mu(t,0,\mathrm{K})\,\mu(t,0,\mathrm{K})\,(T_{s\mathrm{K}} - T_{s\mathrm{U}}), \tag{2.16}$$

$$\bar{B}_s(\mu) \triangleq \sum_{t=1}^{D}\sum_{b=0}^{t-1} \pi_\mu(t,b,\mathrm{U})\,bR_{s\mathrm{U}}\left[1 - q_{ps}^{(\mathrm{I})} - \mu(t,b,\mathrm{U})\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right)\right]. \tag{2.17}$$

In (2.14), $T_{s\mathrm{U}}\bar{W}_s(\mu)$ is the SU throughput attained without FIC/BIC, while the terms $\bar{F}_s(\mu)$ and $\bar{B}_s(\mu)$ account for the throughput gains of FIC and BIC, respectively. Conversely, the PU accrues the throughput $T_p^{(\mathrm{I})}$ if the SU is idle and $T_p^{(\mathrm{A})}$ if the SU accesses the channel, so that (2.12) is given by

$$\bar{T}_p(\mu) = T_p^{(\mathrm{I})} - (T_p^{(\mathrm{I})} - T_p^{(\mathrm{A})})\bar{W}_s(\mu). \tag{2.18}$$

The quantity $(T_p^{(\mathrm{I})} - T_p^{(\mathrm{A})})\bar{W}_s(\mu)$ is referred to as the *PU throughput loss* induced by the secondary access policy μ [19]. The following result follows directly from (2.13), (2.14) and (2.18).

Lemma 2.3.1. *The problem (2.13) is equivalent to*

$$\mu^* = \arg\max_{\mu\in\mathcal{U}} \bar{T}_s(\mu) \text{ s.t. } \bar{W}_s(\mu) \le \min\left\{\frac{(1-q_{pp}^{(\mathrm{I})})\epsilon_{\mathrm{PU}}}{q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})}}, \frac{\mathcal{P}_s^{(\mathrm{th})}}{P_s}\right\} \triangleq \epsilon_{\mathrm{W}}. \tag{2.19}$$

In the next section, we characterize the solution of (2.19). We will need the following definition.

Definition 2.3.1. Let μ be the policy such that secondary access takes place if and only if the PU message is known to SUrx, *i.e.*, $\mu(\mathbf{s}) = 1,\ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}$, $\mu(\mathbf{s}) = 0,\ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{U}}$. We denote the SU access rate achieved by such policy as $\epsilon_{\mathrm{th}} = \bar{W}(\mu)$. The system is in the *low SU access rate regime* if $\epsilon_{\mathrm{W}} \le \epsilon_{\mathrm{th}}$ in (2.19). Otherwise, the system is in the *high SU access rate regime*.

2.4 Optimal Policy

In this section, we characterize in closed form the optimal policy in the low SU access rate regime, and we present an algorithm to derive the optimal policy in the high SU access rate regime.

2.4.1 Low SU Access Rate Regime

The next lemma shows that, in the low SU access rate regime, an optimal policy prescribes that secondary access only takes place in the states where the PU message is known to SUrx, with an equal probability in all such states. It follows that only FIC, and not BIC, is needed in this regime to attain optimal performance.

Lemma 2.4.1. *In the low SU access rate regime $\epsilon_{\mathrm{W}} \leq \epsilon_{\mathrm{th}}$, an optimal policy is given by*[6]

$$\mu^*(\mathbf{s}) = \frac{\epsilon_{\mathrm{W}}}{\epsilon_{\mathrm{th}}},\ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}},\ \mu^*(\mathbf{s}) = 0,\ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{U}}. \tag{2.20}$$

Moreover, $\bar{T}_s(\mu^*) = T_{s\mathrm{K}}\epsilon_{\mathrm{W}}$, $\bar{P}_s(\mu^*) = P_s\epsilon_{\mathrm{W}}$, and $\bar{T}_p(\mu^*) = T_p^{(\mathrm{I})} - (T_p^{(\mathrm{I})} - T_p^{(\mathrm{A})})\epsilon_{\mathrm{W}}$.

Proof. For any policy $\mu \in \mathcal{U}$ obeying the SU access rate constraint $\bar{W}_s(\mu) \leq \epsilon_{\mathrm{W}}$, we have $\bar{T}_s(\mu) \leq \bar{W}_s(\mu)T_{s\mathrm{K}} \leq \epsilon_{\mathrm{W}}T_{s\mathrm{K}}$. The first inequality holds since $\bar{W}_s(\mu)T_{s\mathrm{K}}$ is the long-term throughput achievable when the PU message is known *a priori* at SUrx, which is an upper bound to the performance; the second from the SU access rate constraint. The upper bound $\epsilon_{\mathrm{W}}T_{s\mathrm{K}}$ is achieved by policy (2.20), as can be directly seen by substituting (2.20) in (2.14), (2.15). □

Remark 2.4.1. Note that secondary accesses in states $\mathcal{S}_{\mathrm{U}}$, where the PU message is unknown to SUrx, would obtain a smaller throughput, namely at most $T_{s\mathrm{U}} + p_{s,\mathrm{buf}}R_{s\mathrm{U}} \leq T_{s\mathrm{K}}$, where $T_{s\mathrm{U}}$ is the "instantaneous" throughput and $p_{s,\mathrm{buf}}R_{s\mathrm{U}}$ is the BIC throughput, *possibly* recovered via BIC in a future ARQ retransmission. Therefore, SU accesses in states $\mathcal{S}_{\mathrm{K}}$ are more "cost effective".

2.4.2 High SU Access Rate Regime

In this section, we study the high SU access rate regime in which $\epsilon_{\mathrm{W}} > \epsilon_{\mathrm{th}}$, thus complementing the analysis above for the regime where $\epsilon_{\mathrm{W}} \leq \epsilon_{\mathrm{th}}$. It will be seen that, if $\epsilon_{\mathrm{W}} > \epsilon_{\mathrm{th}}$, unlike in the low SU access rate regime, the SU should generally access the channel also in states $\mathcal{S}_{\mathrm{U}}$ where the PU message is unknown to SUrx in order to achieve the optimal performance. Therefore, both BIC and FIC are necessary to attain optimality. In this section, we derive the optimal policy. We first introduce some necessary definitions and notations.

[6]The optimal policy in the low SU access rate is not unique. In fact, any policy μ such that $\mu(\mathbf{s}) = 0,\ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{U}}$ and $\bar{W}_s(\mu) = \epsilon_{\mathrm{th}}$ is optimal, attaining the same throughput $\bar{T}_s(\mu) = T_{s\mathrm{K}}\epsilon_{\mathrm{th}}$ as (2.20).

Definition 2.4.1 (Secondary access efficiency)**.** We define the *secondary access efficiency* under policy $\mu \in \mathcal{U}$ in state $\mathbf{s} \in \mathcal{S}$ as

$$\eta_{\mu}(\mathbf{s}) = \frac{\frac{\mathrm{d}\bar{T}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}}{\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}}. \tag{2.21}$$

The secondary access efficiency can be interpreted as follows. If the secondary access probability is increased in state $\mathbf{s} \in \mathcal{S}$ by a small amount δ, then the PU throughput loss is increased by an amount equal to $\delta(T_p^{(\mathrm{I})} - T_p^{(\mathrm{A})})\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}$ (from (2.18)), the SU power is increased by an amount equal to $\delta P_s \frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}$ (from (2.14)), and the SU throughput augments or diminishes by an amount equal to $\delta\frac{\mathrm{d}\bar{T}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})}$ (depending on the sign of the derivative). Therefore, $\eta_{\mu}(\mathbf{s})$ yields the rate of increase (or decrease if $\eta_{\mu}(\mathbf{s}) < 0$) of the SU throughput per unit increase of the SU access rate, as induced by augmenting the secondary channel access probability in state $\mathbf{s}$. Equivalently, it measures how efficiently the SU can access the channel in state $\mathbf{s}$, in terms of maximizing the SU throughput gain while minimizing its negative impact on the PU throughput and on the SU power expenditure.

Remark 2.4.2. It is worth noting that the definition of $\eta_{\mu}(\mathbf{s})$ given in Def. 2.4.1 is not completely rigorous. In fact, under a generic policy μ, the Markov chain of the PU-SU system may not be irreducible [23], so that state $\mathbf{s}$ may not be accessible, hence $\pi_{\mu}(\mathbf{s}) = 0$ and $\frac{\mathrm{d}\bar{T}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} = \frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} = 0$. One example is the idle policy $\mu(\mathbf{s}) = 0,\ \forall \mathbf{s}$: since the SU is always idle, the buffer at SUrx is always empty, hence states (t, b, U) with $b > 0$ are never accessed. To overcome this problem, a formal definition is given in Appendix 2.B, by treating the Markov chain of the PU-SU system as the limit of an irreducible Markov chain. $\eta_{\mu}(\mathbf{s})$ is explicitly derived in Lemma 2.7.3 in Appendix 2.B.

We denote the indicator function of state $\mathbf{s}$ as $\delta_{\mathbf{s}} : \mathcal{S} \mapsto \{0, 1\}$, with $\delta_{\mathbf{s}}(\mathbf{s}) = 1$, $\delta_{\mathbf{s}}(\sigma) = 0,\ \forall \sigma \neq \mathbf{s}$. Moreover, we denote the policy at the ith iteration of the algorithm as $\mu^{(i)}$. We are now ready to describe the algorithm that obtains an optimal policy in the high SU access rate regime. An intuitive explanation of the algorithm can be found below.

Algorithm 1 (Derivation of the optimal policy)**.**

1. `INIT`:
   - Let $\mu^{(0)}$ be the policy $\mu^{(0)}(\mathbf{s}) = 0,\ \forall\, \mathbf{s} \in \mathcal{S}_{\mathrm{U}}$, $\mu^{(0)}(\mathbf{s}) = 1,\ \forall\, \mathbf{s} \in \mathcal{S}_{\mathrm{K}}$, and $i = 0$.
   - Let $\mathcal{S}_{\mathrm{idle}}^{(0)} \equiv \{\mathbf{s} \in \mathcal{S} : \mu^{(0)}(\mathbf{s}) = 0\} \equiv \mathcal{S}_{\mathrm{U}}$ be the set of states where the SU is idle.

2. `STAGE i`:

   (a) Compute $\eta_{\mu^{(i)}}(\mathbf{s}), \forall\, \mathbf{s} \in \mathcal{S}_{\text{idle}}^{(i)}$ and let $\mathbf{s}^{(i)} \triangleq \arg\max_{\mathbf{s}\in\mathcal{S}_{idle}^{(i)}} \eta_{\mu^{(i)}}(\mathbf{s})$.

   (b) `STAGE i` $\eta_{\mu^{(i)}}(\mathbf{s}^{(i)}) \leq 0$, `go to STEP 3)`. Otherwise, let $\mu^{(i+1)} = \mu^{(i)} + \delta_{\mathbf{s}^{(i)}}$, $\mathcal{S}_{\text{idle}}^{(i+1)} \equiv \mathcal{S}_{\text{idle}}^{(i)} \setminus \{\mathbf{s}^{(i)}\}$.

   (c) Set $i := i+1$. If $\mathcal{S}_{\text{idle}}^{(i)} \equiv \emptyset$, `go to STEP 3)`. Otherwise, `repeat from STEP 2)`.

3. Let $N = i$, the sequence of states $(\mathbf{s}^{(0)}, \ldots, \mathbf{s}^{(N-1)})$ and of policies $(\mu^{(0)}, \ldots, \mu^{(N-1)})$.

4. `Optimal policy`: given ϵ_{W},

   (a) If $\bar{W}_s(\mu^{(N-1)}) \leq \epsilon_{\text{W}}$, then $\mu^* = \mu^{(N-1)}$.

   (b) Otherwise, $\mu^* = \lambda\mu^{(j)} + (1-\lambda)\mu^{(j+1)}$, where $j \triangleq \max\left\{i : \bar{W}_s\left(\mu^{(i)}\right) \leq \epsilon_{\text{W}}\right\}$ and $\lambda \in (0,1]$ uniquely solves $\bar{W}_s(\lambda\mu^{(j)} + (1-\lambda)\mu^{(j+1)}) = \epsilon_{\text{W}}$.

The algorithm, starting from the optimal policy for the case $\epsilon_{\text{W}} = \epsilon_{\text{th}}$ (Lemma 2.4.1), ranks the states in the set $\mathcal{S}_{\text{U}}$ in decreasing order of secondary access efficiency, and iteratively allocates the secondary access to the state with the highest efficiency, among the states where the SU is idle. The rationale of this step is that secondary access in the most efficient state yields the steepest increase of the SU throughput, per unit increase of the SU access rate or, equivalently, of the PU throughput loss and of the SU power expenditure. The optimality of Algorithm 1 is established in the following theorem.

Theorem 2.4.2. *Algorithm 1 returns an optimal policy for the optimization problem (2.19).*

Proof. See Appendix 2.C. □

2.5 Special Case: degenerate cognitive radio network scenario

We point out that Algorithm 1 determines the optimal policy for a generic set of system parameters. However, the resulting optimal policy does not always have a structure that is easily interpreted. In this section, we consider a special case of the general model discussed so far, a *degenerate cognitive radio network*, where the activity of the PU is unaffected by the transmissions of the SU, *i.e.*, the channel gain between the SU transmitter and the PU receiver is zero.

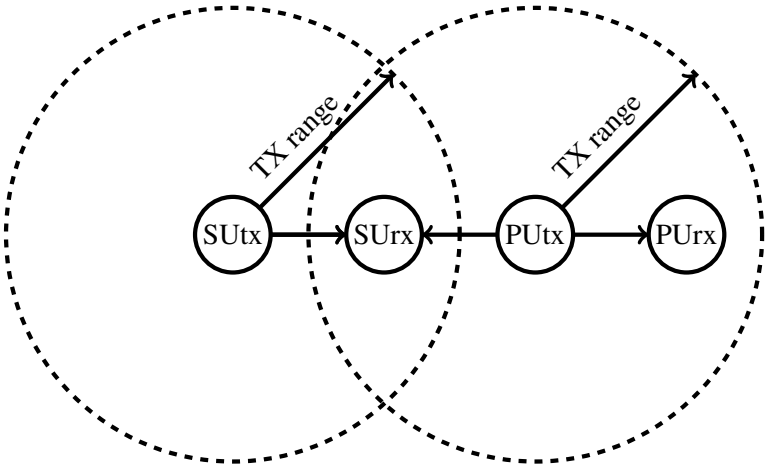


Figure 2.5. *Degenerate cognitive radio network*

Consider the scenario depicted in Fig. 2.5, where PUrx is outside the transmission range of SUtx, whereas SUrx is inside the transmission range of both SUtx and SUrx. In this scenario, the interference produced by SU to PU is negligible. In contrast, the PU produces significant interference at the SU receiver. The SU thus potentially benefits by employing the BIC and FIC mechanisms. We denote this scenario as a *Degenerate cognitive radio network*, and we model it by assuming that the SNR of the interfering link SUtx→PUrx is deterministically equal to zero, *i.e.*, $\gamma_{sp} = 0$. From (2.1), we then have $q_{pp}^{(\mathrm{I})} = q_{pp}^{(\mathrm{A})} \triangleq q_{pp}$, *i.e.*, the outage performance of the PU is unaffected by the activity of the SU, and the primary ARQ process is independent of the secondary access policy. We define

$$\Delta_s \triangleq \frac{T_{s\mathrm{K}} - T_{s\mathrm{U}} - p_{s,\mathrm{buf}} R_{s\mathrm{U}}}{R_{s\mathrm{U}}}. \tag{2.22}$$

From (2.9), it follows that $\Delta_s \geq 0$, with equality if $R_{s\mathrm{U}} = R_{s\mathrm{K}}$. Therefore, $R_{s\mathrm{U}}\Delta_s$ is the marginal throughput gain accrued in the states where the PU message is known to SUrx, over the throughput accrued in the states where the PU message is unknown (instantaneous throughput $T_{s\mathrm{U}}$ plus BIC throughput $p_{s,\mathrm{buf}} R_{s\mathrm{U}}$, possibly recovered in a future ARQ retransmission). The following lemma proves that, if the marginal throughput gain Δ_s is "small", the secondary accesses in the high SU access rate regime in a degenerate cognitive radio network are allocated, in order, to the states in $\mathcal{S}_{\mathrm{K}}$ (Lemma 2.4.1), then to the idle states (t, b, U) in $\mathcal{S}_{\mathrm{U}}$, giving priority to states with low b and t over states with high b and t, respectively. An illustrative example of the optimal policy for this scenario is given in Fig. 2.6.

Lemma 2.5.1. *In the degenerate cognitive radio network scenario with $q_{pp}^{(\mathrm{A})} = q_{pp}^{(\mathrm{I})} = q_{pp}$, if*

$$\Delta_s < \frac{1 - q_{ps}^{(\mathrm{A})}}{q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}} p_{s,\mathrm{buf}}, \tag{2.23}$$

the sequence of policies $(\mu^{(0)}, \ldots, \mu^{(N-1)})$ returned by Algorithm 1 is such that, $\forall i \in \mathbb{N}(0, N-1)$,

$$\mu^{(i)}(\mathbf{s}) = 1, \ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}, \tag{2.24}$$

$$\mu^{(i)}(t, b, \mathrm{U}) = \begin{cases} 1 & b < b^{(i)}(t) \\ 0 & b \geq b^{(i)}(t), \end{cases}, \ \forall (t, b, \mathrm{U}) \in \mathcal{S}_{\mathrm{U}}, \tag{2.25}$$

where $b^{(i)}(t)$ is non-increasing in t and non-decreasing in i, with $b^{(0)}(t) = 0$ and $b^{(N-1)}(t) = \bar{b}_{\max}(t)$, i.e.,

$$\bar{b}_{\max}(t) = b^{(N-1)}(t) \geq \cdots \geq b^{(i)}(t) \geq b^{(i-1)}(t) \geq \cdots \geq b^{(0)}(t) = 0. \tag{2.26}$$

$$b^{(i)}(1) \geq b^{(i)}(2) \geq \cdots \geq b^{(i)}(t-1) \geq b^{(i)}(t) \geq \cdots \geq b^{(i)}(D), \tag{2.27}$$

where

$$\bar{b}_{\max}(t) = \left\lceil \frac{\begin{array}{l} \frac{T_{s\mathrm{U}}}{R_{s\mathrm{U}}} \left[1 - q_{pp}\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) A_0(t+1)\right] \\ + \left(\frac{1 - q_{ps}^{(\mathrm{A})}}{q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}} p_{s,\mathrm{buf}} - \Delta_s\right) q_{pp}\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) A_0(t+1) \end{array}}{\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right)\left(1 - q_{pp}(1 - q_{ps}^{(\mathrm{I})}) A_0(t+1)\right)} \right\rceil - 1 \tag{2.28}$$

and we have defined

$$A_0(\tau) \triangleq \frac{1 - q_{pp}^{D-\tau+1} q_{ps}^{(\mathrm{I})(D-\tau+1)}}{1 - q_{pp} q_{ps}^{(\mathrm{I})}}, \tag{2.29}$$

$$A_1(\tau) \triangleq \frac{1 - q_{pp}^{D-\tau+1}}{1 - q_{pp}}. \tag{2.30}$$

Proof. See Appendix 2.D. □

Remark 2.5.1. Interestingly, this is the same result derived in our work [24] for $D = 2$. However, therein the result was shown to hold for general $q_{pp}^{(\mathrm{A})} \geq q_{pp}^{(\mathrm{I})}$ (not necessarily a degenerate cognitive

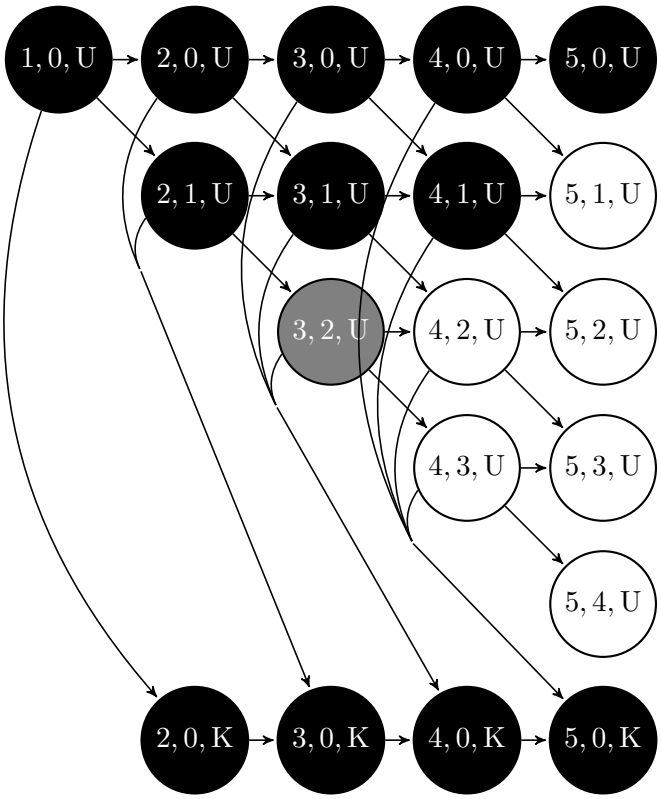


Figure 2.6. *Illustrative example of the structure of the optimal secondary access policy for the degenerate cognitive radio network; the SU is active in the black states, idle in the white ones, and randomly accesses the channel in the gray state; the arrows indicate the possible state transitions (transitions to state* $(1, 0, \mathrm{U})$ *are omitted).*

radio network), whereas Lemma 2.5.1 holds for general D but only for a degenerate cognitive radio network scenario.

The lemma dictates that, in the degenerate cognitive radio network scenario, the SU should restrict its channel accesses to the states corresponding to a low primary ARQ index and small buffer occupancy at the SU receiver. Alternatively, the larger the ARQ index or the buffer occupancy, the smaller the incentive to access the channel. By doing so, the SU maximizes the buffer occupancy in the early HARQ retransmission attempts, and invests in the future BIC recovery. When the primary ARQ state t approaches the deadline D, the SU is incentivized to idle so as to help SUrx to decode the PU message, thus enabling the recovery of the failed SU transmissions from the buffered received signals via BIC, before the ARQ deadline D is reached and the buffer is depleted. Moreover, when the buffer state b grows, since $q_{ps}^{(\mathrm{A})} > q_{ps}^{(\mathrm{I})}$, the instantaneous reward accrued by staying idle ($(1 - q_{ps}^{(\mathrm{I})})bR_{s\mathrm{U}}$) approaches and, at some point, becomes larger than the reward accrued by transmitting ($T_{s\mathrm{U}} + (1 - q_{ps}^{(\mathrm{A})})bR_{s\mathrm{U}}$), hence the incentive to stay idle grows. On the other hand, if Δ_s

is large, then the marginal throughput gain accrued in the states where the PU message is known to SUrx, over the throughput accrued in the states where the PU message is unknown, is large. The SU is thus incentivized to stay idle in the initial ARQ rounds, so as to help SUrx decode the PU message. Therefore, for large Δ_s, the optimal policy may not obey the structure of Lemma 2.5.1.

As a final remark, note that, in the degenerate cognitive radio network scenario, the only limitation to the activity of the SU is the secondary power expenditure $\bar{P}_s(\mu)$, since the primary throughput is unaffected. In the special case $\mathcal{P}_s^{(\mathrm{th})} = P_s$ in (2.13), neither the secondary power expenditure nor the primary throughput degradation limit the activity of the SU, hence the optimal policy solves the unconstrained maximization problem $\mu^* = \arg\max_\mu \bar{T}_s(\mu)$, whose solution follows as a corollary of Lemma 2.5.1.

Corollary 2.5.2. *In the degenerate cognitive radio network scenario, the solution of the unconstrained optimization problem $\mu^* = \arg\max_\mu \bar{T}_s(\mu)$ yields*

$$\mu^*(\mathbf{s}) = 1, \ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}, \tag{2.31}$$

$$\mu^*(t, b, \mathrm{U}) = \begin{cases} 1 & b < \bar{b}_{\max}(t) \\ 0 & b \geq \bar{b}_{\max}(t), \end{cases} \ , \ \forall (t, b, \mathrm{U}) \in \mathcal{S}_{\mathrm{U}}, \tag{2.32}$$

where $\bar{b}_{\max}(t)$ is defined in (2.28).

2.6 Numerical Results

We consider a scenario with Rayleigh fading channels, *i.e.*, the SNR $\gamma_x, \ x \in \{s, p, sp, ps\}$, is an exponential random variable with mean $\mathbb{E}[\gamma_x] = \bar{\gamma}_x$. We consider the following parameters, unless otherwise stated. The average SNRs are set to $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_p = 10$, $\bar{\gamma}_{sp} = 2$. The ARQ deadline is $D = 5$. $R_{s\mathrm{K}}$ is chosen as $R_{s\mathrm{K}} = \arg\max_{R_s} T_{s\mathrm{K}}(R_s)$. The PU rate R_p is chosen as the maximizer of the instantaneous PU throughput under an idle SU, *i.e.*, $R_p = \arg\max_R T_p^{(\mathrm{I})}(R)$. For the rate $R_{s\mathrm{U}}$, we evaluate the two cases $R_{s\mathrm{U}} = R_{s\mathrm{U}}^*$ and $R_{s\mathrm{U}} = R_{s\mathrm{K}}$, where $R_{s\mathrm{U}}^* = \arg\max_{R_s} T_{s\mathrm{U}}(R_s, R_p)$. The former maximizes the instantaneous throughput under interference from the PU, thus neglecting the buffering capability at SUrx; therefore, the choice $R_{s\mathrm{U}} = R_{s\mathrm{U}}^*$ reflects a pessimistic expectation of the ability of SUrx to decode the PU message and to enable BIC. As to the latter, from (2.9) we have $R_{s\mathrm{U}} = R_{s\mathrm{K}} = \arg\max_{R_s} T_{s\mathrm{U}}(R_s, R_p) + p_{s,\mathrm{buf}}(R_s, R_p) R_{s\mathrm{K}}$, hence $R_{s\mathrm{U}} = R_{s\mathrm{K}}$ maximizes the sum of the instantaneous throughput and the future throughput possibly recovered via BIC, thus

PU		
$R_p \simeq 2.52$	$q_{pp}^{(\mathrm{I})} \simeq 0.38$	$q_{pp}^{(\mathrm{A})} \simeq 0.68$
SU, $R_{s\mathrm{U}} = \arg\max_{R_s} T_{s\mathrm{U}}(R_s, R_p)$		
$R_{s\mathrm{U}} = 1.12$ $q_{ps}^{(\mathrm{I})} \simeq 0.61$ $R_{s\mathrm{K}} \simeq 1.91$	$T_{s\mathrm{U}} \simeq 0.59$ $q_{ps}^{(\mathrm{A})} \simeq 0.74$ $T_{s\mathrm{K}} \simeq 1.10$	$p_{s,\mathrm{buf}} = 0.26$
SU, $R_{s\mathrm{U}} = R_{s\mathrm{K}}$		
$R_{s\mathrm{U}} \simeq 1.91$ $q_{ps}^{(\mathrm{I})} \simeq 0.61$ $R_{s\mathrm{K}} \simeq 1.91$	$T_{s\mathrm{U}} \simeq 0.40$ $q_{ps}^{(\mathrm{A})} \simeq 0.88$ $T_{s\mathrm{K}} \simeq 1.10$	$p_{s,\mathrm{buf}} = 0.37$

Table 2.2. *parameters of the SU and PU, for the SNRs* $\bar{\gamma}_s = 5$, $\bar{\gamma}_p = 10$, $\bar{\gamma}_{ps} = 5$, $\bar{\gamma}_{sp} = 2$.

reflecting an optimistic expectation of the ability of SUrx to decode the PU message, which enables BIC. The PU throughput loss constraint is set to $\epsilon_{\mathrm{PU}} = 0.2$, and the constraint on the SU power is set to $\mathcal{P}_s^{(\mathrm{th})} = P_s$ (inactive). The resulting values of the system parameters are listed in Table 2.2.

We consider the following schemes: "FIC/BIC", which employs both FIC and BIC; the optimal "FIC/BIC" policy is derived using Algorithm 1 and Lemma 2.4.1; "FIC only", which does not employ the buffering mechanism (*i.e.*, $B = 0$; the optimal policy is obtained by letting $p_{s\mathrm{buf}} = 0$ in Algorithm 1, *i.e.*, SU transmissions are discarded in case of transmission failure); "no FIC/BIC", which employs neither BIC nor FIC. In this case, the SU message is decoded by leveraging the PU codebook structure [25]; however, possible knowledge of the PU message gained during the decoding operation is only used in the slot where the PU message is acquired, but is neglected in the past/future PU retransmissions. For "no FIC/BIC", the optimal policy consists in accessing the channel with a constant probability in all time-slots, independently of the underlying state, so as to attain the PU throughput loss constraint with equality. "PM known" refers to an ideal scenario where SUrx perfectly knows the current PU message in advance, and removes its interference; specifically, SUtx transmits with rate $R_{s\mathrm{K}}$, thus accruing the throughput $T_{s\mathrm{K}}$ at each secondary access; "PM known" thus yields an upper bound to the performance of any other policy considered.

In Fig. 2.7, we plot the SU throughput versus the PU throughput, obtained by varying the SU access rate constraint ϵ_{W} in (2.19) from 0 to 1. As expected, the best performance is attained by "FIC/BIC", since the joint use of BIC and FIC enables IC at SUrx over the entire sequence of PU retransmissions. "FIC only" incurs a throughput penalty (except in the low SU access rate regime $\bar{T}_p(\mu) \geq 1.37$ where, from Lemma 2.4.1, "FIC/BIC" does not employ BIC), since the SU transmissions which undergo outage due to severe interference from the PU are simply dropped. "no FIC/BIC" incurs a further throughput loss, since possible knowledge about the PU message is not exploited to

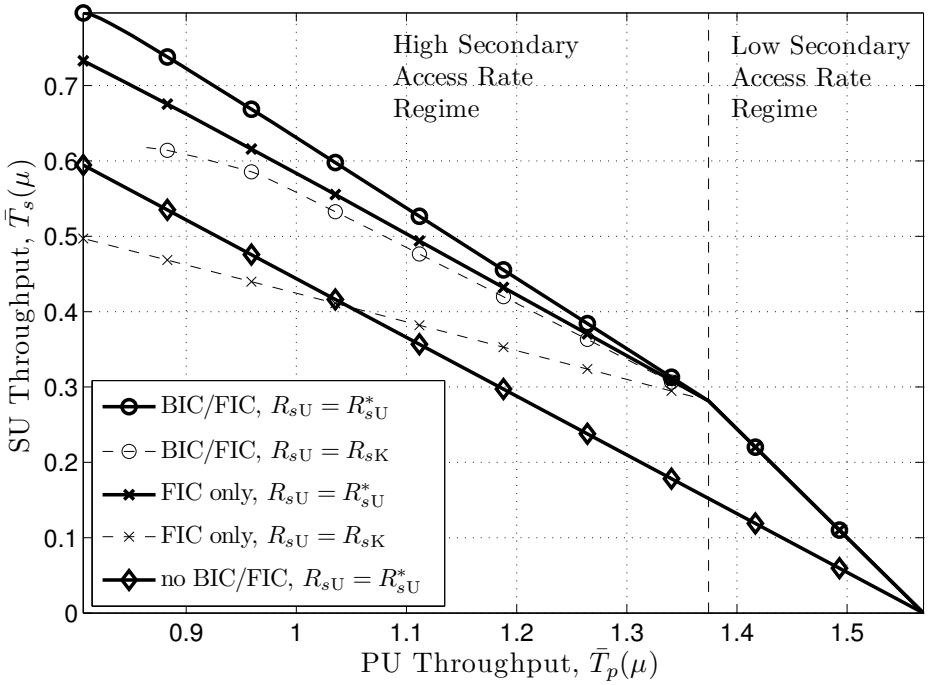


Figure 2.7. *SU throughput vs PU throughput.* $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$. *The other parameters are given in Table 2.2.*

perform IC. Concerning the choice of the transmission rates, we note that the selection $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$ outperforms $R_{s\mathrm{U}} = R_{s\mathrm{K}}$ for the scenario considered. Note that, with $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$, the SU accrues a larger instantaneous throughput ($T_{s\mathrm{U}}$), but FIC and BIC are impaired, since both the buffering probability (2.8), $p_{s,\mathrm{buf}}$, and the probability that SUrx does not successfully decode the PU message, $q^{(\mathrm{A})}_{ps}$, diminish. Hence, in this case the instantaneous throughput maximization has a stronger impact on the performance than enabling FIC/BIC at SUrx.

In Fig. 2.8, we plot the SU throughput versus the SNR ratio $\bar{\gamma}_{sp}/\bar{\gamma}_p$, where $\bar{\gamma}_p = 5$ and $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$. Note that, when $\bar{\gamma}_{sp}/\bar{\gamma}_p = 0$, we obtain the degenerate cognitive radio network scenario, investigated in Sec. 2.5, for which the optimal policy is defined in Corollary 2.5.2. We observe that, for $\bar{\gamma}_{sp}/\bar{\gamma}_p \leq 0.5$, the SU throughput increases. In fact, in this regime the activity of the SU causes little harm to the PU, and the constraint on the PU throughput loss is inactive. The SU thus maximizes its own throughput. As $\bar{\gamma}_{sp}$ increases from 0 to $0.5\bar{\gamma}_p$, the activity of the SU induces more frequent primary ARQ retransmissions, hence there are more IC opportunities available and the SU throughput augments. On the other hand, as $\bar{\gamma}_{sp}$ grows beyond $0.5\bar{\gamma}_p$, the constraint on the PU throughput loss becomes active, secondary accesses become more and more harmful to the PU and take place more and more sparingly, hence the SU throughput degrades.

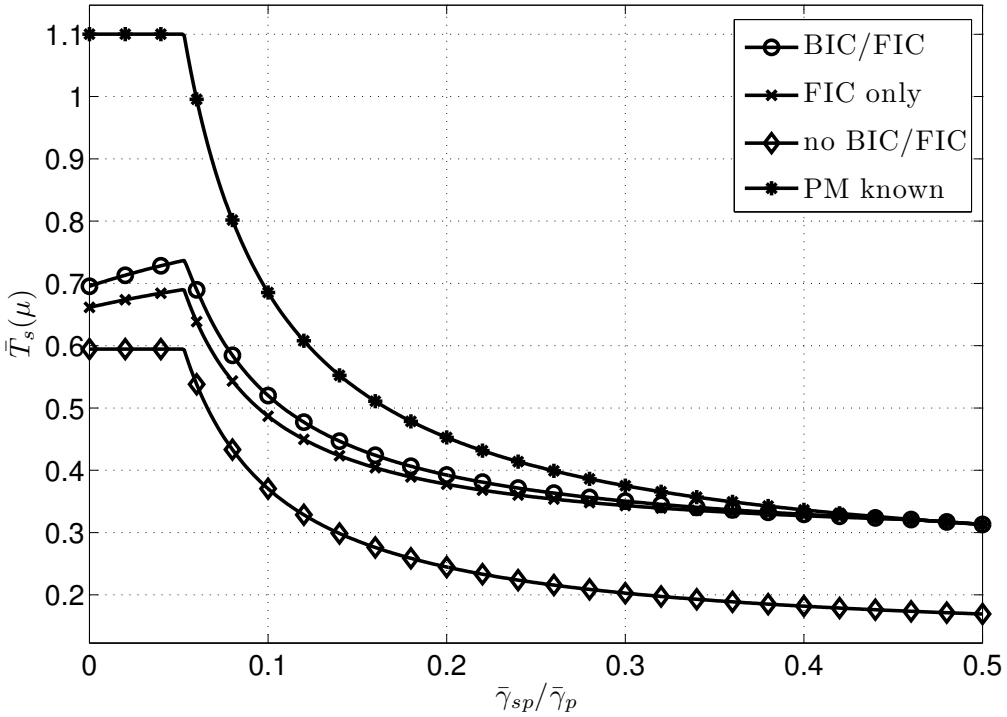


Figure 2.8. *SU throughput vs SNR ratio $\bar{\gamma}_{sp}/\bar{\gamma}_p$. PU throughput loss constraint $\epsilon_{\rm PU} = 0.2$. $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_p = 10$. $R_{s\rm U} = R^*_{s\rm U}$.*

In Fig. 2.9, we plot the SU throughput versus the SNR ratio $\bar{\gamma}_{ps}/\bar{\gamma}_s$, where $\bar{\gamma}_s = 5$ and $R_{s\rm U} = R^*_{s\rm U}$, which is a function of $\bar{\gamma}_{ps}$. We notice that, when $\bar{\gamma}_{ps} = 0$, the upper bound is achieved with equality, since the SU operates under no interference from the PU. The upper bound is approached also for $\bar{\gamma}_{ps} \gg \bar{\gamma}_s$, corresponding to a strong interference regime where, with high probability, SUrx can successfully decode the PU message, remove its interference from the received signal, and then attempt to decode the SU message. The worst performance is attained when $\bar{\gamma}_{ps} \simeq \bar{\gamma}_s/2$. In fact, the interference from the PU is neither weak enough to be simply treated as noise, nor strong enough to be successfully decoded and then removed.

In Fig. 2.10, we plot the SU throughput versus the SU rate ratio $R_{s\rm U}/R_{s\rm K}$, where $R_{s\rm K} \simeq 1.91$ is kept fixed. Clearly, "no FIC/BIC" attains the best performance for $R_{s\rm U} = R^*_{s\rm U}$, which maximizes the throughput $T_{s\rm U}(R_{s\rm U}, R_p)$ achieved when neither FIC nor BIC are used. On the other hand, the performance of "FIC/BIC" is maximized for a slightly larger value of $R_{s\rm U}$. In fact, this value reflects the optimal trade-off between maximizing the throughput $T_{s\rm U}$ ($R_{s\rm U} \simeq 0.59 R_{s\rm K}$ in Fig. 2.11), maximizing the buffering probability, $p_{s,\rm buf}$ ($R_{s\rm U} \to 1$), and minimizing the probability that SUrx does not successfully decode the PU message, $q^{(\rm A)}_{ps}$ ($R_{s\rm U} \to 0$). Finally, "FIC only" is optimized by

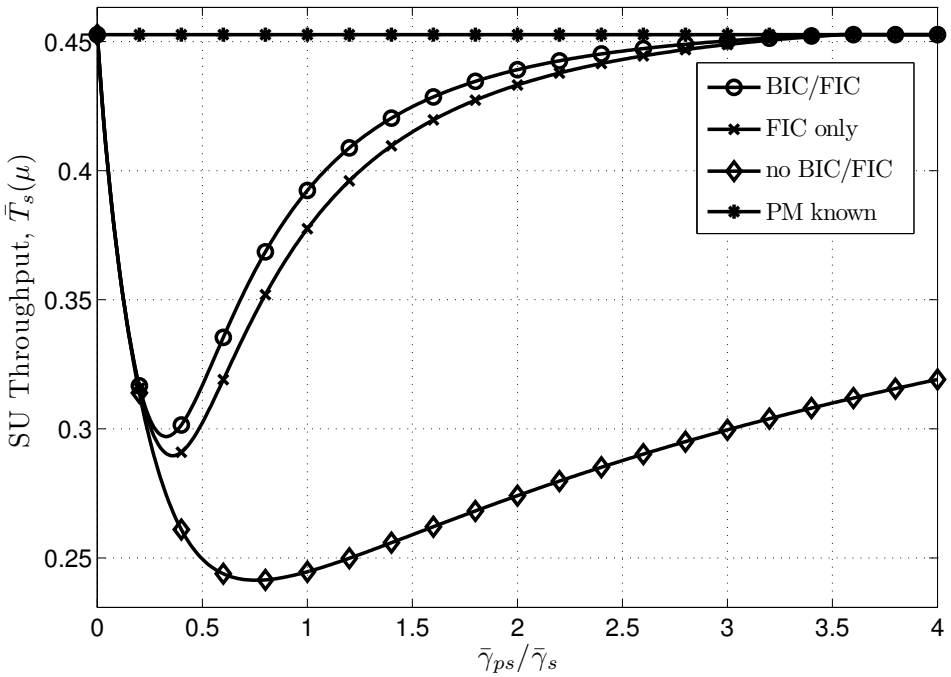


Figure 2.9. *SU throughput vs SNR ratio* $\bar{\gamma}_{ps}/\bar{\gamma}_s$. *PU throughput loss constraint* $\epsilon_{\mathrm{PU}} = 0.2$. $\bar{\gamma}_s = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$. $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$.

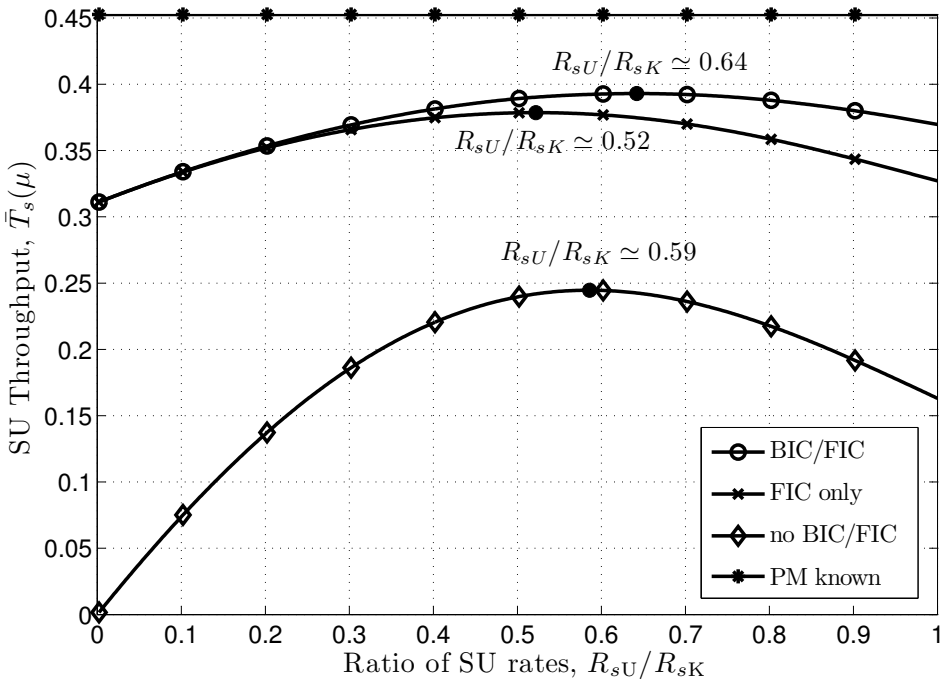


Figure 2.10. *SU throughput vs SU rate ratio* $R_{s\mathrm{U}}/R_{s\mathrm{K}}$. $R_{s\mathrm{K}} \simeq 1.91$ *is kept fixed. PU throughput loss constraint* $\epsilon_{\mathrm{PU}} = 0.2$. $\bar{\gamma}_s = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$, $\bar{\gamma}_{ps} = 5$.

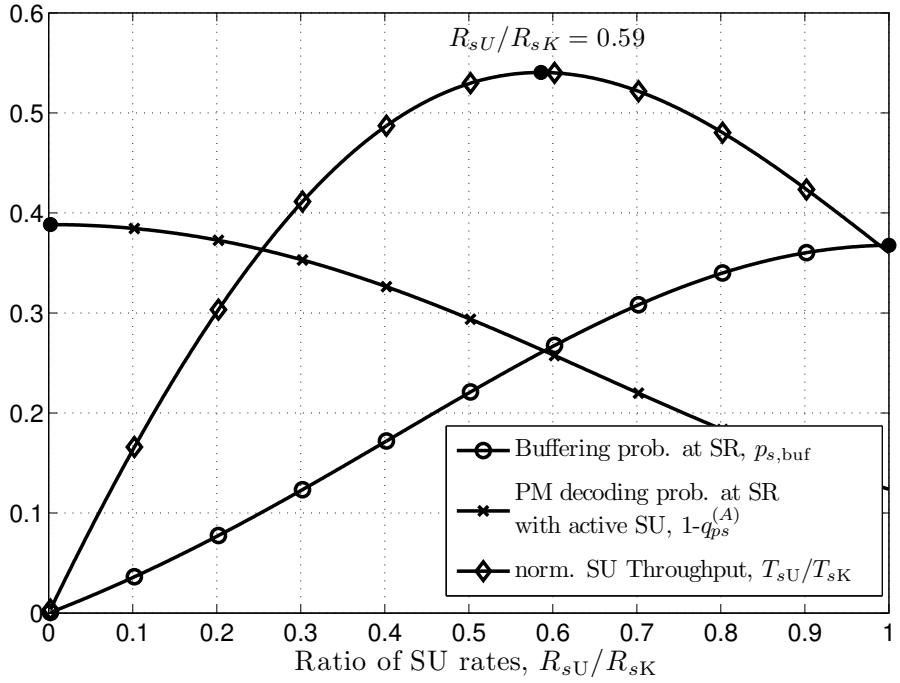


Figure 2.11. *Probabilities* $p_{s,\mathrm{buf}}$, $1 - q_{ps}^{(\mathrm{A})}$ *and normalized SU throughput* $T_{s\mathrm{U}}$ *vs the SU rate ratio* $R_{s\mathrm{U}}/R_{s\mathrm{K}}$. $R_{s\mathrm{K}} \simeq 1.91$ *is kept fixed.* $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$.

$R_{s\mathrm{U}} \simeq 0.52 R_{s\mathrm{K}} < R_{s\mathrm{U}}^*$. Since "FIC only" does not use BIC, this value reflects the optimal trade-off between maximizing $T_{s\mathrm{U}}$ and minimizing $q_{ps}^{(\mathrm{A})}$ ($R_{s\mathrm{U}} \to 0$).

In Fig. 2.12, we plot the SU throughput versus the ARQ deadline D. We notice that, when $D = 1$, all the IC mechanisms considered attain the same performance as "no FIC/BIC". In fact, this is a scenario where the PU does not employ ARQ, hence no redundancy is introduced in the primary transmission process. Interestingly, by employing FIC or BIC, the performance improves as D increases. In fact, the larger D, the more the redundancy introduced by the primary ARQ process, hence the more the opportunities for FIC/BIC at SUrx.

2.7 Conclusions

We have investigated the idea of leveraging the redundancy introduced by the ARQ protocol implemented by a Primary User (PU) to perform Interference Cancellation (IC) at the receiver of a Secondary User (SU) pair: the SU receiver (SUrx), after decoding the PU message, exploits this knowledge to perform *Forward IC* (FIC) in the following ARQ retransmissions and *Backward IC* (BIC) in the previous ARQ retransmissions, corresponding to SU transmissions whose decoding failed due to

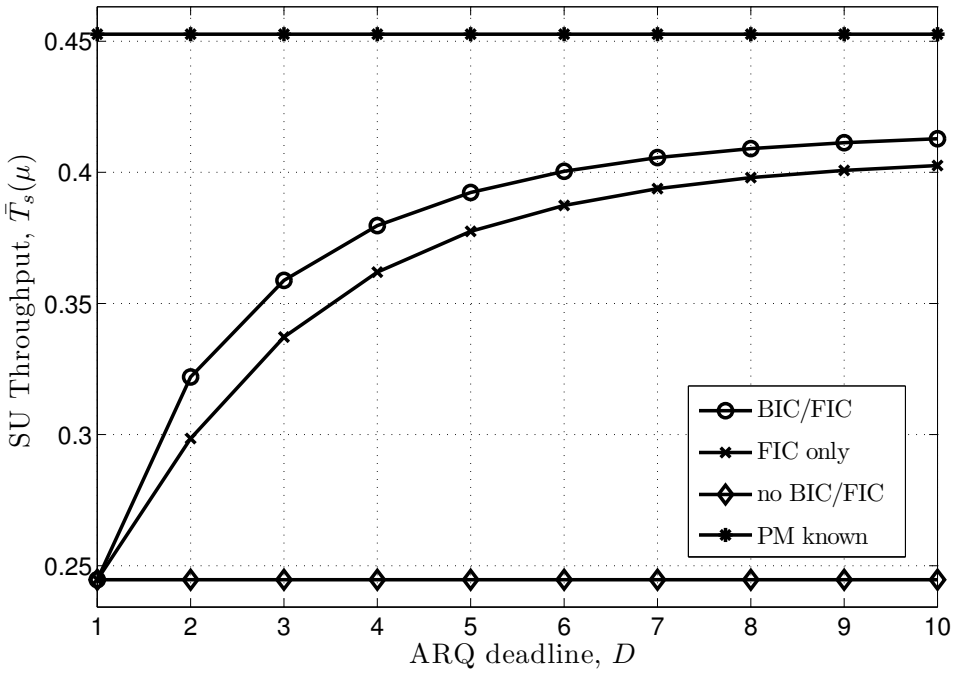


Figure 2.12. *SU throughput vs ARQ deadline D. PU throughput loss constraint $\epsilon_{\mathrm{PU}} = 0.2$. $\bar{\gamma}_s = \bar{\gamma}_{ps} = 5$, $\bar{\gamma}_{sp} = 2$, $\bar{\gamma}_p = 10$. $R_{s\mathrm{U}} = R^*_{s\mathrm{U}}$.*

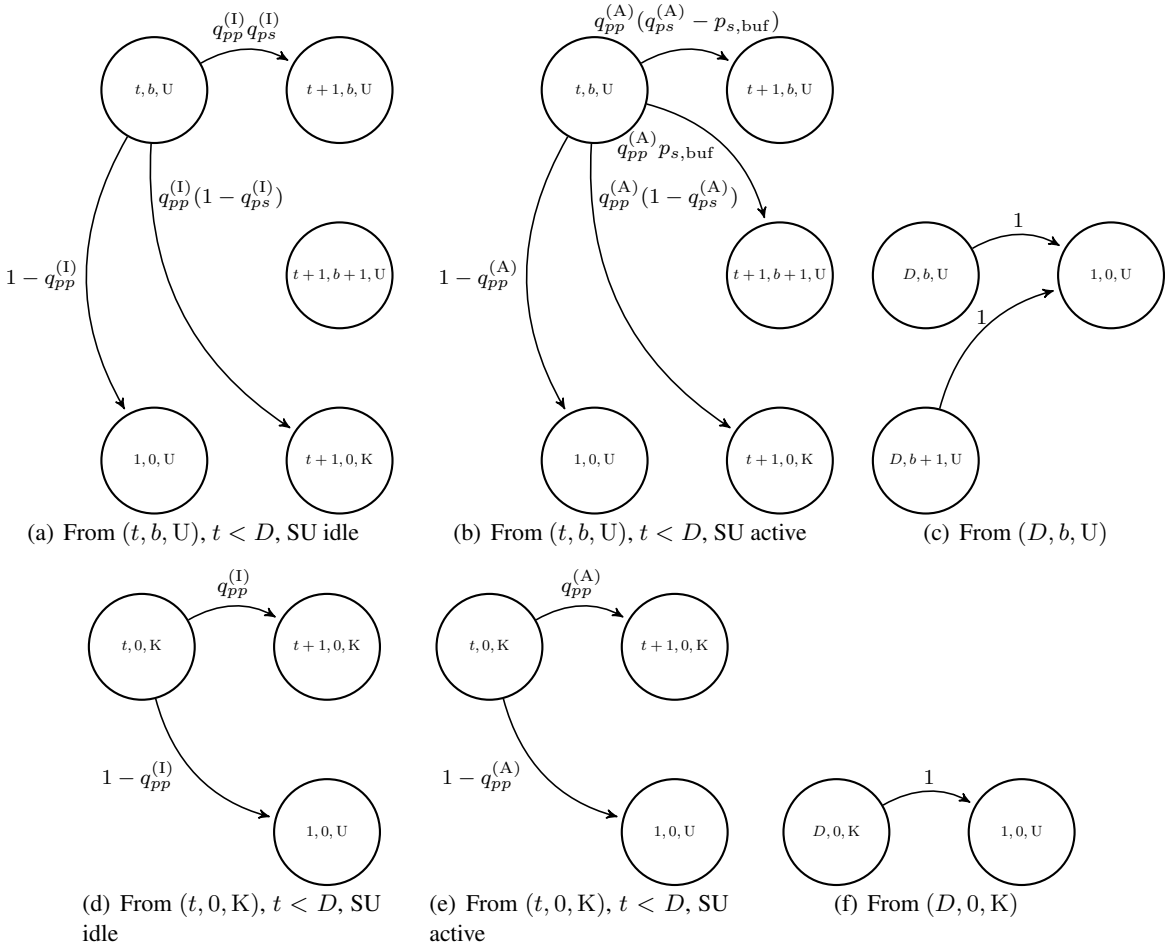
severe interference from the PU. We have employed a stochastic optimization approach to optimize the SU access strategy which maximizes the average long-term SU throughput, under constraints on the average long-term PU throughput degradation and SU power expenditure. We have proved that the SU prioritizes its channel accesses in the states where SUrx knows the PU message, thus enabling FIC, and we have provided an algorithm to optimally allocate additional secondary access opportunities in the states where the PU message is unknown. Finally, we have shown numerically the throughput gain of the proposed schemes.

Appendix 2.A: SU and PU performance metrics and properties

In this appendix, we compute $\bar{T}_s(\mu)$, $\bar{W}_s(\mu)$ and state properties of $\bar{W}_s(\mu)$.

Definition 2.7.1. We define $\mathbf{G}_\mu(t,b,\Phi)$, $\mathbf{V}_\mu(t,b,\Phi)$ and $\mathbf{D}_\mu(t,b,\Phi)$ as the average throughput, the average number of secondary channel accesses and the average number of time-slots, respectively, accrued starting from state (t,b,Φ) until the end of the primary ARQ cycle under policy μ (*i.e.*, until the recurrent state $(1,0,\mathrm{U})$ is reached). Starting from $\mathbf{X}_\mu(D+1,b,\Phi) = 0,\ \forall b, \forall \Phi \in \{\mathrm{U},\mathrm{K}\}$,[7]

[7] We introduce the fictitious state $(D+1,b,\Phi)$ for notational convenience.



(a) From (t, b, U), $t < D$, SU idle

(b) From (t, b, U), $t < D$, SU active

(c) From (D, b, U)

(d) From $(t, 0, \mathrm{K})$, $t < D$, SU idle

(e) From $(t, 0, \mathrm{K})$, $t < D$, SU active

(f) From $(D, 0, \mathrm{K})$

Figure 2.13. *Transition probabilities*

where $\mathbf{X}_\mu$ stands for $\mathbf{G}_\mu$, $\mathbf{V}_\mu$ or $\mathbf{D}_\mu$ (we write $\mathbf{X} \in \{\mathbf{G}, \mathbf{V}, \mathbf{D}\}$), these are defined recursively as, for $t \in \mathbb{N}(1, D)$, $b \in \mathbb{N}(0, t-1)$,

$$\begin{aligned}\mathbf{X}_\mu(t,b,\mathrm{U}) &= x_\mu(t,b,\mathrm{U}) + \Pr_\mu(t+1,b,\mathrm{U}|t,b,\mathrm{U})\mathbf{X}_\mu(t+1,b,\mathrm{U}) \qquad (2.33)\\ &\quad + \Pr_\mu(t+1,b+1,\mathrm{U}|t,b,\mathrm{U})\mathbf{X}_\mu(t+1,b+1,\mathrm{U}) + \Pr_\mu(t+1,0,\mathrm{K}|t,b,\mathrm{U})\mathbf{X}_\mu(t+1,0,\mathrm{K}),\\ \mathbf{X}_\mu(t,0,\mathrm{K}) &= x_\mu(t,0,\mathrm{K}) + \left[q_{pp}^{(\mathrm{I})} + \mu(t,0,\mathrm{K})(q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})})\right]\mathbf{X}_\mu(t+1,0,\mathrm{K}), \qquad (2.34)\end{aligned}$$

where $x_\mu(t,b,\Phi)$ is the cost/reward accrued in state (t,b,Φ) and $\Pr_\mu(\cdot|\cdot)$ is the one-step transition probability, which can be derived with the help of Fig. 2.13 by taking the expectation with respect to the actions *SU idle* (I, with probability $1-\ \mu(t,b,\Phi)$) and *SU active* (A, with probability $\mu(t,b,\Phi)$), yielding

$$\Pr_\mu(t+1,b,\mathrm{U}|t,b,\mathrm{U}) = \mu(t,b,\mathrm{U})q_{pp}^{(A)}\left(q_{ps}^{(A)} - p_{s,\mathrm{buf}}\right) + (1-\mu(t,b,\mathrm{U}))q_{pp}^{(I)}q_{ps}^{(I)}, \qquad (2.35)$$

$$\Pr_\mu(t+1,b+1,\mathrm{U}|t,b,\mathrm{U}) = \mu(t,b,\mathrm{U})q_{pp}^{(A)}p_{s,\mathrm{buf}}, \qquad (2.36)$$

$$\Pr_\mu(t+1,0,\mathrm{K}|t,b,\mathrm{U}) = \mu(t,b,\mathrm{U})q_{pp}^{(A)}\left(1-q_{ps}^{(A)}\right) + (1-\mu(t,b,\mathrm{U}))q_{pp}^{(I)}\left(1-q_{ps}^{(I)}\right). \qquad (2.37)$$

Namely, if $\mathbf{X} = \mathbf{G}$ (throughput), then $x_\mu(t,b,\Phi),\ \Phi \in \{\mathrm{U},\mathrm{K}\}$, is the expected throughput accrued in state (t,b,Φ), and is given by

$$\begin{aligned}x_\mu(t,0,\mathrm{K}) =& \mu(t,0,\mathrm{K})T_{s\mathrm{K}} \triangleq g_\mu(t,0,\mathrm{K}), \qquad (2.38)\\ x_\mu(t,b,\mathrm{U}) =& \left[\mu(t,b,\mathrm{U})(1-q_{ps}^{(\mathrm{A})}) + (1-\mu(t,b,\mathrm{U}))(1-q_{ps}^{(\mathrm{I})})\right]bR_{s\mathrm{U}} + \mu(t,b,\mathrm{U})T_{s\mathrm{U}}\\ \triangleq& g_\mu(t,b,\mathrm{U}), \qquad (2.39)\end{aligned}$$

where the first term in (2.39) accounts for the successful recovery of the b SU messages from the buffered received signals via BIC, when the PU message is decoded by SUrx; if $\mathbf{X} = \mathbf{V}$ (secondary access), then $x_\mu(t,b,\Phi)$ is the SU access probability in state (t,b,Φ), *i.e.*,

$$x_\mu(t,b,\Phi) = \mu(t,b,\Phi) \triangleq v_\mu(t,b,\Phi); \qquad (2.40)$$

finally, if $\mathbf{X} = \mathbf{D}$ (time-slots), then

$$x_\mu(t,b,\Phi) = 1 \triangleq d_\mu(t,b,\Phi), \qquad (2.41)$$

corresponding to one time-slot. Moreover, we define, for $\mathbf{X} \in \{\mathbf{G}, \mathbf{V}, \mathbf{D}\}$,

$$\mathbf{X}'_{\mu}(\mathbf{s}) \triangleq \frac{\mathrm{d}\mathbf{X}'_{\mu}(\mathbf{s})}{\mathrm{d}\mu(\mathbf{s})}. \tag{2.42}$$

The number of visits to state $(1, 0, \mathrm{U})$ up to time-slot n is a renewal process [26]. Each renewal interval (*i.e.*, the ARQ sequence in which the PU attempts to deliver a specific packet) has average duration $\mathbf{D}_{\mu}(1, 0, \mathrm{U})$, over which the expected accrued SU throughput is $\mathbf{G}_{\mu}(1, 0, \mathrm{U})$, and the expected number of secondary channel accesses is $\mathbf{V}_{\mu}(1, 0, \mathrm{U})$. Then, the following lemma directly follows from the strong law of large numbers for renewal-reward processes [26].

Lemma 2.7.1. *The average long-term SU throughput and access rate are given by* $\bar{T}_s(\mu) = \frac{\mathbf{G}_{\mu}(1,0,\mathrm{U})}{\mathbf{D}_{\mu}(1,0,\mathrm{U})}$ *and* $\bar{W}_s(\mu) = \frac{\mathbf{V}_{\mu}(1,0,\mathrm{U})}{\mathbf{D}_{\mu}(1,0,\mathrm{U})}$*, respectively.*

We have the following lemma.

Lemma 2.7.2. *We have*

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} \geq 0, \quad \forall \mathbf{s} \in \mathcal{S},\ \forall \mu \in \mathcal{U}. \tag{2.43}$$

The inequality is strict if and only if state $\mathbf{s}$ *is accessible from* $(1, 0, \mathrm{U})$ *under policy* μ*,* i.e.*,* $\exists\, n > 0$: $\Pr_{\mu}^{(n)}(\mathbf{s}|(1, 0, \mathrm{U})) > 0$*. Moreover, for all* $\mathbf{s} \in \mathcal{S}$ *we have*

$$\mathbf{V}'_{\mu}(\mathbf{s}) - \mathbf{D}'_{\mu}(\mathbf{s})\bar{W}_s(\mu) > 0. \tag{2.44}$$

Proof. If state $\mathbf{s}$ is not accessible from state $(1, 0, \mathrm{U})$ under policy μ, then the steady state distribution satisfies $\pi_{\mu}(\mathbf{s}) = 0$, hence $\bar{W}_s(\mu)$ is unaffected by $\mu(\mathbf{s})$. Otherwise, from Lemma 2.7.1 we have that

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} = \frac{\frac{\mathrm{d}\mathbf{V}_{\mu}(1,0,\mathrm{U})}{\mathrm{d}\mu(\mathbf{s})} - \frac{\mathrm{d}\mathbf{D}_{\mu}(1,0,\mathrm{U})}{\mathrm{d}\mu(\mathbf{s})}\bar{W}_s(\mu)}{\mathbf{D}_{\mu}(1, 0, \mathrm{U})} \propto \mathbf{V}'_{\mu}(\mathbf{s}) - \mathbf{D}'_{\mu}(\mathbf{s})\bar{W}_s(\mu), \tag{2.45}$$

where $\propto$ represents equality up to a positive multiplicative factor, and the right hand side holds since, $\forall \mathbf{X} \in \{\mathbf{V}, \mathbf{D}\}$ and $(t, b, \Phi) \in \mathcal{S}$, $\frac{\mathrm{d}\mathbf{X}_{\mu}(1,0,\mathrm{U})}{\mathrm{d}\mu(t,b,\Phi)} = \Pr_{\mu}^{(t)}(t, b, \Phi|1, 0, \mathrm{U})\, \mathbf{X}'_{\mu}(t, b, \Phi)$.

If $\mathbf{s} \in \mathcal{S}_{\mathrm{K}}$, *i.e.*, $\mathbf{s} = (t, 0, \mathrm{K})$, we have

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(t, 0, \mathrm{K})} \propto \mathbf{V}'_{\mu}(t, 0, \mathrm{K}) - \mathbf{D}'_{\mu}(t, 0, \mathrm{K})\bar{W}_s(\mu) \geq \mathbf{V}'_{\mu}(t, 0, \mathrm{K}) - \mathbf{D}'_{\mu}(t, 0, \mathrm{K}) \triangleq A_{\mu}(t), \tag{2.46}$$

where we have used the fact that $\bar{W}_s(\mu) \leq 1$ and, from (2.34) and (2.41), $\mathbf{D}'_\mu(t,0,\mathrm{K}) = (q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})})\mathbf{D}_\mu(t+1,0,\mathrm{K}) \geq 0$.

We now prove by induction that $A_\mu(t) > 0, \forall\, t \in \mathbb{N}(1,T)$, so that (2.43) and (2.44) follow for $\mathbf{s} \in \mathcal{S}_\mathrm{K}$. From (2.33), for $t < D$, after algebraic manipulation we obtain

$$\begin{aligned} A_\mu(t) &= 1 + (q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})})[\mathbf{V}_\mu(t+1,0,\mathrm{K}) - \mathbf{D}_\mu(t+1,0,\mathrm{K})] \\ &= 1 - q_{pp}^{(\mathrm{A})} + \Pr_\mu(t+2,0,\mathrm{K}|t+1,0,\mathrm{K})A_\mu(t+1). \end{aligned} \tag{2.47}$$

Since $A_\mu(D) = 1 > 0$, we obtain $A_\mu(t) > 0$ by induction.

If $\mathbf{s} \in \mathcal{S}_\mathrm{U}$, *i.e.*, $\mathbf{s} = (t,b,\mathrm{U})$, we have

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(t,b,\mathrm{U})} \propto \mathbf{V}'_\mu(t,b,\mathrm{U}) - \mathbf{D}'_\mu(t,b,\mathrm{U})\bar{W}_s(\mu). \tag{2.48}$$

We prove that $\mathbf{V}'_\mu(t,b,\mathrm{U}) - \mathbf{D}'_\mu(t,b,\mathrm{U})\bar{W}_s(\mu) > 0$ in two steps, so that (2.43) and (2.44) follow for $\mathbf{s} \in \mathcal{S}_\mathrm{U}$. First, we prove that $C_\mu(t,b) \triangleq \mathbf{D}'_\mu(t,b,\mathrm{U}) \geq 0$. Then, since $\bar{W}_s(\mu) \leq 1$, we obtain

$$\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(t,b,0)} \propto \mathbf{V}'_\mu(t,b,\mathrm{U}) - C_\mu(t,b)\bar{W}_s(\mu) \geq \mathbf{V}'_\mu(t,b,\mathrm{U}) - \mathbf{D}'_\mu(t,b,\mathrm{U}) \triangleq B_\mu(t,b). \tag{2.49}$$

Finally, we prove that $B_\mu(t,b) > 0$.

Proof of $C_\mu(t,b) \geq 0$: from (2.33), for $t < D$ we have

$$\begin{aligned} C_\mu(t,b) =& [q_{pp}^{(\mathrm{A})}(1 - q_{ps}^{(\mathrm{A})}) - q_{pp}^{(\mathrm{I})}(1 - q_{ps}^{(\mathrm{I})})]\mathbf{D}_\mu(t+1,0,\mathrm{K}) \\ &+ [q_{pp}^{(\mathrm{A})}(q_{ps}^{(\mathrm{A})} - p_{s,\mathrm{buf}}) - q_{pp}^{(\mathrm{I})}q_{ps}^{(\mathrm{I})}]\mathbf{D}_\mu(t+1,b,\mathrm{U}) + q_{pp}^{(\mathrm{A})}p_{s,\mathrm{buf}}\mathbf{D}_\mu(t+1,b+1,\mathrm{U}). \end{aligned} \tag{2.50}$$

Using the recursions (2.33) and rearranging the terms, we obtain the recursive expression

$$\begin{aligned} C_\mu(t,b) =& \Pr_\mu(t+2,b+2,\mathrm{U}|t+1,b+1,\mathrm{U})C_\mu(t+1,b+1) \\ &+ q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})} + \Pr_\mu(t+2,b,\mathrm{U}|t+1,b,\mathrm{U})C_\mu(t+1,b) \\ &+ \Big[(1 - \mu(t+1,0,\mathrm{K}))q_{pp}^{(\mathrm{I})}(1 - q_{ps}^{(\mathrm{I})}) + \mu(t+1,0,\mathrm{K})q_{pp}^{(\mathrm{A})}(1 - q_{ps}^{(\mathrm{A})})\Big](q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})})\mathbf{D}_\mu(t+2,0,\mathrm{K}). \end{aligned}$$

Since $C_\mu(D,b) = 0,\ \forall\, b \in \mathbb{N}(0,D-1)$, it follows by induction on t that $C_\mu(b,t) \geq 0$.

Proof of $B_\mu(t,b) > 0$: From (2.33), for $t < D$ we obtain the following recursive expression for

$B_\mu(t,b)$, after algebraic manipulation,

$$\begin{aligned} B_\mu(t,b) &= 1 - q_{pp}^{(\mathrm{A})} + \Pr_\mu(t+2, b, \mathrm{U} | t+1, b, \mathrm{U}) B_\mu(t+1, b) \\ &+ \Pr_\mu(t+2, b+2, \mathrm{U} | t+1, b+1, \mathrm{U}) B_\mu(t+1, b+1) \\ &+ \left[(1 - \mu(t+1, 0, \mathrm{K})) q_{pp}^{(\mathrm{I})} (1 - q_{ps}^{(\mathrm{I})}) + \mu(t+1, 0, \mathrm{K}) q_{pp}^{(\mathrm{A})} (1 - q_{ps}^{(\mathrm{A})}) \right] A_\mu(t+1), \end{aligned} \tag{2.51}$$

here $A_\mu(t)$ is defined in (2.46). The result follows by induction, since $B_\mu(D,b) = 1 > 0$ and $A_\mu(t+1) > 0$. □

Appendix 2.B: SU access efficiency

In this appendix, we give a rigorous definition of SU access efficiency, thus complementing Def. 2.4.1. Moreover, in Lemma 2.7.3, we derive it. We recall that $\Pr_\mu^{(n)}(\mathbf{s}|\mathbf{s}_0)$ is the n-step transition probability of the chain from $\mathbf{s}_0$ to $\mathbf{s}$.

Definition 2.7.2. Let $\tilde{\mu} \in \mathcal{U}$ be a policy such that $\exists n > 0 : \Pr_{\tilde{\mu}}^{(n)}(\mathbf{s}|(1,0,\mathrm{U})) > 0$, and $\mu_\upsilon = (1-\upsilon)\mu + \upsilon\tilde{\mu}$, where $\upsilon \in (0,1]$, $\mu \in \mathcal{U}$. We define the *SU access efficiency* under μ in state $\mathbf{s}$ as

$$\eta_\mu(\mathbf{s}) = \lim_{\upsilon \to 0^+} \left. \frac{\frac{\mathrm{d}\bar{T}_s(\mu_\upsilon)}{\mathrm{d}\mu_\upsilon(\mathbf{s})}}{\frac{\mathrm{d}\bar{W}_s(\mu_\upsilon)}{\mathrm{d}\mu_\upsilon(\mathbf{s})}} \right|_{\mu_\upsilon}.$$

Remark 2.7.1. Notice that the condition $\exists\ n > 0 :\ \Pr_{\tilde{\mu}}^{(n)}(\mathbf{s}|(1,0,\mathrm{U})) > 0$ guarantees that state $\mathbf{s}$ is accessible from state $(1,0,\mathrm{U})$ under policy μ_υ, for $\upsilon > 0$. Under this condition, $\frac{\mathrm{d}\bar{W}_s(\mu)}{\mathrm{d}\mu(\mathbf{s})} > 0$ (Lemma 2.7.2 in Appendix 2.A), hence the fraction within the limit is well defined for $\upsilon > 0$ and in the limit $\upsilon \to 0^+$. One such policy $\tilde{\mu}$ is $\tilde{\mu}(\mathbf{s}) = 0.5, \forall \mathbf{s} \in \mathcal{S}$.

Using Lemma 2.7.1 and Def. 2.7.1 in Appendix 2.A and Def. 2.7.2, $\eta_\mu(\mathbf{s})$ can be derived according to the following lemma.

Lemma 2.7.3. *We have* $\eta_\mu(\mathbf{s}) = \frac{\mathbf{G}'_\mu(\mathbf{s}) - \mathbf{D}'_\mu(\mathbf{s})\bar{T}_s(\mu)}{\mathbf{V}'_\mu(\mathbf{s}) - \mathbf{D}'_\mu(\mathbf{s})\bar{W}_s(\mu)}$.

Remark 2.7.2. This is well defined, since $\mathbf{V}'_\mu(\mathbf{s}) - \mathbf{D}'_\mu(\mathbf{s})\bar{W}_s(\mu) > 0$ from Lemma 2.7.2 in Appendix 2.A.

Appendix 2.C: Proof of Theorem 2.4.2

Proof of Theorem 2.4.2. In the first part of the theorem, we prove that, by initializing Algorithm 1 with the idle policy $\mu^{(0)}$, $\mu^{(0)}(\mathbf{s}) = 0,\ \forall \mathbf{s} \in \mathcal{S}$, and with the set of idle states $\mathcal{S}_{\text{idle}}^{(0)} \equiv \mathcal{S}$, we obtain an optimal policy. In the second part of the proof, we prove the optimality of the specific initialization of Algorithm 1 for the high SU access rate regime.

Let $\tilde{\mu}$ be a policy under which all states $\mathbf{s} \in \mathcal{S}$ are accessible from state $(1, 0, \mathrm{U})$, *i.e.*, $\exists\ n > 0 :\ \Pr_{\tilde{\mu}}^{(n)}(\mathbf{s}|(1,0,\mathrm{U})) > 0$. One such policy is $\tilde{\mu}(\mathbf{s}) = 0.5,\ \forall \mathbf{s} \in \mathcal{S}$. Consider a modified Markov Decision Process, parameterized by $\upsilon \in (0,1)$, obtained by applying the policy $(1-\upsilon)\mu + \upsilon\tilde{\mu}$ to the original system, where $\mu \in \mathcal{U}$. Since $\mu, \tilde{\mu} \in \mathcal{U}$ and $\upsilon \in (0,1)$, it follows that $(1-\upsilon)\mu + \upsilon\tilde{\mu} \in \mathcal{U}$. We define $\bar{T}_s(\mu,\upsilon) \triangleq \bar{T}_s((1-\upsilon)\mu + \upsilon\tilde{\mu})$ and $\bar{W}_s(\mu,\upsilon) \triangleq \bar{W}_s((1-\upsilon)\mu + \upsilon\tilde{\mu})$, and we study

$$\mu^{*(\upsilon)} = \arg\max_{\mu\in\mathcal{U}} \bar{T}_s(\mu,\upsilon) \text{ s.t. } \bar{W}_s(\mu,\upsilon) \leq \epsilon_{\mathrm{W}}, \tag{2.52}$$

where the parameter υ is small enough to guarantee a feasible problem, *i.e.*, $\exists\, \mu \in \mathcal{U} : \bar{W}_s(\mu,\upsilon) \leq \epsilon_{\mathrm{W}}$. (2.19) is obtained in the limit $\upsilon \to 0^+$. Notice that, $\forall\ \mu \in \mathcal{U}$, under policy $(1-\upsilon)\mu + \upsilon\tilde{\mu}$, all the states $\mathbf{s} \in \mathcal{S}$ are accessible from state $(1,0,\mathrm{U})$, and the Markov chain is irreducible. Hence, from Lemma 2.7.2 in Appendix 2.A, $\bar{W}_s(\mu,\upsilon)$ is a strictly increasing function of $\mu(\mathbf{s}),\ \forall \mathbf{s} \in \mathcal{S}$. This is an important assumption in the following proof.

Let $\mathcal{D} \subset \mathcal{U}$ be the set of all the deterministic policies, and $\mathcal{G}_\upsilon = \left\{\left(\bar{W}_s(\mu,\upsilon), \bar{T}_s(\mu,\upsilon)\right), \mu \in \mathcal{D}\right\}$. With the help of Fig. 2.14, for any $\mu \in \mathcal{U}$, we have that $\left(\bar{W}_s(\mu,\upsilon), \bar{T}_s(\mu,\upsilon)\right) \in \mathrm{conv}(\mathcal{G}_\upsilon)$, where $\mathrm{conv}(\mathcal{G}_\upsilon)$ is the convex hull of the set $\mathcal{G}_\upsilon$. In particular, for the optimal policy we have $\left(\bar{W}_s(\mu^{*(\upsilon)},\upsilon), \bar{T}_s(\mu^{*(\upsilon)},\upsilon)\right) \in \mathrm{bd}(\mathcal{G}_\upsilon)$, where $\mathrm{bd}(\mathcal{G}_\upsilon)$ denotes the boundary of $\mathrm{conv}(\mathcal{G}_\upsilon)$.

Algorithm 1 determines the sequence of vertices of the polyline $\mathrm{bd}(\mathcal{G}_\upsilon)$ in the limit $\upsilon \to 0^+$ (bold line in Fig. 2.14). For $\upsilon > 0$, starting from the leftmost vertex of $\mathrm{bd}(\mathcal{G}_\upsilon)$, achieved by the *idle* policy $\mu^{(0)}(\mathbf{s}) = 0,\ \forall \mathbf{s} \in \mathcal{S}$ (this follows from the fact that $\bar{W}_s(\mu,\upsilon)$ is a strictly increasing function of $\mu(\mathbf{s})$, hence it is minimized by the idle policy), the algorithm determines iteratively the next vertex of $\mathrm{bd}(\mathcal{G}_\upsilon)$ as the maximizer of the slope

$$\mu^{(i+1)} = \underset{\mu\in\mathcal{D}:\bar{W}_s(\mu,\upsilon)>\bar{W}_s(\mu^{(i)},\upsilon)}{\arg\max} \frac{\bar{T}_s(\mu,\upsilon) - \bar{T}_s(\mu^{(i)},\upsilon)}{\bar{W}_s(\mu,\upsilon) - \bar{W}_s(\mu^{(i)},\upsilon)}. \tag{2.53}$$

Since (2.19) has one constraint, the optimal policy $\mu^{*(\upsilon)}$ is randomized in one state [22], and hence

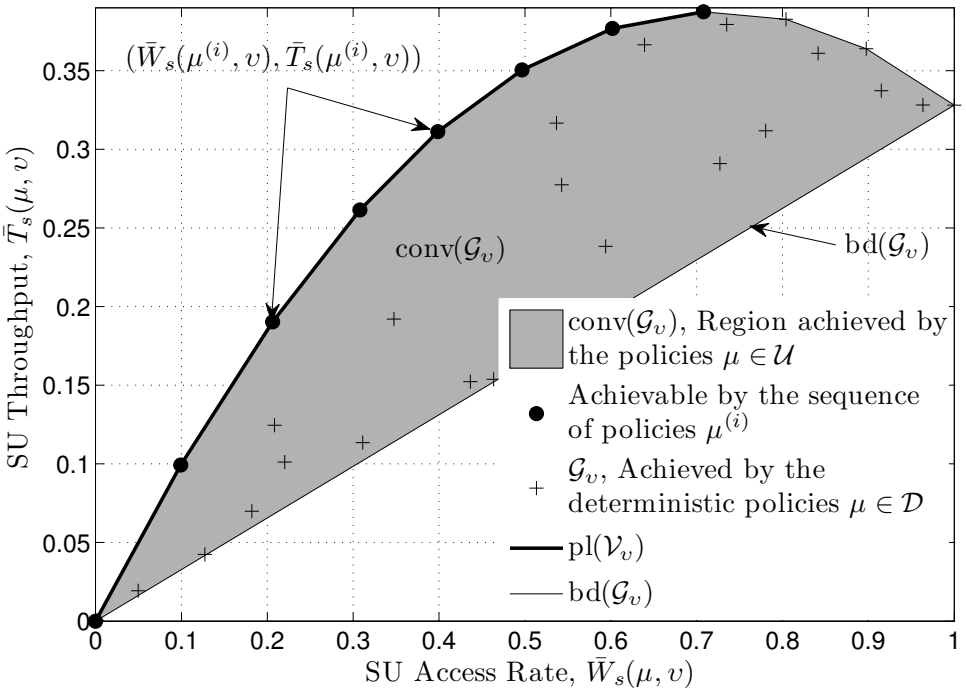


Figure 2.14. *Geometric interpretation of problem (2.52)*

each segment on the boundary $\mathrm{bd}(\mathcal{G}_\upsilon)$ between pairs $(\bar{W}_s(\mu^{(i)},\upsilon),\bar{T}_s(\mu^{(i)},\upsilon))$ achievable with deterministic policies is attained by a policy that is randomized in only one state. It follows that $\mu^{(i)}$ and $\mu^{(i+1)}$ differ in only one state. Moreover, in (2.53) the maximization is over $\mu \in \mathcal{D}$ such that $\bar{W}_s(\mu,\upsilon) > \bar{W}_s(\mu^{(i)},\upsilon)$, *i.e.*, since $\bar{W}_s(\mu,\upsilon)$ is a strictly increasing function of $\mu(\mathbf{s})$ and $\mu^{(i+1)}$ and $\mu^{(i)}$ differ in only one position, $\mu^{(i+1)}$ is obtained from $\mu^{(i)}$ by allocating one more secondary access to a state which is idle under $\mu^{(i)}$. In (2.53), the maximization is thus over $\left\{\mu^{(i)}+\delta_{\mathbf{s}} : \mathbf{s} \in \mathcal{S}^{(i)}_{\mathrm{idle}}\right\}$, and, after algebraic manipulation, $\mu^{(i+1)}$ in (2.53) maximizes

$$\max_{\mathbf{s}\in\mathcal{S}^{(i)}_{\mathrm{idle}}} \frac{\bar{T}_s(\mu^{(i)}+\delta_{\mathbf{s}},\upsilon)-\bar{T}_s(\mu^{(i)},\upsilon)}{\bar{W}_s(\mu^{(i)}+\delta_{\mathbf{s}},\upsilon)-\bar{W}_s(\mu^{(i)},\upsilon)} = \max_{\mathbf{s}\in\mathcal{S}^{(i)}_{\mathrm{idle}}} \eta_{(1-\upsilon)\mu^{(i)}+\upsilon\tilde{\mu}}(\mathbf{s}).$$

Stage i of the algorithm is thus proved. If $\eta_{(1-\upsilon)\mu^{(i)}+\upsilon\tilde{\mu}}(\mathbf{s}) \leq 0$, we have $\bar{W}_s\left(\mu^{(i)}+\delta_{\mathbf{s}},\upsilon\right) > \bar{W}_s\left(\mu^{(i)},\upsilon\right)$ and $\bar{T}_s\left(\mu^{(i)}+\delta_{\mathbf{s}},\upsilon\right) \leq \bar{T}_s\left(\mu^{(i)},\upsilon\right)$. If this condition holds $\forall\, \mathbf{s} \in \mathcal{S}^{(i)}_{\mathrm{idle}}$, any next vertex of the polyline $\mathrm{bd}(\mathcal{G}_\upsilon)$ yields a decrease of the SU throughput and a larger SU access rate, hence a sub-optimal set of policies, and the algorithm stops.

By construction, the algorithm returns a sequence of policies $(\mu^{(i)}, i \in \mathbb{N}(0,N-1))$, characterized by strictly increasing values of the SU throughput and of the SU access rate. The optimal policy belongs to the polyline with vertices $\mathcal{V}_\upsilon \equiv \{(\bar{W}_s(\mu^{(i)},\upsilon),\bar{T}_s(\mu^{(i)},\upsilon)), i \in \mathbb{N}(0,N-1)\}$, denoted by

$\mathrm{pl}(\mathcal{V}_\upsilon)$ in Fig. 2.14. Then, (2.19) becomes equivalent to $T_s^{*(\upsilon)} = \max_{(W_s,T_s)\in\mathcal{V}_\upsilon} T_s$ s.t. $W_s \leq \epsilon_{\mathrm{W}}$, whose solution is given in the last step of Algorithm 1. The result finally follows for $\upsilon \to 0^+$.

To conclude, we prove the initialization of Algorithm 1 for the high SU access rate. Let $(\mu^{(0)},\ldots,\mu^{(N-1)})$ and $(\mathbf{s}^{(0)},\ldots,\mathbf{s}^{(N-1)})$ be the sequence of deterministic policies and of states returned by Algorithm 1, obtained by initializing the algorithm as in the first part of the proof. Let

$$\mathcal{D}_0 \equiv \{\mu \in \mathcal{D} : \mu(t,0,0) = 0 \,\forall\, t \in \mathbb{N}(1,T)\}\,, \tag{2.54}$$

$\tilde{\mathcal{D}}_0 \equiv \{\mu \in \mathcal{D}_0 : \mu(\mathbf{s}) = 1,\ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}\}$, and $N_0 \triangleq \max\{i \in \{0,\ldots,N-1\} : \bar{W}_s(\mu^{(i)}) < \epsilon_{\mathrm{th}}\}$. We prove that $\mu^{(N_0+1)} \in \tilde{\mathcal{D}}_0$, *i.e.*, $\mu^{(N_0+1)}(\mathbf{s}) = 1,\ \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}$. From the definition of $\tilde{\mathcal{D}}_0$ and the construction of the algorithm, it follows that, for $i > N_0$, $\mu^{(i)}(\mathbf{s}) = 1,\ \forall\, \mathbf{s} \in \mathcal{S}_{\mathrm{K}}$. Moreover, from Lemma 2.7.4, $\bar{W}_s(\mu^{(N_0+1)}) = \epsilon_{\mathrm{th}}$. Hence, for the high SU access rate $\epsilon > \epsilon_{th}$, the optimal policy μ^* obeys $\mu^*(\mathbf{s}) = 1, \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}$. Then, letting $\mathcal{U}_1 \equiv \{\mu \in \mathcal{U} : \mu(\mathbf{s}) = 1, \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}}\}$, the optimization problem (2.19) can be restricted to the set of randomized policies $\mu \in \mathcal{U}_1 \subset \mathcal{U}$ when $\epsilon > \epsilon_{th}$. Equivalently, secondary accesses taking place in $\mathcal{S}_{\mathrm{U}}$ can be obtained by initializing the algorithm with $\mu^{(0)}(\mathbf{s}) = 0,\ \mathbf{s} \in \mathcal{S}_{\mathrm{U}}$, $\mu^{(0)}(\mathbf{s}) = 1,\ \mathbf{s} \in \mathcal{S}_{\mathrm{K}}$, $S_{\mathrm{idle}}^{(0)} \equiv \mathcal{S}_{\mathrm{U}}$.

Proof of $\mu^{(N_0+1)} \in \tilde{\mathcal{D}}_0$: We prove by induction that $\mu^{(i)} \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0,\ \forall i \leq N_0$ and $\mu^{(N_0+1)} \in \tilde{\mathcal{D}}_0$. Assume that, for some $i \geq 0$, $\mu^{(j)} \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0,\ \forall j \leq i$. From Lemma 2.7.4, it follows that $N_0 \geq i$. This clearly holds for $i = 0$. We show that this implies that either $\mu^{(i+1)} \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0$, hence $N_0 > i$, thus proving the induction step, or $\mu^{(i+1)} \in \tilde{\mathcal{D}}_0$, hence $N_0 = i$, thus proving the property. The result follows since $N_0 \leq 1 + |\mathcal{S}| < \infty$ (*i.e.*, $i = N_0$ is reached within a finite number of steps).

From Lemma 2.7.5, $\eta_{\mu^{(i)}}(\mathbf{s}) = T_{s\mathrm{K}} > 0, \forall \mathbf{s} \in \mathcal{S}_{\mathrm{K}} \cap \mathcal{S}_{\mathrm{idle}}^{(i)}$ and $\eta_{\mu^{(i)}}(t,0,\mathrm{U}) < T_{s\mathrm{K}}, \forall t \in \mathbb{N}(1,D)$, hence, from the main iteration stage of the algorithm it follows that $\mu^{(i+1)} \in \mathcal{D}_0$. In particular, if $\mu^{(i+1)} \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0$, then $N_0 > i$ from Lemma 2.7.4. On the other hand, if $\mu^{(i+1)} \in \tilde{\mathcal{D}}_0$, then, from Lemma 2.7.4, $N_0 = i$. The property is thus proved. □

Lemma 2.7.4. $\bar{W}_s(\mu) < \epsilon_{\mathrm{th}}, \forall \mu \in \mathcal{D}_0 \setminus \tilde{\mathcal{D}}_0$ *and* $\bar{W}_s(\mu) = \epsilon_{\mathrm{th}}, \forall \mu \in \tilde{\mathcal{D}}_0$.

Proof. Let $\mu \in \tilde{\mathcal{D}}_0$. Since the states (t,b,U) with $b > 0$ are not accessible from $(1,0,\mathrm{U})$ under μ, the transmission probability $\mu(t,b,\mathrm{U})$, $b > 0$, does not affect $\bar{W}_s(\mu)$. Then, from Def. 2.3.1, we have $\bar{W}_s(\mu) = \epsilon_{\mathrm{th}}$.

Let $\mu \in \mathcal{D} \setminus \tilde{\mathcal{D}}_0$. Letting $\mathcal{S}_\mu = \{\mathbf{s} \in \mathcal{S}_{\mathrm{K}} : \mu(\mathbf{s}) = 0\}$, we have that $\mu + \sum_{\mathbf{s}\in\mathcal{S}_\mu} \delta_{\mathbf{s}} \in \tilde{\mathcal{D}}_0$. Finally,

since every $\mathbf{s} \in \mathcal{S}_\mu$ is accessible from $(1, 0, \mathrm{U})$ under μ, and $\mathcal{S}_\mu$ is non-empty, from Lemma 2.7.2 in Appendix 2.A and the previous case, it follows that $\bar{W}_s(\mu) < \bar{W}_s(\mu + \sum_{\mathbf{s}\in\mathcal{S}_\mu} \delta_{\mathbf{s}}) = \epsilon_{\mathrm{th}}$. □

Lemma 2.7.5. *Let $\mu \in \mathcal{U}$ such that $\mu(t, 0, \mathrm{U}) = 0 \ \forall \ t \in \mathbb{N}(1, D)$. Then, $\eta_\mu(t, 0, \mathrm{U}) < T_{s\mathrm{K}}$ and $\eta_\mu(t, 0, \mathrm{K}) = T_{s\mathrm{K}}, \ \forall t$.*

Proof. Let $\mu \in \mathcal{U}$ such that $\mu(t, 0, \mathrm{U}) = 0 \ \forall \ t \in \mathbb{N}(1, D)$. It follows that the states (t, b, U) with $b > 0$ are not accessible, hence their steady state probability satisfies $\pi_\mu(t, b, \mathrm{U}) = 0, \ \forall \ t, \ \forall \ b > 0$. It is then straightforward to show, by using the recursion (2.33), that $\mathbf{G}_\mu(t, 0, \mathrm{U}) = T_{s\mathrm{K}}\mathbf{V}_\mu(t, 0, \mathrm{K})$, $\mathbf{G}_\mu(t, 0, \mathrm{K}) = T_{s\mathrm{K}}\mathbf{V}_\mu(t, 0, \mathrm{K})$ and $\bar{T}_s(\mu) = T_{s\mathrm{K}}\bar{W}_s(\mu)$. Then, using these expressions, the recursion (2.33) and Lemma 2.7.3, we obtain $\eta_\mu(t, 0, \mathrm{K}) = T_{s\mathrm{K}}$ and

$$\eta_\mu(t, 0, \mathrm{U}) = T_{s\mathrm{K}} - \frac{T_{s\mathrm{K}}\mathbf{V}'_\mu(t, 0, \mathrm{U}) - \mathbf{G}'_\mu(t, 0, \mathrm{U})}{\mathbf{V}'_\mu(t, 0, \mathrm{U}) - \mathbf{D}'_\mu(t, 0, \mathrm{U})\bar{W}_s(\mu)}. \tag{2.55}$$

We now prove that $\eta_\mu(t, 0, \mathrm{U}) < T_{s\mathrm{K}}$, which proves the lemma. Equivalently, using Lemma 2.7.2 in Appendix 2.A and (2.33), we prove that

$$T_{s\mathrm{K}}\mathbf{V}'_\mu(t, 0, \mathrm{U}) - \mathbf{G}'_\mu(t, 0, \mathrm{U}) = (T_{s\mathrm{K}} - T_{s\mathrm{U}}) + q^{(\mathrm{A})}_{pp} p_{s,\mathrm{buf}}[T_{s\mathrm{K}}\mathbf{V}_\mu(t, 1, \mathrm{U}) - \mathbf{G}_\mu(t, 1, \mathrm{U})] > 0.$$

Letting

$$M_\mu(t, b) = b(T_{s\mathrm{K}} - T_{s\mathrm{U}}) + q^{(\mathrm{A})}_{pp} p_{s,\mathrm{buf}}[T_{s\mathrm{K}}\mathbf{V}_\mu(t, b, \mathrm{U}) - \mathbf{G}_\mu(t, b, \mathrm{U})] > 0, \ \forall \ t, b \geq 1, \tag{2.56}$$

(2.56) is equivalent to $M_\mu(t, 1) > 0$. We now prove by induction that $M_\mu(t, b) > 0, \ \forall \ t, b \geq 1$, yielding (2.56) as a special case when $b = 1$. For $t = D + 1$ we have $M_\mu(D + 1, b) = b(T_{s\mathrm{K}} - T_{s\mathrm{U}}) > 0$, since $T_{s\mathrm{K}} > T_{s\mathrm{U}}$ and $b \geq 1$. Now, let $t \leq D$ and assume $M_\mu(t + 1, b) > 0$. Using (2.33), after algebraic manipulation we obtain

$$\begin{aligned} M_\mu(t, b) &= b(T_{s\mathrm{K}} - T_{s\mathrm{U}}) - q^{(\mathrm{A})}_{pp} p_{s,\mathrm{buf}} \left[1 - \mu(t, b, \mathrm{U})q^{(\mathrm{A})}_{ps} - (1 - \mu(t, b, \mathrm{U}))q^{(\mathrm{I})}_{ps}\right] bR_{s\mathrm{U}} \\ &+ q^{(\mathrm{A})}_{pp} p_{s,\mathrm{buf}}\mu(t, b, \mathrm{U})(T_{s\mathrm{K}} - T_{s\mathrm{U}}) + \mathrm{Pr}_\mu(t + 1, b, \mathrm{U}|t, b, \mathrm{U})[M_\mu(t + 1, b) - b(T_{s\mathrm{K}} - T_{s\mathrm{U}})] \\ &+ \mathrm{Pr}_\mu(t + 1, b + 1, \mathrm{U}|t, b, \mathrm{U})M_\mu(t + 1, b) - \mathrm{Pr}_\mu(t + 1, b + 1, \mathrm{U}|t, b, \mathrm{U})(b + 1)(T_{s\mathrm{K}} - T_{s\mathrm{U}}). \end{aligned} \tag{2.57}$$

Finally, since $M_\mu(t+1,b) > 0$ by the induction hypothesis, using inequality (2.9) we obtain

$$M_\mu(t,b) > p_{s,\mathrm{buf}} b R_{s\mathrm{U}} \left(1 - q_{pp}^{(\mathrm{A})}\right) + p_{s,\mathrm{buf}} b R_{s\mathrm{U}} (1-\mu(t,b,\mathrm{U})) q_{ps}^{(\mathrm{I})} (q_{pp}^{(\mathrm{A})} - q_{pp}^{(\mathrm{I})}) > 0, \qquad (2.58)$$

which proves the induction step. The lemma is proved. □

Appendix 2.D: Proof of Lemma 2.5.1

Proof of Lemma 2.5.1. Let $\mathcal{D} \subset \mathcal{U}$ be the set of all the deterministic (non-randomized) policies. Let

$$\begin{aligned}\tilde{\mathcal{D}} \equiv \quad & \{\mu \in \mathcal{D} : \mu(t,b,\mathrm{U}) = 1, \forall t, b < b(t);\ \mu(t,b,\mathrm{U}) = 0, \forall t, b \geq b(t);\ \mu(\mathbf{s}) = 1, \mathbf{s} \in \mathcal{S}_{\mathrm{K}}; \\ & \exists\, b(\cdot) : b(t+1) \leq b(t)\ \forall t\}\,.\end{aligned}$$

By inspection, we have that the sequences of policies (2.24) are such that $\mu^{(i)} \in \tilde{\mathcal{D}},\ \forall i \in \mathbb{N}(0, N-1)$. Therefore, the first part of the lemma states that $\mu^{(i)} \in \tilde{\mathcal{D}},\ \forall i \in \mathbb{N}(0, N-1)$. We prove this property by induction. Namely, we show that $\mu^{(i)} \in \tilde{\mathcal{D}} \Rightarrow \mu^{(i+1)} \in \tilde{\mathcal{D}}$. Then, since $\mu^{(0)} \in \tilde{\mathcal{D}}$ (initialization of Algorithm 1) it follows that $\mu^{(i)} \in \tilde{\mathcal{D}},\ \forall i$. Let $\mu^{(i)} \in \tilde{\mathcal{D}}$, *i.e.*, $\mu^{(i)}$ is given by (2.24) for some $b^{(i)}(t)$ non-increasing in t. The set of idle states is then given by

$$\mathcal{S}_{\mathrm{idle}}^{(i)} \equiv \left\{(t,b,\mathrm{U}) \in \mathcal{S}_{\mathrm{U}} : t \in \mathbb{N}(1,D), b \geq b^{(i)}(t)\right\}. \qquad (2.59)$$

We then prove that, under the hypotheses of the lemma, $\eta_{\mu^{(i)}}(t,b,\mathrm{U}) > \eta_{\mu^{(i)}}(t,b+1,\mathrm{U})$ and $\eta_{\mu^{(i)}}(t,b,\mathrm{U}) > \eta_{\mu}(t+1,b,\mathrm{U}),\ \forall (t,b,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)}$. It follows that the SU access efficiency is maximized by the state in the idle set $\mathcal{S}_{\mathrm{idle}}^{(i)}$ with the lowest value of the primary ARQ state t, among the states with the same buffer occupancy b, and with the fewest number of buffered received signals b, among the states with the same primary ARQ state t. Therefore, in the main iteration stage of the algorithm, the SU access efficiency is maximized by $\mathbf{s}^{(i)} = \arg\max_{\mathbf{s} \in \mathcal{S}_{\mathrm{idle}}^{(i)}} \eta_{\mu^{(i)}}(\mathbf{s})$, where $\mathbf{s}^{(i)} = (t,b,\mathrm{U})$ is such that $\tau \geq t$, $\beta \geq b$, $\forall\ (\tau,\beta,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)}$. By inspection, we have that $\mu^{(i+1)} = \mu^{(i)} + \delta_{\mathbf{s}^{(i)}} \in \tilde{\mathcal{D}}$, hence the induction step is proved.

We thus need to prove the induction step, *i.e.*, letting $\mu^{(i)} \in \tilde{\mathcal{D}}$, we show that

$$\begin{aligned}&\eta_{\mu^{(i)}}(t,b,\mathrm{U}) > \eta_{\mu^{(i)}}(t,b+1,\mathrm{U}),\ \forall (t,b,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)}, \\ &\eta_{\mu^{(i)}}(t,b,\mathrm{U}) > \eta_{\mu}(t+1,b,\mathrm{U}),\ \forall (t,b,\mathrm{U}) \in \mathcal{S}_{\mathrm{idle}}^{(i)}. \end{aligned} \qquad (2.60)$$

To this end, note that, in the degenerate cognitive radio network scenario, the primary ARQ process is not affected by the SU access scheme, hence, using the notation in Appendix 2.A, $\mathbf{D}'_{\mu^{(i)}}(t,b,\mathrm{U})=0$. By the definition of SU access efficiency (2.7.3), we thus obtain

$$\eta_{\mu^{(i)}}(t,b,\mathrm{U})=\frac{\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U})}{\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U})}, \tag{2.61}$$

where, using (2.33), (2.35-2.37), (2.39) and (2.40),

$$\begin{aligned}\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U})=&T_{s\mathrm{U}}+\left(q_{ps}^{(\mathrm{I})}-q_{ps}^{(\mathrm{A})}\right)bR_{s\mathrm{U}}+q_{pp}(q_{ps}^{(\mathrm{A})}-p_{s,\mathrm{buf}}-q_{ps}^{(\mathrm{I})})\mathbf{G}_{\mu^{(i)}}(t+1,b,\mathrm{U}) \\ &+q_{pp}p_{s,\mathrm{buf}}\mathbf{G}_{\mu^{(i)}}(t+1,b+1,\mathrm{U})+q_{pp}(q_{ps}^{(\mathrm{I})}-q_{ps}^{(\mathrm{A})})\mathbf{G}_{\mu^{(i)}}(t+1,0,\mathrm{K}),\end{aligned} \tag{2.62}$$

$$\begin{aligned}\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U})=&1+q_{pp}(q_{ps}^{(\mathrm{A})}-p_{s,\mathrm{buf}}-q_{ps}^{(\mathrm{I})})\mathbf{V}_{\mu^{(i)}}(t+1,b,\mathrm{U}) \\ &+q_{pp}p_{s,\mathrm{buf}}\mathbf{V}_{\mu^{(i)}}(t+1,b+1,\mathrm{U})+q_{pp}(q_{ps}^{(\mathrm{I})}-q_{ps}^{(\mathrm{A})})\mathbf{V}_{\mu^{(i)}}(t+1,0,\mathrm{K}).\end{aligned} \tag{2.63}$$

Using the fact that $\mu^{(i)}(\tau,\beta,\mathrm{U})=0,\ \forall\tau\geq t,\beta\geq b$, it can be proved that

$$\mathbf{V}_{\mu^{(i)}}(\tau,\beta,\mathrm{U})=A_1(\tau)-A_0(\tau), \tag{2.64}$$

$$\mathbf{G}_{\mu^{(i)}}(\tau,\beta,\mathrm{U})=(1-q_{ps}^{(\mathrm{I})})\beta R_{s\mathrm{U}}A_0(\tau)+T_{s\mathrm{K}}(A_1(\tau)-A_0(\tau)), \tag{2.65}$$

$$\mathbf{V}_{\mu^{(i)}}(\tau,0,\mathrm{K})=A_1(\tau), \tag{2.66}$$

$$\mathbf{G}_{\mu^{(i)}}(\tau,0,\mathrm{K})=T_{s\mathrm{K}}A_1(\tau), \tag{2.67}$$

where $A_0(\cdot)$ and $A_1(\cdot)$ are defined in (2.29) and (2.30), respectively. The expressions (2.64-2.67) can be easily verified by induction, starting from $\tau=D+1$ backward. In fact, for $\tau=D+1$, we have $A_0(D+1)=A_1(D+1)=0$, hence we obtain $\mathbf{V}_{\mu^{(i)}}(D+1,\beta,\mathrm{U})=\mathbf{G}_{\mu^{(i)}}(D+1,\beta,\mathrm{U})=\mathbf{V}_{\mu^{(i)}}(D+1,0,\mathrm{K})=\mathbf{G}_{\mu^{(i)}}(D+1,0,\mathrm{K})=0$, which is consistent with Def. 2.7.1. The induction step can be proved by inspection, using the recursive expression (2.33) and the fact that $\mu(\tau,\beta,\mathrm{U})=0,\ \forall\tau\geq t,\beta\geq b$. Substituting the expressions (2.64-2.67) in (2.62) and (2.63), we obtain

$$\begin{aligned}\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U})=&T_{s\mathrm{U}}+\left(q_{ps}^{(\mathrm{I})}-q_{ps}^{(\mathrm{A})}\right)bR_{s\mathrm{U}}\left[1-q_{pp}(1-q_{ps}^{(\mathrm{I})})A_0(t+1)\right] \\ &+q_{pp}p_{s,\mathrm{buf}}(1-q_{ps}^{(\mathrm{I})})R_{s\mathrm{U}}A_0(t+1)+q_{pp}(q_{ps}^{(\mathrm{I})}-q_{ps}^{(\mathrm{A})})T_{s\mathrm{K}}A_0(t+1),\end{aligned} \tag{2.68}$$

$$\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U})=1-q_{pp}(q_{ps}^{(\mathrm{A})}-q_{ps}^{(\mathrm{I})})A_0(t+1). \tag{2.69}$$

Proof of $\eta_{\mu^{(i)}}(t,b,0) > \eta_{\mu^{(i)}}(t,b+1,0)$

By substituting (2.68) and (2.69) in (2.61), and noticing that $\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) = \mathbf{V}'_{\mu^{(i)}}(t,b+1,\mathrm{U})$ from (2.69) and $\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) > 0$ (from Lemma 2.7.2 with $\mathbf{D}'_{\mu}(\mathbf{s}) = 0$), the condition $\eta_{\mu^{(i)}}(t,b,0) > \eta_{\mu^{(i)}}(t,b+1,0)$ is equivalent to $\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) > \mathbf{G}'_{\mu^{(i)}}(t,b+1,\mathrm{U})$, which is readily verified from (2.68), since

$$\begin{aligned}\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) - \mathbf{G}'_{\mu^{(i)}}(t,b+1,\mathrm{U}) &= \left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) R_{s\mathrm{U}} \left[1 - q_{pp}(1-q_{ps}^{(\mathrm{I})})A_0(t+1)\right] \\ &> \left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) \frac{1-q_{pp}}{1-q_{pp}q_{ps}^{(\mathrm{I})}} R_{s\mathrm{U}} > 0, \end{aligned} \tag{2.70}$$

where the first inequality follows from the fact that $A_0(t+1) < \frac{1}{1-q_{pp}q_{ps}^{(\mathrm{I})}}$, the second from $q_{ps}^{(\mathrm{I})} < q_{ps}^{(\mathrm{A})}$.

Proof of $\eta_{\mu^{(i)}}(t,b,0) > \eta_{\mu}(t+1,b,0)$

Since $\mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) > 0$, the condition $\eta_{\mu^{(i)}}(t,b,0) > \eta_{\mu}(t+1,b,0)$ is equivalent to

$$\begin{aligned}&\mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U}) \left(\mathbf{V}'_{\mu^{(i)}}(t+1,b,\mathrm{U}) - \mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U})\right) \\ &> \mathbf{V}'_{\mu^{(i)}}(t,b,\mathrm{U}) \left(\mathbf{G}'_{\mu^{(i)}}(t+1,b,\mathrm{U}) - \mathbf{G}'_{\mu^{(i)}}(t,b,\mathrm{U})\right). \end{aligned} \tag{2.71}$$

Using (2.68) and (2.69), after algebraic manipulation we obtain the equivalent condition

$$\left(1-q_{ps}^{(\mathrm{A})}\right) p_{s,\mathrm{buf}} + \left(1-q_{ps}^{(\mathrm{A})}\right)\left(q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}\right) b + \left(q_{ps}^{(\mathrm{I})} - q_{ps}^{(\mathrm{A})}\right) \Delta_s > 0, \tag{2.72}$$

where we have used the fact that $T_{s\mathrm{K}} = \Delta_s R_{s\mathrm{U}} + T_{s\mathrm{U}} + p_{s,\mathrm{buf}} R_{s\mathrm{U}}$. Since we require this condition to hold $\forall b \geq 0$ and the left hand expression is minimized by $b = 0$, the condition (2.72) should be satisfied for $b = 0$, yielding the equivalent condition $\Delta_s < \frac{1-q_{ps}^{(\mathrm{A})}}{q_{ps}^{(\mathrm{A})} - q_{ps}^{(\mathrm{I})}} p_{s,\mathrm{buf}}$, which is an hypothesis of the lemma.

It is thus proved that the sequence of policies returned by Algorithm 1 has the structure defined by (2.24), where $b^{(i)}(t)$ satisfies the inequality (2.26). Moreover, the inequality (2.27) holds since, by the algorithm construction, $\mu^{(i+1)}$ is obtained from $\mu^{(i)}$ by "activating" one additional state from the set of idle states $\mathcal{S}_{\mathrm{idle}}^{(i)}$.

The second part of the lemma states that $b^{(N-1)}(t) = \bar{b}_{\max}(t)$, where $\bar{b}_{\max}(t)$ is given by (2.28). This is a consequence of the fact that Algorithm 1 stops if the SU access efficiency becomes non-

positive, *i.e.*, $\eta_{\mu^{(i)}}(\mathbf{s}) \leq 0, \forall \mathbf{s} \in \mathcal{S}_{\text{idle}}^{(i)}$. From (2.61), this condition is equivalent to $\mathbf{G}'_{\mu^{(i)}}(t, b, \mathrm{U}) \leq 0$, $\forall (t, b, \mathrm{U}) \in \mathcal{S}_{\text{idle}}^{(i)}$. By using (2.68) and by solving $\mathbf{G}'_{\mu^{(i)}}(t, b, \mathrm{U}) \leq 0$ with respect to b, the result follows. □

Chapter 3

Optimal Management Policies for Energy Harvesting Wireless Sensor Networks

3.1 Introduction

In the previous chapter, we have addressed the problem of secondary access in cognitive radio networks, using stochastic optimization tools to determine the optimal access policy of secondary users in the network. We have shown that the exploitation of side information about the licensed (primary) users, *e.g.*, the HARQ protocol and the current primary message, jointly with the optimization of the available resources, can help alleviate the spectrum scarcity problem, thus enabling the coexistence of primary and secondary users and improving the spectral efficiency.

While the previous chapter dealt with a spectrum scarcity problem, in this chapter we face the problem of energy scarcity, with special emphasis on Wireless Sensor Networks (WSN), composed of miniaturized devices with sensing and communication capabilities [27]. Recent technological advances and enhancements of consumer electronics have led to the widespread diffusion of WSNs. One key requirement of such networks is a prolonged and unsupervised sensor operation over time, which poses the problem of their energy autonomy.

The use of non-rechargeable batteries is currently widespread for powering WSN sensor nodes. However, in many WSN applications, autonomous operation is required and factors such as the sheer number of nodes or inaccessibility render battery replacement unrealistic and cost-prohibitive [28], hence lifetime and *energy scarcity* becomes a critical concern in the design of WSNs [29–31]. Recent advances in the field of small-scale Energy Harvesting (EH) will enable the sensor to use ambient

energy absorbed, for instance, from solar, wind, piezo-electric, thermal or RF sources [32–36], to power its circuitry and to perform data sensing, processing and communication tasks. By relying on a potentially unlimited energy reservoir (ambient energy), the EH approach, combined with an intelligent use of the local energy storage, is envisioned to greatly prolong the WSN operating life [37], and could, in principle, lead to perpetual operation, thus alleviating the energy scarcity problem in WSNs.

In contrast to battery-operated sensors (without EH capability), where energy efficiency and conservation are crucial to prolong lifetime, in EH powered Sensors (EHSs) the energy supply is potentially unlimited, but its availability is random and intermittent over time. Present technologies require a local energy storage element to filter out the fluctuations in the EH process, as data sensing and processing, transmission/reception tasks, and higher layer operations (*e.g.*, routing) rely on a continuous and stable energy reserve. The objective thus shifts from lifetime maximization and energy conservation to the optimal management of the harvested energy, so as to provide a stable energy supply to the sensor node by minimizing the deleterious impact of energy depletion. However, one critical aspect of WSN design is that sensor nodes typically have limited processing capabilities. Therefore, a special focus should be dedicated to the design of energy management policies with small computational overhead.

In this chapter, we are concerned with a fundamental question: how should statistical information on the ambient energy supply be exploited in order to optimize EHS operation? We consider a WSN consisting of multiple EHSs, as depicted in Fig. 3.1, which judiciously report data of varying importance to a Fusion Center (FC). Practical examples of this setting include: temperature sensors, where higher temperature readings, being indicators of overheating or fire, are more important; sensors which act as a relay of different priority packets in a wireless network [38]; data transmission over a fading channel, where the number of bits which can be reliably transmitted depend on the instantaneous channel realization. Energy is harvested from an ambient source modeled by a two-state Markov chain, where "GOOD" and "BAD" correspond to an abundance and scarcity of ambient energy, respectively, and is stored in a rechargeable battery. Given that data transmission incurs an energy cost, our objective is to characterize *low-complexity* energy management policies, which achieve near-optimal performance in terms of the average long-term importance of the reported data, at a fraction of the complexity. Initially, we focus on the operation of a single EH device. Then, we address the problem of multiaccess in a WSN.

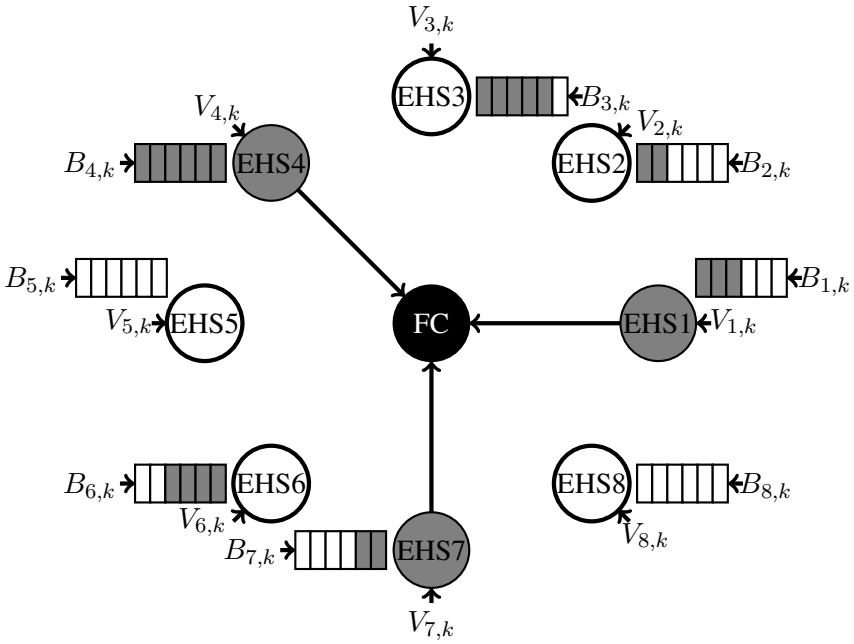


Figure 3.1. *Energy Harvesting Wireless Sensor Network (EH-WSN)*

The issue of energy management for solar-powered EHSs and RFIDs has previously been addressed in [39] and [40], respectively, primarily from a numerical standpoint. [41] derived the policy which maximizes the long-term detection probability of a random event and, in [42], a similar problem was considered in the context of body sensor networks. [43] studied data transmission for a two-state fading channel, and properties of the policy that maximizes the long-term discounted throughput were derived. In [44,45], policies which stabilize the data queue of an EHS with random data arrivals were proposed and analyzed. [46–48] derived policies that maximize the data throughput of the EHS by a deadline, relying, however, on the assumption that energy arrivals (and also the channel fading profile in [46]) are known beforehand. Other related works include [49], which explored activation policies in a network of EHSs, and [50], which derived power management algorithms for EHSs with battery inefficiencies. The contribution of our work with respect to, *e.g.*, [41–43, 45], is to explicitly take into account the impact of a finite battery capacity on the performance and its interplay with the EH process. The problem of maximizing the average long-term importance of the reported data for a replenishable sensor is formulated in [51], for a continuous-time model with Poisson EH and data processes. In addition, we introduce time-correlation in the energy supply, and investigate its impact on the performance. [38] investigates the relaying of packets of different priorities in a network of energy-limited sensors, but does not account for EH capability.

Despite the intense research effort in the design of optimal energy management policies for a sin-

gle EH device, *e.g.*, see [41,42,44,45], the problem of analyzing and modeling the interaction among multiple EH nodes in a network has not received much attention so far. Some notable exceptions are [52, 53], which address the design of Medium Access Control (MAC) protocols for EH-WSNs, focusing on TDMA and (dynamic) framed Aloha; [54], which develops efficient energy management policies that stabilize the data queues, as well as efficient MAC policies; [55], which focuses on the design of a MAC protocol for multi-hop EH-WSN that can achieve high throughput and fairness, using a probabilistic polling mechanism that adapts to changing energy harvesting rates or node densities to manage packet collisions and channel contention; [49], which addresses the problem of how sensor nodes should be activated dynamically so as to optimize the sensing and event detection performance of the network.

3.1.1 Contributions

Within this framework, we first analyze the performance of a single EHS. Specifically, we derive analytically the performance of a Balanced Policy (BP), which adapts the transmission probability based *only* on the harvesting state, but not on the current energy level in the battery, such that, on average, energy harvesting and consumption are balanced. Numerical results demonstrate that the optimal BP performs very well with respect to the globally optimal policy, and the gap between the two is reduced even further if the sensor is forced to transmit when the battery is fully charged. The main implication of these results is that near-optimal performance can be achieved with simple adaptation to the ambient energy supply, *without* precise knowledge of the energy stored in the sensor battery at any given time.

A key result is that the EHS performance is heavily dependent on the *power-to-depletion* ρ, defined as the power that a fully charged battery can supply over a BAD EH period, such that, on average, it is depleted at the end of the period. Essentially, ρ captures the ability of the battery to absorb the ambient energy fluctuations and provide a stable energy supply to the sensor circuitry. In the spirit of the "offline" optimization framework of [46–48], we also study an EHS with a deterministic and periodic energy supply and show that the optimal policy depends exclusively on ρ. Based on this analysis, we propose a heuristic BP which is shown to perform well in the original stochastic model.

Then, we extend the model to include multiple EHSs in an EH-WSN, which randomly access the wireless channel to transmit data packets of random importance to a common fusion center. Assuming that data transmission incurs an energy cost and simultaneous transmission from multiple EHSs

causes collision and packet loss, we study the problem of designing optimal random access policies, so as to maximize the network utility, defined as the average long-term aggregate network importance of the data packets successfully reported to the fusion center. Due to the generally non-convex structure of the optimization problem, we resort to approximate solutions. In particular, we use a mathematical artifice based on a game theoretic formulation of the multiaccess problem, where each sensor node is a player which attempts to selfishly maximize the network utility. We characterize the *Symmetric Nash Equilibrium* (SNE) of this game, where all the sensor nodes employ the same policy, and we provide an algorithm to compute it. Moreover, we propose low-complexity policies which only loosely depend on the exact amount of energy available in the battery, while attaining near-optimal data reporting performance.

3.1.2 Structure of the chapter

This chapter is organized as follows. In Sec. 3.2, we describe the system model for the scenario with a single EHS. The optimization problem and the formal policy definitions are presented in Sec. 3.3, followed by the analysis of the BP in Sec. 3.4. Sec. 3.5 is devoted to the analysis of a continuous-time, deterministic model, and the discussion of its connection to the stochastic model. Numerical results for the scenario with a single EHS are presented in Sec. 3.6. Then, we shift to the analysis of a multiaccess problem in EH-WSN. The multiaccess model is presented in Sec. 3.7. Sec. 3.8 defines the control policies and states the optimization problem, which is further developed in Sec. 3.9. In Sec. 3.10, we design low-complexity policies, which are suitable for practical implementation. In Sec. 3.11, we present some numerical results for the multiaccess scenario. Finally, Sec. 3.12 summarizes our main conclusions. The proofs of the theorems and lemmas are provided in the appendices at the end of the chapter.

We close this section with a note on the notation employed throughout this chapter: $\bar{x} = 1 - x$ is the complement of $x \in [0, 1]$ and $\chi(\cdot)$ is the indicator function; random variables are denoted by uppercase letters, and their values by lowercase ones. The list of symbols is provided in Table 3.1.

3.2 System Model: single EHS

The block diagram of a wireless EHS is shown in Fig. 3.2. The energy harvesting unit collects ambient energy, which is stored in a battery (or super-capacitor) and then used to power the sensing apparatus and the RF circuitry. A processing unit, *e.g.*, a micro-controller, manages the energy con-

Table 3.1. *List of symbols*

λ_G	EH probability in GOOD EH state
β	Average EH rate
p_G	Transition probability from GOOD to GOOD EH state
p_B	Transition probability from BAD to BAD EH state
$\pi_A(G)$	Probability of GOOD EH state
$\pi_A(B)$	Probability of BAD EH state
D_G	Average duration of GOOD EH period
D_B	Average duration of BAD EH period
$\gamma = \frac{D_G}{D_B} = \frac{\pi_A(G)}{\pi_A(B)}$	Ratio of average durations of GOOD and BAD EH periods
$e_{\max}$	Battery capacity
$\rho = \frac{e_{\max}}{D_B}$	Power-to-depletion
$\theta \in \{0, 1\}$	Overflow Avoidance parameter
η	Transmission probability induced by threshold policy μ

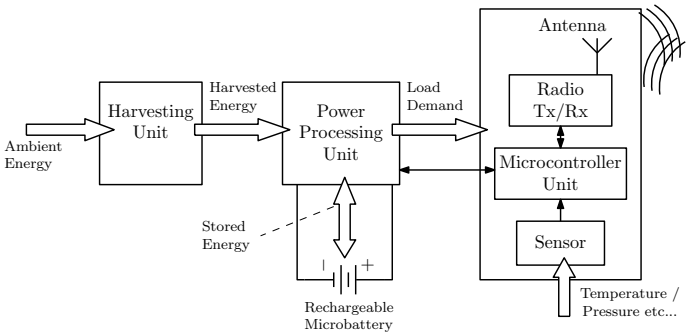


Figure 3.2. *Block diagram of an EHS*

sumption of the EHS. The sensing apparatus collects data and measurements from the sensing field, which are collected in data packets to be reported to FC. We consider a slotted-time system, where slot k is the time interval $[kT, kT + T)$, $k \in \mathbb{Z}^+$, and T is the slot duration. At each time instant k, the EHS has a new data packet to send to FC with importance V_k. We assume that a stringent delay requirement is enforced at the EHS: the packet is either sent to FC over the interval $[kT, kT + \delta T)$, where $\delta \in (0, 1]$ is the duty cycle,[1] or it is dropped. Note that typical WSN applications are loss tolerant, since sensing data exhibit redundancy and correlation over space and time.

The EHS battery is modeled by a buffer. As in previous work [41,42,57], we assume that each position in the buffer can hold one energy quantum and that the transmission of one data packet requires the expenditure of one energy quantum.[2] The maximum number of quanta that can be stored, *i.e.*, the battery capacity, is $e_{\max}$ and the set of possible energy levels is denoted by $\mathcal{E} = \{0, 1, \ldots, e_{\max}\}$. At

[1] $\delta \in (0, 1]$ models a typical characteristic of EHS systems (see, *e.g.*, [56]): the energy to perform a given task (transmit a packet) is spent much faster than it is collected. Note that the value of δ has no impact on the subsequent analysis.

[2] We only consider the energy expenditure associated with RF transmission.

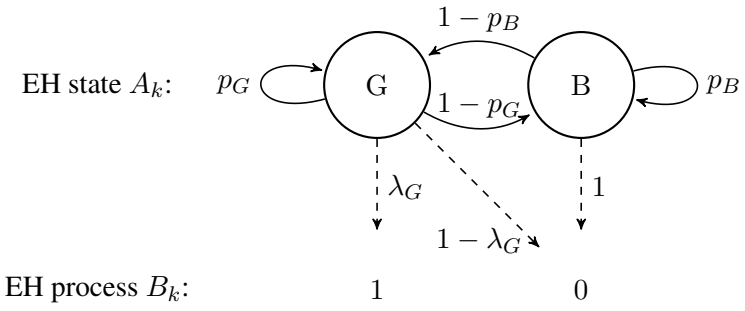


Figure 3.3. *Energy Harvesting process*

time $k+1$, $k \in \mathbb{Z}^+$, the amount of energy in the buffer is

$$E_{k+1} = \min\left\{E_k - Q_k + B_k, e_{\max}\right\}, \tag{3.1}$$

where $\{B_k\}$ is the *energy arrival process* and $\{Q_k\}$ is the *action process*. $Q_k = 1$ if the current data packet is transmitted, which results in the expenditure of one energy quantum, and $Q_k = 0$ otherwise. B_k models the randomness in the energy harvested in slot k. We assume that $B_k \in \{0, 1\}$, *i.e.*, either one energy quantum is harvested, or no energy is harvested at all. Moreover, the energy harvested in time-slot k can be used only in a later time-slot. As a consequence, if the battery is depleted, *i.e.*, $E_k = 0$, then $Q_k = 0$. We model the underlying EH process $\{A_k\}$ as a two-state Markov chain, with state space $\{\mathrm{G}, \mathrm{B}\}$, where G and B denote the GOOD and BAD harvesting states, respectively, as depicted in Fig. 3.3. If $A_k = \mathrm{G}$ (GOOD state), then $B_k = 1$ with probability λ_G, where $\lambda_G \in (0, 1]$, and $B_k = 0$ with probability $1 - \lambda_G$; if $A_k = \mathrm{B}$ (BAD state), then $B_k = 0$. When $\lambda_G < 1$, energy is harvested at a slower rate than it is consumed for data transmission: on average, $1/\lambda_G$ time-slots are required to harvest one energy quantum in the GOOD state. We denote the transition probabilities of $\{A_k\}$ from G to G and from B to B as $p_G = \Pr(A_k = \mathrm{G}|A_{k-1} = \mathrm{G})$ and $p_B = \Pr(A_k = \mathrm{B}|A_{k-1} = \mathrm{B})$, respectively. The steady-state distribution of $\{A_k\}$ is thus

$$\pi_A(\mathrm{G}) = \frac{\bar{p}_B}{\bar{p}_B + \bar{p}_G}, \;\; \pi_A(\mathrm{B}) = \frac{\bar{p}_G}{\bar{p}_B + \bar{p}_G}. \tag{3.2}$$

The average durations of the GOOD and BAD EH periods are denoted by D_G and D_B, respectively, and their ratio by $\gamma = D_G/D_B$. Simple calculations yield that $D_G = 1/\bar{p}_G$, $D_B = 1/\bar{p}_B$ and $\gamma = \pi_A(\mathrm{G})/\pi_A(\mathrm{B})$. Finally, since one energy quantum is harvested with probability λ_G in every

GOOD time-slot, the average EH rate, *i.e.*, the average long-term amount of energy harvested by the EH unit in one time-slot, is

$$\beta = \lim_{K\to\infty} \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} B_k\right] = \lambda_G \pi_A(\mathrm{G}), \tag{3.3}$$

where $\beta \in (0,1)$. Note that β, γ and λ_G are related as

$$\beta = \frac{\lambda_G \gamma}{\gamma + 1}. \tag{3.4}$$

We now formally define the events of energy *outage* and *overflow*.

Definition 3.2.1 (Outage)**.** In slot k, *energy outage* occurs if $E_k = 0$.

Definition 3.2.2 (Overflow)**.** In slot k, *energy overflow* occurs if $(E_k = e_{\max}) \cap (B_k = 1) \cap (Q_k = 0)$.

Under energy outage, no transmissions can be performed, *i.e.*, $Q_k = 0$. Energy overflow occurs when a harvested energy quantum ($B_k = 1$) cannot be stored due to a fully charged battery ($E_k = e_{\max}$) in an idle time-slot ($Q_k = 0$), and is thus lost.

The state of the EHS at time k is given by $(\boldsymbol{S}_k, V_k)$, where $\boldsymbol{S}_k = (E_k, A_{k-1}) \in \mathcal{S}$ is the joint energy level and EH state, with $\mathcal{S} = \mathcal{E} \times \{\mathrm{G}, \mathrm{B}\}$, and $V_k \in \mathbb{R}^+$ is the importance value of the current data packet. We model V_k as a continuous random variable with probability density function (pdf) $f_V(v),\ v \geq 0$, with support $(0, +\infty)$, and assume that $\{V_k\}$ are i.i.d. Note that, at time k, the EHS controller can infer the posterior distribution of A_{k-1}, $\Pr(A_{k-1} = a | B_0, \ldots, B_{k-1})$ for $a \in \{\mathrm{G}, \mathrm{B}\}$, from the observation of the EH process $\{B_0, \ldots, B_{k-1}\}$. In fact, $\Pr(A_{k-1} = a | B_0, \ldots, B_{k-1})$ can be computed recursively as

$$\begin{aligned} &\Pr(A_{k-1} = a | B_0, \ldots, B_{k-1}) \\ &= \frac{\Pr(B_{k-1}|A_{k-1} = a) \sum_{a_0} \Pr(A_{k-1} = a | A_{k-2} = a_0) \Pr(A_{k-2} = a_0 | B_0, \ldots, B_{k-2})}{\sum_{a_1} \Pr(B_{k-1}|A_{k-1} = a_1) \sum_{a_0} \Pr(A_{k-1} = a_1 | A_{k-2} = a_0) \Pr(A_{k-2} = a_0 | B_0, \ldots, B_{k-2})}, \end{aligned} \tag{3.5}$$

where $\Pr(A_{k-2} = a_0 | B_0, \ldots, B_{k-2})$ is the posterior distribution of A_{k-2}, given the EH sequence $B_0, \ldots, B_{k-2}$, computed in the previous time-slot. The state A_{k-1} can then estimated from the posterior distribution (3.5). For example, the Maximum-A-Posteriori (MAP) criterion yields

$$\hat{A}_{k-1} = \arg\max_a \Pr(A_{k-1} = a | B_0, \ldots, B_{k-1}). \tag{3.6}$$

In this thesis, we assume that perfect knowledge of A_{k-1} is available at the EHS controller, and leave the problem of estimating A_{k-1} as future work.

3.3 Optimization Problem and Policy Definitions

3.3.1 Optimization Problem

Given $\boldsymbol{s}_k = (e, a) \in \mathcal{S}$ and $V_k = v \in \mathbb{R}^+$, the policy μ implemented by the controller in Fig. 3.2 is defined by the probability $\mu(1; e, a, v)$ of transmitting the data packet in slot k. The respective probability of discarding the data packet is $\mu(0; e, a, v) = 1 - \mu(1; e, a, v)$.[3] Given an initial state $\boldsymbol{S}_0 \in \mathcal{S}$, the average long-term importance of the reported data (from now on referred to as *average reward* for brevity) under policy μ is

$$G(\mu; \boldsymbol{S}_0) = \lim_{K\to\infty} \inf \frac{1}{K} \mathbb{E}\left[\sum_{k=0}^{K-1} Q_k V_k \,\middle|\, \boldsymbol{S}_0 \right]. \tag{3.7}$$

The expectation in (3.7) is taken with respect to $\{B_k, A_k, Q_k, V_k\}$, where, at each instant k, Q_k is drawn according to policy μ and depends on the state (E_k, A_{k-1}, V_k), and E_k is given by (3.1).

The optimization problem at hand is to determine the optimal policy μ^* such that

$$\mu^* = \arg\max_{\mu} G(\mu; \boldsymbol{S}_0). \tag{3.8}$$

We now establish that μ^* has a threshold structure with respect to the data importance.

Lemma 3.3.1. *For each state $(e, a) \in \mathcal{S}$, there exists a threshold $v^*_{\text{th}}(e, a)$ such that*

$$\mu^*(1; e, a, v) = \begin{cases} 1, & v \geq v^*_{\text{th}}(e, a), \\ 0, & v < v^*_{\text{th}}(e, a). \end{cases} \tag{3.9}$$

Proof. See Appendix 3.A. □

Intuitively, Lemma 3.3.1 states that, for a given transmission probability budget $\mathbb{E}_V[\mu(1; e, a, V)]$, the optimal policy prioritizes the transmission of high over low importance data. As a consequence, we henceforth only consider policies with the structure defined in (3.9). For a threshold policy μ, the

[3] For the sake of maximizing an average long-term reward function of the state and action processes, it is sufficient to consider only stationary policies depending on the present state [22].

transmission probability in state (e, a) is

$$\eta(e,a) = \mathbb{E}_V[\mu(1;e,a,V)] = \bar{F}_V(v_{\text{th}}(e,a)), \tag{3.10}$$

where $\bar{F}_V(v), v \geq 0$, is the complementary cumulative distribution function (ccdf) of the importance value process. The expected reported data importance in state (e, a) is $g(\eta(e, a))$, where $g(x), x \in [0, 1]$, is a function defined as

$$g(x) = \mathbb{E}_V\left[\chi\left(V \geq \bar{F}_V^{-1}(x)\right) V\right] = \int_{\bar{F}_V^{-1}(x)}^{\infty} \nu f_V(\nu) \mathrm{d}\nu, \tag{3.11}$$

and $\bar{F}_V^{-1}(x)$ denotes the inverse of $\bar{F}_V(v)$. In words, $g(x)$ is the expected accrued reward when only the data with importance above the threshold $v = \bar{F}_V^{-1}(x)$ is reported. The function $g(x)$ has the following properties, which are stated without proof.

Lemma 3.3.2. *The function $g(x)$ is strictly increasing, strictly concave in x and $g'(x) = \bar{F}_V^{-1}(x)$, with $\lim_{x\to 0} g'(x) = +\infty$.*

From (3.9) and (3.10), it is seen that the mapping between a threshold policy μ and its respective $v_{\text{th}}(\cdot)$ and $\eta(\cdot)$ is one-to-one. Moreover, due to the independence between (A_k, B_k) and V_k, the transition probabilities of the time-homogeneous Markov chain $\{\boldsymbol{S}_k\}$ are governed by η. Therefore, in the remainder of the chapter, we refer to a threshold policy μ in terms of its corresponding transmission probability function $\eta(e, a), (e, a) \in \mathcal{S}$.

3.3.2 Policy Definitions

For the sake of mathematical tractability and without loss of optimality in (3.8), we only consider the set of policies that result in an average reward independent of the initial state $\boldsymbol{S}_0$.

Definition 3.3.1. The set $\mathcal{U}$ of admissible policies is defined as

$$\mathcal{U} = \{\eta : \eta(0,a) = 0,\ \eta(e_{\max},a) \in (0,1],\ \eta(e,a) \in (0,1),\ e = 1,\ldots,e_{\max}-1,\ \forall a \in \{\mathrm{G},\mathrm{B}\}\}.$$

It can be shown that the Markov chain $\{(E_k, A_{k-1})\}$ under policy $\eta \in \mathcal{U}$ has a unique closed communicating class. Hence, there exists a unique steady-state distribution, $\pi_\eta(e, a), (e, a) \in \mathcal{S}$,

independent of $\boldsymbol{S}_0$ [18]. From (3.7), for any $\eta \in \mathcal{U}$, we have

$$G(\eta) = \lim_{K\to\infty} \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} \chi\left(V_k \geq \bar{F}_V^{-1}(\eta(E_k, A_{k-1}))\right) V_k \,\middle|\, \boldsymbol{S}_0\right]$$
$$= \sum_{e=1}^{e_{\max}} \sum_{a\in\{\mathrm{G,B}\}} \pi_\eta(e,a) g(\eta(e,a)). \tag{3.12}$$

The optimization problem (3.8) over the class of admissible policies is stated as

$$\eta^* = \arg\max_{\eta\in\mathcal{U}} G(\eta). \tag{3.13}$$

The optimal policy η^* can be found numerically using the Policy Iteration Algorithm (PIA) for infinite horizon, average cost-per-stage problems [1,58]. In general, η^* is a function of the EH state A_{k-1} *and* the energy available in the battery, E_k. This implies a high implementation complexity for three reasons: the controller must make decisions based on the energy level, which may be too computationally intensive for the ultra-low power electronics typically found in practical EHSs (for example, PIA requires to update iteratively the transmission probability $\eta(e,a)$ for each value of the energy level $e \in \mathcal{E}$ and of the EH state $a \in \times\{\mathrm{G},\mathrm{B}\}$); the transmission probability for each state needs to be stored in an $2 \times e_{\max}$ look-up table, which takes up an amount of memory proportional to the size of the battery; and knowledge of E_k might be hard to obtain or imprecise at best [59, 60]. Motivated by these observations, we focus on the low-complexity Balanced Policy (BP), defined below.

Definition 3.3.2. A BP is any policy $\eta \in \mathcal{U}$ such that, for $a \in \{\mathrm{G},\mathrm{B}\}$,

$$\eta(e,a) = \begin{cases} \eta_a, & e \in \{1,2,\ldots,e_{\max}-1\}, \\ \theta + \bar{\theta}\eta_a, & e = e_{\max}, \end{cases} \tag{3.14}$$

where $\theta \in \{0,1\}$ is the Overflow Avoidance (OA) parameter and η_G and η_B are such that

$$\pi_A(\mathrm{G})\eta_G + \pi_A(\mathrm{B})\eta_B = \beta. \tag{3.15}$$

If $\theta = 0$, the transmission probability of the BP depends only on the EH state, *i.e.*, it is η_G in the GOOD state and η_B in the BAD state. If $\theta = 1$, the sensor always transmits when the battery is fully charged, thus avoiding energy overflow (Def. 3.2.2). OA introduces a mild dependence of the BP on the energy level, since the controller is required to know when the battery is fully charged.

According to (3.15), the BP "balances" the average energy consumption rate (left hand side of (3.15)) with the average EH rate (right hand side of (3.15)), if the impact of energy outage and overflow due to the finite battery capacity is neglected. Alternatively, since $\gamma = D_G/D_B = \pi_A(\mathrm{G})/\pi_A(\mathrm{B})$ and $\beta = \lambda_G \pi_A(\mathrm{G})$, (3.15) is equivalent to $D_G(\lambda_G - \eta_G) = D_B \eta_B$, *i.e.*, under the BP, an equilibrium amongst the recharge/discharge phases is achieved, in the sense that the expected energy recharge over the GOOD EH period, $D_G(\lambda_G - \eta_G)$, equals the expected energy discharge over the BAD EH period, $D_B \eta_B$.

From (3.14) and (3.15), it is seen that a BP is uniquely defined by the parameters (η_G, θ), where $\eta_G \in (\max\{\lambda_G - \gamma^{-1}, 0\}, \lambda_G)$ and $\theta \in \{0, 1\}$. In the remainder of the chapter, we thus refer to a BP η in terms of its corresponding pair (η_G, θ). The next section is devoted to the derivation of the average reward under the BP and the characterization of the optimal BP.

3.4 Performance Analysis of the BP

The main theoretical result of this section is a closed-form expression for the average reward of the BP and is presented in Theorem 3.4.1. The proof involves a crafty manipulation of the steady-state equations of the Markov chain (E_k, A_{k-1}) and is found in Appendix 3.B. The complicated general expression hardly lends itself to interpretation. We thus consider an asymptotic regime where energy arrivals are highly correlated *and* the battery capacity is very large. In this regime, we derive the average reward and its main properties (Theorem 3.4.3), and characterize the optimal BP (Lemma 3.4.4).

Theorem 3.4.1. *The average reward of the BP* (η_G, θ) *is*

$$\begin{aligned} G(\eta) =& (\pi_A(\mathrm{G}) - \pi_\eta(0, \mathrm{G}))g(\eta_G) + (\pi_A(\mathrm{B}) - \pi_\eta(0, \mathrm{B}))g(\eta_B) \\ &+ \theta \pi_\eta(e_{\max}, \mathrm{G}) \left(g(1) - g(\eta_G)\right), \end{aligned} \tag{3.16}$$

where

$$\begin{bmatrix} \pi_\eta(0, \mathrm{G}) \\ \pi_\eta(0, \mathrm{B}) \end{bmatrix} = \mathbf{Z}\mathbf{J}^{e_{\max}-1}\mathbf{t}(\theta)\pi_\eta(e_{\max}, \mathrm{G}), \tag{3.17}$$

$$\pi_\eta(e_{\max}, \mathrm{G}) = \left(\begin{bmatrix} 1 & 1 \end{bmatrix} \mathbf{Z}\mathbf{J}^{e_{\max}-1}\mathbf{t}(\theta) + \sum_{e=0}^{e_{\max}-1} \begin{bmatrix} 1 & 1 \end{bmatrix} \mathbf{J}^e \mathbf{t}(\theta) - \theta \frac{\bar{\eta}_G}{\eta_G} \right)^{-1} \tag{3.18}$$

and

$$\mathbf{t}(\theta) = \begin{bmatrix} \theta\frac{1}{\eta_G} + \bar{\theta} \\ \bar{\theta}\frac{\gamma^{-1}\bar{\eta}_G}{D_B\eta_B+\bar{\eta}_B} \end{bmatrix}, \tag{3.19}$$

$$\mathbf{Z} = \frac{1}{\lambda_G}\begin{bmatrix} \eta_G\bar{\lambda}_G & \eta_B\bar{\lambda}_G \\ \gamma^{-1}\eta_G & \eta_B[\gamma^{-1}\bar{\lambda}_G + (D_B-1)\lambda_G] \end{bmatrix}, \tag{3.20}$$

$$\mathbf{J} = \frac{\eta_B/\lambda_G}{\bar{\eta}_B + \eta_B(D_B-\gamma^{-1})}\begin{bmatrix} \frac{\eta_B}{\bar{\eta}_G} \\ 1 \end{bmatrix}\left[\frac{\eta_G\gamma^{-1}}{\eta_B}; D_B-1\right] + \frac{\bar{\lambda}_G}{\lambda_G}\begin{bmatrix} \frac{\eta_G}{\bar{\eta}_G} & \frac{\eta_B}{\bar{\eta}_G}\frac{1}{\bar{\eta}_B+\eta_B(D_B-\gamma^{-1})} \\ 0 & -\frac{\eta_B(D_B-1-\gamma^{-1})}{\bar{\eta}_B+\eta_B(D_B-\gamma^{-1})} \end{bmatrix}. \tag{3.21}$$

Proof. See Appendix 3.B. □

The interpretation of (3.16) is as follows. The terms $(\pi_A(\mathrm{G}) - \pi_\eta(0,\mathrm{G}))g(\eta_G)$ and $(\pi_A(\mathrm{B}) - \pi_\eta(0,\mathrm{B}))g(\eta_B)$ are the average rewards accrued in the GOOD and BAD states, respectively, where the terms $\pi_\eta(0,\mathrm{G})$ and $\pi_\eta(0,\mathrm{B})$ account for the performance loss due to energy outage events. The last term $\theta\pi_\eta(e_{\max},\mathrm{G})\,(g(1)-g(\eta_G))$ accounts for the impact of OA, *i.e.*, the fact that, in state $(e_{\max},\mathrm{G})$, a data packet is always transmitted irrespective of its value.

In general, $G(\eta)$ does not admit a simple expression, due to the presence of the matrix exponential $\mathbf{J}^e$ in (3.17) and (3.18). However, a simple expression can be obtained when $\lambda_G = 1$, *i.e.*, when one energy quantum is always harvested in the GOOD state.

Lemma 3.4.2. *If $\lambda_G = 1$, the average reward of a BP η with parameters (η_G,θ) is*

$$\begin{aligned} G(\eta) =& \beta g(\eta_G) + \bar{\beta}\frac{e_{\max}-\eta_G-\theta\bar{\eta}_G}{e_{\max}+\eta_B(D_B-1)-\theta\bar{\eta}_G}g(\eta_B) \\ &+ \theta\beta\frac{1+\eta_B(D_B-1)-\theta\bar{\eta}_G}{e_{\max}+\eta_B(D_B-1)-\theta\bar{\eta}_G}\left(g(1)-g(\eta_G)\right). \end{aligned} \tag{3.22}$$

Proof. Letting $\lambda_G = 1$ in (3.21), the second term becomes zero and we can verify that $\mathbf{J}^n = \mathbf{J}$, $\forall n \geq 1$. Therefore, (3.17)-(3.18) can be readily computed and Eq. (3.22) follows from (3.16). □

In order to understand (3.22), let us focus on a simple Non-Adaptive BP (NABP) which always transmits with probability β and performs no OA, *i.e.*, $(\eta_G,\theta) = (\beta,0)$. In this case,

$$G(\eta) = g(\beta)\left(\beta + \bar{\beta}\frac{e_{\max}-\beta}{e_{\max}+\beta(D_B-1)}\right). \tag{3.23}$$

Moreover, let $e_{\max} \gg 1$. This is typical of real EHS deployments, *e.g.* in [56], the capacitance is

much larger than the energy of a transmit pulse. In slow-changing environments, it is also reasonable to assume that $D_B \gg 1$. In this setting, we can approximate $G(\eta)$ in (3.23) as

$$G(\eta) \simeq g(\beta)\left(\beta + \bar{\beta}\frac{\rho}{\rho+\beta}\right), \tag{3.24}$$

where we have defined $\rho \triangleq \frac{e_{\max}}{D_B}$. Note that ρ is the normalized power[4] that can be continuously supplied from a fully charged battery over a BAD period, such that, on average, the battery is empty at the end of the period; we thus name ρ the *power-to-depletion*. Its effect on the performance of NABP can be explained as follows. In an ideal scenario with infinite battery capacity, NABP is optimal, owing to the concavity of function $g(x)$ (Lemma 3.3.2). $D_B\beta$ is the expected energy amount drawn from the battery during the BAD period, and $D_G(\lambda_G - \beta) = D_B\beta$ is the expected energy amount by which the battery is recharged over the GOOD period. If a system with *finite* battery capacity is operated under this policy, ρ captures the extent to which the battery can absorb the fluctuations in the EH process. If $\rho \gg \beta$, *i.e.*, the power-to-depletion is much larger than the transmission probability, the battery has a "large" capacity and can sustain a constant energy consumption rate β, rarely being subject to outage and overflow events. In contrast, if $\rho \ll \beta$, the battery has a "small" capacity, hence it is deeply discharged over the BAD EH period, and fully recharged over the GOOD EH period. The performance is thus severely affected by energy outage and overflow, as can be seen from (3.24): letting $\rho \to 0$, $G(\eta) \approx \beta g(\beta)$, *i.e.*, transmitting with constant probability β achieves only a fraction β of the theoretical upper bound $g(\beta)$. This indicates that, for $\rho \ll \beta$, adaptation to the EH state is critical to achieve good performance.

In order to study the impact of ρ on the performance of a general BP, we focus on the asymptotic regime $D_B \to \infty$, $D_G \to \infty$ and $e_{\max} \to \infty$, where the ratios $\rho = e_{\max}/D_B$ and $\gamma = D_G/D_B$ (hence $\beta = \lambda_G\gamma/(1+\gamma)$) are kept *fixed*. This regime corresponds to a scenario of extreme correlation in the EH process, where the GOOD and BAD periods are much longer than a time slot, and the battery capacity is much larger than an energy quantum. For the BP (η_G, θ), we denote the asymptotic average reward as $G^{(\infty)}(\eta_G, \theta; \rho) = \lim_{D_B\to\infty} G(\eta)$. In Theorem 3.4.3, we derive $G^{(\infty)}(\eta_G, \theta; \rho)$ and characterize its main properties.

[4]Note that ρ has units of [energy quanta]/[time-slots], hence it represents a normalized power.

Theorem 3.4.3. *The asymptotic average reward for the BP (η_G, θ) is*

$$\begin{aligned} G^{(\infty)}(\eta_G, \theta; \rho) =& \pi_A(\mathrm{G}) g(\eta_G) + \pi_A(\mathrm{B}) \frac{\rho}{\rho + \eta_B} g(\eta_B) \\ &+ \theta \pi_A(\mathrm{G}) \frac{\lambda_G - \eta_G}{1 - \eta_G} \frac{\eta_B}{\rho + \eta_B} (g(1) - g(\eta_G)). \end{aligned} \tag{3.25}$$

Moreover:

1. $G^{(\infty)}(\eta_G, 1; \rho) > G^{(\infty)}(\eta_G, 0; \rho)$*;*

2. $G^{(\infty)}(\eta_G, \theta; \rho)$ *is an increasing function of* ρ*, with*

$$\lim_{\rho \to \infty} G^{(\infty)}(\eta_G, \theta; \rho) = \pi_A(\mathrm{G}) g(\eta_G) + \pi_A(\mathrm{B}) g(\eta_B), \tag{3.26}$$

$$\lim_{\rho \to 0} G^{(\infty)}(\eta_G, \theta; \rho) = \pi_A(\mathrm{G}) \left[g(\eta_G) + \theta \frac{\lambda_G - \eta_G}{1 - \eta_G} (g(1) - g(\eta_G)) \right]. \tag{3.27}$$

Proof. See Appendix 3.C. □

Eq. (3.25) is a generalization of (3.24) to any BP (η_G, θ) and any $\lambda_G \leq 1$. Property 1) shows that OA increases the (asymptotic) average reward for any η_G and ρ. Intuitively, without OA, part of the energy is lost due to overflow, whereas, with OA, all the harvested energy is used towards data transmission. Property 2) generalizes our previous observations on the performance of NABP: for any BP (η_G, θ), the (asymptotic) average reward increases with ρ, *i.e.*, as the battery capacity becomes larger with respect to D_B. For $\rho \to \infty$, there is no outage nor overflow, which explains the limit in (3.26). In contrast, for $\rho \to 0$, the battery is almost surely led to outage in the BAD state and, in the long term, reward is only accrued in the GOOD state; hence the limit in (3.27).

Having derived the asymptotic average reward for any BP, we now characterize the optimal BP in the asymptotic regime. Let $\eta_G^*(\theta; \rho) = \arg\max_{\eta_G} G^{(\infty)}(\eta_G, \theta; \rho)$.

Lemma 3.4.4. *The optimal BP in the asymptotic regime, $\eta_G^*(\theta; \rho)$, is the unique solution of*

$$L(\eta_G, \theta; \rho) = 0 \tag{3.28}$$

in $\eta_G \in (\max\{\lambda_G - \gamma^{-1}, 0\}, \lambda_G)$*, where*

$$L(\eta_G, \theta; \rho) = \left(1 + \frac{\eta_B}{\rho}\right)^2 g'(\eta_G) + \frac{g(\eta_B)}{\rho} - \left(1 + \frac{\eta_B}{\rho}\right) g'(\eta_B) \tag{3.29}$$
$$- \theta \frac{\eta_B}{\rho \bar{\eta}_G^2} \left(\bar{\eta}_G + \bar{\lambda}_G + \frac{\eta_B}{\rho}\bar{\lambda}_G\right)(g(1) - g(\eta_G)) - \theta \frac{\lambda_G - \eta_G}{1 - \eta_G} \frac{\eta_B}{\rho} \left(1 + \frac{\eta_B}{\rho}\right) g'(\eta_G).$$

Moreover:

1. $\eta_G^*(\theta; \rho) \in (\beta, \lambda_G)$*;*

2. $\eta_G^*(1; \rho) < \eta_G^*(0; \rho)$*;*

3. $\eta_G^*(\theta; \rho)$ *is a decreasing function of* ρ*, for* $\rho \geq \beta$*, with*

$$\lim_{\rho \to \infty} \eta_G^*(\theta; \rho) = \beta, \tag{3.30}$$

$$\lim_{\rho \to 0} \eta_G^*(\theta; \rho) = \lambda_G. \tag{3.31}$$

Proof. See Appendix 3.D. □

The main implication of Lemma 3.4.4 is that the optimal BP in the asymptotic regime can be easily found numerically: from property 1), we know that η_G^* lies in (β, λ_G). Moreover, $L(\eta_G, \theta; \rho)$ is a decreasing function of η_G, with $L(\beta, \theta; \rho) > 0$ and $L(\lambda_G, \theta; \rho) < 0$ (see Appendix 3.D). Hence, $L(\eta_G, \theta; \rho) = 0$ can be solved using the *bisection method* [61].

Property 2) reveals that the optimal BP without OA is more "aggressive" in the GOOD state than the optimal BP with OA. In other words, since there is no protection from overflow, the policy itself tries to minimize energy spillover by forcing a higher consumption rate when energy is available.

Finally, property 3) provides yet further insight as to the characteristics of $\eta_G^*(\theta; \rho)$. In the limit $\rho \to \infty$, the battery capacity is large enough that transmitting with constant probability β is optimal. From (3.26), it is seen that $G^{(\infty)} \to g(\beta)$, *i.e.*, the upper bound is achieved. In contrast, when $\rho \to 0$, the battery capacity is so small relative to the time scale of the EH process that it is optimal to use all the energy as it is being harvested during the GOOD state, *i.e.*, $\eta_G^* \to \lambda_G$. In this case, (3.26) yields $G^{(\infty)} \to \pi_A(\mathrm{G}) g(\lambda_G)$. As ρ takes values from 0 to ∞, η_G^* decreases from λ_G to β (correspondingly, $\eta_B^* = \gamma(\lambda_G - \eta_G^*)$ increases from 0 to β), and the optimal $G^{(\infty)}$ increases from $\pi_A(\mathrm{G}) g(\lambda_G)$ to $g(\beta)$.

3.5 Continuous-Time Model with Deterministic EH process

In this section, we investigate a continuous-time model with a deterministic and periodic EH process. We refer to this model as CDM, to differentiate it from the discrete-time stochastic model of Sec. 3.2, which, in this section, we denote as DSM. The motivation for considering CDM is twofold. In the asymptotic regime $D_B, D_G, e_{\max} \to \infty$ with $\rho = e_{\max}/D_B$ and $\gamma = D_G/D_B$ fixed, DSM resembles a continuous-time-energy model, in that the time-slot duration is much smaller than the time scale of the EH process and the battery capacity is much larger than an energy quantum. Moreover, in CDM, the controller has perfect knowledge of the EH profile, similar to the offline optimization problems addressed in [47,48]. It is thus of interest to derive further insight on DSM using CDM and also to draw parallels with [47,48].

Adhering to the notation in Sec. 3.2, in CDM, the battery capacity is denoted by $e_{\max}$. The EH process is periodic with period $D_G + D_B$, where D_G and D_B denote the (deterministic) durations of the GOOD and BAD periods. Mathematically, letting t be a time counter, when $t \in \mathcal{T}_G$, the EH state is GOOD and, when $t \in \mathcal{T}_B$, it is BAD, where $\mathcal{T}_G = \cup_{k\in\mathbb{Z}}\mathcal{T}_G(k)$, $\mathcal{T}_G(k) = [k(D_G + D_B), k(D_G + D_B) + D_G)$ and $\mathcal{T}_B = \cup_{k\in\mathbb{Z}}\mathcal{T}_B(k)$, $\mathcal{T}_B(k) = [k(D_G + D_B) + D_G, (k+1)(D_G + D_B))$ denote the sets of GOOD and BAD time intervals, respectively. During the GOOD period, energy is harvested with rate λ_G, and, during the BAD period, no energy is harvested; therefore, as in DSM, the average EH rate is $\beta = \lambda_G D_G/(D_G + D_B)$, and we denote the average long-term fraction of time spent in the GOOD and BAD EH periods as $\pi_A(\mathrm{G}) = \frac{D_G}{D_G+D_B}$ and $\pi_A(\mathrm{B}) = \frac{D_B}{D_G+D_B}$, respectively, with $\gamma = D_G/D_B = \pi_A(\mathrm{G})/\pi_A(\mathrm{B})$. A policy for CDM is defined by the energy drawing rate $\eta^{(\mathrm{CDM})}(t) \in [0, 1]$, which specifies the rate according to which energy is drawn from the battery. In particular, if the battery is empty, then $\eta^{(\mathrm{CDM})}(t) = 0$; if it is full for $t \in \mathcal{T}_G$, then $\eta^{(\mathrm{CDM})}(t) \geq \lambda_G$, so that no energy is lost due to overflow. The energy level at time t, $E(t)$, is thus given by

$$\begin{aligned} E(t) &= \min\left\{E(k(D_G + D_B)) + \textstyle\int_{k(D_G+D_B)}^{t}(\lambda_G - \eta^{(\mathrm{CDM})}(\tau))\mathrm{d}\tau, e_{\max}\right\}, && t \in \mathcal{T}_G(k), \\ E(t) &= E(k(D_G + D_B) + D_G) - \textstyle\int_{k(D_G+D_B)+D_G}^{t}\eta^{(\mathrm{CDM})}(\tau)\mathrm{d}\tau, && t \in \mathcal{T}_B(k). \end{aligned} \tag{3.32}$$

Since the EH process is periodic, it is sufficient to consider a periodic policy

$$\eta^{(\mathrm{CDM})}(t + k(D_G + D_B)) = \eta^{(\mathrm{CDM})}(t), \forall t \in [0, D_G + D_B),\ \forall k \in \mathbb{Z}.$$

Therefore, the following boundary conditions must hold for $k \in \mathbb{Z}$ (after, possibly, an initial transient phase, whose impact on the average long-term reward is negligible)

$$E(k(D_G + D_B)) = e_L, \;\; E(k(D_G + D_B) + D_G) = e_H, \tag{3.33}$$

i.e., during the GOOD period, the battery is recharged from e_L to e_H and, during the BAD period, it is discharged from e_H to e_L. By definition, $e_L, e_H \in [0, e_{\max}]$ and $e_L \leq e_H$.

We define the instantaneous reward in CDM as $g(\eta^{(\mathrm{CDM})}(t))$. Note that we employ the same mathematical reward function as DSM for the purpose of comparing the two models later in Lemma 3.5.2. However, it is emphasized that the physical meaning of $g(x)$ is different for each model: in CDM, the argument x is the rate according to which energy is drawn from the battery, and $g(x)$ is the corresponding instantaneous reward; in DSM, x is the transmission probability and $g(x)$, defined in (3.11), is the corresponding expected data importance. With these remarks in place, the average long-term reward in CDM is

$$G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})}) = \frac{\int_0^{D_G} g(\eta^{(\mathrm{CDM})}(\tau))\mathrm{d}\tau + \int_{D_G}^{D_G+D_B} g(\eta^{(\mathrm{CDM})}(\tau))\mathrm{d}\tau}{D_B + D_G}. \tag{3.34}$$

The optimal policy in CDM is the solution of

$$\eta^{(\mathrm{CDM})*} = \arg \max_{\eta^{(\mathrm{CDM})}} G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})}). \tag{3.35}$$

The following lemma determines $\eta^{(\mathrm{CDM})*}$ and the respective optimal average reward. As in DSM, let $\rho = e_{\max}/D_B$ be the power-to-depletion in CDM.

Lemma 3.5.1. *The optimal policy for CDM is*

$$\eta^{(\mathrm{CDM})*}(t) = \begin{cases} \lambda_G - \frac{1}{\gamma}\min\{\beta, \rho\}, & t \in \mathcal{T}_G, \\ \min\{\beta, \rho\}, & t \in \mathcal{T}_B, \end{cases} \tag{3.36}$$

and the optimal average reward

$$G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})*}) = \pi_A(\mathrm{G}) g\left(\lambda_G - \frac{1}{\gamma}\min\{\beta, \rho\}\right) + \pi_A(\mathrm{B}) g\left(\min\{\beta, \rho\}\right). \tag{3.37}$$

Proof. See Appendix 3.E. □

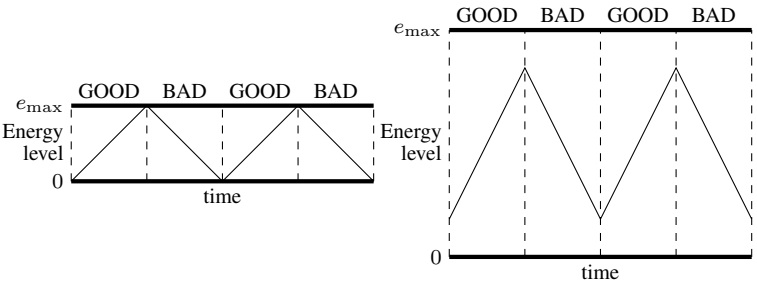


Figure 3.4. *Operation of CDM under the optimal policy. Cases* $\rho = e_{\max}/D_B < \beta$ *(left) and* $\rho = e_{\max}/D_B \geq \beta$ *(right).*

Lemma 3.5.1 distinguishes two regimes of operation in CDM which are determined exclusively by the relation between ρ and β. As illustrated in Fig. 3.4, if $\rho \geq \beta$, energy can be drawn with constant rate β and the optimal average reward is thus $g(\beta)$. If $\rho < \beta$, energy is drawn with rate $\lambda_G - \rho/\gamma = \lambda_G - e_{\max}/D_G > \beta$ during the GOOD phase, and with rate $e_{\max}/D_B < \beta$ during the BAD phase, *i.e.*, the battery is completely recharged and discharged over each cycle ($e_H = e_{\max}$, $e_L = 0$). Under the prism of [47, 48], the optimal energy expenditure curve in the interval $[k, k+1](D_G + D_B)$ is the unique minimum-length curve that lies in the feasible energy "tunnel" defined by the energy arrival curve, and its downward-shifted version by $e_{\max}$. If $\rho \geq \beta$, the slope of the expenditure curve is constant and equal to β, whereas, if $\rho < \beta$, it is $\lambda_G - e_{\max}/D_G$ during the GOOD phase and $e_{\max}/D_B < \beta$ during the BAD phase.

Note that, both in Lemma 3.4.4 and in Lemma 3.5.1, the value of ρ essentially determines the optimal policy. The main difference is that in CDM the EH profile is completely known, thus the energy consumption rate can be optimally adjusted to avoid outage and overflow. In contrast, in DSM energy outage may occur, and energy may also be wasted (if OA is not employed) due to the randomness in the energy arrivals. In the following lemma, we formalize these intuitive remarks by comparing the asymptotic average reward for DSM, $G^{(\infty)}(\eta_G, \theta; \rho)$, with $G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})*})$.

Lemma 3.5.2. *For any BP* (η_G, θ) *in DSM,*

$$G^{(\infty)}(\eta_G, \theta; \rho) \leq G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})*}). \tag{3.38}$$

Proof. See Appendix 3.F. □

We close this section by proposing the following suboptimal policy for DSM, based on the optimal

policy found for CDM in Lemma 3.5.1:

$$\eta(e,\mathrm{G}) = \lambda_G - \frac{\min\{\beta,\rho\}}{\gamma}, \quad \eta(e,\mathrm{B}) = \min\{\beta,\rho\},\ \forall e \in \mathcal{E}\setminus\{0\}. \tag{3.39}$$

This is a BP for DSM, as can be verified from Def. 3.3.2, and its performance can thus be evaluated analytically from Theorem 3.4.1.

3.6 Numerical Results: single EHS

In this section, we present numerical performance results for the scenario with a single EHS, for the following policies:

- *Optimal policy (OP)*, obtained numerically via the PIA [1];
- *Optimal BP with OA (OBP-OA) and without OA (OBP)*, determined by solving (3.28) for $\theta = 1$ and $\theta = 0$ using the bisection method [61];
- *Heuristic BP (HBP)*, defined in (3.39);
- *Non-Adaptive BP (NABP)*, the BP with $(\eta_G, \theta) = (\beta, 0)$;
- *Greedy Policy (GP)*, which always transmits when there is energy in the buffer.

The average reward of OP is computed numerically via the PIA [1]; those of OBP-OA, OBP, HBP and NABP can be computed analytically from Theorem 3.4.1. By definition, the average reward of GP is $\beta g(1)$.

For the purposes of this section, we let $V_k = \log_2(1+\Lambda H_k)$, where H_k is exponentially distributed with unit mean and $\Lambda > 0$ is a scaling parameter. This choice of V_k corresponds to the information rate achievable on a Rayleigh fading channel with gain H_k, where Λ is the average receive SNR, and the transmitter and receiver have full channel state information, so that the former can perform rate adaptation, whereas the latter can employ coherent detection [20]. The ccdf of the data importance is $\bar{F}_V(v) = \Pr\left(H_k \geq \frac{2^v-1}{\Lambda}\right) = \exp\left\{-\frac{2^v-1}{\Lambda}\right\}$. From (3.11),

$$g(x) = \int_{-\ln x}^{\infty} \log_2(1+\Lambda h)e^{-h}\mathrm{d}h = x\log_2(1-\Lambda\ln x) + \log_2(e)e^{\frac{1}{\Lambda}}E_1(\Lambda^{-1}-\ln x), \tag{3.40}$$

where $E_1(x)$ is the exponential integral function [62], defined as $E_1(t) = \int_t^\infty e^{-\tau}/\tau \mathrm{d}\tau$, and computed in Matlab using `expint(t)`. Note that (3.40) is a generalization of [63, Eq. (17)] and [64,

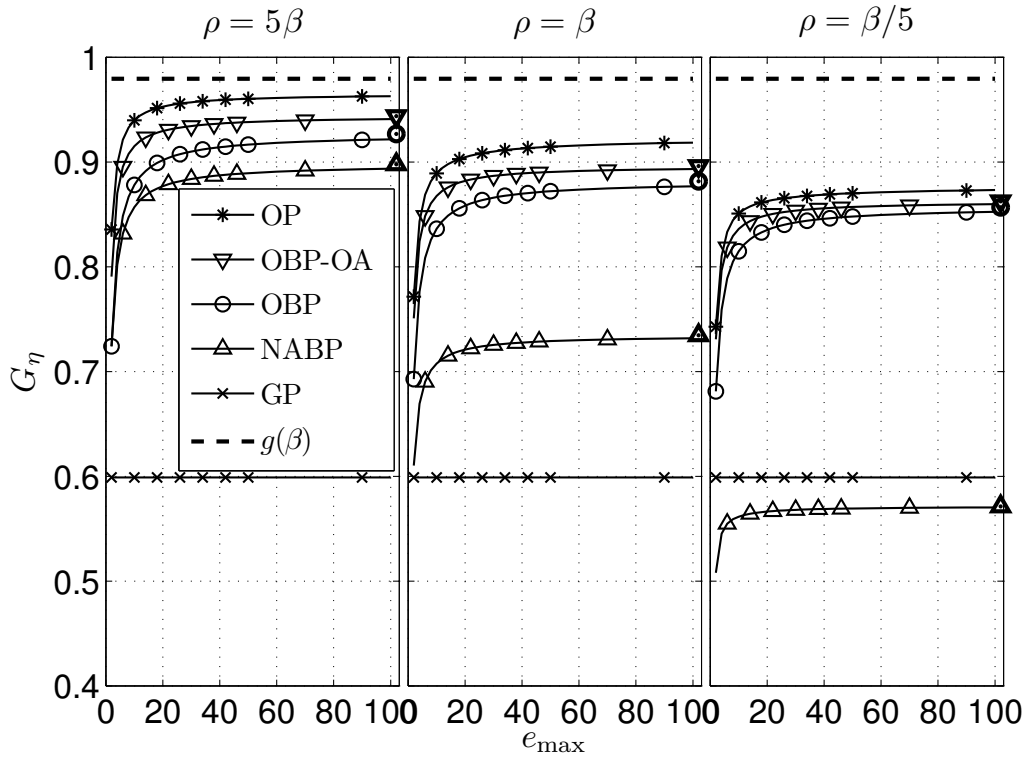


Figure 3.5. *Average reward as a function of $e_{\max}$ for the considered policies. The respective asymptotic average rewards, obtained from (3.25), are plotted with bold markers in the right side of each subplot. The performance is heavily dependent on ρ and only mildly affected by the absolute value of $e_{\max}$. ($\beta = 0.25$, $\lambda_G = 0.5$ and SNR$= 8$dB)*

Eq. (5)] for $x \geq 1$. Unless otherwise stated, we let $\beta = 0.25$ and $\lambda_G = 0.5$; hence, from (3.4), $\gamma = 1$. Moreover, we set $\Lambda = 6.31$, which corresponds to an average SNR of 8dB.

The numerical results provided in this section are derived for representative values of the system parameters. However, we have verified that the following observations hold for broader parameter ranges and pdf of the data importance process.

In Fig. 3.5, we plot the average reward as a function of $e_{\max}$, for $\rho \in \{0.2\beta, \beta, 5\beta\}$. For each value of ρ, D_B is determined as $D_B = e_{\max}/\rho$ and D_G as $D_G = \gamma D_B$. The asymptotic average reward (3.25) is shown with a bold marker in the right side of each subplot. Note that, for all policies, $G(\eta)$ quickly approaches the asymptotic value, *i.e.*, for $e_{\max} \gtrsim 20$, and displays a constant behavior as a function of $e_{\max}$. This suggests that the absolute value of $e_{\max}$ only mildly affects the system performance in the range $e_{\max} \gtrsim 20$. In general, the performance of all policies except GP improves with increasing ρ, and approaches more closely the upper bound $g(\beta)$. It is seen that OBP incurs only a small performance degradation with respect to OP: within 6%, for all values of ρ and $e_{\max} \geq 12$.

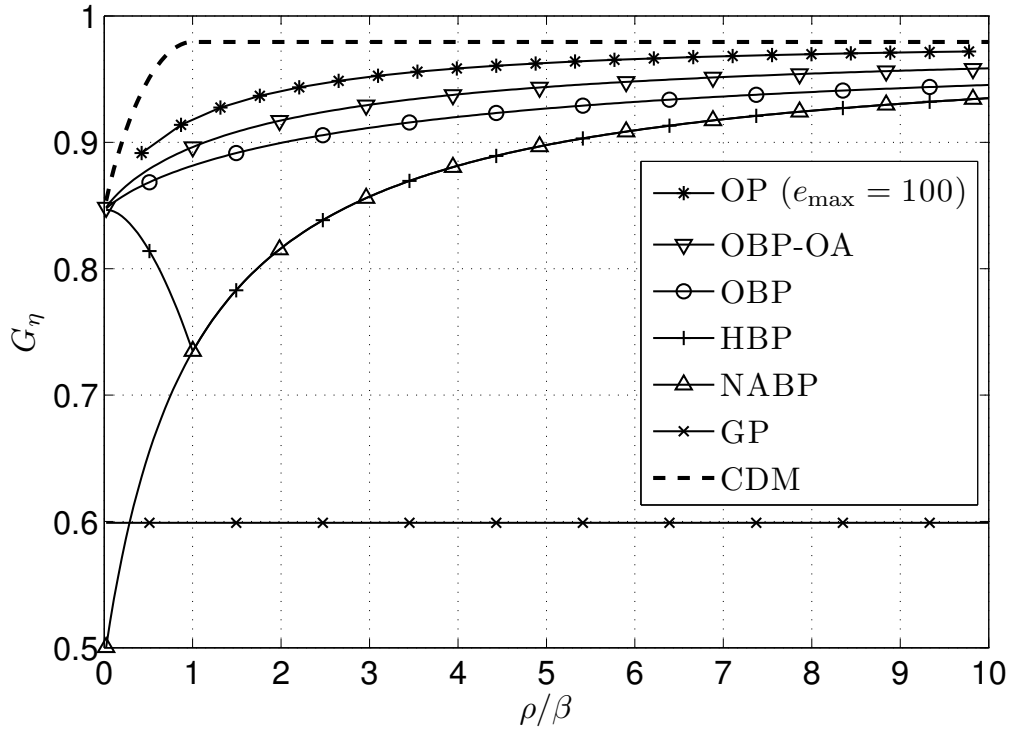


Figure 3.6. *Asymptotic average reward (3.25) vs. ρ/β. ($\beta = 0.25$, $\lambda_G = 0.25$ and SNR$= 8$dB)*

OA reduces the gap even further: within 3% of OP, for $e_{\max} \geq 4$. As discussed in Sec. 3.4, NABP, which does not adapt the transmission probability to the EH state, approaches OBP for large values of ρ, but incurs a significant performance loss for small values of ρ ($\sim 35\%$ compared to OP).

In Fig. 3.6, we plot the asymptotic average reward (3.25) for the considered policies as a function of ρ/β, and, in Fig. 3.7, the respective η_G for the BPs (in Fig. 3.7, we do not plot OP and GP, since the former is also a function of the energy level in the battery, whereas the latter transmits with probability one whenever energy is available). For OP in particular, an approximation of the asymptotic average reward is obtained from the PIA for $e_{\max} = 100$. In Fig. 3.6, we also plot the curve for CDM, which is an upper bound to the asymptotic average reward achieved by any BP, as proved in Lemma 3.5.2. We note that OBP and OBP-OA are within 5% and 2.5% of OP, respectively, for all values considered. HBP, proposed in Sec. 3.5, attains close to optimal performance for very large values of ρ/β and for $\rho/\beta \to 0$. This behavior is explained in Fig. 3.7, where it is seen that the transmission probability of HBP approaches that of OBP for these ranges of ρ. However, for ρ in the vicinity of β, HBP incurs a performance loss, which serves to illustrate the unsuitability of offline policies in a random setting. NABP performs poorly for small values of ρ (60% loss compared to OP) and approaches OBP (and

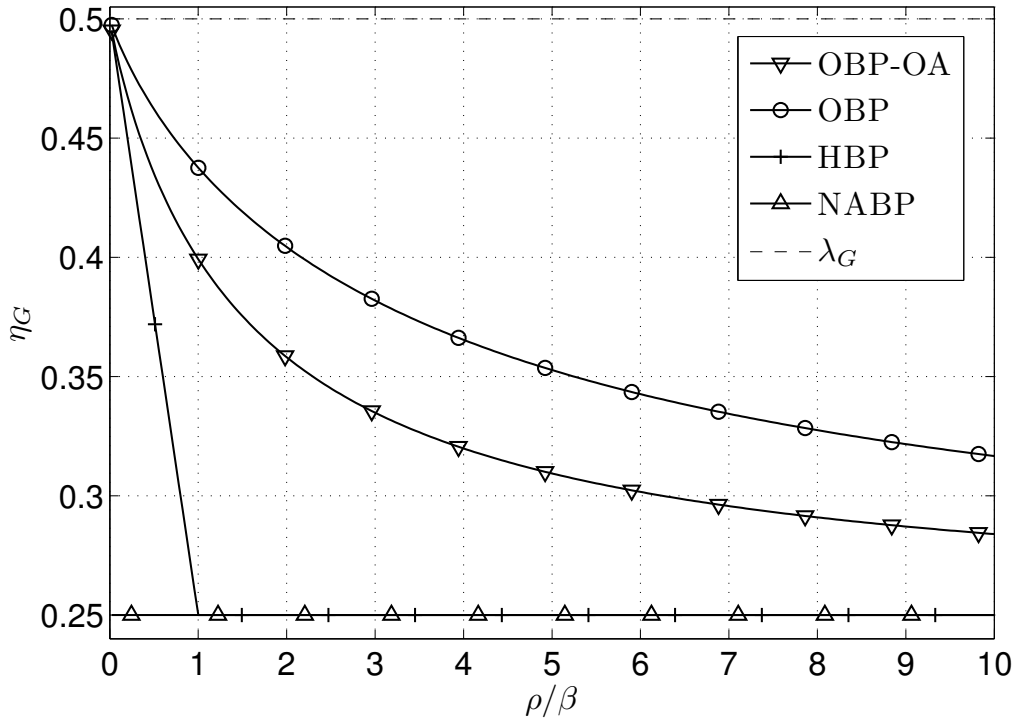


Figure 3.7. *Transmission probability in the GOOD EH state vs. ρ/β corresponding to Fig. 3.6. HBP resembles OBP for small and large ρ/β; this explains the behavior of the respective reward curves in Fig. 3.6. ($\beta = 0.25$, $\lambda_G = 0.5$, SNR$= 8$dB)*

the upper bound $g(\beta)$) for large values of ρ. The properties of the optimal transmission probability for OBP and OBP-OA, derived in Lemma 3.4.4, are confirmed in Fig. 3.7.

Overall, OBP performs so well with respect to OP because it adjusts the transmission probability in the BAD state to avoid outage and in the GOOD state to avoid overflow, which are the main factors that compromise the performance in a finite-capacity system. If they are avoided, then close-to-optimal performance can be achieved, without exact knowledge of the energy level in the battery at any given time. As the power-to-depletion ρ increases, the battery becomes more and more resilient to the randomness in the ambient energy, and the adaptation is less crucial. As shown in Figs. 3.6 and 3.7, in the limit of large ρ, it becomes optimal to transmit with constant probability β irrespective of the state of the EH process.

In Fig. 3.8, we explore the impact of the SNR Λ on the asymptotic average reward (3.25). HBP is not plotted since it is identical to NABP when $\rho = \beta$. As previously, we confirm that OBP-OA and OBP achieve near-optimal performance (within 3% and 5% of OP, respectively). A more significant degradation is incurred by NABP ($17-23\%$) and GP ($18-44\%$). Interestingly, for small SNR values,

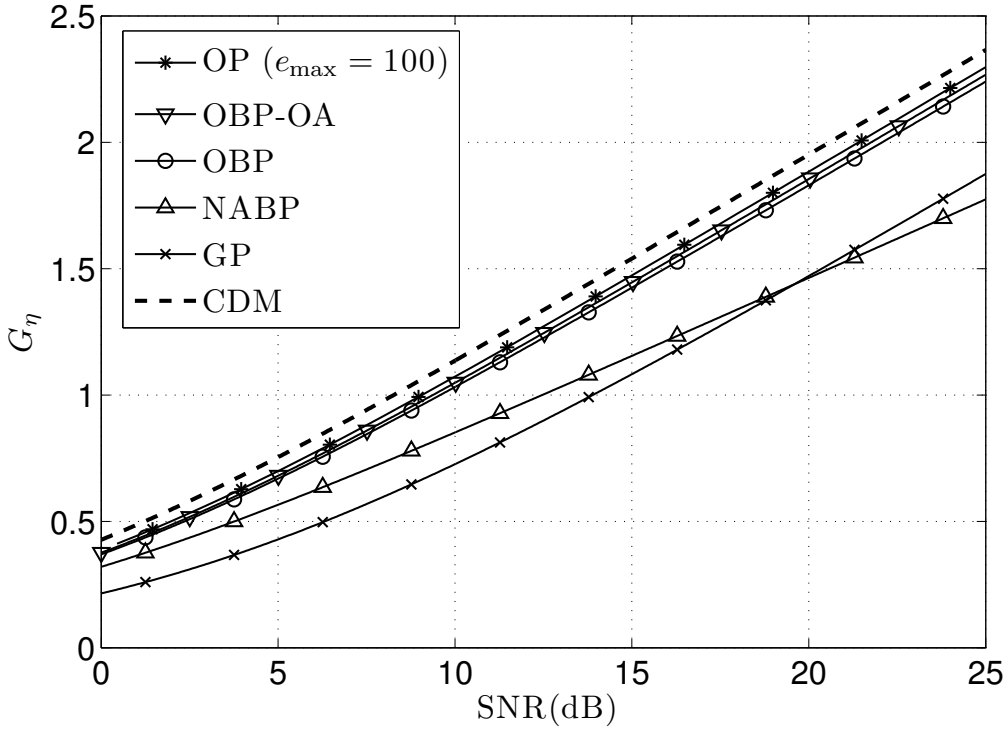


Figure 3.8. *Asymptotic average reward (3.25) as a function of the SNR* Λ. *(*$\beta = 0.25$, $\lambda_G = 0.5$, $\rho = \beta$*)*

NABP outperforms GP, and the trend is reversed for high SNR values.

Finally, in Fig. 3.9, we plot the asymptotic average reward (3.25) as a function of λ_G, setting $\rho = 0.25$ and $\gamma = 1$. Since $\beta = \lambda_G\gamma/(1+\gamma) = \lambda_G/2$, note that β also increases linearly with λ_G. For all policies, the average reward increases with λ_G, since β increases, *i.e.*, energy becomes more abundant. In agreement with all previous results, the performance loss of OBP-OA and OBP with respect to OP is small, within 2.5% and 4%, respectively, for all the considered values of λ_G. In contrast, while NABP approaches OP for small λ_G, its performance significantly degrades for large λ_G (20% of OP for $\lambda_G \to 1$), and is outperformed by GP for $\lambda_G \gtrsim 0.9$. The trend is explained by noting that, as λ_G (equivalently β) increases, the battery is driven more often to outage and overflow due to the fact that the power-to-depletion ρ is fixed, whereas the transmission probability ($\eta_G = \eta_B = \beta$ for NABP) increases.

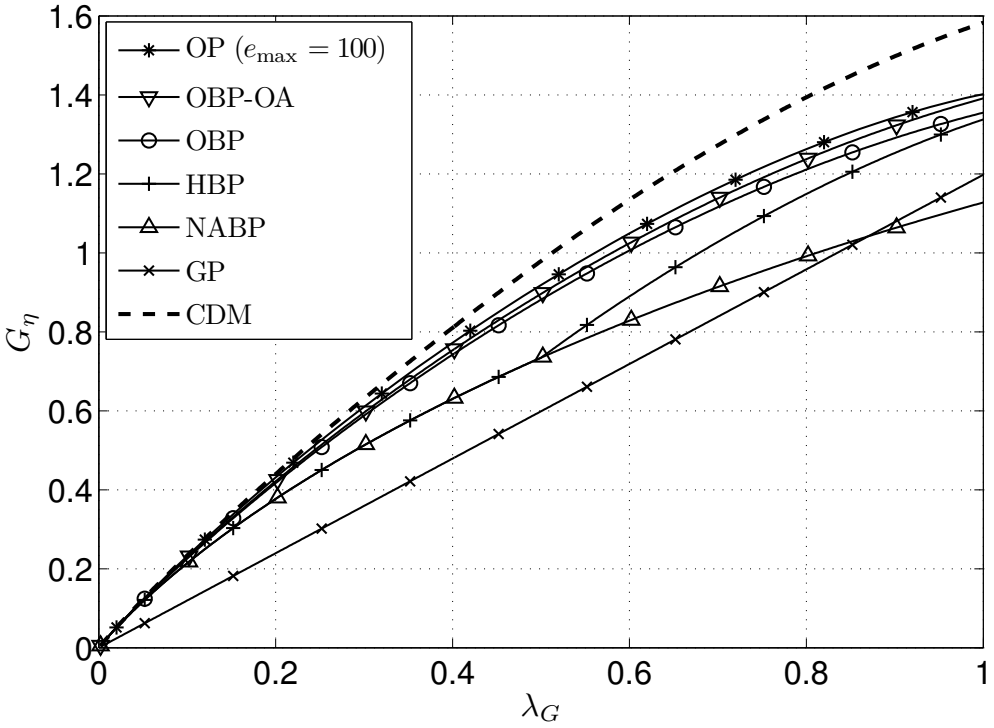


Figure 3.9. *Asymptotic average reward (3.25) vs.* λ_G*. (*$\gamma = 1$*,* $\rho = 0.25$*, SNR=* 8dB*)*

3.7 Multiaccess model: multiple EHSs

We now consider a network of U EHSs, which communicate concurrently via a shared wireless link with a FC, as depicted in Fig. 3.1. Regarding the operation of each single EHS, we refer to Sec. 3.2. In particular, we denote the following quantities related to EHS u at time k:

- $V_{u,k} \in \mathbb{R}^+$: importance of the current data packet;
- $B_{u,k} \in \{0, 1\}$: harvested energy;
- $Q_{u,k} \in \{0, 1\}$: action of the EHS controller (transmit or idle);
- $E_{u,k} \in \mathcal{E}$: energy level in the battery, governed by

$$E_{u,k+1} = \min\left\{E_{u,k} - Q_{u,k} + B_{u,k}, e_{\max}\right\}. \tag{3.41}$$

Unlike the single EHS scenario, we model the EH process as i.i.d. ($p_G = 1 - p_B$), where β denotes the probability to harvest one energy quantum in one time-slot. As a consequence, $B_{u,k}$ is independent

of the underlying EH state $A_{u,k}$, hence $A_{u,k}$ need not to be taken into account. Nevertheless, the following results can be generalized to the scenario where the EH process is a two-state Markov chain. Moreover, we define

$$\mathbf{B}_k = (B_{1,k}, B_{2,k}, \ldots, B_{U,k}) \in \{0,1\}^U, \tag{3.42}$$

$$\mathbf{E}_k = (E_{1,k}, E_{2,k}, \ldots, E_{U,k}) \in \mathcal{E}^U, \tag{3.43}$$

$$\mathbf{Q}_k = (Q_{1,k}, Q_{2,k}, \ldots, Q_{U,k}) \in \{0,1\}^U, \tag{3.44}$$

$$\mathbf{V}_k = (V_{1,k}, V_{2,k}, \ldots, V_{U,k}) \in [\mathbb{R}^+]^U. \tag{3.45}$$

The data importance $V_{u,k}$ and the EH arrival $B_{u,k}$ are assumed to be statistically independent across the EHSs and over time.

Regarding the interaction between the EHSs in the network, we assume a collision model, *i.e.*, if EHS u transmits in time-slot k, the packet is successfully delivered to FC if and only if all the other EHSs remain idle. As in the single EHS scenario, the data packet is discarded if a collision occurs or the EHS decides to remain idle.

3.8 Policy Definition and Optimization Problem

The state of the system at time k is given by $(\mathbf{E}_k, \mathbf{V}_k)$, However, each EHS is assumed to have only *local* knowledge about the state of the system. Namely, EHS u, at time k, only knows its own energy level and data importance $(E_{u,k}, V_{u,k})$, but does not know the energy level and data importance of the other EHSs in the network. As a result, the decision of EHS u on whether to transmit or remain idle is based solely on $(E_{u,k}, V_{u,k})$. In particular, as proved for the single EHS scenario (Lemma 3.3.1), the following threshold policy is optimal:

$$Q_{u,k} = \begin{cases} 1, & V_{u,k} \geq v_{\mathrm{th},u}(E_{u,k}), \\ 0, & V_{u,k} < v_{\mathrm{th},u}(E_{u,k}), \end{cases} \tag{3.46}$$

where $v_{\mathrm{th},u}(e)$ is some importance threshold, and is a function of the energy level e. As in the single EHS scenario, we denote by $\eta_u(e)$ the corresponding transmission probability of EHS u in energy level e, induced by the random importance $V_{u,k}$, and by $g(\eta_u(e))$ the expected data importance reported by EHS u to FC in state e, assuming that all the other EHSs remain idle (no collisions occur).

In the following, we refer to η_u as the policy of EHS u. Moreover, we denote the aggregate policy used by all the EHSs in the network as $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \mu_U)$.

Given an initial state of the energy levels $\mathbf{E}_0 = \mathbf{e}_0 \in \mathcal{E}^U$, we denote the average long-term importance of the data reported by EHS u to FC, under the aggregate policy $\boldsymbol{\eta}$, as

$$R_{\boldsymbol{\eta}}^{(u)}(\mathbf{e}_0) = \liminf_{K\to\infty} \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} Q_{u,k}V_{u,k}\prod_{i\neq u}(1-Q_{i,k}) \,\middle|\, \mathbf{e}_0\right]$$
$$= \liminf_{K\to\infty} \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} g(\eta_u(E_{u,k}))\prod_{i\neq u}(1-\eta_i(E_{i,k})) \,\middle|\, \mathbf{e}_0\right]. \tag{3.47}$$

The expectations above are taken with respect to $\{\mathbf{B}_k, \mathbf{Q}_k, \mathbf{V}_k\}$ where, at each instant k, $Q_{i,k}$ is given by (3.46) for appropriate threshold $v_{\mathrm{th},i}(E_{i,k})$, and $E_{i,k}$ evolves according to (3.41). In the last step, we have used the fact that $Q_{i,k}$ only depends on $(E_{i,k}, V_{i,k})$, and $V_{i,k}$ is i.i.d. across the EHSs, hence

$$\mathbb{E}\left[Q_{u,k}V_{u,k}\prod_{i\neq u}(1-Q_{i,k}) \,\middle|\, \mathbf{E}_k\right] = \mathbb{E}\left[Q_{u,k}V_{u,k}|E_{u,k}\right]\prod_{i\neq u}\left(1-\mathbb{E}\left[Q_{i,k}|E_{i,k}\right]\right)$$
$$= g(\eta_u(E_{u,k}))\prod_{i\neq u}(1-\eta_i(E_{i,k})).$$

The term $Q_{u,k}\prod_{i\neq u}(1-Q_{i,k}) = 1$ if and only if EHS u transmits the current data packet, and all the other EHSs remain idle, so that no collision occurs and the transmission is successful. Moreover, we define the average long-term aggregate importance of the reported data (from now on referred to as *network utility* for brevity) as

$$R_{\boldsymbol{\eta}}(\mathbf{e}_0) = \sum_{u=1}^{U} R_{\boldsymbol{\eta}}^{(u)}(\mathbf{e}_0). \tag{3.48}$$

The objective is to design control policies $\boldsymbol{\eta}$ which maximize the network utility, *i.e.*,

$$\boldsymbol{\eta}^* = \arg\max_{\boldsymbol{\eta}} R_{\boldsymbol{\eta}}(\mathbf{e}_0). \tag{3.49}$$

However, in order to guarantee fairness among the EHSs in the network, we consider only symmetric control policies, *i.e.*, all the EHSs employ the same policy $\eta_u = \eta,\ \forall u$. The optimization in (3.49) is then restricted to such symmetric policies, yielding

$$\eta^* = \arg\max_{\eta} R_{(\eta,\eta,\ldots,\eta)}(\mathbf{e}_0). \tag{3.50}$$

The optimization in (3.50) is carried out in the next section.

It can be shown that, since $g(x)$ is strictly concave, the optimal policy $\boldsymbol{\eta}^*$ is unique and belongs to the set of *admissible* policies $\mathcal{U}$ that result in an average reward independent of the initial state $\mathbf{e}_0$, as defined below.

Definition 3.8.1. The set $\mathcal{U}$ of admissible policies is defined as

$$\mathcal{U} = \{\eta : \eta(0) = 0, \eta(e_{\max}) \in (0, 1], \eta(e) \in (0, 1), e \neq 0, e_{\max}\}.$$

It can be shown that the Markov chain $\{\mathbf{E}_k\}$ under the aggregate policy $\boldsymbol{\eta} \in \mathcal{U}^U$ is irreducible. Hence, there exists a unique steady-state distribution, $\pi_{\boldsymbol{\eta}}(\mathbf{e}), \mathbf{e} \in \mathcal{E}^U$, independent of $\mathbf{e}_0$ [18]. From (3.47), we thus obtain

$$R_{\boldsymbol{\eta}}^{(u)} = \sum_{\mathbf{e} \in \mathcal{E}^U} \pi_{\boldsymbol{\eta}}(\mathbf{e}) g(\eta_u(e_u)) \prod_{i \neq u} (1 - \eta_i(e_i)). \tag{3.51}$$

Moreover, since the action $Q_{u,k}$ is based only on $(E_{u,k}, V_{u,k})$ and does not depend on $(E_{i,k}, V_{i,k}),\ i \neq u$, and harvesting is i.i.d. across EHSs, in the steady state regime, the energy level of EHS u is independent of the energy levels of all the other EHSs, so that we can write $\pi_{\boldsymbol{\eta}}(\mathbf{e}) = \prod_u \pi_{\eta_u}(e_u)$, where $\pi_{\eta_u}(e_u)$ is the steady state distribution of the energy level of EHS u, $\{E_{u,k}\}$, which is characterized in the following lemma.

Lemma 3.8.1. *The steady state distribution of the energy level $E_{u,k}$ under policy $\eta_u \in \mathcal{U}$ is given by*

$$\pi_{\eta_u}(e) = \prod_{i=0}^{e-1} W_{\eta_u}(i) \pi_{\eta_u}(0) = \frac{1}{\prod_{i=e}^{e_{\max}-1} W_{\eta_u}(i)} \pi_{\eta_u}(e_{\max}), \tag{3.52}$$

where we have defined

$$W_{\eta_u}(i) = \frac{\beta \bar{\eta}_u(i)}{\bar{\beta} \eta_u(i+1)},\ i = 0, 1, \ldots, e_{\max} - 1, \tag{3.53}$$

and

$$\pi_{\eta_u}(0) = \frac{1}{1 + \sum_{e=0}^{e_{\max}-1} \prod_{i=0}^{e} W_{\eta_u}(i)}. \tag{3.54}$$

Proof. With the help of Fig. 3.10, the balance equation $\pi_{\eta_u}(e-1)\beta\bar{\eta}_u(e-1) = \pi_{\eta_u}(e)\bar{\beta}\eta_u(e)$, for

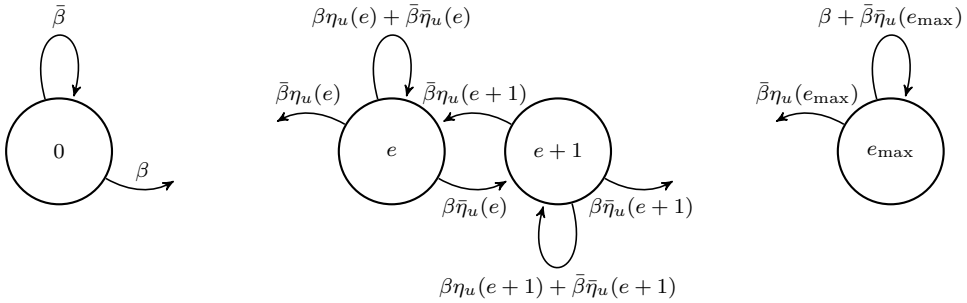


Figure 3.10. *Markov chain and transition probabilities of energy level $E_{u,k}$*

$1 \leq e \leq e_{\max}$, yields

$$\pi_{\eta_u}(e) = W_{\eta_u}(e-1)\pi_{\eta_u}(e-1). \tag{3.55}$$

The expression (3.52) is then obtained by induction, and (3.54) after normalization. □

Letting

$$G(\eta_u) = \sum_{e=1}^{e_{\max}} \pi_{\eta_u}(e) g(\eta_u(e)), \qquad P(\eta_i) = \sum_{e=1}^{e_{\max}} \pi_{\eta_i}(e)\eta_i(e), \tag{3.56}$$

we can rewrite (3.51) as

$$R_{\boldsymbol{\eta}}^{(u)} = G(\eta_u) \prod_{i \neq u} (1 - P(\eta_i)). \tag{3.57}$$

Eq. (3.57) can be interpreted as follows. $G(\eta_u)$ is the average reward of EHS u, assuming that all the other EHSs remain idle, so that no collisions occur. $P(\eta_i)$ is the average long-term transmission probability of EHS i, so that $\prod_{i \neq u}(1 - P(\eta_i))$ is the steady-state probability that all the EHSs, except u, remain idle. From (3.48), the network utility under the aggregate policy $\boldsymbol{\eta}$ then becomes

$$R_{\boldsymbol{\eta}} = \sum_{u=1}^{U} G(\eta_u) \prod_{i \neq u} (1 - P(\eta_i)). \tag{3.58}$$

In the symmetric scenario with $\eta_u = \eta, \ \forall u$, which is the main focus of this work, (3.58) becomes

$$R_{\boldsymbol{\eta}} = UG(\eta)(1 - P(\eta))^{U-1}. \tag{3.59}$$

The optimization problem (3.50) over the class of admissible and symmetric policies is stated as

$$\eta^* = \arg\max_{\eta \in \mathcal{U}} UG(\eta)(1 - P(\eta))^{U-1}, \tag{3.60}$$

and is carried out in the next section.

3.9 Optimization and Analysis

The optimization problem (3.60) when $U = 1$ can be solved by using the *Policy Iteration Algorithm* (PIA) [1] (Algorithm 3 with $\lambda = 0$ in this section). However, in general, when $U > 1$ (3.60) cannot be recast as a convex optimization problem, hence we resort to approximate solutions. In particular, in order to determine a local optimum of (3.60), we use a mathematical artifice based on a game theoretic formulation of the multiaccess problem considered in this work: we model the optimization problem as a game, where it is assumed that each EHS, say u, is a player which attempts to maximize the common payoff (3.58) with respect to its own policy η_u.[5] We proceed as follows. We first characterize the general Nash Equilibrium (NE). Then, we study the existence of the *Symmetric NE* (SNE) for this game, *i.e.*, such that all EHSs employ the same policy $\eta_u^* = \eta^*,\ \forall u$, and have no incentive to deviate from it. In Theorem 3.9.2, we show that the SNE is unique, and we also provide Algorithm 2 to compute it. In Theorem 3.9.3, we prove that the SNE, and thus the policy returned by Algorithm 2, represents a local optimum of the original optimization problem (3.60).

If a NE exists for this game (not necessarily symmetric), defined by the policy profile $\boldsymbol{\eta}^* = (\eta_1^*, \eta_2^*, \ldots, \eta_U^*)$, then it solves

$$\begin{aligned}\eta_u^* &= \arg\max_{\eta_u \in \mathcal{U}} \left[G(\eta_u) \prod_{i \neq u} (1 - P(\eta_i^*)) + (1 - P(\eta_u)) \sum_{n \neq u} G(\eta_n^*) \prod_{i \neq n,u} (1 - P(\eta_i^*)) \right] \\ &= \arg\max_{\eta_u \in \mathcal{U}} \left[G(\eta_u) - P(\eta_u) \sum_{n \neq u} \frac{G(\eta_n^*)}{1 - P(\eta_n^*)} \right],\ \forall u,\end{aligned} \tag{3.61}$$

where, in the last step, we have removed positive multiplicative factors and additive terms independent of η_u, which do not affect the optimization problem. In particular, we are interested in characterizing

[5] We point out that this formulation is only a mathematical artifice to determine the optimal policy, which is then followed by all EHSs (which are not assumed to behave strategically).

the SNE. Then, by further imposing $\eta_u^* = \eta^*,\ \forall u$, in (3.61), we obtain

$$\eta^* = \arg\max_{\eta \in \mathcal{U}} \left[G(\eta) - \Lambda(\eta^*) P(\eta) \right], \tag{3.62}$$

where we have defined

$$\Lambda(\eta) = (U-1)\frac{G(\eta)}{1-P(\eta)}. \tag{3.63}$$

Note that η^* defined in (3.62) is *simultaneously* optimal for all the EHSs, *i.e.*, any unilateral deviation of a single EHS from the SNE η^* yields a smaller network utility $R_{\boldsymbol{\eta}}$. The interpretation of (3.62) is as follows. $G(\eta)$ is the reward when the network contains only one user, so that the unique EHS has no constraint on the collisions caused to other users in the network. The term $\Lambda(\eta^*)$ is interpreted as a Lagrange multiplier constant associated to a constraint on the transmission probability of each EHS, so as to limit the collisions to the other EHSs in the network. The overall objective function is thus interpreted as the maximization of the individual reward of each user, with constraint on the average transmission probability to limit collisions, which are deleterious to network performance. Interestingly, the Lagrange multiplier (3.63) increases with the number of EHSs U, so that, the larger the network size, the more stringent the constraint on the average transmission probability of each EHS. In order to carry out (3.62), we solve the more general optimization problem, for $\lambda \geq 0$,

$$\eta^{(\lambda)} = \arg\max_{\eta \in \mathcal{U}} \left[G(\eta) - \lambda P(\eta) \right]. \tag{3.64}$$

The following properties of $\eta^{(\lambda)}$ can be proved, which follow from the fact that $g(x)$ is a strictly concave function of x (other properties are provided in Theorem 3.10.1):

Proposition 3.9.1. *1.* $\eta^{(\lambda)}$ *is uniquely defined,* i.e.*,*

$$G(\eta^{(\lambda)}) - \lambda P(\eta^{(\lambda)}) > G(\eta) - \lambda P(\eta),\ \forall \eta \neq \eta^{(\lambda)}; \tag{3.65}$$

2. $\eta^{(\lambda)}$ *is continuous in* λ*;*

3. $\eta^{(\lambda)} \in \mathrm{int}(\mathcal{U})$*, where* $\mathrm{int}(\mathcal{U})$ *denotes the interior of* $\mathcal{U}$*;*

4. $0 < P(\eta^{(\lambda)}) \leq \beta$*,* $0 < G(\eta^{(\lambda)}) < g(P(\eta^{(\lambda)})) \leq g(\beta)$.

Remark 3.9.1. The first property is a consequence of the fact that (3.64) can be recast as a convex optimization problem, where the objective function is *strictly* concave. The second property follows from the strict concavity of $g(x)$. The third property follows from the fact that $\eta^{(\lambda)}(e) \in (0,1),\ \forall e \neq 0$, as proved in Theorem 3.10.1. The last property is a consequence of the fact that, for any policy $\eta \in \mathcal{U}$, the average long-term transmission probability cannot be larger than the average harvesting rate β, since the transmission of one packet requires the expenditure of one energy quantum. Moreover, from the strict concavity of $g(x)$, using Jensen's inequality we obtain $G(\eta) < g(P(\eta)) \leq g(\beta)$, since $g(x)$ is an increasing function of x and $P(\eta) \leq \beta$.

By comparing (3.62) and (3.64), we then obtain that η^* is optimal for (3.62) if and only if $\eta^* = \eta^{(\lambda^*)}$, for $\lambda^* \geq 0$ and $\Lambda(\eta^{(\lambda^*)}) = \lambda^*$. The following theorem proves the existence and uniqueness of such η^*.

Theorem 3.9.2. *There exists a unique $\eta^* \in \mathcal{U}$ solution of (3.62),* i.e.*, $\exists!\ \eta^* \in \mathcal{U}$ such that*

$$G(\eta^*) - \Lambda(\eta^*)P(\eta^*) > G(\eta) - \Lambda(\eta^*)P(\eta),\ \forall \eta \neq \eta^*, \eta \in \mathcal{U}.$$

Moreover, $P(\eta^*) \leq \min\{\beta, \frac{1}{U}\}$.

Proof. The existence and uniqueness of η^* solution of (3.62) is proved by using Lemma 3.12.2 in Appendix 3.G. In fact, $h(\lambda) = \Lambda(\eta^{(\lambda)}) - \lambda$ is a continuous decreasing function of λ (since $\Lambda(\eta^{(\lambda)})$ is continuous non-increasing), with limits $h(0) = \Lambda(\eta^{(0)}) > 0$ and $\lim_{\lambda\to\infty} h(\lambda) = -\infty$, hence there exists a unique $\lambda^* \in (0, +\infty)$ such that $h(\lambda^*) = 0$, *i.e.* $\Lambda(\eta^{(\lambda^*)}) = \lambda^*$, which guarantees that $\eta^{(\lambda^*)}$ is optimal for (3.62).

We now prove that $P(\eta^{(\lambda^*)}) \leq \min\{\beta, \frac{1}{U}\}$. From Prop. 3.9.1, we have $P(\eta^{(\lambda^*)}) \leq \beta$, hence it is sufficient to prove that $P(\eta^{(\lambda^*)}) \leq \frac{1}{U}$. This is trivially true if $P(\eta^{(0)}) \leq \frac{1}{U}$, since $P(\eta^{(\lambda)})$ is a non-increasing function of λ (Lemma 3.12.1). Now, assume that $P(\eta^{(0)}) > \frac{1}{U}$. Then, since $\lim_{\lambda\to\infty} P(\eta^{(\lambda)}) = 0$, there exists $\hat{\lambda} \in (0, \infty)$ such that $P(\eta^{(\hat{\lambda})}) = \frac{1}{U}$. For such $\hat{\lambda}$, from (3.64) we have

$$G(\eta^{(\hat{\lambda})}) - \hat{\lambda}\frac{1}{U} = G(\eta^{(\hat{\lambda})}) - \hat{\lambda}P(\eta^{(\hat{\lambda})}) = \max_{\eta\in\mathcal{U}} G(\eta) - \hat{\lambda}P(\eta) > G(0) - \hat{\lambda}P(0) = 0, \qquad (3.66)$$

and, using (3.63) and the fact that $P(\eta^{(\hat{\lambda})}) = \frac{1}{U}$,

$$UG(\eta^{(\hat{\lambda})}) - \hat{\lambda} = (U-1)\frac{G(\eta^{(\hat{\lambda})})}{1-1/U} - \hat{\lambda} = \Lambda(\eta^{(\hat{\lambda})}) - \hat{\lambda} > 0.$$

Therefore, we obtain $\Lambda(\eta^{(\hat{\lambda})}) > \hat{\lambda}$. Since $\Lambda(\eta^{(\lambda)}) - \lambda$ is a decreasing function of λ (Lemma 3.12.2) and $\Lambda(\eta^{(\lambda^*)}) - \lambda^* = 0$, necessarily $\hat{\lambda} < \lambda^*$. Finally, using Lemma 3.12.1, we obtain $P(\eta^{(\hat{\lambda})}) = \frac{1}{U} \geq P(\eta^{(\lambda^*)})$, since $P(\eta^{(\lambda)})$ is a non-increasing function of λ. The second part of the theorem is thus proved. □

We have the following result.

Theorem 3.9.3. *The SNE η^* in (3.62) is a* local *optimum for the optimization problem (3.60).*

Proof. Since η^* is globally optimal for the optimization problem (3.64) when $\lambda = \Lambda(\eta^*)$, and $\eta^* \in \mathrm{int}(\mathcal{U})$ from Prop. 3.9.1, then the gradient with respect to η, $\Delta_\eta(\cdot)$, of the objective function in (3.62), computed in η^*, is equal to zero, and its Hessian with respect to η, $\mathbf{H}_\eta(\cdot)$, computed in η^*, is semidefinite negative. More precisely, since $g(x)$ is a strictly concave function of x, it can be proved that the Hessian of the objective function in (3.62), computed in η^*, is *negative* definite, *i.e.*, for the SNE η^* we have

$$[\Delta_\eta(G(\eta)) - \Lambda(\eta^*)\Delta_\eta(P(\eta))]_{\eta=\eta^*} = \mathbf{0}, \tag{3.67}$$

$$[\mathbf{H}_\eta(G(\eta)) - \Lambda(\eta^*)\mathbf{H}_\eta(P(\eta))]_{\eta=\eta^*} \prec 0. \tag{3.68}$$

On the other hand, the gradient of (3.59) is given by

$$\Delta_\eta\left(R_{\boldsymbol{\eta}}\right) = U(1-P(\eta))^{U-1}\Delta_\eta\left(G(\eta)\right) - U(U-1)G(\eta)(1-P(\eta))^{U-2}\Delta_\eta\left(P(\eta)\right). \tag{3.69}$$

The Hessian matrix of (3.59) is then obtained by further computing the gradient of each component of (3.69), yielding

$$\begin{aligned}\mathbf{H}_\eta\left(R_{\boldsymbol{\eta}}\right) &= U(U-1)(U-2)G(\eta)(1-P(\eta))^{U-3}\Delta_\eta\left(P(\eta)\right)\Delta_\eta(P(\eta))^T \\ &- U(U-1)(1-P(\eta))^{U-2}\Delta_\eta\left(P(\eta)\right)\Delta_\eta(G(\eta))^T - U(U-1)G(\eta)(1-P(\eta))^{U-2}\mathbf{H}_\eta\left(P(\eta)\right) \\ &- U(U-1)(1-P(\eta))^{U-2}\Delta_\eta\left(G(\eta)\right)\Delta_\eta(P(\eta))^T + U(1-P(\eta))^{U-1}\mathbf{H}_\eta\left(G(\eta)\right).\end{aligned} \tag{3.70}$$

By computing (3.69) under the SNE η^*, and by using (3.63) and substituting (3.67) in (3.69), we then obtain $[\Delta_\eta(R_{\boldsymbol{\eta}})]_{\eta=\eta^*} = \mathbf{0}$. Moreover, since $[\mathbf{H}_\eta(G(\eta))]_{\eta=\eta^*} \prec \Lambda(\eta^*)\,[\mathrm{H}_\eta(P(\eta))]_{\eta=\eta^*}$ from (3.68), substituting (3.68) in (3.70) we obtain

$$[\mathbf{H}_\eta(R_{\boldsymbol{\eta}})]_{\eta=\eta^*} \prec -U^2(U-1)G(\eta^*)(1-P(\eta))^{U-3}\left[\Delta_\eta(P(\eta))\Delta(P(\eta))^T\right]_{\eta=\eta^*} \preceq 0,$$

where we have used the fact that the product of the column vector $\Delta_\eta(P(\eta))$ by its transpose is semidefinite positive. Therefore, $[\mathbf{H}_\eta(R_{\boldsymbol{\eta}})]_{\eta=\eta^*} \prec 0$ and $[\Delta_\eta(R_{\boldsymbol{\eta}})]_{\eta=\eta^*} = \mathbf{0}$, hence η^* is a local optimum for (3.60). □

To conclude, we present an algorithm to determine the SNE η^* in (3.62), hence, from Theorem 3.9.3, a local optimum of (3.60). In particular, letting $h(\lambda) = \Lambda(\eta^\lambda) - \lambda$, we employ the bisection method [61] to compute the unique λ^* such that $h(\lambda^*) = 0$, which determines the SNE η^* as $\eta^* = \eta^{(\lambda^*)}$. We use the fact that, from Lemma 3.12.2 in Appendix 3.G, $h(\lambda)$ is a continuous decreasing function of λ, with $h(0) > 0$ and $\lim_{\lambda\to\infty} = -\infty$, so that, if $h(\lambda) > 0$ (respectively, $h(\lambda) < 0$) for some λ, then necessarily $\lambda < \lambda^*$ ($\lambda > \lambda^*$). We need upper and lower bounds to λ^*, denoted as $\lambda_{\max}$ and $\lambda_{\min}$, respectively, so that $\lambda_{\min} \le \lambda^* \le \lambda_{\max}$. These bounds are then iteratively updated and refined, by testing the sign of $h(\lambda)$ for the new $\lambda = (\lambda_{\min} + \lambda_{\max})/2$, until the desired accuracy is attained. The initialization of the lower bound is chosen as $\lambda_{\min} = 0$. As to the upper bound, note that $P(\eta^*) \le \min\{\beta, \frac{1}{U}\}$ from Prop. 3.9.1, hence $G(\eta^*) < g(P(\eta^*)) \le g(\min\{\beta, \frac{1}{U}\})$. Therefore, from (3.63) we obtain

$$\lambda^* = \Lambda(\eta^{(\lambda^*)}) < \min\left\{\frac{U-1}{1-\beta}g(\beta), Ug\left(\frac{1}{U}\right)\right\} = \lambda_{\max}. \tag{3.71}$$

Algorithm 2 (Bisection method)**.**

1. `INIT`: accuracy $\epsilon > 0$, $\lambda_{\min} = 0$ and $\lambda_{\max}$ as in (3.71);
2. `MAIN`: $\lambda := (\lambda_{\min} + \lambda_{\max})/2$; `DETERMINE` $\eta^{(\lambda)}$ using the PIA (Algorithm 3); `COMPUTE` $h(\lambda) = \Lambda(\eta^{(\lambda)}) - \lambda$.
   - `IF` $|h(\lambda)| < \epsilon$, `RETURN` the optimal policy $\eta^{(\lambda)}$;
   - `IF` $h(\lambda) > \epsilon$, `UPDATE` $\lambda_{\min} := \lambda$ and $\lambda_{\max} := \min\{\lambda_{\max}, \Lambda(\eta^{(\lambda)})\}$; `REPEAT MAIN`;
   - `IF` $h(\lambda) < -\epsilon$, `UPDATE` $\lambda_{\max} := \lambda$ and $\lambda_{\min} := \max\{\lambda_{\min}, \Lambda(\eta^{(\lambda)})\}$; `REPEAT MAIN`;

3. `RETURN` optimal policy $\eta^* = \eta^{(\lambda)}$.

Remark 3.9.2. Note that the `UPDATE` step updates both $\lambda_{\min}$ and $\lambda_{\max}$. This is because $h(\lambda)$ is a decreasing function of λ and $\Lambda(\eta^{(\lambda)})$ is a non-increasing function of λ (Lemma 3.12.2), hence, if $h(\lambda) > 0$, then $\lambda < \lambda^* = \Lambda^{(\eta^{(\lambda^*)})} \leq \Lambda^{(\eta^{(\lambda)})}$, and therefore $\lambda < \lambda^* \leq \min\{\lambda_{\max}, \Lambda^{(\eta^{(\lambda)})}\}$. Similarly, if $h(\lambda) < 0$, then $\lambda > \lambda^* = \Lambda^{(\eta^{(\lambda^*)})} \geq \Lambda^{(\eta^{(\lambda)})}$, and therefore $\lambda > \lambda^* \geq \max\{\lambda_{\min}, \Lambda^{(\eta^{(\lambda)})}\}$.

We now present the PIA [1] to compute $\eta^{(\lambda)}$. Starting from an initial policy η, this algorithm iteratively computes the *Policy Evaluation* and *Policy Improvement* steps, until convergence. In the *Policy Evaluation* step, the metric $G(\eta) - \lambda P(\eta)$ is computed under the current policy η, and the *relative value function* $v_\eta : \mathcal{E} \mapsto \mathbb{R}$ is determined as the unique solution of

$$v_\eta(0) = 0, \tag{3.72}$$

$$v_\eta(e) - \sum_{j \in \mathcal{E}} \Pr_\eta(E_{u,k+1} = j | E_{u,k} = e) v_\eta(j) = g(\eta(e)) - \lambda\eta(e) - G(\eta) + \lambda P(\eta), \ \forall e \in \mathcal{E}. \tag{3.73}$$

We have the following lemma.

Lemma 3.9.4. *The relative value function under policy η is given by*

$$v_\eta(e) = \sum_{j=0}^{e-1} \frac{\sum_{i=0}^{j} \pi_\eta(i)[G(\eta) - \lambda P(\eta) - g(\eta(i)) + \lambda\eta(i)]}{\pi_\eta(j)\beta\bar{\eta}(j)}. \tag{3.74}$$

Proof. Trivially, (3.74) is consistent with (3.72) for $e = 0$. Moreover, from (3.73) for $e = 0$, using the fact that $\eta(0) = 0$ and $v_\eta(0) = 0$, we obtain

$$v_\eta(1) = \frac{G(\eta) - \lambda P(\eta)}{\beta}, \tag{3.75}$$

yielding (3.74) for $e = 1$. From (3.73) for $e = 1, 2, \ldots, e_{\max} - 1$, using the expression of the transition probabilities of $E_{u,k}$ (see Fig. 3.10), we have

$$\begin{aligned} &v_\eta(e) - \bar{\beta}\eta(e)v_\eta(e-1) - (\beta\eta(e) + \bar{\beta}\bar{\eta}(e))v_\eta(e) - \beta\bar{\eta}(e)v_\eta(e+1) \\ =&g(\eta(e)) - \lambda\eta(e) - G(\eta) + \lambda P(\eta). \end{aligned} \tag{3.76}$$

By solving the above equation with respect to $v_\eta(e+1)$, we obtain

$$v_\eta(e+1)=\frac{-g(\eta(e))+\lambda\eta(e)+G(\eta)-\lambda P(\eta)-\bar{\beta}\eta(e)v_\eta(e-1)+(\beta\bar{\eta}(e)+\bar{\beta}\eta(e))v_\eta(e)}{\beta\bar{\eta}(e)}. \quad (3.77)$$

Finally, the expression (3.74) for $e>1$ is proved by induction: assuming that $v_\eta(e)$ and $v_\eta(e-1)$ are given by the expression (3.74), and substituting these expressions in (3.77), it can be proved, using the expression of the steady state distribution $\pi_\eta(\cdot)$, given in Lemma 3.8.1, that $v_\eta(e+1)$ is also given by the expression (3.74), thus proving the induction step. Since we have proved that (3.74) holds for $e=0$ and $e=1$, the induction is proved. □

In the *Policy Improvement* step, an improved policy $\hat{\eta}$ is determined by solving, for each $e \in \mathcal{E}\setminus\{0\}$, the convex optimization problem

$$\hat{\eta}(e)=\arg\max_{\tilde{\eta}(e)\in[0,1]} g(\tilde{\eta}(e))-\lambda\tilde{\eta}(e)+\sum_{j\in\mathcal{E}}\Pr_{\tilde{\eta}}(E_{u,k+1}=j|E_{u,k}=e)v_\eta(j), \quad (3.78)$$

or equivalently, using the expression of the transition probabilities of $E_{u,k}$, we obtain

$$\hat{\eta}(e)=\arg\max_{x\in[0,1]} g(x)-\lambda x-\bar{\beta}x\delta_\eta(e)-\beta x\delta_\eta(e+1),\ \forall e\in\mathcal{E}\setminus\{0\}, \quad (3.79)$$

where we have defined $\delta_\eta(e_{\max}+1)=0$ and, for $e\in\mathcal{E}\setminus\{e_{\max}\}$,

$$\delta_\eta(e)=v_\eta(e)-v_\eta(e-1)=\frac{\sum_{i=0}^{e-1}\pi_\eta(i)[G(\eta)-\lambda P(\eta)-g(\eta(i))+\lambda\eta(i)]}{\pi_\eta(e-1)\beta\bar{\eta}(e-1)}. \quad (3.80)$$

Owing to the concavity of $g(x)$, the optimal $\hat{\eta}(e)$ can be found by using the bisection method [61]. We are now ready to state the PIA to solve the optimization problem (3.64).

Algorithm 3 (PIA). 1. `INIT`: $\eta_{\text{old}}\in\mathcal{U}$; accuracy ϵ_{PIA};

2. `Policy evaluation`: compute, using Lemma 3.8.1 and (3.56),

$$\delta_{\eta_{\text{old}}}(e)=\frac{\sum_{i=0}^{e-1}\pi_{\eta_{\text{old}}}(i)[G(\eta_{\text{old}})-\lambda P(\eta_{\text{old}})-g(\eta_{\text{old}}(i))+\lambda\eta_{\text{old}}(i)]}{\pi_{\eta_{\text{old}}}(e-1)\beta\bar{\eta}_{\text{old}}(e-1)},\ e=1,2,\ldots,e_{\max}+1;$$

3. `Policy improvement`: determine a new policy, η_{new}, as the solution of

$$\eta_{\text{new}}(e)=\arg\max_{x\in[0,1]} g(x)-\lambda x-\bar{\beta}x\delta_\eta(e)-\beta x\delta_{\eta_{\text{old}}}(e+1),\ \forall e\in\mathcal{E}\setminus\{0\}; \quad (3.81)$$

4. `Termination test`: If $|G(\eta_{\text{new}}) - G(\eta_{\text{old}})| < \epsilon_{\text{PIA}}$, return the policy $\eta^{(\lambda)} = \eta_{\text{new}}$; otherwise, let $\eta_{\text{old}} := \eta_{\text{new}}$ and repeat from step 2).

3.10 Low Complexity Policies

Note that the PIA may be computationally intensive for the ultra-low power electronics typically found in practical EHSs. In this section, we design low-complexity policies, which are then shown to achieve near-optimal performance. In particular, we use two different approaches. The first approach, developed in Sec. 3.10.1, is based on replacing the PIA in the main iteration stage of Algorithm 2, which is used to compute the optimal policy $\eta^{(\lambda)}$ maximizing $G(\eta) - \lambda P(\eta)$, with a computationally less intensive algorithm, which determines an approximation $\tilde{\eta}^{(\lambda)}$ of the optimal policy $\eta^{(\lambda)}$. The second approach, developed in Sec. 3.10.2, is based on the approximation of large battery capacity.

3.10.1 Low complexity approximation of $\eta^{(\lambda)}$

The following theorem presents structural properties of the optimal policy $\eta^{(\lambda)}$, which are then used to design a low-complexity policy $\tilde{\eta}^{(\lambda)}$.

Theorem 3.10.1. *$\eta^{(\lambda)}$ has the following properties:*

P1) *$\eta^{(\lambda)}$ is a strictly increasing function of $e \in \mathcal{E}$.*

P2) *$\eta^{(\lambda)} \in (\eta_{\text{low}}, \eta_{\text{up}}), \forall e \in \mathcal{E}\backslash\{0\}$, where, letting $x^* = \arg\max_x [g(x) - \lambda x]$, $\eta_{\text{low}} \in (0, \min\{\beta, x^*\})$, $\eta_{\text{up}} \in (\min\{\beta, x^*\}, x^*)$ uniquely solve*

$$g(\eta_{\text{low}}) + (1 - \eta_{\text{low}}) g'(\eta_{\text{low}}) = \lambda + \frac{g(\min\{\beta, x^*\}) - \lambda \min\{\beta, x^*\}}{\beta}, \tag{3.82}$$

$$g(\eta_{\text{up}}) - \eta_{\text{up}} g'(\eta_{\text{up}}) = g(\min\{\beta, x^*\}) - \lambda \min\{\beta, x^*\}. \tag{3.83}$$

Proof. See Appendix 3.H. □

Remark 3.10.1. P1) of Theorem 3.10.1 states the intuitive fact that the more energy available in the buffer, the higher the incentive to transmit. The larger transmission probability in the high energy states reflects the incentive to minimize the impact of energy overflow. In contrast, the smaller transmission probability in the low energy states aims to minimize the impact of energy outage.

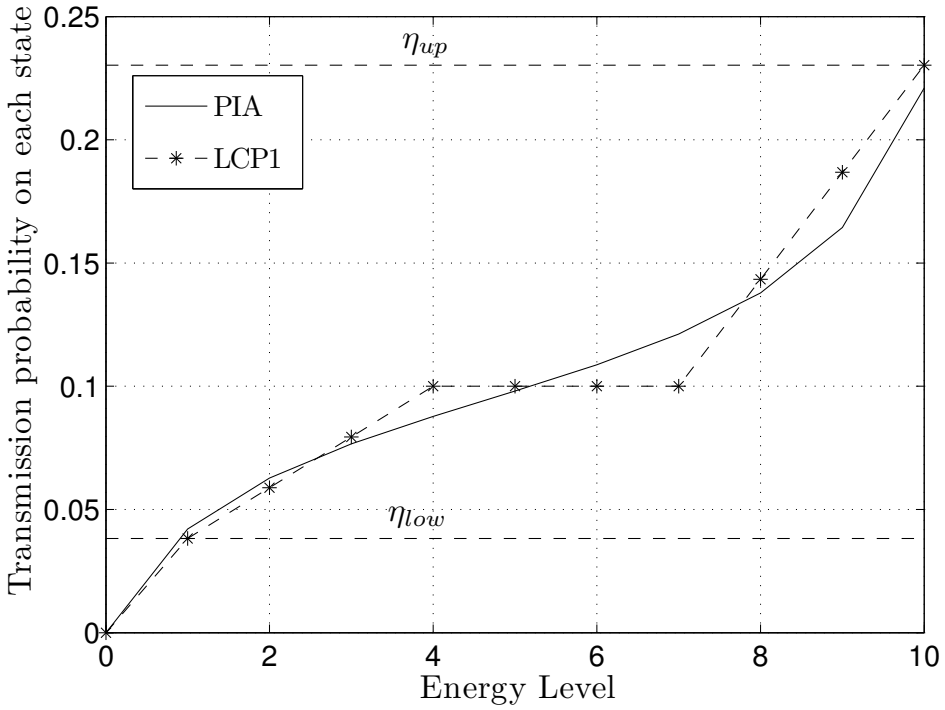


Figure 3.11. *Comparison between the low complexity policy (3.84) (LCP1) and the optimal policy (PIA) computed with Algorithm 3.* $\lambda = 1$, $e_{\max} = 10$, $\beta = 0.1$, $V_{u,k}$ *is exponentially distributed.*

Based on Theorem 3.10.1 and the fact that transmitting with constant probability $\min\{\beta, x^*\}$ is asymptotically optimal for large battery capacity (see Lemmas 3.10.2 and 3.10.3), we construct a heuristic policy which (a) is conservative when energy is low, (b) transmits with probability $\min\{\beta, x^*\}$ in the middle-energy levels, (c) is aggressive when the battery capacity is approached. In particular, we choose the following low-complexity policy $\tilde{\eta}^{(\lambda)}(e)$ (LCP1), depicted in Fig. 3.11 along with the optimal policy $\eta^{(\lambda)}(e)$, obtained with the PIA:

$$\tilde{\eta}^{(\lambda)}(e) = \begin{cases} \frac{4-e}{3}\eta_{\text{low}} + \frac{e-1}{3}\min\{\beta, x^*\}, & 1 \leq e \leq 3 \\ \min\{\beta, x^*\}, & 4 \leq e \leq e_{\max} - 3 \\ \frac{e_{\max}-e}{3}\min\{\beta, x^*\} + \frac{e+3-e_{\max}}{3}\eta_{\text{up}}, & e_{\max} - 2 \leq e \leq e_{\max}. \end{cases} \tag{3.84}$$

The terms η_{low} and η_{up} can be computed from Theorem 3.10.1 *e.g.*, using the bisection method [61].

In Fig. 3.11, we notice a good match between the low-complexity and the optimal policies. Moreover, in Fig. 3.12, the curves $G(\eta) - \lambda P(\eta)$ versus λ perfectly match under the two policies.

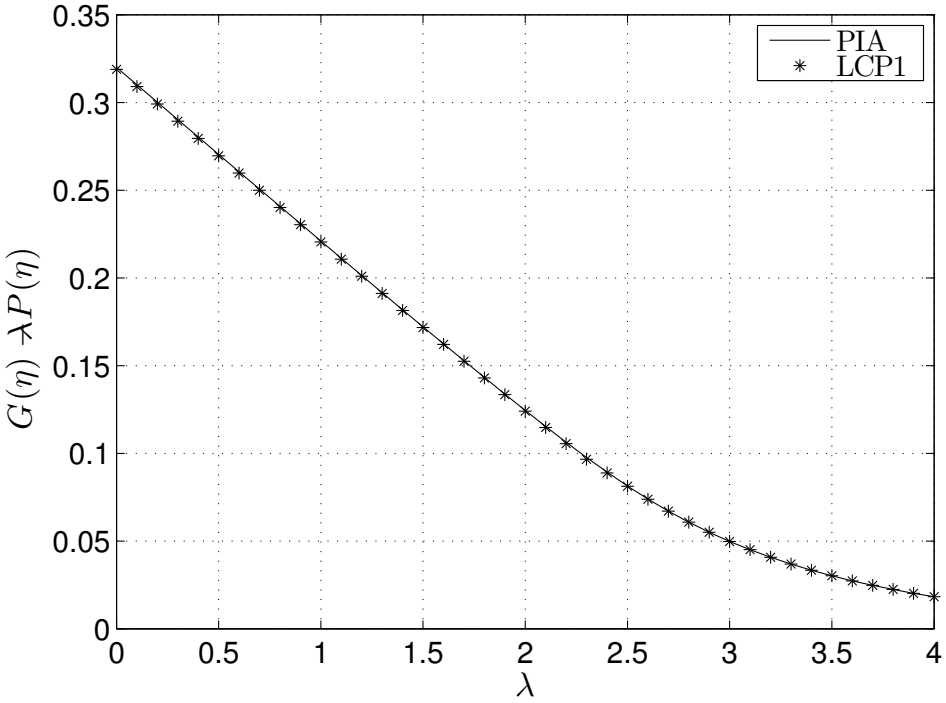


Figure 3.12. *$G(\eta) - \lambda P(\eta)$ as a function of λ, computed under the low complexity policy (LCP1) and the optimal policy (PIA). $\lambda = 1$, $e_{\max} = 10$, $\beta = 0.1$, $V_{u,k}$ is exponentially distributed.*

3.10.2 Large battery capacity approximation

In this section, we design a low complexity policy based on the approximation of large battery capacity. The approximation is based on the following lemma, which follows from property 4) of Prop. 3.9.1.

Lemma 3.10.2. *We have the following upper bounds to the network utility (3.59):*

$$R_{\boldsymbol{\eta}} < Ug(P(\eta))(1 - P(\eta))^{U-1} \leq Ug(p^*)(1 - p^*)^{U-1}, \tag{3.85}$$

where $p^* = \beta$ if $\beta < 1/U$ and $g'(\beta)(1-\beta) - (U-1)g(\beta) > 0$; otherwise, p^* uniquely solves

$$g'(p^*)(1 - p^*) - (U - 1)g(p^*) = 0, \text{ for } p^* \in (0, 1/U). \tag{3.86}$$

Proof. The first inequality in (3.85) directly follows from property 4) of Prop. 3.9.1, since $G(\eta^{(\lambda)}) < g(P(\eta^{(\lambda)}))$. The second inequality is obtained by maximization over $\eta \in \mathcal{U}$. In particular, since

$P(\eta) \in (0, \beta]$ from property 4) of Prop. 3.9.1, we have

$$\max_{\eta \in \mathcal{U}} Ug(P(\eta))(1 - P(\eta))^{U-1} = \max_{p \in (0,\beta]} Ug(p)(1-p)^{U-1}. \tag{3.87}$$

Let $p^* = \arg\max_{p \in (0,\beta]} [Ug(p)(1-p)^{U-1}]$. The derivative of $Ug(p)(1-p)^{U-1}$ with respect to p is positive if and only if

$$h(p) \triangleq g'(p)(1-p) - (U-1)g(p) > 0. \tag{3.88}$$

$h(p)$ is a decreasing function of p, with $\lim_{p \to 0} h(p) = +\infty$ and $h(1/U) = -(U-1)[g(1/U) - 1/Ug'(1/U)] < 0$ (since $g(x)$ is concave increasing and $g(0) = 0$), therefore we have the following cases:

- if $h(\beta) > 0$ (necessarily, $\beta < 1/U$), then $p^* = \beta$;
- otherwise, p^* uniquely solves $h(p^*) = 0$.

The lemma is thus proved. □

Based on this result, we propose the following policy, which is then shown to be asymptotically optimal for large battery capacity:

$$\tilde{\eta}(e) = p^*, \ \forall e \in \mathcal{E} \setminus \{0\}. \tag{3.89}$$

It can be verified, by using the expression of the steady state distribution in Lemma 3.8.1 that, under policy $\tilde{\eta}$,

$$G(\tilde{\eta}) = \frac{1 - \left(\frac{\bar{\beta}p^*}{\beta(1-p^*)}\right)^{e_{\max}}}{1 - \frac{p^*}{\beta}\left(\frac{\bar{\beta}p^*}{\beta(1-p^*)}\right)^{e_{\max}}} g(p^*) \to g(p^*), \tag{3.90}$$

$$P(\tilde{\eta}) = \frac{1 - \left(\frac{\bar{\beta}p^*}{\beta(1-p^*)}\right)^{e_{\max}}}{1 - \frac{p^*}{\beta}\left(\frac{\bar{\beta}p^*}{\beta(1-p^*)}\right)^{e_{\max}}} p^* \to p^*, \tag{3.91}$$

where the limit holds for $e_{\max} \to \infty$, and we have used the fact that $p^* \in (0, \beta]$, so that the exponential terms decay to zero (if $p^* < \beta$; the expressions of $G(\tilde{\eta})$ and $P(\tilde{\eta})$ when $p^* = \beta$ are obtained in the limit $p^* \to \beta$, yielding the same limiting behavior for $e_{\max} \to \infty$). The following lemma is a consequence of this limiting behavior for large battery capacity.

Lemma 3.10.3. *The policy $\tilde{\eta}$ defined in (3.89) is asymptotically optimal for $e_{\max} \to \infty$,* i.e.,

$$\lim_{e_{\max}\to\infty} R_{\tilde{\boldsymbol{\eta}}} = \lim_{e_{\max}\to\infty} \max_{\eta} R_{\boldsymbol{\eta}}. \tag{3.92}$$

Proof. We have

$$R_{\tilde{\boldsymbol{\eta}}} \leq \max_{\eta} R_{\boldsymbol{\eta}} < Ug(p^*)(1-p^*)^{U-1}, \tag{3.93}$$

where the second inequality follows from Lemma 3.10.2. Computing the limit for $e_{\max} \to \infty$ and using (3.90) and (3.91), we obtain

$$Ug(p^*)(1-p^*)^{U-1} = \lim_{e_{\max}\to\infty} R_{\tilde{\boldsymbol{\eta}}} \leq \lim_{e_{\max}\to\infty} \max_{\eta} R_{\boldsymbol{\eta}}. \tag{3.94}$$

The lemma is thus proved by combining (3.93) and (3.94). □

3.11 Numerical Results

In this section, we present some numerical results. We model $V_{u,k}$ as an exponential random variable with unit mean, with pdf $f_V(v) = e^{-v},\ v \geq 0$. From (3.10) and (3.11), we obtain

$$g(x) = x(1 - \ln x). \tag{3.95}$$

In Fig. 3.13, we plot the network utility (3.48) under the policy (3.62), computed using Algorithm 2, We consider different scenarios differing in the battery capacity $e_{\max} \in \{1, 10\}$ and the EH rate $\beta \in \{1/U, 0.1, 0.01\}$, as a function of the number of EHSs in the network U. In particular, when $\beta = 1/U$ the total expected energy harvested by the network in one time-slot is 1. Interestingly, the network utility increases with the number of EHSs U. This behavior is due to the strict concavity of $g(x)$, such that a diminishing return is associated to a larger transmission probability x. Therefore, the smaller the number of EHSs U, the more the transmission opportunities for each EHS, but the smaller the marginal gain, so that the network utility decreases. Clearly, the scenario with $e_{\max} = 10$ outperforms the scenario with $e_{\max} = 1$. Note that, for $U < 10$, the best performance is attained in the case $\beta = 1/U$, since more energy is available to the EHSs. In contrast, for $U > 10$ and $e_{\max} = 10$, the best performance is attained in both cases $\beta = 1/U$ and $\beta = 0.1$, despite a larger energy availability

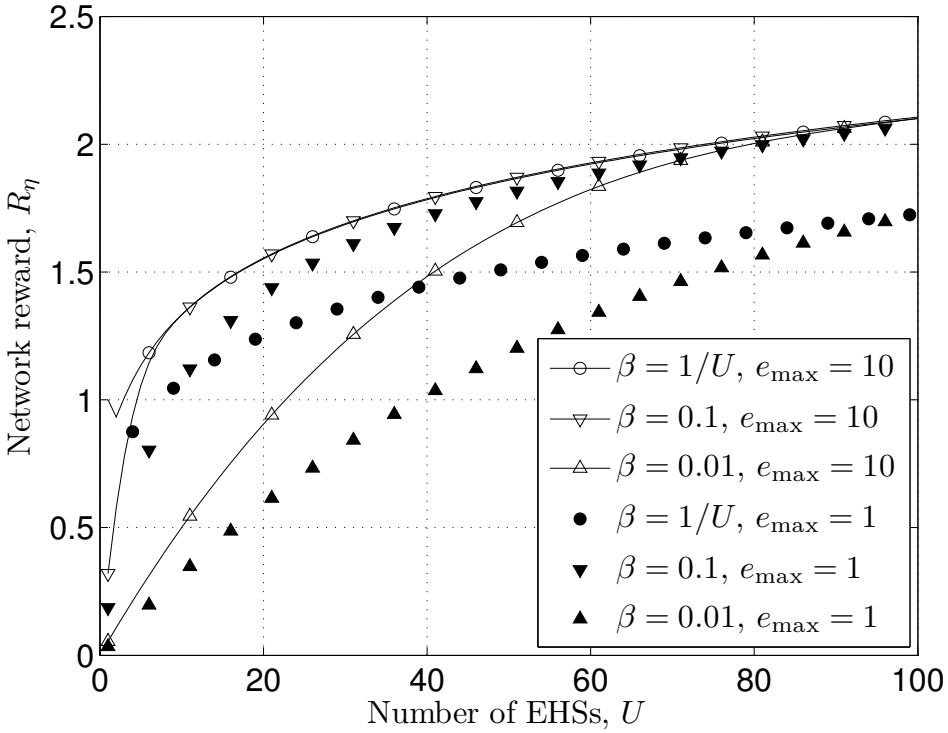


Figure 3.13. *Network utility (3.48) for different values of the battery capacity* $e_{\max} \in \{1, 10\}$ *and for different EH rates* $\beta \in \{1/U, 0.1, 0.01\}$*, as a function of the number of EHSs in the network* U*.*

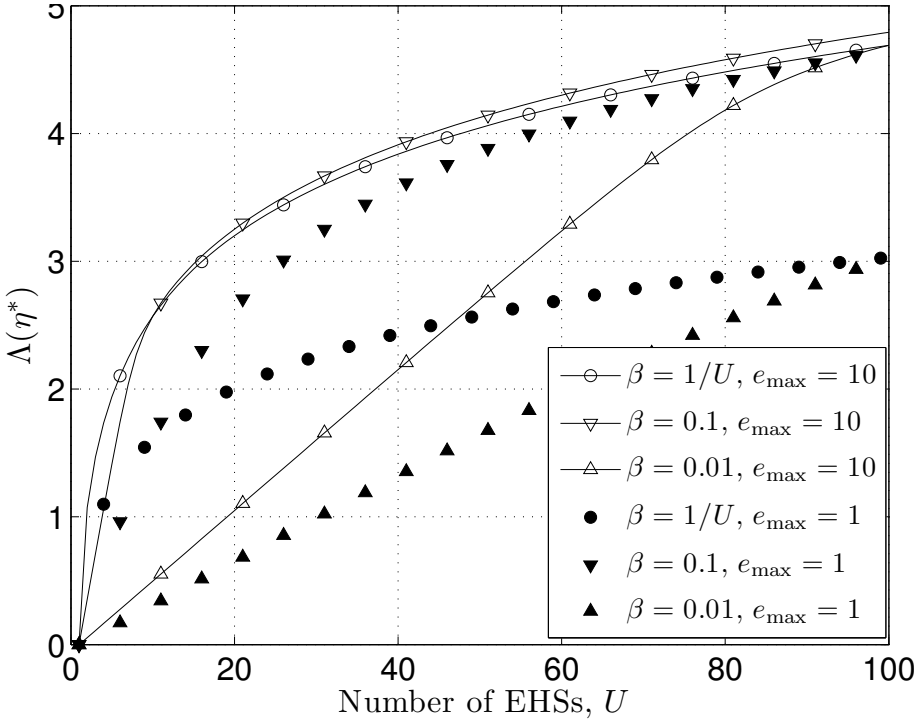


Figure 3.14. $\Lambda(\eta^*)$ *under the SNE* η^* *for different values of the battery capacity* $e_{\max} \in \{1, 10\}$ *and for different EH rates* $\beta \in \{1/U, 0.1, 0.01\}$*, as a function of the number of EHSs in the network* U*.*

in the latter case. This is due to the fact that, as proved in Theorem 3.9.2, under the optimal policy η^*, $P(\eta^*) \leq \min\{\beta, 1/U\} = 1/U$, hence the performance bottleneck is due to the number of EHSs in the network, rather than the energy availability. Clearly, in the case $\beta = 0.1$, a large amount of energy cannot be employed to perform data transmission, and is lost via overflow. This amounts to $\beta - P(\eta^*) \geq \beta - 1/U$. In contrast, when $\beta = 0.01$, we have $P(\eta^*) \leq \min\{\beta, 1/U\} = \beta$ for all values of U considered, hence the performance bottleneck is energy availability.

A different trend is observed when $e_{\max} = 1$. In this case, for $U > 10$, the scenario $\beta = 0.1$ outperforms $\beta = 1/U$. This is a consequence of the fact that, when $e_{\max} = 1$, whenever an EHS transmits, its battery is emptied, hence it enters a recharge phase, with expected duration $1/\beta = U$, during which the EHS is inactive. In contrast, the recharge phase in the scenario $\beta = 0.1$ is much faster, and the EHS becomes quickly available for data transmission.

In Fig. 3.14, we plot the optimal $\lambda^* = \Lambda(\eta^*)$ versus the number of EHSs U. We notice that, the larger the number of EHSs, the larger λ^*. In fact, for a given policy η, the larger U, the more frequent the collisions. A larger λ^* thus balances this phenomenon by penalizing the average transmission probability $P(\eta)^*$ in (3.62), and in turn forces each EHS to transmit more sparingly, so as to accommodate the transmissions of more nodes in the network. Moreover, the larger the EH rate β, the larger λ^*. In fact, the larger β, the larger the energy availability for each EHS, which could, in principle, transmit more frequently and, at the same time, cause more collisions. The effect of a larger β (having more transmissions, hence more collisions in the system) is thus balanced by a larger λ^*, which penalizes high transmission probabilities.

In Fig. 3.15, we compare the network utility computed under the optimal policy (PIA), obtained using Algorithm 2, and the following low-complexity policies:

- LCP1, obtained using Algorithm 2; however, the PIA in the main iteration stage of the algorithm is replaced with policy (3.84), studied in Sec. 3.10.1;
- LCP2, given by (3.89) and studied in Sec. 3.10.2; LCP2 transmits with a fixed probability p^* (see Lemma 3.10.2) whenever the battery is non-empty.

We notice that LCP1 closely approaches the performance of the optimal policy, for all values of $e_{\max}$ and U considered. On the other hand, LCP2 incurs a performance degradation for small number of EHSs and small battery capacity. Nonetheless, the larger the battery capacity, the closer LCP2 to PIA. This behavior is consistent with Lemma 3.10.3, since LCP2 is asymptotically optimal for large battery capacity. Regarding the impact of the number of EHSs U on the performance, we observe

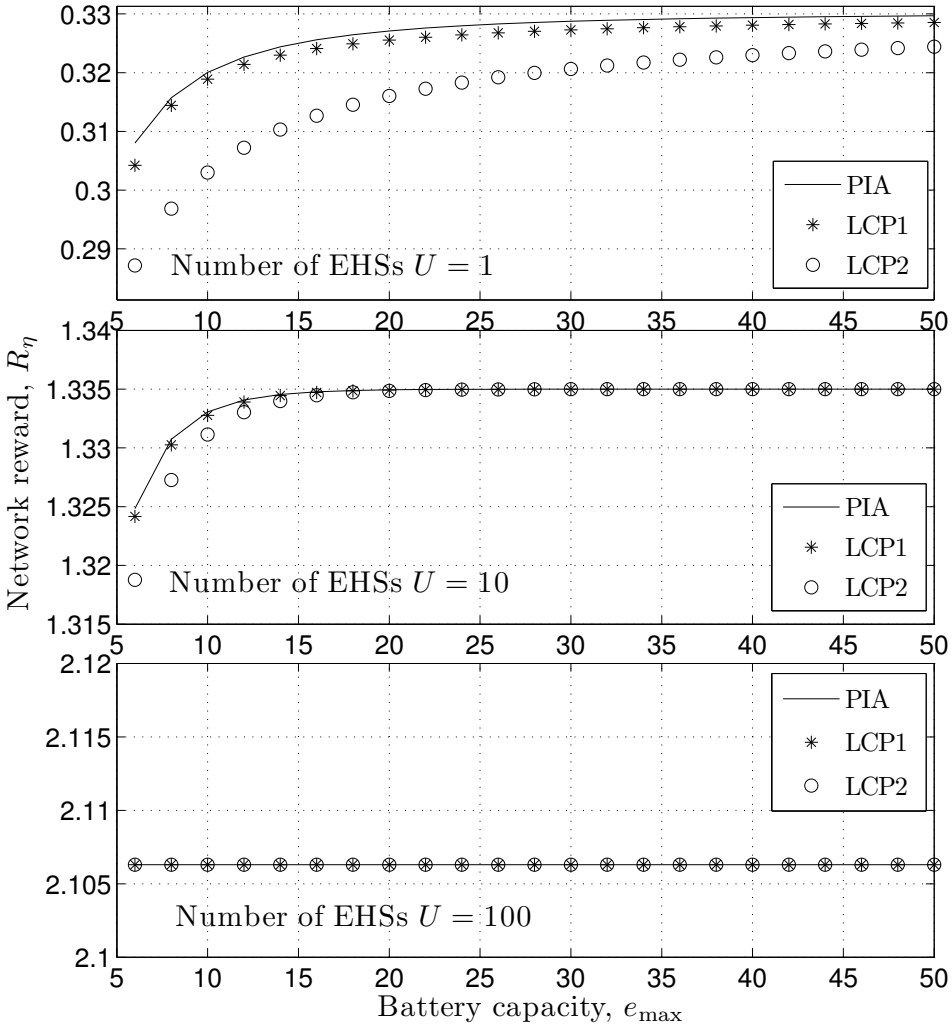


Figure 3.15. *Comparison of optimal policy, computed with the PIA, and the low complexity policies LCP1 and LCP2.* $\beta = 0.1$

that, the larger U, the closer LCP2 to PIA, and the smaller the impact of the battery capacity on the performance (of all policies). This can be explained by noticing that the transmission probability of LCP2 satisfies $p^* < 1/U$. Therefore, the larger U, the less frequent the transmissions, and the more the energy availability. In particular, when $1/U \leq \beta$, *i.e.*, $U \geq 10$, then $p^* < \beta$, hence the energy recharge process is faster than the energy consumption process. It follows that the battery is seldom discharged, and energy outage events seldom occur, so that the finite battery capacity has a negligible impact on the performance. A similar behavior holds for the optimal policy, as discussed in the comments to Fig. 3.13.

3.12 Conclusions

In this chapter, we have considered the general problem of optimizing the data reporting performance of an EH-WSN in the presence of a stochastic ambient energy source. The choice of the performance metric was the aggregate average long-term importance of the reported data, which reflects *both* the value of reporting high-importance data *and* the demand for perpetual operation, typically required in practice. For the single EHS scenario, due to the high implementation complexity of the globally optimal policy, we considered low-complexity balanced policies that solely adapt to the EH state, such that, on average, energy consumption and harvesting are balanced. For the broad range of parameter values considered in this work, the optimal BP was shown to attain near-optimal performance; the gap with respect to the globally optimal policy is narrowed even further, if a simple overflow avoidance scheme is employed. We have demonstrated that the salient parameter is the power-to-depletion ρ, which essentially reflects the ability of the battery to filter out the fluctuations in the harvested energy. Intuitively, as ρ decreases, the adaptation of the transmission probability to the energy supply becomes more critical.

We have then studied the problem of multiaccess in a EH-WSN, where each EHS randomly accesses the shared wireless channel to report data of varying importance to the common fusion center. Simultaneous transmission by multiple EHSs incurs a collision and data loss. We have studied the problem of designing random access policies so as to maximize the overall network utility, defined as the average long-term aggregate network importance of the data packets successfully reported to the fusion center. It is shown that the interaction among multiple EHSs in the network introduces new dynamics in the system, which are not accounted for in the single EHS model. In particular, for small number of EHSs, few collisions occur, hence the bottleneck of the system is energy availability, due to the random and limited energy supply. On the other hand, for large number of EHSs, frequent collisions occur, hence the bottleneck of the system is multiaccess rather than energy availability.

Overall, our results and analysis are encouraging for practical EHS design, as they indicate that near-optimal data reporting performance can be achieved with low-complexity policies that simply adapt to the state of the ambient energy source (single EHS scenario), or that only loosely depend on the energy level in the battery (multiple EHSs scenario).

Appendix 3.A: Proof of Lemma 3.3.1

Proof of Lemma 3.3.1. Let μ be a stationary randomized policy, and $\mathcal{R}_\mu$ be the set of stationary randomized policies which induce the same transmission probability as μ with respect to the data importance random variable, *i.e.*, $\mathbb{E}_V\left[\tilde{\mu}(1;e,a,V)\right] = \mathbb{E}_V\left[\mu(1;e,a,V)\right]$, $\forall (e,a) \in \mathcal{S}$, $\forall \tilde{\mu} \in \mathcal{R}_\mu$. Then, since $\mu \in \mathcal{R}_\mu$, from (3.7) we have

$$G_\mu(\boldsymbol{S}_0) \leq \max_{\tilde{\mu}\in\mathcal{R}_\mu} G_{\tilde{\mu}}(\boldsymbol{S}_0). \tag{3.96}$$

We now show that the maximizer of the right hand side of (3.96) has a threshold structure with respect to the data importance. From (3.7), for any $\tilde{\mu} \in \mathcal{R}_\mu$, we have

$$G_{\tilde{\mu}}(\boldsymbol{S}_0) = \sum_{(e,a)\in\mathcal{S}} \pi_{\tilde{\mu}}(e,a;\boldsymbol{S}_0)\mathbb{E}_V[\tilde{\mu}(1;e,a,V)V], \tag{3.97}$$

where, for each $(e,a) \in \mathcal{S}$, we have defined the steady state distribution of $\{(E_k, A_{k-1})\}$

$$\pi_{\tilde{\mu}}(e,a;\boldsymbol{S_0}) = \lim_{K\to\infty} \inf \frac{1}{K} \sum_{k=0}^{K-1} \Pr_{\tilde{\mu}}(E_k = e, A_{k-1} = a|\boldsymbol{S}_0),$$

and we have used the fact that V_k is independent of (E_k, A_{k-1}) and, given $(E_k, A_{k-1}, V_k) = (e,a,v)$, $Q_k = 1$ with probability $\tilde{\mu}(1;e,a,v)$, independently of $\boldsymbol{S}_0$. From the independence between V_k and $\{(E_i, A_{i-1}), i = 0,\ldots,k\}$, $\forall k$, it can be proved by induction on k that the probability $\Pr_{\tilde{\mu}}(E_k = e, A_{k-1} = a|\boldsymbol{S}_0)$ depends on $\tilde{\mu}$ only through the expectation $\eta(e,a) = \mathbb{E}_V[\tilde{\mu}(1;e,a,V)]$, which is common to all $\tilde{\mu} \in \mathcal{R}_\mu$. Then, since $\mathbb{E}_V[\tilde{\mu}(1;e,a,V)] = \eta(e,a)$, $\forall \tilde{\mu} \in \mathcal{R}_\mu$, it follows that the steady state distribution of $\{(E_k, A_{k-1})\}$ is the same for all $\tilde{\mu} \in \mathcal{R}_\mu$, *i.e.*, $\pi_{\tilde{\mu}}(e,a;\boldsymbol{S_0}) = \pi_\mu(e,a;\boldsymbol{S_0})$. Therefore, from (3.96) and (3.97) we obtain

$$G_\mu(\boldsymbol{S}_0) \leq \max_{\tilde{\mu}\in\mathcal{R}_\mu} G(\tilde{\mu},\boldsymbol{S}_0) = \sum_{(e,a)\in\mathcal{S}} \pi_\mu(e,a;\boldsymbol{S}_0)\mathbb{E}_V[\mu^*(1;e,a,V)V], \tag{3.98}$$

where, for each $(e,a) \in \mathcal{S}$, $\mu^*(1;e,a,\cdot)$ is defined as

$$\tilde{\mu}^*(1;e,a,\cdot) = \underset{\tilde{\mu}(e,a,\cdot):\mathbb{R}^+\mapsto[0,1]}{\arg\max} \mathbb{E}_V[\tilde{\mu}(1;e,a,V)V], \text{ s.t. } \mathbb{E}_V[\tilde{\mu}(1;e,a,V)] = \eta(e,a). \tag{3.99}$$

Since (3.99) is a convex optimization problem, it can be solved using the Lagrangian method [65],

which yields the threshold structure in (3.9).

□

Appendix 3.B: Proof of Theorem 3.4.1

Proof of Theorem 3.4.1. Let us consider a BP η with parameters (η_G, θ). With the help of Fig. 3.16.a), we have the following steady-state equation in state $(e_{\max}, \mathrm{B})$:

$$\begin{aligned}\pi_\eta(e_{\max},\mathrm{B}) &= \pi_\eta(e_{\max},\mathrm{B})p_B(1-\eta(e_{\max},\mathrm{B})) + \pi_\eta(e_{\max},\mathrm{G})\bar{p}_G(1-\eta(e_{\max},\mathrm{G}))\\ &= \pi_\eta(e_{\max},\mathrm{B})p_B\bar{\theta}\bar{\eta}_B + \pi_\eta(e_{\max},\mathrm{G})\bar{p}_G\bar{\theta}\bar{\eta}_G, \end{aligned} \tag{3.100}$$

where the last expression follows from the definition of BP in (3.14). Then, substituting $p_B = 1 - D_B^{-1}$ and $p_G = 1 - \gamma^{-1}D_B^{-1}$, and letting $\mathbf{t}(\theta)$ as in (3.19), after algebraic manipulation we obtain

$$\begin{bmatrix} \pi_\eta(e_{\max},\mathrm{G}) \\ \pi_\eta(e_{\max},\mathrm{B}) \end{bmatrix} = (\theta\eta_G + \bar{\theta})\mathbf{t}(\theta)\pi_\eta(e_{\max},\mathrm{G}). \tag{3.101}$$

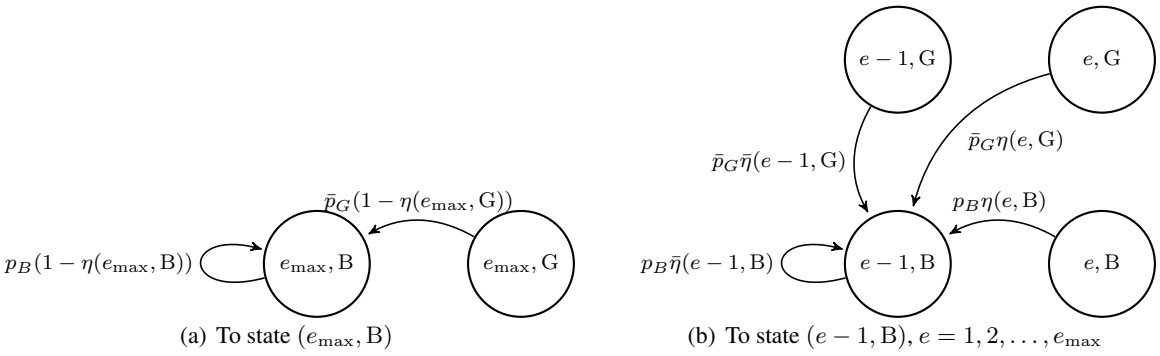
In the long-term, the frequency of transitions from energy level $e-1$ to e and from e to $e-1$ must be the same. Therefore, with the help of Fig. 3.17, for $e \in \{1,\dots,e_{\max}\}$ we have

$$\begin{aligned}&\pi_\eta(e-1,\mathrm{G})p_G\lambda_G\bar{\eta}(e-1,\mathrm{G}) + \pi_\eta(e-1,\mathrm{B})\bar{p}_B\lambda_G\bar{\eta}(e-1,\mathrm{B})\\ &= \pi_\eta(e,\mathrm{G})\left(p_G\bar{\lambda}_G + \bar{p}_G\right)\eta(e,\mathrm{G}) + \pi_\eta(e,\mathrm{B})\left(\bar{p}_B\bar{\lambda}_G + p_B\right)\eta(e,\mathrm{B}).\end{aligned} \tag{3.102}$$

Moreover, with the help of Fig. 3.16.b), we have the following steady-state equation in state $(e-1,\mathrm{B})$, for $e \in \{1,\dots,e_{\max}\}$:

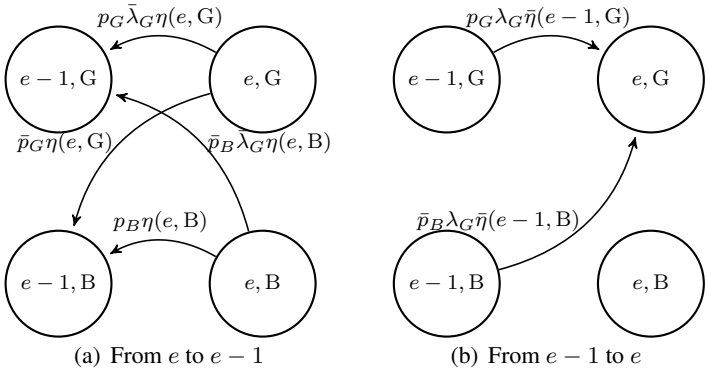
$$\begin{aligned}\pi_\eta(e-1,\mathrm{B}) =&\pi_\eta(e-1,\mathrm{B})p_B\bar{\eta}(e-1,\mathrm{B}) + \pi_\eta(e-1,\mathrm{G})\bar{p}_G\bar{\eta}(e-1,\mathrm{G})\\ &+ \pi_\eta(e,\mathrm{B})p_B\eta(e,\mathrm{B}) + \pi_\eta(e,\mathrm{G})\bar{p}_G\eta(e,\mathrm{G}).\end{aligned} \tag{3.103}$$

Solving the system of equations (3.102), (3.103) with respect to $\pi_\eta(e-1,\mathrm{G})$, $\pi_\eta(e-1,\mathrm{B})$, for



(a) To state $(e_{\max},\mathrm{B})$ (b) To state $(e-1,\mathrm{B})$, $e=1,2,\ldots,e_{\max}$

Figure 3.16. *Transition probabilities to state* $(e-1,\mathrm{B})$



(a) From e to $e-1$ (b) From $e-1$ to e

Figure 3.17. *Transition probabilities from* e *to* $e-1$ *and from* $e-1$ *to* e

$e \in \{1,\ldots,e_{\max}\}$ we obtain

$$\begin{bmatrix} \pi_\eta(e-1,\mathrm{G}) \\ \pi_\eta(e-1,\mathrm{B}) \end{bmatrix} = \mathbf{K}(e-1) \begin{bmatrix} \pi_\eta(e,\mathrm{G}) \\ \pi_\eta(e,\mathrm{B}) \end{bmatrix}, \tag{3.104}$$

where we have defined the 2×2 matrix

$$\mathbf{K}(e-1) = \begin{bmatrix} K_{GG}(e-1) & K_{GB}(e-1) \\ K_{BG}(e-1) & K_{BB}(e-1) \end{bmatrix}, \tag{3.105}$$

with components

$$K_{GG}(e-1) = \eta(e,\mathrm{G})\frac{(D_B\bar{\lambda}_G + \gamma^{-1}\lambda_G)\eta(e-1,\mathrm{B}) + \bar{\lambda}_G\bar{\eta}(e-1,\mathrm{B})}{\bar{\eta}(e-1,\mathrm{G})\lambda_G[D_B\eta(e-1,\mathrm{B}) + \bar{\eta}(e-1,\mathrm{B}) - \eta(e-1,\mathrm{B})\gamma^{-1}]},$$
$$K_{GB}(e-1) = \eta(e,\mathrm{B})\frac{\bar{\eta}(e-1,\mathrm{B})\bar{\lambda}_G + \eta(e-1,\mathrm{B})[D_B - \lambda_G]}{\bar{\eta}(e-1,\mathrm{G})\lambda_G[D_B\eta(e-1,\mathrm{B}) + \bar{\eta}(e-1,\mathrm{B}) - \eta(e-1,\mathrm{B})\gamma^{-1}]},$$
$$K_{BG}(e-1) = \frac{\gamma^{-1}\eta(e,\mathrm{G})}{\lambda_G[D_B\eta(e-1,\mathrm{B}) + \bar{\eta}(e-1,\mathrm{B}) - \eta(e-1,\mathrm{B})\gamma^{-1}]},$$
$$K_{BB}(e-1) = \frac{\eta(e,\mathrm{B})[\gamma^{-1}\bar{\lambda}_G + (D_B-1)\lambda_G]}{\lambda_G[D_B\eta(e-1,\mathrm{B}) + \bar{\eta}(e-1,\mathrm{B}) - \eta(e-1,\mathrm{B})\gamma^{-1}]}. \tag{3.106}$$

In particular, substituting the expression of the BP in Def. 3.3.2, we have $\mathbf{K}(0) = \mathbf{Z}$, where $\mathbf{Z}$ is defined in (3.20), $\mathbf{K}(e) = \mathbf{J}$, $\forall e \in \{1,2,\ldots,e_{\max}-2\}$, where $\mathbf{J}$ is defined in (3.21), and $\mathbf{K}(e_{\max}-1) = \theta\mathbf{J}\begin{bmatrix} \frac{1}{\eta_G} & 0 \\ 0 & \frac{1}{\eta_B} \end{bmatrix} + \bar{\theta}\mathbf{J}$. Using (3.104), (3.101), and the fact that $\mathbf{K}(e_{\max}-1)\mathbf{t}(\theta) = \mathbf{J}\mathbf{t}(\theta)(\bar{\theta} + \theta/\eta_G)$, we then obtain

$$\begin{bmatrix} \pi_\eta(e,\mathrm{G}) \\ \pi_\eta(e,\mathrm{B}) \end{bmatrix} = \mathbf{J}^{e_{\max}-e}\mathbf{t}(\theta)\pi_\eta(e_{\max},\mathrm{G}), \text{ for } e \in \{1,\ldots,e_{\max}-1\}, \tag{3.107}$$

and $[\pi_\eta(0,\mathrm{G});\pi_\eta(0,\mathrm{B})]^T = \mathbf{Z}\,[\pi_\eta(1,\mathrm{G});\pi_\eta(1,\mathrm{B})]^T$, yielding (3.17). $\pi_\eta(e_{\max},\mathrm{G})$ is finally obtained by the law of total probability $\sum_{e=0}^{e_{\max}}[\pi_\eta(e,\mathrm{G}) + \pi_\eta(e,\mathrm{B})] = 1$, yielding (3.18).

The average reward under the BP directly follows by substituting the expressions of the BP and of the steady state distribution in (3.12), and using the fact that, by marginalization over the battery state, $\sum_{e=0}^{e_{\max}}\pi_\eta(e,a) = \pi_A(a)$, $\forall a \in \{\mathrm{G},\mathrm{B}\}$, and, when overflow avoidance is employed ($\theta = 1$), $\pi_\eta(e_{\max},\mathrm{B}) = 0$ from (3.101). □

Appendix 3.C: Proof of Theorem 3.4.3

Proof of Theorem 3.4.3. Proof of (3.25): In this proof, the notation $f(D_B) = h(D_B) + \mathcal{O}(D_B^n)$ is equivalent to $\limsup_{D_B\to\infty}\left|\frac{f(D_B)-h(D_B)}{D_B^n}\right| < \infty$. If $f(D_B)$ is a matrix, this definition applies to each component.

It can be shown that the eigenvalue decomposition of matrix $\mathbf{J}$ in (3.21) is given by $\mathbf{J} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$,

where, letting $\nu = \frac{\eta_G \bar{\lambda}_G}{\bar{\eta}_G \lambda_G} \left(1 - \frac{1}{1+\eta_B(D_B - 1 - \gamma^{-1})}\right)$,

$$\mathbf{V} = \begin{bmatrix} \gamma & 1 \\ 1 & 0 \end{bmatrix} + D_B^{-1} \begin{bmatrix} 0 & \frac{\bar{\lambda}_G - \eta_B}{\eta_B} \\ 0 & -\frac{\bar{\eta}_G \eta_G}{\eta_B^2} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix}. \tag{3.108}$$

Since $\mathbf{V}\mathbf{V}^{-1} = \mathbf{I}_2$, where $\mathbf{I}_2$ is the 2×2 identity matrix, it can be shown that

$$\mathbf{V}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -\gamma \end{bmatrix} + D_B^{-1} \frac{1}{\eta_B^2} \begin{bmatrix} \bar{\eta}_G \eta_G \\ -\bar{\lambda}_G \eta_B + \eta_B^2 - \gamma \bar{\eta}_G \eta_G \end{bmatrix} \begin{bmatrix} 1 & -\gamma \end{bmatrix} + \mathcal{O}(D_B^{-2}). \tag{3.109}$$

Moreover, from (3.19) and (3.20) we have

$$\mathbf{t}(\theta) = \left(\theta \frac{1}{\eta_G} + \bar{\theta}\right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + D_B^{-1} \bar{\theta} \frac{\gamma^{-1} \bar{\eta}_G}{\eta_B} \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \mathcal{O}(D_B^{-2}), \tag{3.110}$$

$$\mathbf{Z} = D_B \eta_B \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} + \mathcal{O}(1). \tag{3.111}$$

Now, letting $e_{\max} = \rho D_B$, where ρ is fixed, we have

$$\mathbf{Z}\mathbf{J}^{e_{\max}-1}\mathbf{t}(\theta) \overset{\text{(a)}}{=} \mathbf{Z}\mathbf{V} \begin{bmatrix} 1 & 0 \\ 0 & \nu^{e_{\max}-1} \end{bmatrix} \mathbf{V}^{-1}\mathbf{t}(\theta) \overset{\text{(b)}}{=} \mathbf{Z}\mathbf{V} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{V}^{-1}\mathbf{t}(\theta) + \mathcal{O}(D_B^{-1}), \tag{3.112}$$

where (a) follows from the eigenvalue decomposition of $\mathbf{J}$, (b) follows from the fact that $\lim_{D_B \to \infty} \nu \in (0, 1)$, hence $\nu^{e_{\max}-1} = \mathcal{O}(D_B^{-2})$. From (3.109) and (3.110), we have that

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{V}^{-1}\mathbf{t}(\theta) = D_B^{-1} \frac{\bar{\eta}_G}{\eta_B^2} (\theta + \bar{\theta}\lambda_G) + \mathcal{O}(D_B^{-2}). \tag{3.113}$$

Moreover, from (3.108) and (3.111) we obtain

$$\mathbf{Z}\mathbf{V} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = D_B \eta_B \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \mathcal{O}(1). \tag{3.114}$$

Then, substituting (3.113) and (3.114) in (3.112), we obtain

$$\mathbf{Z}\mathbf{J}^{e_{\max}-1}\mathbf{t}(\theta) = \frac{\bar{\eta}_G}{\eta_B}(\theta + \bar{\theta}\lambda_G)\begin{bmatrix} 0 \\ 1 \end{bmatrix} + \mathcal{O}(D_B^{-1}). \tag{3.115}$$

Using a similar approach, it can be proved that

$$\begin{aligned}\sum_{e=0}^{e_{\max}-1}\begin{bmatrix} 1 & 1 \end{bmatrix}\mathbf{J}^e\mathbf{t}(\theta) &= \begin{bmatrix} 1 & 1 \end{bmatrix}\mathbf{V}\begin{bmatrix} e_{\max} & 0 \\ 0 & \frac{1-\nu^{e_{\max}}}{1-\nu} \end{bmatrix}\mathbf{V}^{-1}\mathbf{t}(\theta) \\ &= \rho(\gamma+1)\frac{\bar{\eta}_G}{\eta_B^2}(\theta + \bar{\theta}\lambda_G) + \frac{\gamma\bar{\eta}_G\lambda_G}{\eta_B}\left(\theta\frac{1}{\eta_G} + \bar{\theta}\right) + \mathcal{O}(D_B^{-1}).\end{aligned} \tag{3.116}$$

Then, substituting (3.115) and (3.116) in (3.18), we obtain

$$\pi_\eta(e_{\max}, \mathrm{G}) = \pi_A(\mathrm{G})\frac{\lambda_G - \eta_G}{1-\eta_G}\frac{\eta_B}{\rho+\eta_B}\left(\theta + \bar{\theta}\frac{1}{\lambda_G}\right) + \mathcal{O}(D_B^{-1}), \tag{3.117}$$

where we have used the fact that $\left(1 + \mathcal{O}(D_B^{-1})\right)^{-1} = 1 + \mathcal{O}(D_B^{-1})$. Then, from (3.17) and (3.115), we obtain

$$\pi_\eta(0, \mathrm{G}) = \mathcal{O}(D_B^{-1}), \tag{3.118}$$

$$\pi_\eta(0, \mathrm{B}) = \frac{\bar{\eta}_G}{\eta_B}(\theta + \bar{\theta}\lambda_G)\pi_\eta(e_{\max}, \mathrm{G}) + \mathcal{O}(D_B^{-1}) = \pi_A(\mathrm{B})\frac{\eta_B}{\rho+\eta_B} + \mathcal{O}(D_B^{-1}). \tag{3.119}$$

The asymptotic expression (3.25) is finally obtained by substituting (3.117), (3.118) and (3.119) in (3.16), and letting $D_B \to \infty$, where the terms $\mathcal{O}(D_B^{-1}) \to 0$ for $D_B \to \infty$.

Proof of property 1): The inequality $G^{(\infty)}(\eta_G, 1; \rho) > G^{(\infty)}(\eta_G, 0; \rho)$ is proved by direct inspection of (3.25), since $g(1) > g(\eta_G)$ and $\lambda_G > \eta_G$.

Proof of property 2): We want to prove that $G^{(\infty)}(\eta_G, \theta; \rho)$ is an increasing function of ρ. Equivalently, $\frac{\mathrm{d}G^{(\infty)}(\eta_G,\theta;\rho)}{\mathrm{d}\rho} > 0$. The derivative of (3.25) with respect to ρ is given by

$$\begin{aligned}\frac{\mathrm{d}G^{(\infty)}(\eta_G, \theta; \rho)}{\mathrm{d}\rho} &= \pi_A(\mathrm{B})\frac{\eta_B}{(\rho+\eta_B)^2}\left[g(\eta_B) - \eta_B g(1)\right] + \bar{\theta}\pi_A(\mathrm{B})\frac{\eta_B^2}{(\rho+\eta_B)^2}g(1) \\ &\quad + \theta\pi_A(\mathrm{G})\eta_B\frac{\lambda_G - \eta_G}{1-\eta_G}\frac{1}{(\rho+\eta_B)^2}\left[g(\eta_G) - \eta_G g(1)\right] > 0,\end{aligned}$$

where we have used the fact that, since $g(x)$ is a concave function of x and $g(0) = 0$ (Lemma 3.3.2), $g(x) > xg(1)$. The limits for $\rho \to \infty$ and $\rho \to 0$ are finally obtained by computing the corresponding

limit in (3.25).

□

Appendix 3.D: Proof of Lemma 3.4.4

Proof of Lemma 3.4.4. Proof of (3.28) and property 1) We first prove that the optimal BP $(\eta_G^*(\theta;\rho),\theta)$ uniquely solves (3.28). After algebraic manipulation, we find that the derivative of $G^{(\infty)}(\eta_G,\theta;\rho)$ in (3.25) with respect to η_G is positive if and only if $L(\eta_G,\theta;\rho) > 0$, where $L(\eta_G,\theta;\rho)$ is given in (3.29). Moreover,

$$\begin{aligned}\frac{\mathrm{d}L(\eta_G,\theta;\rho)}{\mathrm{d}\eta_G} \propto& \rho g''(\eta_G) + g''(\eta_G)\frac{\eta_B}{\bar{\eta}_G}(\theta\bar{\lambda}_G + \bar{\theta}\bar{\eta}_G) + \gamma\rho g''(\eta_B) \\ &- 2\theta\frac{\bar{\lambda}_G^2\gamma}{\bar{\eta}_G^3}(g(\eta_G) + \bar{\eta}_G g'(\eta_G) - g(1)) - \bar{\theta}\gamma 2g'(\eta_G) < 0, \end{aligned} \tag{3.120}$$

where $\propto$ denotes proportionality up to a positive multiplicative factor, and the inequality holds since $g(\eta_G) + \bar{\eta}_G g'(\eta_G) - g(1) > 0$ and $g'' < 0$, from the concavity of $g(x)$. Therefore, $L(\eta_G,\theta;\rho)$ is a decreasing function of η_G. Moreover,

$$\begin{aligned} L(\lambda_G,\theta;\rho) =& g'(\lambda_G) - g'(0) < 0, \qquad \text{for } \eta_G = \lambda_G,\ \eta_B = 0, \\ L(\beta,\theta;\rho) =& \bar{\theta}\frac{g(\beta)}{\rho} + \bar{\theta}\frac{\beta}{\rho}\left(1+\frac{\beta}{\rho}\right)g'(\beta) + \theta\frac{\beta}{\rho\bar{\beta}^2}\left(\bar{\beta} + \bar{\lambda}_G + \frac{\beta}{\rho}\bar{\lambda}_G\right)(g(\beta) + \bar{\beta}g'(\beta) - g(1)) \\ &+ \theta\frac{1}{\rho}(g(\beta) - \beta g'(\beta)) > 0, \qquad \text{for } \eta_G = \eta_B = \beta. \end{aligned}$$

We conclude that there exists a unique $\eta_G \in (\beta,\lambda_G)$ which maximizes $G^{(\infty)}(\eta_G,\theta;\rho)$, obtained as the unique solution of $L(\eta_G,\theta;\rho) = 0$.

Proof of property 2) Since $L(\eta_G,1;\rho) < L(\eta_G,0;\rho)$, it follows that, under the optimal BP,

$$L(\eta_G^*(0;\rho),0;\rho) = 0 = L(\eta_G^*(1;\rho),1;\rho) < L(\eta_G^*(1;\rho),0;\rho).$$

Since $L(\eta_G,0;\rho)$ is a decreasing function of η_G, it follows that $\eta_G^*(1;\rho) < \eta_G^*(0;\rho)$.

Proof of property 3): We now prove that $\eta_G^*(\theta;\rho)$ is a decreasing function of ρ, for $\rho \geq \beta$, *i.e.*,

$\frac{\mathrm{d}\eta_G^*(\theta;\rho)}{\mathrm{d}\rho} < 0$. For the optimal BP $\eta_G^*(\theta;\rho)$, we have $L(\eta_G^*(\theta;\rho),\theta;\rho) = 0$, hence

$$\frac{\mathrm{d}L(\eta_G^*(\theta;\rho),\theta;\rho)}{\mathrm{d}\rho} = \left[\frac{\mathrm{d}L(\eta_G,\theta;\rho)}{\mathrm{d}\rho} + \frac{\mathrm{d}L(\eta_G,\theta;\rho)}{\mathrm{d}\eta_G}\frac{\mathrm{d}\eta_G^*(\theta;\rho)}{\mathrm{d}\rho}\right]_{\eta_G=\eta_G^*(\theta;\rho)} = 0.$$

Then, since $\frac{\mathrm{d}L(\eta_G,\theta;\rho)}{\mathrm{d}\eta_G} < 0$ from (3.120), $\frac{\mathrm{d}\eta_G^*(\theta;\rho)}{\mathrm{d}\rho} < 0$ is equivalent to $\frac{\mathrm{d}L(\eta_G,\theta;\rho)}{\mathrm{d}\rho}\Big|_{\eta_G=\eta_G^*(\theta;\rho)} < 0$. Moreover, since $L(\eta_G^*(\theta;\rho),\theta;\rho) = 0$, we have

$$\frac{\mathrm{d}L(\eta_G,\theta;\rho)}{\mathrm{d}\rho}\bigg|_{\eta_G=\eta_G^*(\theta;\rho)} = \left(1+\frac{\eta_B}{\rho}\right)^2 \frac{\mathrm{d}\left[L(\eta_G,\theta;\rho)\cdot\frac{\rho^2}{(\rho+\eta_B)^2}\right]}{\mathrm{d}\rho}\bigg|_{\eta_G=\eta_G^*(\theta;\rho)}.$$

Finally, from (3.29) and since $\eta_G = \eta_G^*(\theta;\rho) \in (\beta,\lambda_G)$ and $\eta_B = \gamma(\lambda_G-\eta_G) \in (0,\beta)$, we have

$$\begin{aligned}(\rho+\eta_B)^3 \frac{\mathrm{d}\left[L(\eta_G,\theta;\rho)\cdot\frac{\rho^2}{(\rho+\eta_B)^2}\right]}{\mathrm{d}\rho} &= -\bar{\theta}\,(\rho-\eta_B)\,g(\eta_B) - \bar{\theta}\,(\rho+\eta_B)\,\eta_B g'(\eta_B)\\ &- \theta[g(\eta_B)-\eta_B g'(\eta_B)]\,(\rho-\eta_B) - 2\theta[g'(\eta_B)-g'(\eta_G)]\rho\eta_B\\ &- \theta\frac{\eta_B}{\bar{\eta}_G^2}(g(\eta_G)+\bar{\eta}_G g'(\eta_G)-g(1))\left[\bar{\eta}_G(\rho-\eta_B)+\bar{\lambda}_G(\rho+\eta_B)\right] < 0,\end{aligned}$$

where the inequality holds for $\rho \geq \beta$ (which implies $\rho \geq \eta_B$, since $\eta_B \leq \beta$), and we have used the fact that $g(x)$ is a concave increasing function of x with $g(0) = 0$ (Lemma 3.3.2), hence $g(\eta_B) - \eta_B g'(\eta_B) > 0$, $g'(\eta_B) > g'(\eta_G)$ (since $\eta_B < \eta_G$) and $g(\eta_G) + \bar{\eta}_G g'(\eta_G) - g(1) > 0$. Equivalently, $\frac{\mathrm{d}L(\eta_G,\theta;\rho)}{\mathrm{d}\rho}\Big|_{\eta_G=\eta_G^*(\theta;\rho)} < 0$ and $\frac{\mathrm{d}\eta_G^*(\theta;\rho)}{\mathrm{d}\rho} < 0$.

In the limit $\rho \to \infty$, we have $\lim_{\rho\to\infty} L(\eta_G,\theta;\rho) = g'(\eta_G) - g'(\eta_B)$, which is equal to zero if and only if $\eta_G^*(\theta;\infty) = \beta$, proving (3.30). For $\rho \to 0$, we have

$$\lim_{\rho\to 0}\rho^2 L(\eta_G,\theta;\rho) = \bar{\theta}\eta_B^2 g'(\eta_G) + \theta\frac{\eta_B^2}{\bar{\eta}_G^2}\bar{\lambda}_G(g(\eta_G)+\bar{\eta}_G g'(\eta_G)-g(1)) > 0. \tag{3.121}$$

Hence, for $\rho \to 0$, the asymptotic reward is a strictly increasing function of η_G, for $\eta_G \in (\beta,\lambda_G)$, and is maximized by $\eta_G = \lambda_G$, proving (3.31). □

Appendix 3.E: Proof of Lemma 3.5.1

Proof of Lemma 3.5.1. From the concavity of $g(x)$ and Jensen's inequality [65], (3.34) implies that

$$G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})}) \leq \frac{D_G g\left(\frac{1}{D_G}\int_0^{D_G}\eta^{(\mathrm{CDM})}(\tau)\mathrm{d}\tau\right) + D_B g\left(\frac{1}{D_B}\int_0^{D_B}\eta^{(\mathrm{CDM})}(\tau)\mathrm{d}\tau\right)}{D_G + D_B},$$

which is attained with equality if and only if the energy drawing rates are constant over the GOOD and BAD periods, *i.e.*, $\eta^{(\mathrm{CDM})}(t) = \eta_G, \forall t \in \mathcal{T}_G$ and $\eta^{(\mathrm{CDM})}(t) = \eta_B, \forall t \in \mathcal{T}_B$, where η_G and η_B are constants in $(0, 1)$. Substituting in (3.34), we obtain

$$G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})}) = \pi_A(\mathrm{G})g(\eta_G) + \pi_A(\mathrm{B})g(\eta_B). \tag{3.122}$$

Moreover, from (3.32) and (3.33), η_B and η_G are related to e_L and e_H by

$$e_H = \min\left\{e_L + D_G(\lambda_G - \eta_G), e_{\max}\right\}, \;\; e_L = e_H - D_B\eta_B. \tag{3.123}$$

Note that a policy such that $e_L + D_G(\lambda_G - \eta_G) > e_{\max}$ incurs energy overflow, hence it is strictly sub-optimal. This can be shown by defining an improved policy $\hat{\eta}^{(\mathrm{CDM})}$, with $\hat{\eta}^{(\mathrm{CDM})}(t) = \hat{\eta}_G > \eta_G, \; \forall t \in \mathcal{T}_G$, where $\hat{\eta}_G$ is the unique solution of $e_L + D_G(\lambda_G - \hat{\eta}_G) = e_{\max}$. Under the new policy, we have

$$G_{\mathrm{CDM}}(\hat{\eta}^{(\mathrm{CDM})}) = \pi_A(\mathrm{G})g(\hat{\eta}_G) + \pi_A(\mathrm{B})g(\eta_B) > \pi_A(\mathrm{G})g(\eta_G) + \pi_A(\mathrm{B})g(\eta_B) = G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})}).$$

We thus only consider η_G such that $e_L + D_G(\lambda_G - \eta_G) \leq e_{\max}$. From (3.123), letting $\Delta = e_H - e_L$, η_G and η_B are then given by

$$\eta_G = \lambda_G - \frac{\Delta}{D_G}, \;\; \eta_B = \frac{\Delta}{D_B}. \tag{3.124}$$

Note that $\Delta \in [0, \min\{\lambda_G D_G, D_B, e_{\max}\}]$ since, during the GOOD EH period, the battery cannot be recharged by more than $\lambda_G D_G$ and, during the BAD EH period, it cannot be discharged by more than D_B. Substituting (3.124) in (3.122), we obtain

$$G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})}) = \pi_A(\mathrm{G})g\left(\lambda_G - \gamma^{-1}\frac{\Delta}{D_B}\right) + \pi_A(\mathrm{B})g\left(\frac{\Delta}{D_B}\right). \tag{3.125}$$

We now maximize the right hand side with respect to $\Delta \in [0, \min\{D_G, D_B, e_{\max}\}]$. We have that $\frac{\mathrm{d}}{\mathrm{d}\Delta} G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})}) > 0$ if and only if $\Delta < \beta D_B$. Therefore, $G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})})$ is maximized by

$$\Delta^* = \min\{\beta D_B, \lambda_G D_G, D_B, e_{\max}\} = D_B \min\{\beta, \lambda_G \gamma, 1, \rho\} = D_B \min\{\beta, \rho\}, \tag{3.126}$$

where the last equality follows from the fact that $\beta < \min\{\lambda_G \gamma, 1\}$. The optimal energy drawing rates η_G and η_B are obtained by substituting (3.126) in (3.124) and (3.125), thus proving (3.36) and (3.37). □

Appendix 3.F: Proof of Lemma 3.5.2

Proof of Lemma 3.5.2. We consider a BP for DSM with parameters (η_G, θ). Without loss of generality, we assume that $\eta_G \in (\beta, \lambda_G)$, *i.e.*, $\eta_B \in (0, \beta)$, since the optimal BP, which maximizes the asymptotic average reward (3.25), satisfies this condition (Property 1) of Lemma 3.4.4). If $\rho \geq \beta$, then $G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})*}) = g(\beta)$ and (3.38) holds from Jensen's inequality and Lemma 3.3.2. If $\rho < \beta$, let

$$Z(\eta_G) = \begin{cases} \pi_A(\mathrm{G}) \frac{\rho}{\eta_B} g(\eta_G) + \pi_A(\mathrm{G}) \frac{\eta_B - \rho}{\eta_B} g(\lambda_G) + \pi_A(\mathrm{B}) \frac{\rho}{\eta_B} g(\eta_B) & \text{if } \eta_G < \lambda_G - \frac{\rho}{\gamma} \ (\eta_B > \rho), \\ \pi_A(\mathrm{G}) g(\eta_G) + \pi_A(\mathrm{B}) g(\eta_B) & \text{if } \eta_G \geq \lambda_G - \frac{\rho}{\gamma} \ (\eta_B \leq \rho). \end{cases} \tag{3.127}$$

We have that $Z(\eta_G) \leq G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})*})$. This can be proved by applying Jensen's inequality to $Z(\eta_G)$ when $\eta_G < \lambda_G - \frac{\rho}{\gamma}$, and by using the fact that $Z(\eta_G)$ is a decreasing function of η_G, for $\eta_G \geq \lambda_G - \frac{\rho}{\gamma}$. From this and property 1) of Theorem 3.4.3, it follows that

$$G^{(\infty)}(\eta_G, \theta; \rho) - G_{\mathrm{CDM}}(\eta^{(\mathrm{CDM})*}) \leq G^{(\infty)}(\eta_G, 1; \rho) - Z(\eta_G). \tag{3.128}$$

We finally prove that the right hand side above is negative. In fact, if $\eta_G \geq \lambda_G - \frac{\rho}{\gamma}$, from (3.127) and Lemma 3.3.2 we obtain

$$G^{(\infty)}(\eta_G, 1; \rho) - Z(\eta_G) = -\pi_A(\mathrm{B}) \frac{\eta_B}{\rho + \eta_B} \left[(g(\eta_B) - \eta_B g(1)) + \frac{\eta_B}{\bar{\eta}_G} (g(\eta_G) - \eta_G g(1)) \right] < 0.$$

On the other hand, if $\eta_G < \lambda_G - \frac{\rho}{\gamma}$, we obtain

$$G^{(\infty)}(\eta_G, 1; \rho) - Z(\eta_G) = -\frac{\rho^2}{\rho + \eta_B}\pi_A(\mathrm{B})\left[\frac{g(\eta_B)}{\eta_B} - \frac{g(\eta_G)}{\eta_G}\right] \tag{3.129}$$
$$-\frac{\rho^2}{\rho + \eta_B}\pi_A(\mathrm{B})\frac{\gamma\lambda_G}{\eta_B}\left[\frac{g(\eta_G)}{\eta_G} - \frac{g(\lambda_G)}{\lambda_G}\right] - \frac{\eta_B^2}{\rho + \eta_B}\pi_A(\mathrm{B})\left[\frac{g(\lambda_G) - g(\eta_G)}{\lambda_G - \eta_G} - \frac{g(1) - g(\eta_G)}{1 - \eta_G}\right] < 0.$$

where the inequality holds from the concavity of $g(x)$ (Lemma 3.3.2), which implies $(g(y) - g(x))/(y - x) > (g(z) - g(x))/(z - x)$, for any $x < y < z$, and from the fact that $0 < \eta_B < \eta_G < \lambda_G \leq 1$. The lemma is thus proved. □

Appendix 3.G: Useful lemmas for the multiaccess model

Lemma 3.12.1. *$P(\eta^{(\lambda)})$ is a non-increasing function of λ, for $\lambda \geq 0$, with limits $P(\eta^{(0)}) > 0$ and $\lim_{\lambda\to\infty} P(\eta^{(\lambda)}) = 0$.*

Proof. Assume by contradiction that $\lambda_1 > \lambda_2$ and $P(\eta^{(\lambda_2)}) < P(\eta^{(\lambda_1)})$. Then we have

$$\begin{aligned} &G(\eta^{(\lambda_2)}) - \lambda_2 P(\eta^{(\lambda_2)}) \geq G(\eta^{(\lambda_1)}) - \lambda_1 P(\eta^{(\lambda_1)}) + (\lambda_1 - \lambda_2)P(\eta^{(\lambda_1)}) \\ >&G(\eta^{(\lambda_1)}) - \lambda_1 P(\eta^{(\lambda_1)}) + (\lambda_1 - \lambda_2)P(\eta^{(\lambda_2)}) \geq G(\eta^{(\lambda_2)}) - \lambda_1 P(\eta^{(\lambda_2)}) + (\lambda_1 - \lambda_2)P(\eta^{(\lambda_2)}) \\ =&G(\eta^{(\lambda_2)}) - \lambda_2 P(\eta^{(\lambda_2)}), \end{aligned} \tag{3.130}$$

where the first inequality follows from $G(\eta^{(\lambda_2)}) - \lambda_2 P(\eta^{(\lambda_2)}) \geq G(\eta^{(\lambda_1)}) - \lambda_2 P(\eta^{(\lambda_1)})$ (since $\eta^{(\lambda_2)}$ maximizes $G(\eta) - \lambda_2 P(\eta)$), the second inequality follows from the hypothesis, the last inequality from the fact that $G(\eta^{(\lambda_1)}) - \lambda_1 P(\eta^{(\lambda_1)}) \geq G(\eta^{(\lambda_2)}) - \lambda_1 P(\eta^{(\lambda_2)})$ (since $\eta^{(\lambda_1)}$ maximizes $G(\eta) - \lambda_1 P(\eta)$). It follows that $G(\eta^{(\lambda_2)}) - \lambda_2 P(\eta^{(\lambda_2)}) > G(\eta^{(\lambda_2)}) - \lambda_2 P(\eta^{(\lambda_2)})$, yielding a contradiction. The lemma is thus proved. □

Lemma 3.12.2. *$\Lambda(\eta^{(\lambda)})$ is a continuous, non-increasing function of λ, for $\lambda \geq 0$, with limits $\Lambda(\eta^{(0)}) \in (0, \infty)$ and $\lim_{\lambda\to\infty} \Lambda(\eta^{(\lambda)}) = 0$.*

Proof. For $\lambda = 0$ we have $\eta^{(0)} = \arg\max_{\eta\in\mathcal{U}} G(\eta)$. Then, we obtain

$$\Lambda(\eta^{(0)}) = (U - 1)\frac{G(\eta^{(0)})}{1 - P(\eta^{(0)})} \in (0, \infty). \tag{3.131}$$

$\Lambda(\eta^{(0)})$ is positive and bounded since $0 < G(\eta^{(0)}) \leq g(1) < \infty$ and $P(\eta^{(0)}) \leq \bar{b} < 1$. On the other hand, for $\lambda \to \infty$, we have $\eta^{(\lambda)} \to 0$, hence $G(\eta^{(\lambda)}) \to 0$, $P(\eta^{(\lambda)}) \to 0$ and $\lim_{\lambda\to\infty} \Lambda(\eta^{(\lambda)}) = 0$.

To conclude, we prove that $\Lambda(\eta^{(\lambda)})$ is a non-increasing function of λ, *i.e.*, $\Lambda(\eta^{(\lambda_1)}) \leq \Lambda(\eta^{(\lambda_2)})$ for $\lambda_1 > \lambda_2 \geq 0$. Using (3.63), this is true if and only if

$$G(\eta^{(\lambda_2)})(1 - P(\eta^{(\lambda_1)})) - G(\eta^{(\lambda_1)})(1 - P(\eta^{(\lambda_2)})) \geq 0.$$

Equivalently, by rearranging the terms,

$$\begin{aligned} A(\lambda_1, \lambda_2) \triangleq &(G(\eta^{(\lambda_2)}) - \lambda_2 P(\eta^{(\lambda_2)}))(1 - P(\eta^{(\lambda_1)})) - (G(\eta^{(\lambda_1)}) - \lambda_1 P(\eta^{(\lambda_1)}))(1 - P(\eta^{(\lambda_2)})) \\ &+ \lambda_2 P(\eta^{(\lambda_2)})(1 - P(\eta^{(\lambda_1)})) - \lambda_1 P(\eta^{(\lambda_1)})(1 - P(\eta^{(\lambda_2)})) \geq 0. \end{aligned}$$

Using the fact that $\eta^{(\lambda_2)}$ is optimal for (3.64) when $\lambda = \lambda_2$, hence $G(\eta^{(\lambda_2)}) - \lambda_2 P(\eta^{(\lambda_2)}) \geq G(\eta^{(\lambda_1)}) - \lambda_2 P(\eta^{(\lambda_1)})$ a sufficient condition which guarantees that $A(\lambda_1, \lambda_2) \geq 0$ is that

$$\begin{aligned} &(G(\eta^{(\lambda_1)}) - \lambda_2 P(\eta^{(\lambda_1)}))(1 - P(\eta^{(\lambda_1)})) - (G(\eta^{(\lambda_1)}) - \lambda_1 P(\eta^{(\lambda_1)}))(1 - P(\eta^{(\lambda_2)})) \\ &+\lambda_2 P(\eta^{(\lambda_2)})(1 - P(\eta^{(\lambda_1)})) - \lambda_1 P(\eta^{(\lambda_1)})(1 - P(\eta^{(\lambda_2)})) \geq 0. \end{aligned}$$

After rearranging the terms, it can be readily verified that this is equivalent to

$$(P(\eta^{(\lambda_2)}) - P(\eta^{(\lambda_1)}))[G(\eta^{(\lambda_1)}) + \lambda_2(1 - P(\eta^{(\lambda_1)}))] \geq 0,$$

which clearly holds from $G(\eta^{(\lambda_1)}) + \lambda_2(1 - P(\eta^{(\lambda_1)})) > 0$ and Lemma 3.12.1. The lemma is thus proved. □

Appendix 3.H: Proof of Theorem 3.10.1

Proof of Theorem 3.10.1. For notational conciseness, we let $Z_\lambda(\eta) = G(\eta) - \lambda P(\eta)$ and $z_\lambda(x) = g(x) - \lambda x$. From Lemma 3.3.2, $z_\lambda(x)$ is a strictly concave function of x (but not necessarily increasing).

After algebraic manipulation, the derivative of $Z_\lambda(\eta)$ with respect to $\eta(1)$ is given by

$$\begin{aligned}\frac{\mathrm{d}Z_\lambda(\eta)}{\mathrm{d}\eta(1)} &\propto z_\lambda(\eta(1)) + (1-\eta(1))z'_\lambda(\eta(1)) - \frac{Z_\lambda(\eta)}{\beta} \\ &> z_\lambda(\eta(1)) + (1-\eta(1))z'_\lambda(\eta(1)) - \frac{z_\lambda(\min\{\beta, x^*\})}{\beta} \triangleq L(\eta(1)),\end{aligned}$$

where the second step follows from the fact that $Z_\lambda(\eta) < z_\lambda(\min\{\beta, x^*\})$ and $x^* = \arg\max_x z_\lambda(x)$. From the concavity of $z_\lambda(x)$ it can be shown that $L(x)$ is a decreasing function of x, with $\lim_{x\to 0^+} L(x) > 0$ (since $z_\lambda(x) = g'(x) - \lambda \to \infty$ for $x \to 0$) and $L(\min\{\beta, x^*\}) < 0$. In fact, if $\beta < x^*$ we have

$$L(\min\{\beta, x^*\}) = z_\lambda(\beta) + (1-\beta)z'_\lambda(\eta(1)) - \frac{z_\lambda(\beta)}{\beta} = -\frac{\bar{\beta}}{\beta}\left(g(\beta) - \beta g'(\beta)\right) < 0, \qquad (3.132)$$

where in the second step we have used the fact that $z_\lambda(x) = g(x) - \lambda x$ and $g(x) - xg'(x) > 0$, from the concavity of $g(x)$. On the other hand, if $\beta \geq x^*$, we have

$$L(x^*) = (1-x^*)z'_\lambda(x^*) - z_\lambda(x^*)\frac{\bar{\beta}}{\beta} = -z_\lambda(x^*)\frac{\bar{\beta}}{\beta} < 0, \qquad (3.133)$$

where we have used the fact that x^* maximizes $z_\lambda(x)$, hence $z'_\lambda(x^*) = 0$ and $z_\lambda(x^*) > 0$. Therefore, there exists a unique $\eta_{\text{low}} \in (0, \min\{\beta, x^*\})$ that solves $L(\eta_{\text{low}}) = 0$ (equivalent to (3.82)). Then, for all $\eta(1) \leq \eta_{\text{low}}$ we have $L(\eta(1)) \geq 0$, hence $\frac{\mathrm{d}Z_\lambda(\eta)}{\mathrm{d}\eta(1)} > 0$, which proves that $\eta(1) \leq \eta_{\text{low}}$ is strictly suboptimal.

Similarly, after algebraic manipulation, the derivative of $Z_\lambda(\eta)$ with respect to $\eta(e_{\max})$ is given by

$$\begin{aligned}\frac{\mathrm{d}Z_\lambda(\eta)}{\mathrm{d}\eta(e_{\max})} &\propto -z_\lambda(\eta(e_{\max})) + \eta(e_{\max})z'_\lambda(\eta(e_{\max})) + Z_\lambda(\eta) \\ &< -z_\lambda(\eta(e_{\max})) + \eta(e_{\max})z'_\lambda(\eta(e_{\max})) + z_\lambda(\min\{\beta, x^*\}) \triangleq U(\eta(e_{\max})).\end{aligned}$$

Since $z_\lambda(x)$ is concave, it can be proved that $U(x)$ is a decreasing function of x. Moreover, if $\beta < x^*$,

$$U(\min\{\beta, x^*\}) = \beta z'_\lambda(\beta) > 0, \qquad U(x^*) = -z_\lambda(x^*) + z_\lambda(\beta) < 0.$$

On the other hand, if $\beta \geq x^*$, then $U(x^*, \beta) = 0$. Therefore, there exists a unique $\eta_{\text{up}} \in (\min\{\beta, x^*\}, x^*)$ (in particular, $\eta_{\text{up}} = x^*$ if $\beta \geq x^*$) that solves $U(\eta_{\text{up}}) = 0$ (equivalent to (3.83)). Then, for all

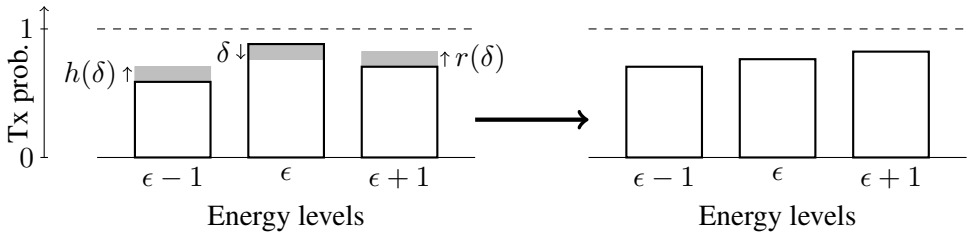


Figure 3.18. *Transmission transfer technique*

$\eta(e_{\max}) \geq \eta_{\text{up}}$ we have $U(\eta(e_{\max})) \leq 0$, hence $\frac{\mathrm{d}Z_\lambda(\eta)}{\mathrm{d}\eta(e_{\max})} < 0$. It follows that $\eta(e_{\max}) \geq \eta_{\text{up}}$ is strictly suboptimal.

We now prove P1) by contradiction, by using a similar technique employed in [19]. In particular, since we have proved that $\eta(e_{\max}) \geq \eta_{\text{up}}$ (and, in particular, $\eta(e_{\max}) \geq x^*$) is strictly suboptimal, we assume that $\eta(e_{\max}) < x^*$. It follows that $z'_\lambda(\eta(e_{\max})) > 0$. Let $\eta \in \mathcal{U}$ be a generic transmission policy such that $\eta(e_{\max}) < x^*$, which violates P1). Then, there exists $\epsilon \in \{1, \ldots, e_{\max} - 1\}$ such that

$$\eta_0(\epsilon - 1) < \eta_0(\epsilon) \geq \eta_0(\epsilon + 1). \tag{3.134}$$

Note that P1) is violated since $\eta_0(\epsilon) \geq \eta_0(\epsilon + 1)$, *i.e.*, η_0 is not strictly increasing from ϵ to $\epsilon + 1$. With the help of Fig. 3.18, we now define a new transmission policy, η_δ,[6] parameterized by $\delta > 0$, as:

$$\eta_\delta(e) = \begin{cases} \eta_0(e), & e \in \mathcal{E} \setminus \{\epsilon - 1, \epsilon, \epsilon + 1\}, \\ \eta_0(\epsilon - 1) + h(\delta), & e = \epsilon - 1, \\ \eta_0(\epsilon) - \delta, & e = \epsilon \\ \eta_0(\epsilon + 1) + r(\delta), & e = \epsilon + 1. \end{cases}$$

Intuitively, policy η_δ is constructed from the original policy η_0 by *transferring* some transmissions from energy state ϵ to states $(\epsilon+1)$ and $(\epsilon-1)$, whereas transmissions in all other states are unaffected. The functions $r(\delta) > 0$ and $h(\delta) \geq 0$ are uniquely defined as follows. If $\epsilon > 1$, the transfer of transmissions is done so as to preserve the steady state distribution of visiting the *lower* energy states $\{0, \ldots, \epsilon - 2\}$ and the *higher* energy states $\{\epsilon + 2, \epsilon + 3, \ldots, e_{\max}\}$. On the other hand, if $\epsilon = 1$, $h(\delta) = 0$ and $r(\delta)$ is chosen so as to preserve the steady state distribution of visiting the *higher* energy states $\{3, \ldots, e_{\max}\}$. By using this technique, on the one hand, the new policy η_δ partially corrects

[6]With a slight abuse of notation, in this proof we use the subscript δ as a parameter of the policy η_δ, whereas the subscript i in η_i is used in Sec. 3.7 and in the following sections to indicate EHS i.

the violation of P1), by diminishing the gap $\eta(\epsilon) - \eta(\epsilon+1)$ by a quantity $\delta + r(\delta) > 0$; on the other hand, the perturbation on the steady state distribution is confined only to the states $\{\epsilon-1, \epsilon, \epsilon+1\}$, thus simplifying the analysis. Formally,

1. if $\epsilon = 1$, let $h(\delta) = 0$ and $r(\delta)$ such that $\pi_{\eta_\delta}(e_{\max}) = \pi_{\eta_0}(e_{\max}), \forall \delta < \kappa$;

2. if $\epsilon > 1$, let $h(\delta)$ and $r(\delta)$ be such that

$$\begin{cases} \pi_{\eta_\delta}(e_{\max}) = \pi_{\eta_0}(e_{\max}) \\ \pi_{\eta_\delta}(0) = \pi_{\eta_0}(0) \end{cases}, \forall \delta < \kappa, \tag{3.135}$$

where $0 < \kappa \ll 1$ is an arbitrarily small constant, which guarantees an admissible policy $\eta_\delta \in \mathcal{U}$.

Then, we prove that $\frac{\mathrm{d}Z_\lambda(\eta_\delta)}{\mathrm{d}\delta}\Big|_{\delta=0} > 0$. It follows that there exists $\kappa > 0$ such that $Z_\lambda(\eta_\delta) > Z_\lambda(\eta_0),\ \forall \delta \in (0, \kappa)$, hence η_0 is strictly sub-optimal. By contradiction, any policy violating P1) is strictly suboptimal, hence the property is proved.

Note that the policy η_δ is unaffected in states $e \in \{0, 1, \ldots, \epsilon-2\} \cup \{\epsilon+2, \epsilon+3, \ldots, e_{\max}\}$, *i.e.*, $\eta_\delta(e) = \eta_0(e)$. Therefore, using (3.52), for $e \geq \epsilon + 2$ it can be shown that

$$\pi_{\eta_\delta}(e) = \frac{1}{\prod_{i=e}^{e_{\max}-1} W_{\eta_\delta}(i)} \pi_{\eta_\delta}(e_{\max}) = \frac{1}{\prod_{i=e}^{e_{\max}-1} W_{\eta_0}(i)} \pi_{\eta_0}(e_{\max}) = \pi_{\eta_0}(e), \tag{3.136}$$

hence the steady-state distribution of visiting states $e \geq \epsilon + 2$ is unaffected by policy η_δ (not only state $e = e_{\max}$). Similarly, for $\epsilon > 1$ and $e \leq \epsilon - 2$, we have

$$\pi_{\eta_\delta}(e) = \prod_{i=0}^{e-1} W_{\eta_\delta}(i) \pi_{\eta_\delta}(0) = \prod_{i=0}^{e-1} W_{\eta_0}(i) \pi_{\eta_0}(0) = \pi_{\eta_0}(e), \tag{3.137}$$

so that the steady-state distribution of visiting states $e \leq \epsilon - 2$ is unaffected by policy η_δ (not only state $e = 0$). Therefore, the perturbation in the steady-state distribution, induced by policy η_δ, is confined to states $\{\epsilon - 1, \epsilon.\epsilon + 1\}$ only, hence the average reward under policy η_δ is given by

$$\begin{aligned} Z_\lambda(\eta_\delta) = Z_\lambda(\eta_0) - \pi_{\eta_0}(\epsilon-1) z_\lambda(\eta_0(\epsilon-1)) - \pi_{\eta_0}(\epsilon) z_\lambda(\eta_0(\epsilon)) - \pi_{\eta_0}(\epsilon+1) z_\lambda(\eta_0(\epsilon+1)). \\ + \pi_{\eta_\delta}(\epsilon-1) z_\lambda(\eta_0(\epsilon-1) + h(\delta)) + \pi_{\eta_\delta}(\epsilon) z_\lambda(\eta_0(\epsilon) - \delta) + \pi_{\eta_\delta}(\epsilon+1) z_\lambda(\eta_0(\epsilon+1) + r(\delta)). \end{aligned}$$

By computing the derivative in zero of $Z_\lambda(\eta_\delta)$ with respect to δ, we obtain

$$\left.\frac{\mathrm{d}Z_\lambda(\eta_\delta)}{\mathrm{d}\delta}\right|_{\delta=0} = \pi_{\eta_0}(\epsilon-1)z'_\lambda(\eta_0(\epsilon-1))h'(0) - \pi_{\eta_0}(\epsilon)z'_\lambda(\eta_0(\epsilon)) + \pi_{\eta_0}(\epsilon+1)z'_\lambda(\eta_0(\epsilon+1))r'(0)$$
$$+ \left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon-1)}{\mathrm{d}\delta}\right|_{\delta=0} z_\lambda(\eta_0(\epsilon-1)) + \left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon)}{\mathrm{d}\delta}\right|_{\delta=0} z_\lambda(\eta_0(\epsilon)) + \left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon+1)}{\mathrm{d}\delta}\right|_{\delta=0} z_\lambda(\eta_0(\epsilon+1)). \quad (3.138)$$

The derivative of the steady state distribution is computed as follows. Using (3.52) and the fact that $\pi_{\eta_\delta}(e_{\max}) = \pi_{\eta_0}(e_{\max})$ and $W_{\eta_\delta}(i) = W_{\eta_0}(i)$ for $i < \epsilon - 2$ and $i > \epsilon + 1$, we obtain

$$\pi_{\eta_\delta}(\epsilon+1) = \left(\chi(\epsilon = e_{\max}-1) + \chi(\epsilon < e_{\max}-1)\frac{W_{\eta_0}(\epsilon+1)}{W_{\eta_\delta}(\epsilon+1)}\right)\pi_{\eta_0}(\epsilon+1) \quad (3.139)$$
$$\pi_{\eta_\delta}(\epsilon) = \frac{W_{\eta_0}(\epsilon)}{W_{\eta_\delta}(\epsilon)}\left(\chi(\epsilon = e_{\max}-1) + \chi(\epsilon < e_{\max}-1)\frac{W_{\eta_0}(\epsilon+1)}{W_{\eta_\delta}(\epsilon+1)}\right)\pi_{\eta_0}(\epsilon) \quad (3.140)$$
$$\pi_{\eta_\delta}(\epsilon-1) = \frac{W_{\eta_0}(\epsilon-1)}{W_{\eta_\delta}(\epsilon-1)}\frac{W_{\eta_0}(\epsilon)}{W_{\eta_\delta}(\epsilon)}\left[1 - \chi(\epsilon < e_{\max}-1)\left(1 - \frac{W_{\eta_0}(\epsilon+1)}{W_{\eta_\delta}(\epsilon+1)}\right)\right]\pi_{\eta_0}(\epsilon-1). \quad (3.141)$$

Then, using (3.53) and the structure of the perturbed policy η_δ in (3.135), the derivative of (3.139), (3.140) and (3.141) with respect to δ is given by

$$\left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon+1)}{\mathrm{d}\delta}\right|_{\delta=0} = \chi(\epsilon < e_{\max}-1)\frac{1}{\bar{\eta}_0(\epsilon+1)}r'(0)\pi_{\eta_0}(\epsilon+1), \quad (3.142)$$
$$\left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon)}{\mathrm{d}\delta}\right|_{\delta=0} = \left(\frac{r'(0)}{\eta_0(\epsilon+1)\left(1-\chi(\epsilon < e_{\max}-1)\eta_0(\epsilon+1)\right)} - \frac{1}{\bar{\eta}_0(\epsilon)}\right)\pi_{\eta_0}(\epsilon), \quad (3.143)$$
$$\left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon-1)}{\mathrm{d}\delta}\right|_{\delta=0} = \left(\frac{r'(0)}{\eta_0(\epsilon+1)\left(1-\chi(\epsilon < e_{\max}-1)\eta_0(\epsilon+1)\right)} - \frac{1}{\eta_0(\epsilon)\bar{\eta}_0(\epsilon)} + \frac{1}{\bar{\eta}_0(\epsilon-1)}h'(0)\right)$$
$$\times \pi_{\eta_0}(\epsilon-1). \quad (3.144)$$

The terms $h'(0)$ and $r'(0)$ in (3.138), (3.142), (3.143) and (3.144) are computed as follows. If $\epsilon > 1$, η_δ is such that $\pi_{\eta_\delta}(e_{\max}) = \pi_{\eta_0}(e_{\max})$ and $\pi_{\eta_\delta}(0) = \pi_{\eta_0}(0)$. Therefore, from (3.52) with $e = e_{\max} - 1$ we obtain

$$\prod_{i=0}^{e_{\max}-1} W_{\eta_\delta}(i) = \prod_{i=0}^{e_{\max}-1} W_{\eta_0}(i). \quad (3.145)$$

By computing the derivative of each side of the above expression with respect to δ, we obtain

$$\sum_{k=0}^{e_{\max}-1} \frac{\left.\frac{\mathrm{d}W_{\eta_\delta}(k)}{\mathrm{d}\delta}\right|_{\delta=0}}{W_{\eta_0}(k)} = 0, \quad (3.146)$$

and using the structure of the policy η_δ in (3.135), we obtain, after algebraic manipulation,

$$h'(0) = \frac{\bar{\eta}_0(\epsilon-1)\eta_0(\epsilon-1)}{\bar{\eta}_0(\epsilon)\eta_0(\epsilon)} - r'(0)\frac{\bar{\eta}_0(\epsilon-1)\eta_0(\epsilon-1)}{\eta_0(\epsilon+1)\left(1-\chi(\epsilon<e_{\max}-1)\eta_0(\epsilon+1)\right)}. \tag{3.147}$$

Note that the above expression holds also for $\epsilon = 1$. In fact, in this case we have $\eta_0(\epsilon-1) = 0$, hence we obtain $h'(0) = 0$, which is consistent with the fact that $h(\delta) = 0$ for $\delta > 0$. Moreover, by normalization we have $\sum_e \pi_{\eta_\delta}(e) = \sum_e \pi_{\eta_0}(e) = 1$. Using the fact that $\pi_{\eta_\delta}(e) = \pi_{\eta_0}(e)$ for $e < \epsilon - 1$ and $e_\epsilon + 1$, we then obtain

$$\pi_{\eta_\delta}(\epsilon-1) + \pi_{\eta_\delta}(\epsilon) + \pi_{\eta_\delta}(\epsilon+1) = \pi_{\eta_0}(\epsilon-1) + \pi_{\eta_0}(\epsilon) + \pi_{\eta_0}(\epsilon+1). \tag{3.148}$$

From (3.52), using the fact that $\pi_{\eta_\delta}(e_{\max}) = \pi_{\eta_0}(e_{\max})$, (3.148) yields

$$\begin{aligned}&\frac{1}{\prod_{i=\epsilon-1}^{e_{\max}-1} W_{\eta_\delta}(i)} + \frac{1}{\prod_{i=\epsilon}^{e_{\max}-1} W_{\eta_\delta}(i)} + \frac{1}{\prod_{i=\epsilon+1}^{e_{\max}-1} W_{\eta_\delta}(i)}\\ &= \frac{1}{\prod_{i=\epsilon-1}^{e_{\max}-1} W_{\eta_0}(i)} + \frac{1}{\prod_{i=\epsilon}^{e_{\max}-1} W_{\eta_0}(i)} + \frac{1}{\prod_{i=\epsilon+1}^{e_{\max}-1} W_{\eta_0}(i)}.\end{aligned} \tag{3.149}$$

By computing the derivative of each side above, using (3.53) and the structure of policy η_δ in (3.135), we obtain, after algebraic manipulation,

$$\begin{aligned}&r'(0)\frac{1-\chi(\epsilon=e_{\max}-1)\eta_0(\epsilon+1)}{\eta_0(\epsilon+1)\bar{\eta}_0(\epsilon+1)W_{\eta_0}(\epsilon)}\left(1+\frac{1}{W_{\eta_0}(\epsilon-1)}+\chi(\epsilon<e_{\max}-1)\eta_0(\epsilon+1)W_{\eta_0}(\epsilon)\right)\\ &-\frac{1}{W_{\eta_0}(\epsilon)}\frac{1}{\eta_0(\epsilon)\bar{\eta}_0(\epsilon)}\left(\eta_0(\epsilon)+\frac{1}{W_{\eta_0}(\epsilon-1)}\right)+h'(0)\frac{1}{W_{\eta_0}(\epsilon-1)}\frac{1}{W_{\eta_0}(\epsilon)}\frac{1}{\bar{\eta}_0(\epsilon-1)}=0.\end{aligned} \tag{3.150}$$

Finally, by replacing (3.147) into (3.150), solving for $r'(0)$, and then substituting $r'(0)$ into (3.147) to obtain $h'(0)$, we obtain

$$r'(0) = \frac{1-\chi(\epsilon<e_{\max}-1)\left(\beta+\bar{\beta}\eta_0(\epsilon+1)\right)}{\beta+(\bar{\beta}-\beta\chi(\epsilon<e_{\max}-1))\eta_0(\epsilon)}\frac{\eta_0(\epsilon+1)}{\bar{\eta}_0(\epsilon)}, \tag{3.151}$$

$$h'(0) = \frac{\beta}{\eta_0(\epsilon)}\frac{\bar{\eta}_0(\epsilon-1)\eta_0(\epsilon-1)}{\beta+(\bar{\beta}-\beta\chi(\epsilon<e_{\max}-1))\eta_0(\epsilon)}. \tag{3.152}$$

Note that both $h'(0) \geq 0$ and $r'(0) > 0$, which is consistent with the fact that $h(\delta) \geq 0$ and $r(\delta) > 0$.

Substituting these expressions into (3.142)-(3.144), we obtain

$$\left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon+1)}{\mathrm{d}\delta}\right|_{\delta=0} = \chi(\epsilon < e_{\max}-1)\frac{\beta}{\beta+(1-2\beta)\eta_0(\epsilon)}\pi_{\eta_0}(\epsilon), \tag{3.153}$$

$$\left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon)}{\mathrm{d}\delta}\right|_{\delta=0} = \frac{\bar{\beta}-\chi(\epsilon < e_{\max}-1)\beta}{\beta+(\bar{\beta}-\chi(\epsilon < e_{\max}-1)\beta)\eta_0(\epsilon)}\pi_{\eta_0}(\epsilon), \tag{3.154}$$

$$\left.\frac{\mathrm{d}\pi_{\eta_\delta}(\epsilon-1)}{\mathrm{d}\delta}\right|_{\delta=0} = -\frac{\bar{\beta}}{\beta+(\bar{\beta}-\chi(\epsilon < e_{\max}-1)\beta)\eta_0(\epsilon)}\pi_{\eta_0}(\epsilon). \tag{3.155}$$

Finally, by substituting the above expressions in (3.138), we obtain

$$\begin{aligned}\left.\frac{\mathrm{d}Z_\lambda(\eta_\delta)}{\mathrm{d}\delta}\right|_{\delta=0} &\propto -\bar{\beta}\left[z_\lambda(\eta_0(\epsilon-1)) - \eta_0(\epsilon-1)z'_\lambda(\eta_0(\epsilon-1))\right] - \beta z'_\lambda(\eta_0(\epsilon)) \\ &+ (\bar{\beta}-\beta\chi(\epsilon < e_{\max}-1))\left[z_\lambda(\eta_0(\epsilon)) - \eta_0(\epsilon)z'_\lambda(\eta_0(\epsilon))\right] + \chi(\epsilon = e_{\max}-1)\frac{\beta}{\bar{\beta}}z'_\lambda(\eta_0(\epsilon+1)) \\ &+ \chi(\epsilon < e_{\max}-1)\beta\left[z_\lambda(\eta_0(\epsilon+1)) + z'_\lambda(\eta_0(\epsilon+1))\bar{\eta}_0(\epsilon+1)\right] \triangleq w(\eta_0(\epsilon-1),\eta_0(\epsilon),\eta_0(\epsilon+1)).\end{aligned} \tag{3.156}$$

We now show that the $w(\eta_0(\epsilon-1),\eta_0(\epsilon),\eta_0(\epsilon+1)) > 0$, thus proving P1). The derivative of $w(\eta_0(\epsilon-1),\eta_0(\epsilon),\eta_0(\epsilon+1))$ with respect to $\eta_0(\epsilon-1)$ is negative, since $z_\lambda(x)$ is strictly concave. Then, since $\eta(\epsilon-1) < \eta(\epsilon)$, we obtain

$$\begin{aligned}&w(\eta_0(\epsilon-1),\eta_0(\epsilon),\eta_0(\epsilon+1)) > w(\eta_0(\epsilon),\eta_0(\epsilon),\eta_0(\epsilon+1)) \\ =&-\beta\chi(\epsilon < e_{\max}-1)[z_\lambda(\eta_0(\epsilon)) + \bar{\eta}_0(\epsilon)z'_\lambda(\eta_0(\epsilon))] - \beta\chi(\epsilon = e_{\max}-1)z'_\lambda(\eta_0(\epsilon)) \\ &+\chi(\epsilon = e_{\max}-1)\frac{\beta}{\bar{\beta}}z'_\lambda(\eta_0(\epsilon+1)) + \chi(\epsilon < e_{\max}-1)\beta\left[z_\lambda(\eta_0(\epsilon+1)) + z'_\lambda(\eta_0(\epsilon+1))\bar{\eta}_0(\epsilon+1)\right].\end{aligned} \tag{3.157}$$

Similarly, the derivative of $w(\eta_0(\epsilon),\eta_0(\epsilon),\eta_0(\epsilon+1))$ with respect to $\eta_0(\epsilon)$ is positive. Then, since $\eta(\epsilon) \geq \eta(\epsilon+1)$, we obtain

$$\begin{aligned}&w(\eta_0(\epsilon-1),\eta_0(\epsilon),\eta_0(\epsilon+1)) > w(\eta_0(\epsilon),\eta_0(\epsilon),\eta_0(\epsilon+1)) \geq w(\eta_0(\epsilon+1),\eta_0(\epsilon+1),\eta_0(\epsilon+1)) \\ =&\chi(\epsilon = e_{\max}-1)\frac{\beta^2}{\bar{\beta}}z'_\lambda(\eta_0(\epsilon+1)) \geq 0,\end{aligned} \tag{3.158}$$

where the last inequality follows from the fact that, if $\epsilon < e_{\max}-1$, then $w(\eta_0(\epsilon+1),\eta_0(\epsilon+1),\eta_0(\epsilon+1)) = 0$, and, if $\epsilon = e_{\max}-1$, then $w(\eta_0(\epsilon+1),\eta_0(\epsilon+1),\eta_0(\epsilon+1)) > 0$ since $\eta_0(\epsilon+1) < x^*$ and $z'_\lambda(\eta_0(\epsilon+1)) > 0$. P1) is thus proved.

Finally, P2) is proved by combining P1) with the fact that $\eta(1) \leq \eta_{\mathrm{low}}$ and $\eta(e_{\max}) \geq \eta_{\mathrm{up}}$ are

strictly suboptimal, yielding, for the optimal policy,

$$0 < \eta_{\text{low}} < \eta(1) < \eta(2) < \cdots < \eta(e_{\max}) < \eta_{\text{up}} \leq x^*. \tag{3.159}$$

□

Chapter 4

Battery-lifetime maximization in Energy Harvesting Wireless Sensor Networks

4.1 Introduction

In the previous chapter, we have investigated optimal energy management policies for energy harvesting devices. A common assumption employed in the previous models and in the literature is that the rechargeable battery used to store the incoming ambient energy, and from which energy is drawn to power the device, is *ideal* and not subject to degradation phenomena, *i.e.*, it can operate perpetually without incurring a performance degradation.

In reality, batteries involve more complex mechanisms than just storing and drawing energy on-demand and without side effects. The focus of this chapter is on degradation effects, which cause the storage capability of a battery to diminish over time, *depending* on how the battery is used [66]. Degradation phenomena due to deep discharge are particularly strong for Lithium-Ion (Li-Ion) batteries, which represent the reference case of rechargeable batteries in consumer electronics. Importantly, it is recognized that the deeper the discharge of the battery, the *faster* the degradation. Thus, for example, an appropriate approach to enhancing the battery lifetime could be to have very frequent and shallow discharge periods, compatibly with the operating constraints of the network and the intermittent nature of the ambient energy supply. In contrast, performing deep discharge cycles, *e.g*, in time intervals during which ambient energy is scarcely available, should be avoided as it is detrimental to battery lifetime.

In an Energy Harvesting system, the ambient energy source often provides most of the energy

within certain periods of time, during which the on-board battery is recharged. In the remaining periods, little or no energy is available from the source, and the on-board battery is partially or totally discharged, depending on the load demand. The charge/discharge process of the battery is called *cycling*, and the percentage amount D of energy withdrawn from the battery during discharge, with respect to its nominal capacity, is termed *Depth of Discharge* (DoD). In a photovoltaic scavenger, for instance, battery cycling is determined on a daily basis by the availability of solar energy. Other energy sources, such as RF, thermal or mechanical may present different trends. In general, the target application and deployment scenario of the WSN play an important role in determining the cycling period and its degree of randomness. Denoting with C_0 the nominal battery capacity in milliampere-hours (mAh) and with $E(N_{\mathrm{cyc}})$ the total energy delivered by the battery after N_{cyc} cycles at DoD D, one might expect that

$$E(N_{\mathrm{cyc}}) = N_{\mathrm{cyc}} \cdot C_0 \cdot D. \tag{4.1}$$

Two fundamental facts, however, complicate the deceptively simple scenario implied by (4.1). First, a rechargeable battery has a finite *cycle life*, *i.e.*, it cannot cycle indefinitely due to irreversible degradation mechanisms, which ultimately reduce C_0 to unrecoverable levels [66]. Manufacturers typically define the battery cycle life N_{cyc} as the number of cycles a battery delivers at DoD $D = 1$ before C_0 drops below a given threshold, *e.g.*, 80% or 50% of the initial value [67]. Secondly, the foregoing degradation process is strongly dependent on *how* the battery is cycled. More precisely, shallow DoDs result in a slower degradation of C_0 and ultimately in increased cycle life [66, 68–70]. For instance, a microbattery rated with $N_{\mathrm{cyc}} = 100$ cycles at 100% DoD may last up to $N_{\mathrm{cyc}} = 1000$ cycles at 20% DoD, indicating that roughly twice the energy is extracted from the battery in the latter case [67]. A simple heuristic model for the N_{cyc} vs. D dependence, which captures the ongoing battery degradation, is

$$N_{\mathrm{cyc}}(D) = N_{\mathrm{cyc},0} \cdot e^{\alpha(1-D)}, \tag{4.2}$$

where $N_{\mathrm{cyc},0}$ represents the cycle life at 100% DoD, and α is a characteristic constant of the battery. Exponential-based models like (4.2) have been proven to be a good fit for data from a rather wide range of battery chemistries and sizes [68–71]. Eq. (4.2) may therefore be taken as representative also for microbatteries targeted for low-power equipment. Note, however, that different $N_{\mathrm{cyc}}(D)$ re-

lationships could be employed depending on the available experimental data and the desired accuracy.

Acknowledging the degradation of the battery capacity and the dependence of N_{cyc} on D open up intriguing options for more advanced energy-aware policies, which are the main focus of this work, and represent an important step towards the realistic characterization of rechargeable batteries and, by extension, of WSNs and their management strategies. In this chapter, the foregoing qualitative discussion is formulated within the framework of a stochastic model which captures the essential features of the problem, such as source pseudo-periodicity, battery cycling and cycle life vs DoD dependence found in commercial microbatteries.

Remarkably, a strong suit of the approach taken in this chapter is to join two different perspectives, namely, those of microelectronics and network engineering. Microelectronic characterizations of batteries often give a very detailed parametric description but fail to provide a behavioral analysis over time and in a broader context. Conversely, network models may be entirely flawed if they do not properly account for a correct physical characterization. In this sense, we aim at bridging the gap between these two approaches.

In the literature, a limited number of works attempted to model realistic battery imperfections and non-idealities, and their impact on the performance of harvesting based devices and networks. In this context, the offline model considered in [72], where energy arrivals are known non-causally at the controller, includes battery leakage effect, and accommodates also the degradation of the battery capacity over time; however, it assumes that battery degradation is deterministic and not influenced by the charge/discharge policy; in contrast, we explicitly model this interaction. [73] models the non-linearity between the energy storage level and the power delivered by a battery. [74] presents a stochastic model to capture the recovery effect of electrochemical cells, based on which efficient battery management policies can be designed.

4.1.1 Contributions

We propose a stochastic Markov chain framework, suitable for policy optimization, which captures the degradation status of the battery and its interplay with the energy management policy, which determines the discharge/recharge process of the battery. Based on this stochastic model, we develop a stochastic optimization problem which accounts explicitly for the trade-off between battery lifetime and Quality of Service (QoS) of the EHS. We prove a general result of Markov chains, which

exploits the timescale separation between the communication time-slot of the device and the battery degradation process, and enables an efficient optimization.

The battery degradation parameters of the stochastic model are then extrapolated from manufacturer-provided data [67], based on the exponential battery degradation model (4.2). We show that this model fits well the behavior of real batteries for what concerns their storage capacity degradation over time. We demonstrate that a degradation-aware policy significantly improves the lifetime of the sensor compared to "greedy" policies, while guaranteeing the minimum required QoS. Finally, a simple heuristic policy, which never discharges the battery below a given threshold, is shown to achieve close-to-optimal performance in terms of battery lifetime.

4.1.2 Structure of the chapter

This chapter is organized as follows. In Sec. 4.2, we present the general stochastic framework and define the optimization problem, which is further developed in Sec. 4.3. In Sec. 4.4, we extrapolate the battery degradation probabilities from experimental data and models available in the literature. In Sec. 4.5, we provide numerical results. Sec. 4.6 concludes the chapter. The proof of the main theorem is provided in the appendix at the end of this chapter.

4.2 System Model

We consider a generalization of the single Energy Harvesting Sensor (EHS) model of the previous chapter. However, unlike it, the following model does not account for the importance of the current data packet V_k, *i.e.*, the importance is assumed constant over time.

Time is slotted, where slot k is the time interval $[kT, kT+T)$, $k \in \mathbb{Z}^+$, and T is the slot duration. The battery is modeled by a buffer with nominal capacity C_0, and is uniformly quantized to a number of energy levels, using a quantization step (energy quantum) $\Delta c \ll C_0$. The maximum number of quanta that can be stored at the nominal capacity is $e_{\max} = \left\lfloor \frac{C_0}{\Delta c} \right\rfloor$ and the set of possible energy levels is denoted by $\mathcal{E} = \{0, 1, \ldots, e_{\max}\}$. Due to the aforementioned battery degradation mechanisms, the nominal battery capacity $e_{\max}$ is not always entirely available, but rather decreases over time. Let $E_{\max}(k)$ be the battery capacity at time k, with $E_{\max}(k) \leq E_{\max}(k-1)$ and $E_{\max}(0) = e_{\max}$. Denote the (quantized) energy level of the battery at time k as E_k. The evolution of E_k is given by

$$E_{k+1} = \min\left\{[E_k - Q_k]^+ + B_k, E_{\max}(k+1)\right\}, \tag{4.3}$$

where $[x]^+ = \max\{x, 0\}$ and:

- $\{B_k\}$ is the energy harvesting process, taking values in $\mathcal{B} \triangleq \{0, 1, \ldots, B\}$. We define an underlying *energy harvesting state* process $\{A_k\}$, and we model it as an irreducible stationary Markov chain with transition probabilities $p_A(a_1|a_0) \triangleq \Pr(A_{k+1} = a_1|A_k = a_0)$ and steady state distribution $\pi_A(a)$, taking values in a finite state space $\mathcal{A}$. Given $A_k \in \mathcal{A}$, the energy harvest B_k is drawn from $\mathcal{B}$ according to the distribution $p_B(b|a) \triangleq \Pr(B_k = b|A_k = a)$. Then, we denote the average harvesting rate as $\beta \triangleq \sum_{a\in\mathcal{A}} \pi_A(a) \sum_{b\in\mathcal{B}} b p_B(b|a)$. We assume that a new energy quantum harvested in slot k can only be used in a later slot.

- $\{Q_k\}$ is the action process, which is governed by the EHS controller, as detailed in Sec. 4.2.1, and takes values in $\mathcal{Q} \triangleq \{0\} \cup \{Q_{\min}, \ldots, Q_{\max}\}$. $Q_{\min}$ and $Q_{\max}$ represent a minimum and maximum load requirements, respectively. Action $Q_k = 0$ accounts for the possibility to remain idle in time-slot k, due to either a controller's decision or energy outage.

We model the battery degradation process, which causes the battery capacity $E_{\max}(k)$ to diminish irreversibly over time, as follows. We define the *battery health state*, H_k, taking values in $\mathcal{H} \equiv \{0, 1, \ldots, H_{\max}\}$, where $H_{\max} > 0$. For a given $H_k = h$, the battery capacity at time k, *i.e.*, the total amount of energy delivered by a fully charged battery over a discharge phase, is given by

$$E_{\max}(k) = \left\lfloor \frac{h}{H_{\max}} e_{\max} \right\rfloor, \tag{4.4}$$

and the set of available energy levels is denoted by $\mathcal{E}(h) = \left\{0, 1, \ldots, \left\lfloor \frac{h}{H_{\max}} e_{\max} \right\rfloor\right\}$. We assume that $\{\text{History up to time } k-1\} \to (H_k, E_k) \to H_{k+1}$ forms a Markov chain, *i.e.*, H_{k+1} is independent of the history up to time $k-1$, given (H_k, E_k). We denote the transition probability from health state $H_k = h$ to health state $H_{k+1} = h - 1$ as

$$p_H(h; e) \triangleq \Pr(H_{k+1} = h - 1|H_k = h, E_k = e). \tag{4.5}$$

Moreover, $\Pr(H_{k+1} = \tilde{h}|H_k = h, E_k = e) = 0$ if $\tilde{h} \notin \{h-1, h\}$, $\forall e \in \mathcal{E}(h)$, so that no transition is possible between two non-consecutive or to a higher health state. As a consequence, the probability of remaining in health state h is $1 - p_H(h; e)$. We further make the following assumptions on $p_H(h; e)$:

Assumption 1. a) $p_H(h; e) > 0,\ \forall h \in \mathcal{H}, e \in \mathcal{E}(h)$,

b) $p_H(h; e) \ll 1,\ \forall h \in \mathcal{H}, e \in \mathcal{E}(h)$,

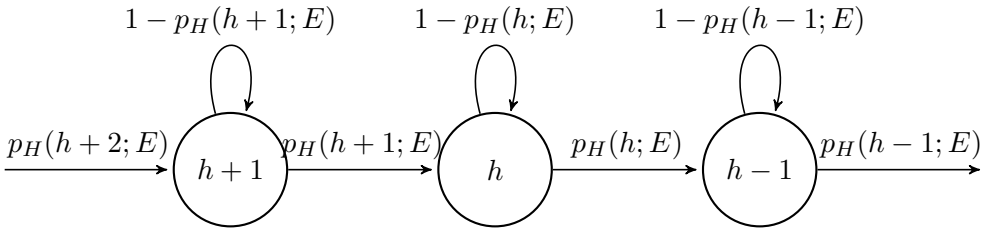


Figure 4.1. *Transition probabilities of health state H_k, which depend on the current energy level $E_k = E$*

c) $p_H(h_1; e_1) \geq p_H(h_2; e_2),\ \forall h_2 \geq h_1,\ e_2 \geq e_1$.

Ass. 1.a) implies that the battery health state will eventually reach state $H_k = 0$, so that the lifetime, defined in Def. 4.2.1 in Sec. 4.2.1, is finite; Ass. 1.b) expresses the fact that aging processes taking place in the battery operate over time scales that are much longer than the cycling period and the communication time-slot of the EHS; Ass. 1.c) means that the more discharged and degraded the battery, the faster the battery degradation process [66].

At time k, $\boldsymbol{S}_k = (E_k, H_k, A_{k-1})$ is the EHS state, taking values in the state space $\mathcal{S} \equiv \mathcal{E} \times \mathcal{H} \times \mathcal{A}$. In practice, $\boldsymbol{S}_k$ should be inferred and estimated from measurements of the battery energy level, capacity, and input energy flows. For simplicity, we assume that $\boldsymbol{S}_k$ is perfectly known to the EHS controller. Note that the harvesting state A_k is unknown at time k, as reflected by state $\boldsymbol{S}_k$, since B_k has not been observed yet, hence A_k can only be inferred from the a-priori transition probability $p_A(A_k|A_{k-1})$. On the other hand, the posterior distribution of A_{k-1} can be inferred recursively from the observed harvesting sequence $\{B_0, \ldots, B_{k-1}\}$, as in (3.5) of the previous chapter. For example, for a solar harvesting source, we may have $\mathcal{A} = \{\text{day}, \text{night}\}$. The state $A_{k-1} \in \mathcal{A}$ may then be estimated as, for appropriate choice of the threshold λ_{th} and of the window N,

$$\hat{A}_{k-1} = \begin{cases} \text{day} & \text{if } \frac{1}{N}\sum_{i=k-N}^{k-1} B_i > \lambda_{\text{th}}, \\ \text{night} & \text{otherwise.} \end{cases} \tag{4.6}$$

4.2.1 Policy definition and Optimization problem

Given $\boldsymbol{S}_k = (E_k, H_k, A_{k-1})$, the EHS controller determines $Q_k \in \mathcal{Q}$ at time k according to a given *policy* μ_{H_k}. Formally, μ_{H_k} is a probability measure on the action space $\mathcal{Q}$, parameterized by the state (E_k, A_{k-1}), *i.e.*, $\mu_h(q; e, a)$ is the probability of requesting q energy quanta from the battery, when operating in state $\boldsymbol{S}_k = (E_k, H_k, A_{k-1}) = (e, h, a)$.[1] Under any policy μ, the state process

[1]For the sake of maximizing a long-term average reward function of the state and action processes, it is sufficient to consider only state-dependent stationary policies [1].

$\{\boldsymbol{S}_k\}$ is a Markov chain, so that the whole decision problem is a Markov Decision Process [1].

The instantaneous reward accrued in time-slot k, in state $\boldsymbol{S}_k = (E_k, H_k, A_{k-1})$ under action Q_k, is defined as

$$g(Q_k, E_k) = \begin{cases} 0, & Q_k > E_k, \\ g^*(Q_k), & Q_k \leq E_k, \end{cases} \tag{4.7}$$

where $g^*(Q_k)$ is a concave increasing function of Q_k with $g^*(0) = 0$.[2] When the amount of energy requested by the controller exceeds that available in the battery (case $Q_k > E_k$), the task cannot be successfully completed, and the battery is depleted while no reward is earned.

We define the hitting times of the health states as

$$K_h = \min\{k \geq 0 : H_k = h\},\ h \in \mathcal{H}. \tag{4.8}$$

K_h is a random variable, which depends on the realization of $\{(B_k, Q_k, H_k)\}$. Given an initial state $\boldsymbol{S}_0 = (E_0, H_{\max}, A_{-1})$ and a policy μ, we define the *total average reward* $G_\mu^{\text{tot}}(h, \boldsymbol{S}_0)$, the *battery lifetime* $T_\mu(h, \boldsymbol{S}_0)$ and the *average reward per time-slot* $G_\mu(h, \boldsymbol{S}_0)$ of health state h as

$$G_\mu^{\text{tot}}(h, \boldsymbol{S}_0) = \mathbb{E}\Bigg[\sum_{k=K_h}^{K_{h-1}-1} g(Q_k, E_k) \,\Bigg|\, \boldsymbol{S}_0 \Bigg], \tag{4.9}$$

$$T_\mu(h, \boldsymbol{S}_0) = \mathbb{E}\left[K_{h-1} - K_h \mid \boldsymbol{S}_0 \right], \tag{4.10}$$

$$G_\mu(h, \boldsymbol{S}_0) = \frac{G_\mu^{\text{tot}}(h, \boldsymbol{S}_0)}{T_\mu(h, \boldsymbol{S}_0)}, \tag{4.11}$$

where the expectation is taken with respect to $\{(B_k, A_k, H_k, Q_k)\}$ and Q_k is drawn according to μ. In particular, $G_\mu^{\text{tot}}(h, \boldsymbol{S}_0)$ is the expected cumulative reward earned over health state h; $T_\mu(h, \boldsymbol{S}_0)$ is the expected number of time-slots spent in health state h; and $G_\mu(h, \boldsymbol{S}_0)$ represents the expected reward per time-slot accrued in health state h.

With these definitions at hand, let $\mathcal{G}^*$ be a minimum QoS requirement, which is met in health state h if $G_\mu(h, \boldsymbol{S}_0) \geq \mathcal{G}^*$. We give the following definition.

Definition 4.2.1. (Battery Lifetime) If $G_\mu(H_{\max}, \boldsymbol{S}_0) \geq \mathcal{G}^*$, the battery lifetime $T_\mu(\mathcal{G}^*, \boldsymbol{S}_0)$ under

[2] Note that such choice of a concave increasing reward function models many cases of interest, and is widely used in the literature, *e.g.*, see [46].

policy μ is defined as

$$T_\mu(\mathcal{G}^*, \boldsymbol{S}_0) = \sum_{h \geq h_\mu^*} T_\mu(h, \boldsymbol{S}_0), \tag{4.12}$$

$$\text{where } h_\mu^* = \max\{h : G_\mu(h, \boldsymbol{S}_0) < \mathcal{G}^*\} + 1 \tag{4.13}$$

is the index of the lowest health state in which the QoS is met. Otherwise, $T_\mu(\mathcal{G}^*, \boldsymbol{S}_0) = 0$.

The condition $G_\mu(H_{\max}, \boldsymbol{S}_0) \geq \mathcal{G}^*$ guarantees that the problem is feasible; otherwise, the lifetime is zero as there is no satisfactory reward even in the healthiest state. The lifetime is defined such that the QoS requirement $\mathcal{G}^*$ is guaranteed at each health state $h \geq h_\mu^*$, *i.e.*, $G_\mu(h, \boldsymbol{S}_0) \geq \mathcal{G}^*$. In particular, the QoS constraint inherently assumes that the battery degradation processes taking place in the battery operate over time scales which are much longer than the communication time-slot (Ass. 1.b)), so that the system approaches a steady state operation in each health state. For the lower health state $h_\mu^* - 1$, we have $G_\mu(h_\mu^* - 1, \boldsymbol{S}_0) < \mathcal{G}^*$, *i.e.*, the EHS can no longer sustain the QoS requirement, and battery failure is declared. Note that a QoS requirement on each health state $h \geq h_\mu^*$ is stricter than an *average* QoS requirement over the entire lifetime, defined as $\sum_{h \geq h_\mu^*} G_\mu^{\text{tot}}(h, \boldsymbol{S}_0) / \sum_{h \geq h_\mu^*} T_\mu(h, \boldsymbol{S}_0)$. The latter may induce policies that exhibit wide performance variability across the health states, as made clear in the following example.

Example 2. *Consider a system with $\mathcal{G}^* = 1.5$ and $H_{\max} = 2$, and a policy μ such that*

$$G_\mu(h, \boldsymbol{S}_0) = h,\ T_\mu(h, \boldsymbol{S}_0) = 10^6,\ \forall h \in \{0, 1, 2\}. \tag{4.14}$$

Then, according to Def. 4.2.1, we have $T_\mu(\mathcal{G}^*, \boldsymbol{S}_0) = 10^6$, since the QoS $\mathcal{G}^*$ can be supported only at health state 2. However, an average *QoS of*

$$\frac{G_\mu^{\text{tot}}(2, \boldsymbol{S}_0) + G_\mu^{\text{tot}}(1, \boldsymbol{S}_0)}{T_\mu(2, \boldsymbol{S}_0) + T_\mu(1, \boldsymbol{S}_0)} = 1.5 = \mathcal{G}^* \tag{4.15}$$

can be supported over a time-interval of duration 2×10^6, which is twice as long as $T_\mu(\mathcal{G}^*, \boldsymbol{S}_0)$, despite the fact that a poor performance is attained in health state 1.

The optimization problem at hand is to determine the optimal μ^* such that the battery lifetime is

maximized, under a given constraint on the minimum QoS $\mathcal{G}^*$, *i.e.*,

$$\mu^* = \arg\max_{\mu} T_\mu(\mathcal{G}^*, \boldsymbol{S}_0) = \arg\max_{\mu} \sum_{h \geq h^*_\mu} T_\mu(h, \boldsymbol{S}_0), \tag{4.16}$$

where h^*_μ is given in (4.13). The solution to (4.16) is carried out in the next section.

4.3 Optimization

In this section, we develop problem (4.16), showing that it can be recast as an independent Linear Program (LP) on each health state, under Ass. 1.b) on $p_H(h; e)$. The solution to the optimization problem relies on the timescale separation between the communication time-slot of the EHS and the battery degradation process, *i.e.*, the EHS achieves a steady state operation in each health state. In this light, we give the following definition.

Definition 4.3.1. (Steady State distribution of the non-absorbed chain) Assume that the EHS operates indefinitely at health state $h \in \mathcal{H}$ without being absorbed by the lower health state, *i.e.*, $p_H(h; e) = 0, \forall e \in \mathcal{E}(h)$. Denote the steady state distribution of $(e, a) \in \mathcal{E}(h) \times \mathcal{A}$ in health state h under policy μ_h as[3]

$$\pi^h_{\mu_h}(e, a) = \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K-1} P^{(k)}(e, a | \boldsymbol{S}_0), \tag{4.17}$$

where $\boldsymbol{S}_0 = (E_0, h, A_{-1})$ is the initial state and

$$P^{(k)}(e, a | \boldsymbol{S}_0) = \Pr\left(E_k = e, A_{k-1} = a | \boldsymbol{S}_0, p_H(h; \cdot) = 0\right).$$

We define the following quantities.

Definition 4.3.2. (Approximate reward per stage and lifetime of health state h)

$$\hat{G}_{\mu_h}(h) = \sum_{(e,a) \in \mathcal{E}(h) \times \mathcal{A}} \pi^h_{\mu_h}(e, a) \mathbb{E}_{\mu_h(\cdot; e, a)} \left[g(Q, e)\right], \tag{4.18}$$

$$\hat{T}_{\mu_h}(h) = \left(\sum_{(e,a) \in \mathcal{E}(h) \times \mathcal{A}} \pi^h_{\mu_h}(e, a) p_H(h; e) \right)^{-1}, \tag{4.19}$$

[3]We assume that μ_h induces a Markov chain with a single closed communicating class, so that $\pi^h_{\mu_h}(e, a)$ exists and is independent of $\boldsymbol{S}_0$ [23].

where $\mathbb{E}_{\mu_h(\cdot;e,a)}\left[g(Q,e)\right] = \sum_{q\in\mathcal{Q}} \mu_h(q;e,a)g(q,e)$ is the expected reward in state (e,a).

Remark: Note that $\pi^h_{\mu_h}$ in (4.17) is computed under the assumption that the EHS operates indefinitely in health state h, *i.e.*, $p_H(h;e) = 0,\ \forall e$, whereas the term $p_H(h;e)$ in (4.19) is the actual degradation probability. $\hat{G}_{\mu_h}(h)$ can be interpreted as the long-term average reward per time-slot in health state h, whereas $\hat{T}_{\mu_h}(h)^{-1}$ can be interpreted as the long-term average probability of making a transition to the lower health state $h-1$. Such observations are formalized in the following theorem.

Theorem 4.3.1. *Let* $p^*_H(h) = \max_e p_H(h;e)$. *For* $p^*_H(h) \to 0$,

$$G_\mu(h,\boldsymbol{S}_0) = \hat{G}_{\mu_h}(h) + \mathcal{O}(p^*_H(h)), \tag{4.20}$$

$$T_\mu(h,\boldsymbol{S}_0) = \hat{T}_{\mu_h}(h) + \mathcal{O}(1), \tag{4.21}$$

where $f(x) = \mathcal{O}(v(x))$ *for* $x \to 0$ *denotes a quantity such that* $\limsup_{x\to 0}\left|\frac{f(x)}{v(x)}\right| < +\infty$.

Proof. The proof is provided in Appendix 4.A as a general result of Markov chains. □

From Theorem 4.3.1, when $\max_e p_H(h;e) \ll 1$, the duration of health state h, $T_\mu(h,\boldsymbol{S}_0)$, can be approximated by $\hat{T}_{\mu_h}(h)$, up to a bounded additive factor. Since $T_\mu(h,\boldsymbol{S}_0) \to +\infty$ for $\max_e p_H(h;e) \to 0$ (in fact, the smaller $\max_e p_H(h;e)$, the less likely the health process to be absorbed by the lower health state $h-1$, hence the longer the amount of time spent in health state h), (4.21) is a good approximation. On the other hand, the average reward per time-slot in health state h, $G_\mu(h,\boldsymbol{S}_0)$, can be approximated by $\hat{G}_{\mu_h}(h)$ up to an additive factor, which decays to zero at least as quickly as $\max_e p_H(h;e)$. Both approximations are independent of the initial state $\boldsymbol{S}_0$, and solely depend on the steady state distribution (4.17) induced by policy μ_h, which is approached in each health state. Since $\max_e p_H(h;e) \ll 1$ by Ass. 1.b), we use Theorem 4.3.1 and replace (4.20-4.21) in (4.12), yielding

$$T_\mu(\mathcal{G}^*,\boldsymbol{S}_0) \simeq \sum_{h\geq h^*_\mu} \hat{T}_{\mu_h}(h), \text{ where } h^*_\mu = \max\left\{h : \hat{G}_{\mu_h}(h) < \mathcal{G}^*\right\} + 1. \tag{4.22}$$

Finally, substituting (4.22) in (4.16), we obtain the approximation

$$\mu^* = \arg\max_\mu \sum_{h\geq h^*_\mu} \hat{T}_{\mu_h}(h). \tag{4.23}$$

Note that $\hat{T}_{\mu_h}(h)$ and $\hat{G}_{\mu_h}(h)$ are independent of the policy $\mu_{\tilde{h}}$ for $\tilde{h} \neq h$. Therefore, (4.23) can be solved independently for each health state h, yielding the following algorithm.

Algorithm 4. • `INIT`: set $h = H_{\max}$, `REP=true`

• `WHILE REP=true AND h>0 SOLVE`

$$\mu_h^* = \arg\min_{\mu_h} \sum_{(e,a)\in\mathcal{E}(h)\times\mathcal{A}} \pi^h_{\mu_h}(e,a) p_H(h;e) \tag{4.24}$$
$$\text{s.t.} \sum_{(e,a)\in\mathcal{E}(h)\times\mathcal{A}} \pi^h_{\mu_h}(e,a) \left(\mathbb{E}_{\mu_h(\cdot;e,a)}\left[g(Q,e)\right] - \mathcal{G}^*\right) \geq 0.$$

If the problem is infeasible, set `REP=false`, $h^*_{\mu^*} = h+1$. If it is feasible and $h = 1$, set $h^*_{\mu^*} = 1$. Otherwise, update $h := h-1$. `END WHILE`

• `RETURN` the optimal policy $\mu^* = (\mu_h^*)_{h\geq h^*_{\mu^*}}$, with lifetime $T_{\mu^*}(\mathcal{G}^*, \boldsymbol{S}_0) \simeq \sum_{h\geq h^*_{\mu^*}} \hat{T}_{\mu_h^*}(h)$.

Remark: Step 2) is equivalent to

$$\mu_h^* = \arg\max_{\mu_h} \hat{T}_{\mu_h}(h), \;\; \text{s.t. } \hat{G}_{\mu_h}(h) \geq \mathcal{G}^*, \tag{4.25}$$

and is obtained by substituting the expressions of $\hat{T}_{\mu_h}(h)$ and $\hat{G}_{\mu_h}(h)$ (see Def. 4.3.2) in (4.25). It can be solved numerically via standard stochastic optimization tools, such as LP [1]. Thus, the optimal policy μ_h^* maximizes the lifetime (equivalently, it minimizes the long-term probability of battery degradation to the lower health state $h-1$) with a constraint on the minimum average QoS. Step 2) also determines $h^*_{\mu^*}$ in (4.13), for the optimal policy μ^*. Finally, in step 3) the optimal policy is found by concatenating the sub-policies μ_h^* for $h \geq h^*_{\mu^*}$, and the corresponding lifetime (4.2.1) is computed using (4.21). The main advantage of this approach over a standard approach which solves the original optimization problem (4.16) jointly is that (4.16) is decomposed into a sequence of independent sub-problems (4.24) for each health state h, thus reducing the overall computational burden.

4.4 Extrapolation of the Degradation Probabilities from Experimental Data

The battery degradation probabilities can be evaluated from manufacturer-provided data [67] by employing the deterministic, continuous time battery degradation model (4.2). These probabilities should be denoted as $p_H(h;e)$, depending on the health h and the energy e, as in (4.5). The dependence of $p_H(h;e)$ on h is quite difficult to capture; however, in our numerical evaluations we found that its effect is generally very mild. Even by neglecting it entirely, one can still obtain a very good match with manufacturer data. Therefore, we drop any dependence on h and we denote the degradation probability as $p_H(e)$, *i.e.*, just depending on e.

In Sec. 4.4.1, (4.2) is used to simulate an experiment where the battery is cyclically discharged and recharged at a given DoD until its capacity degrades to a fraction of the nominal capacity. First, the number of cycles as a function of the DoD and of the battery degradation rate function is derived. Then, the battery degradation rate function is found by matching the theoretical curve for the number of cycles to manufacturer data and the exponential model (4.2). In Sec. 4.4.2, the $p_H(e)$'s are found by matching the *deterministic* degradation times derived in Sec. 4.4.1 with the *average* degradation times in the proposed stochastic, discrete time model.

4.4.1 Deterministic Degradation Model

We employ model (4.2) for the relationship between number of cycles and DoD, where the constants $N_{\mathrm{cyc},0}$ and $\alpha > 0$ depend on the specific battery model employed. In particular, $N_{\mathrm{cyc}}(D)$ is counted until the battery capacity degrades to a fraction $x \in (0,1)$ of the initial capacity (*e.g.*, $x \in \{0.5, 0.8\}$), so that, in general, $N_{\mathrm{cyc},0}$ and α also depend on x.

Herein, we assume that the degradation process is a function of the instantaneous energy level of the battery only, as discussed in the introduction to this section, and is described by the *rate of capacity degradation function* $\rho(e\Delta c/C_0)$ (in $\mathrm{mAh/s}$) at the energy level $e\Delta c \in [0, C_0]$, where C_0 is the nominal capacity, Δc is the energy quantum and e is the energy level normalized to the quantum Δc. Then, if the battery operates at energy level $e\Delta c$ for δ seconds, its capacity degrades by $\delta\rho(e\Delta c/C_0)$ mAh. Moreover, we conjecture that, for proper coefficients $\theta > 0$, $\zeta > 0$,

$$\rho(e\Delta c/C_0) = \zeta e^{\theta(1-e\Delta c/C_0)}. \tag{4.26}$$

In the following analysis, and by simulation in Sec. 4.5, we show that this choice fits well the exponential model (4.2) for typical values of D (*e.g.*, $D \in [0.2, 1]$). Let $C_n,\ n \geq 0$ be the battery capacity at the beginning of the nth discharge/recharge cycle. In the nth cycle, the battery discharges from C_n to $C_n - C_0 D$ (with DoD D), and it then recharges from $C_n - C_0 D$ to C_{n+1}. Note that $C_{n+1} \leq C_n$, *i.e.*, the capacity at the end of the nth cycle cannot be larger than at the beginning of the cycle, due to irreversible degradation mechanisms.

The battery degradation in the nth cycle as a function of ρ and D is denoted by $\Delta_\rho(D, C_n) = C_n - C_{n+1}$. Assuming that $\Delta_\rho(D, C_n) \ll 2DC_0$, *i.e.*, the battery degradation is much smaller than the amount of energy exchanged by the battery over each cycle (this is a good approximation for typical values of D), and the discharge/recharge current is I, the duration of the nth discharge/recharge cycle is denoted by $T_n = [2DC_0 - \Delta_\rho(D, C_n)]/I \simeq 2DC_0/I$. The energy level over the nth cycle, $E_n(t)\Delta c$, where $t \in (0, T_n)$, evolves as

$$\text{Discharge phase: } E_n(t)\Delta c = C_n - It,\ t \in (0, T_n/2), \tag{4.27}$$

$$\text{Recharge phase: } E_n(t)\Delta c = C_n - DC_0 + I(t - T_n/2),\ t \in (T_n/2, T_n). \tag{4.28}$$

Moreover, due to the ongoing degradation, the instantaneous battery capacity in the nth cycle, denoted by $C_n(t),\ t \in (0, T_n)$, obeys

$$\frac{\mathrm{d}C_n(t)}{\mathrm{d}t} = C_n'(t) = -\rho\left(\frac{E_n(t)\Delta c}{C_0}\right),\ t \in (0, T_n), \tag{4.29}$$

with the boundary conditions $C_n(0) = C_n$, $C_n(T_n) = C_{n+1}$. By integrating the energy flows in one cycle, we then have

$$C_{n+1} = C_n + \int_0^{T_n/2} C_n'(\tau)\mathrm{d}\tau + \int_{T_n/2}^{T_n} C_n'(\tau)\mathrm{d}\tau, \tag{4.30}$$

and, substituting (4.29) in (4.30) and using the expression of ρ given in (4.26) and those for $E_n(t)$ given in (4.27) and (4.28) for the two integrals, we obtain

$$\Delta_\rho(D, C_n) = \frac{2C_0\zeta}{I\theta} e^{\theta(1 - C_n/C_0)}(e^{\theta D} - 1). \tag{4.31}$$

$N_{\mathrm{cyc}}(D)$ is equivalently defined as $N_{\mathrm{cyc}}(D) = \min\{n : C_n < xC_0\}$, since the number of cycles is counted until the battery capacity degrades to a fraction x of the nominal capacity. Herein, based

on the fact that the battery capacity slowly degrades from the nominal value C_0 to the target xC_0, and that the number of cycles to obtain a small capacity degradation $\mathrm{d}C \ll C_0$ from $C \in (0, C_0]$ to $C - \mathrm{d}C$ are $\mathrm{d}C/\Delta_\rho(D, C)$, we approximate $N_{\mathrm{cyc}}(D)$ with the integral expression

$$N_{\mathrm{cyc}}(D) \simeq \int_{xC_0}^{C_0} \frac{1}{\Delta_\rho(D, C)} \mathrm{d}C. \tag{4.32}$$

Substituting (4.31) in (4.32), we thus obtain

$$N_{\mathrm{cyc}}(D) = \left(\frac{I}{2\zeta} \frac{1 - e^{-\theta(1-x)}}{1 - e^{-\theta D}} \right) e^{-\theta D}. \tag{4.33}$$

Note that the term within the parentheses is a decreasing function of D, hence we obtain

$$N_{\mathrm{cyc}}(D) \geq \frac{I}{2\zeta} \frac{1 - e^{-\theta(1-x)}}{1 - e^{-\theta}} e^{-\theta D} \triangleq \hat{N}_{\mathrm{cyc}}(D), \tag{4.34}$$

where equality holds for $D = 1$. Finally, by approximating $N_{\mathrm{cyc}}(D)$ with its lower bound $\hat{N}_{\mathrm{cyc}}(D)$ and by matching this expression to the exponential model (4.2), yields

$$\alpha = \theta \text{ and } \zeta = \frac{I}{2N_{\mathrm{cyc},0}} \frac{1 - e^{-\alpha(1-x)}}{e^{\alpha} - 1} \text{ in (4.26).}$$

Remark 4.4.1. Note that the approximation (4.32) does not follow the exponential model (4.2). In particular, for $D \to 0$, in (4.32) we have $N_{\mathrm{cyc}}(D) \to \infty$. This is due to the fact that, in the derivation of (4.32), we have assumed that $\Delta_\rho(D, C_n) \ll 2DC_0$, *i.e.*, the DoD D is *large* with respect to the battery degradation in each cycle. However, this is a good approximation for typical values of D which the exponential model (4.2) has been fitted to [68–71], *e.g.*, $D \in [0.2, 1]$.

4.4.2 Stochastic Degradation Model

Based on the deterministic battery degradation model analyzed in the previous section, we now derive the degradation probabilities $p_H(e)$ for the stochastic model. To this end, we compute the deterministic time it takes for the battery to degrade from health state h, with capacity $\frac{h}{H_{\max}} C_0$, to the next lower health state $h - 1$, with capacity $\frac{h-1}{H_{\max}} C_0$. Then, we relate the *deterministic* degradation times to the *average* degradation times in the discrete-time stochastic model, and derive the corresponding transition probability.

Assume that the battery operates indefinitely at energy level $e\Delta c$ in the deterministic model stud-

ied in Sec. 4.4.1. The initial battery capacity is $C(0) = \frac{h}{H_{\max}} e_{\max} \Delta c$. From (4.29), the battery capacity as a function of time is given by $C(t) = C(0) - \rho(e\Delta c/C_0)t$ and degrades to the next health state with capacity $\frac{h-1}{H_{\max}} e_{\max} \Delta c$ over a time-interval of duration

$$T_{\det}(e) = \frac{e_{\max} \Delta c}{H_{\max} \rho(e\Delta c/C_0)}. \tag{4.35}$$

On the other hand, in the stochastic, discrete-time model, assuming that the battery operates indefinitely at energy level e, measured in energy quanta, the *average* amount of time (in s) it takes for the battery to degrade to the lower health state is

$$T_{\text{stoc}}(e) = \frac{\Delta t}{p_H(e)}, \tag{4.36}$$

where Δt is the time-slot duration. By forcing $T_{\text{stoc}}(e) = T_{\det}(e)$, we finally obtain the relation

$$p_H(e) = \gamma \exp\left\{\alpha\left(1 - \frac{e}{e_{\max}}\right)\right\}, \tag{4.37}$$

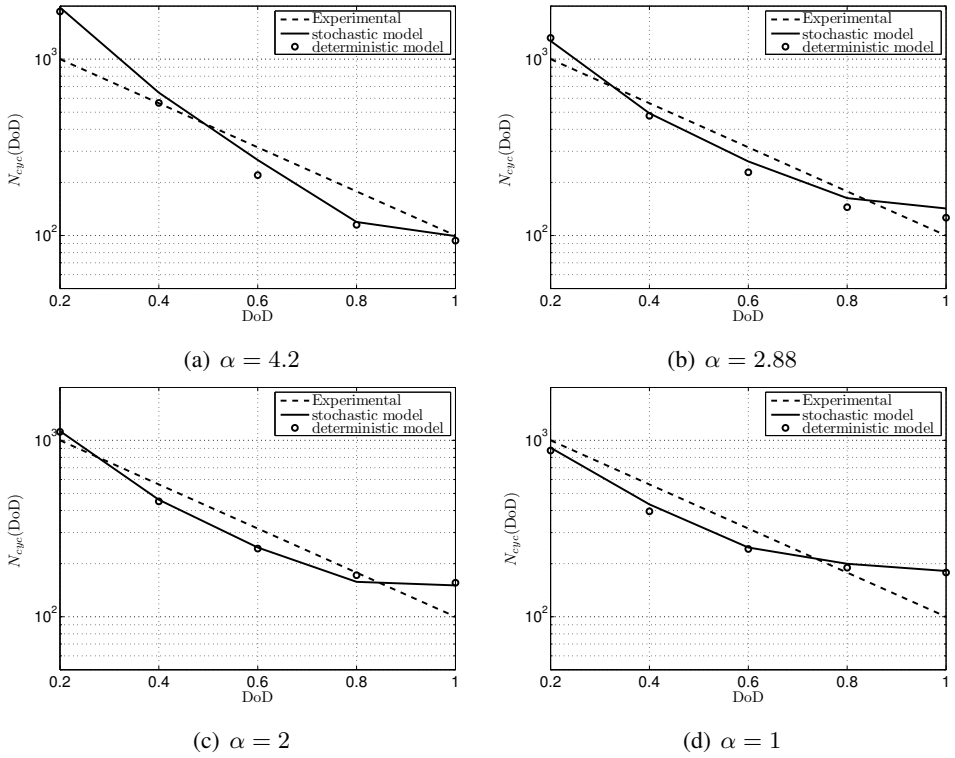
where $\gamma = \frac{\Delta t H_{\max}}{\Delta c e_{\max}} \zeta$ is a dimensionless constant. We note that (4.37) obeys Ass. 1.a) (as long as $\gamma \neq 0$) and Ass. 1.c) (since $\alpha > 0$). Moreover, if $\gamma \ll 1$, also Ass. 1.b) holds.

Remark 4.4.2. It is worth noting that the absolute value of γ does not affect the solution of the optimization problem (4.24), which, under the relationship (4.37), becomes

$$\begin{aligned} \mu_h^* = \arg\min_{\mu_h} & \sum_{(e,a)\in\mathcal{E}(h)\times\mathcal{A}} \pi_{\mu_h}^h(e,a) \exp\left\{\alpha\left(1 - \frac{e}{e_{\max}}\right)\right\} \\ \text{s.t.} & \sum_{(e,a)\in\mathcal{E}(h)\times\mathcal{A}} \pi_{\mu_h}^h(e,a) \left(\mathbb{E}_{\mu_h(\cdot;e,a)}\left[g(Q,e)\right] - \mathcal{G}^*\right) \geq 0. \end{aligned}$$

4.5 Numerical Results

In this section, we present some numerical results. In particular, we validate the proposed stochastic framework to model the battery degradation process, and we assess the performance of the proposed lifetime aware policies in terms of maximizing the battery lifetime, while guaranteeing a target QoS to the system. We consider a battery with capacity $e_{\max} = 500$ energy levels and $H_{\max} = 50$ health states. The parameter α, which determines the degradation probabilities $p_H(e)$ in (4.37), is obtained by interpolating the data-sheet values in [67] of Li-Ion rechargeable micro batteries, which



(a) $\alpha = 4.2$ (b) $\alpha = 2.88$

(c) $\alpha = 2$ (d) $\alpha = 1$

Figure 4.2. *Number of cycles versus DoD. The curve for the stochastic model is obtained by averaging the number of cycles over 10 iterations.*

may be envisioned for applications in WSNs. In particular, we refer to the battery type MS920SE, which is declared to provide 100 cycles at 100% DoD until the battery capacity degrades to 50% of the initial capacity C_0, and 1000 cycles at 20% DoD. Assuming the exponential relationship (4.2) yields $N_{\mathrm{cyc},0} = 100$ and $\alpha \simeq 2.88$, from which we compute the degradation probabilities $p_H(e)$, given by (4.37). As discussed in Sec. 4.4.2, the constant γ in (4.37) does not affect the optimization problem (4.24), hence we choose a small value $\gamma = 2.5 \cdot 10^{-5}$ so as to satisfy Ass. 1.b) and Theorem 4.3.1.

In Fig. 4.2, we validate the proposed stochastic model against the experimental curve (4.2) for the $N_{\mathrm{cyc}}(D)$ versus DoD D dependence, for the battery model considered. In particular, these curves are obtained by cyclically discharging and recharging the battery with different values of the DoD D. The curves associated with the *stochastic* model are obtained by employing the stochastic model proposed in this chapter to generate the health state process $\{H_k\}$, which determines the battery capacity via (4.4). The curves associated with the *deterministic* model, instead, are obtained by employing the deterministic degradation model developed in Sec. 4.4.1 to generate the battery degradation process. The number of cycles for a specific value of the DoD D and a specific model are counted until the

capacity degrades to 50% of the initial capacity C_0. We notice that there is a good match between the deterministic and stochastic models, which gives evidence of the fact that the proposed Markov model captures the fundamental behavior of real batteries for what concerns their storage capacity degradation over time. Moreover, the stochastic model exhibits a good fit to the experimental curve, which validates our analysis in Sec. 4.4. The value $\alpha = 2.88$ best matches the experimental curve (we have verified that $\alpha = 2.88$ minimizes the mean square error with respect to the experimental curve, in the logarithmic domain).

In the following figures, the underlying energy harvesting process $\{A_k\}$ is modeled as a two state Markov chain with state space $\mathcal{A} = \{\mathrm{G}, \mathrm{B}\}$ and transition probabilities $p_A(\mathrm{G}|\mathrm{G}) = p_A(\mathrm{B}|\mathrm{B}) = 0.96$, where G and B denote the "good" and "bad" harvesting states, respectively. In the "bad" state ($A_k = \mathrm{B}$), no energy is harvested, *i.e.*, $B_k = 0$; in the "good" state ($A_k = \mathrm{G}$), the harvested energy is $B_k = 20$ deterministically. The average harvesting rate is thus given by $\beta = 10$. In this case, we have a one-to-one mapping between A_k and B_k, so that, by measuring B_k, the state A_k is known exactly.

We employ the reward function $g^*(Q_k) = \log_2(1 + \sigma Q_k/\beta)$, with $\sigma = 10$, which models the Shannon capacity of the static Gaussian channel, where σ is an SNR scaling parameter [20]. The action space is $\mathcal{Q} = \{0, \ldots, 20\}$.

We consider the *Constant Load Lifetime Unaware Policy* (CLLUP), which supports a constant load of β energy quanta, irrespective of the energy level available in the battery, and remains idle under energy outage. This policy does not require communication between the EHS controller and the power processing unit (Fig. 3.2), since the current energy need not be known.

Moreover, we consider the *Lifetime Unaware Policy* (LUP), which greedily maximizes the average long-term reward (4.18) for the actual value of the battery capacity, without taking into account the impact of the policy on the battery lifetime. It is found via the Policy Iteration algorithm [1] as the solution of

$$\mu_h^* = \arg\max_{\mu_h} \hat{G}_{\mu_h}(h), \ \forall h \in \mathcal{H}. \tag{4.38}$$

This policy requires full knowledge of the current energy level, hence communication between the EHS controller and the power processing unit.

Finally, we consider the following policies, which explicitly take into account battery lifetime:

- *Lifetime Aware Optimal Policy* (LAOP): this is the optimal policy solution of problem (4.16), found via Algorithm 4.

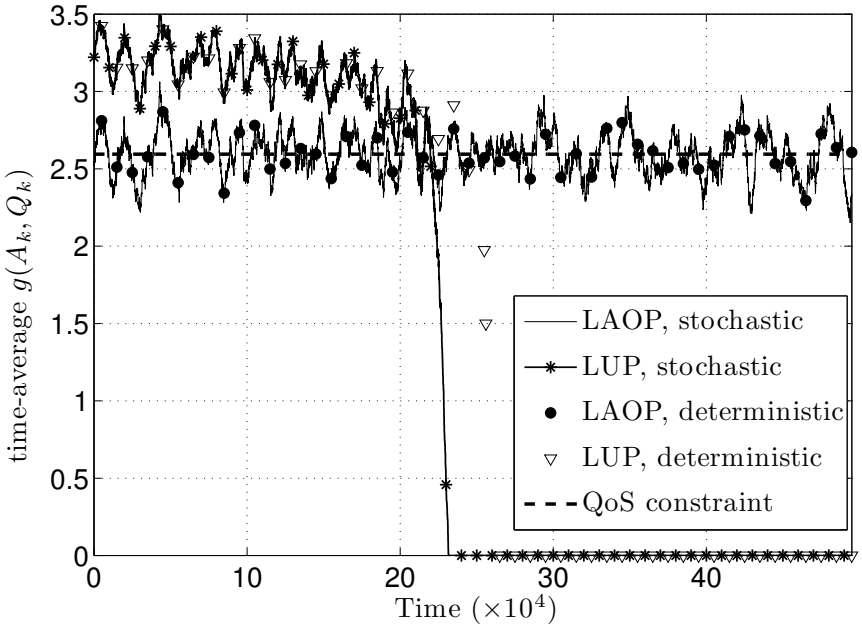


Figure 4.3. *Comparison via simulation of stochastic and deterministic degradation models. Each point in the curve is obtained by a moving-average window of* 5000 *time-slots. QoS requirement* $\mathcal{G}^* = 2.59$ *(corresponding to* 80% *of the maximum reward* $\max_{\mu_{H_{\max}}} \hat{G}_{\mu_{H_{\max}}}(H_{\max}) \simeq 3.24$ *in the maximum health state).*

- *Constant Load Lifetime Aware Policy* (CLLAP): This policy supports a constant load of β energy quanta, equal to the average harvesting rate, when the battery energy level is above a given DoD, and remains idle otherwise. If the battery capacity degrades to a value such that the required DoD cannot be supported anymore, battery failure is declared.

In the following plots, for a given policy and QoS $\mathcal{G}^*$, the battery lifetime is computed according to (4.12), using standard results on absorbing Markov Chains, see [23]. The corresponding *minimum reward* supported by policy μ over the battery lifetime is defined as $G_{\min}(\mu, \mathcal{G}^*) = \min_{h \geq h^*_\mu} G_\mu(h, \boldsymbol{S}_0)$, where h^*_μ and $G_\mu(h, \boldsymbol{S}_0)$ are defined in (4.13) and (4.11), respectively. The minimum reward represents the average reward per slot (averaged over a timescale much larger than the communication time-scale, but smaller than the battery degradation process) that is *guaranteed* over the entire battery lifetime.

To further validate the stochastic model proposed in this chapter, in Figs. 4.3 and 4.4 we plot the result of a simulation, where the battery degradation process follows either the stochastic model of Sec. 4.2, or the deterministic model of Sec. 4.4.1. However, notice that, in the latter case, the term *deterministic* is only referred to the fact that, in each time-slot, the battery capacity degrades by a deterministic quantity, which depends on the energy level, as in Sec. 4.4.1. On the other hand, the

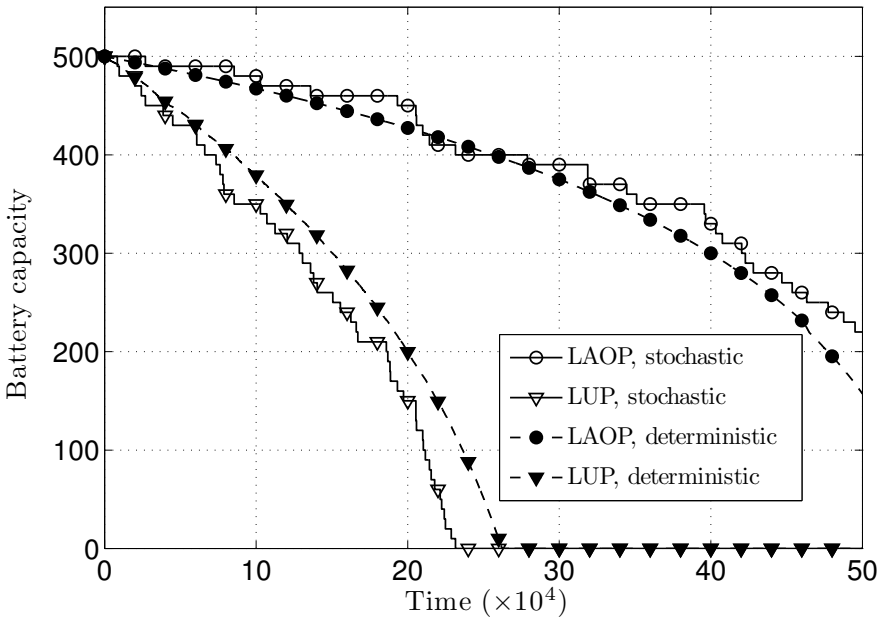


Figure 4.4. *Capacity degradation under the stochastic and deterministic degradation models. QoS requirement* $\mathcal{G}^* = 2.59$.

energy level is a stochastic process, induced by the stochastic energy arrival and decision processes. In particular, in Fig. 4.3, we plot the moving average curve associated with the reward sequence $\{g(Q_k, E_k)\}$, and, in Fig. 4.4, we plot the time-sequence of the battery capacity. We notice a good match between the curves associated with the deterministic and stochastic models. Moreover, as shown in Fig. 4.3, LUP achieves a larger reward than LAOP in the time-horizon $[0, 20 \times 10^4]$, where the battery capacity is larger than ~ 150 (Fig. 4.4). This is because LUP exploits all the available energy levels to earn the maximum reward, by performing deep charge/discharge cycles. However, such behavior quickly deteriorates the battery capacity, which decays to zero much faster than LAOP. In contrast, LAOP performs close to the QoS requirement, and it intelligently manages the battery to prolong its lifetime. Finally, notice that the time-average reward sequence exhibits fluctuations around its mean. This is due to the stochastic energy harvesting supply.

In Fig. 4.5, we plot the minimum reward $G_{\min}(\mu, \mathcal{G}^*)$ versus the corresponding battery lifetime normalized to the maximum lifetime, which is defined as the lifetime when the battery is always fully charged, so that battery degradation mechanisms are slower, according to our extrapolated model (4.37) and Ass. 1.c). We note that, for a given minimum guaranteed QoS (a value in the y-axis of the figure), LAOP achieves a significant gain in terms of battery lifetime with respect to the "greedy" policy LUP, which does not take into consideration battery degradation mechanisms. In particular,

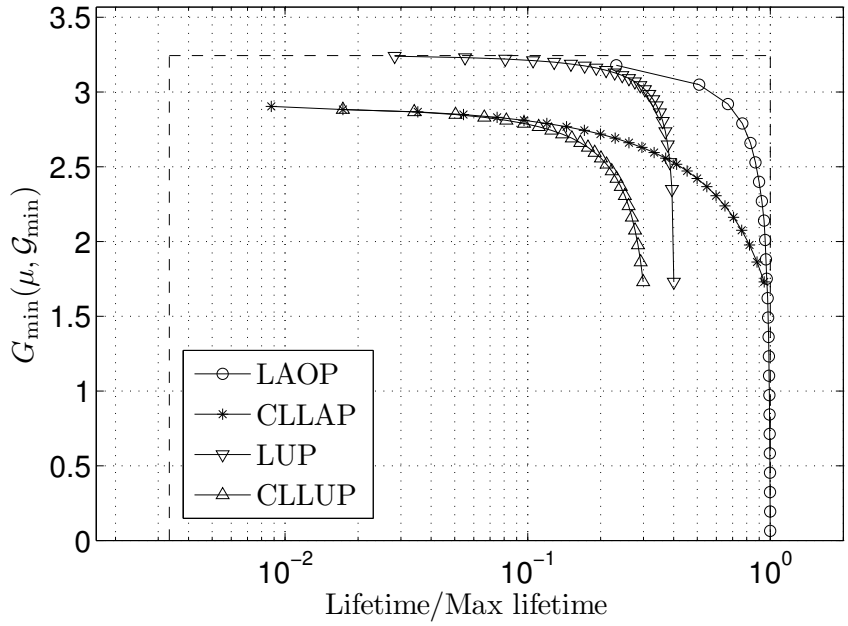


Figure 4.5. *Minimum reward over the battery lifetime versus normalized lifetime. The dashed lines represent the minimum and maximum lifetime and the maximum reward* $\max_{\mu_{H_{\max}}} \hat{G}_{\mu_{H_{\max}}}(H_{\max})$.

the lifetime is increased by a factor ~ 2.5. The same observation holds when comparing CLLAP and CLLUP. Moreover, although CLLAP incurs a loss with respect to LAOP, it provides a good heuristic to enhance the battery lifetime, that is, battery lifetime can be significantly increased by allowing only shallow battery discharges, and by avoiding battery discharge below a predetermined DoD value. Finally, for all policies, the longer the lifetime, the smaller the minimum reward attained. This is due to the inherent trade-off between lifetime and QoS. Namely, the battery lifetime is maximized by performing shallow charge/discharge cycles, which in turn considerably limits the usable energy levels, thus impairing the ability of the battery to filter out the fluctuations in the intermittent energy harvesting process, and to provide a satisfactory QoS over time. Conversely, the QoS is maximized by performing deep battery discharges, *e.g.*, during a long period of energy shortage, which inevitably shortens battery lifetime. This behavior is not captured by the models commonly used in the literature, which assume a perpetual battery operation, *e.g.*, [41, 42, 44, 45, 57, 60].

In Fig. 4.6, we plot the lifetime of each health state $h \in \mathcal{H}$, defined in (4.10) (lines). We also plot the lifetime approximation (4.19) (markers). We notice that the exact lifetime expression (4.10) is closely approximated by (4.19), as proved in Theorem 4.3.1 when $\max_q p_H(h; e) \ll 1$. Moreover, LAOP maximizes the lifetime of all health states. In fact, LAOP is found using Algorithm 4, which, in step 2), determines the optimal policy which minimizes, on each health state $h \in \mathcal{H}$, the steady

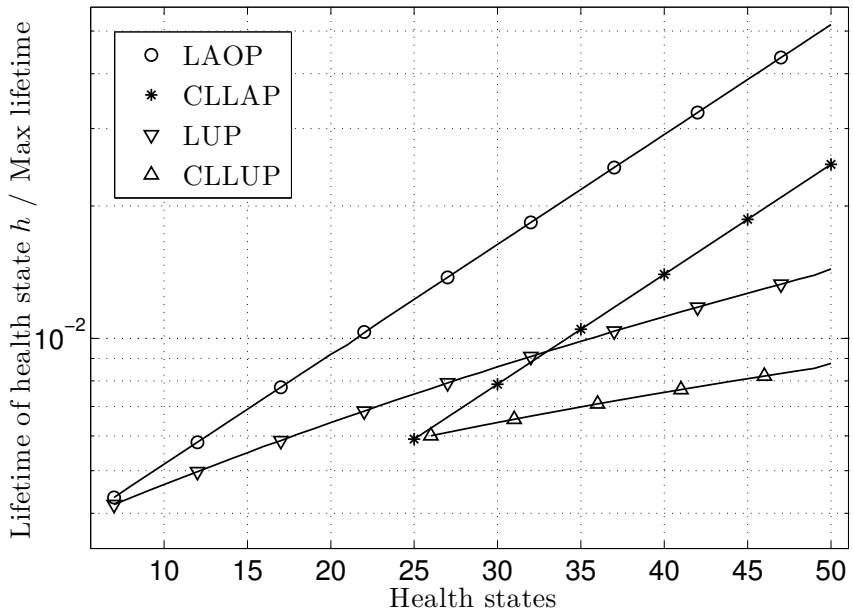


Figure 4.6. *Normalized lifetime of each health state. Exact lifetime (4.10) (lines). Approximation (4.19) (markers). QoS requirement* $\mathcal{G}^* = 2.59$.

state probability of degradation (equivalently, it maximizes the lifetime of health state h), subject to a QoS constraint $\mathcal{G}^*$. Conversely, a much shorter lifetime is attained by LUP in each health state, since this policy greedily maximizes the reward, without taking into account its impact on the battery degradation. Similar considerations hold for CLLAP and CLLUP. In general, the more degraded the battery, the faster the degradation. This behavior is consistent with Ass. 1.c).

Finally, in Fig. 4.7, we plot the cumulative steady state distribution of the energy levels, for the maximum health state $H_{\max}$, for LUP and LAOP, for different QoS requirements (corresponding, in sequence, to 80%, 84%, 88%, 92% and 96% of the maximum reward $\max_{\mu_{H_{\max}}} \hat{G}_{\mu_{H_{\max}}}(H_{\max}) \simeq 3.24$ in the maximum health state). We note that the steady state distribution of LUP, which does not take into account the ongoing battery degradation mechanisms, is spread over all the battery energy levels. In particular, this policy operates for a significant amount of time at low energy levels, thus inducing a fast battery degradation. Conversely, LAOP spreads the steady state distribution over the upper energy levels only, thus slowing down battery degradation. Moreover, the larger the QoS requirement, the more spread the steady state distribution under LAOP over lower energy levels. This is because deeper discharge cycles need to be performed, in order to meet a higher QoS requirement.

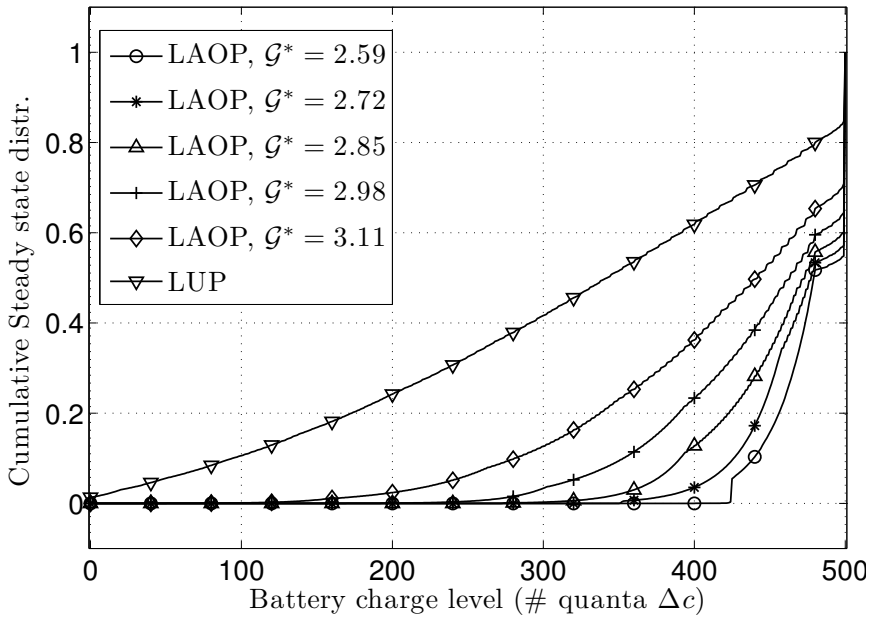


Figure 4.7. *Cumulative steady state distribution of energy levels at the maximum health state* $H_{\max}$.

4.6 Conclusions

We have analyzed the impact of battery management policies on the irreversible degradation of the storage capacity of realistic batteries, affecting the lifetime of harvesting based Wireless Sensor Networks. We have proposed a general framework, based on Markov chains and suitable for policy optimization, which captures the degradation status of the battery. The proposed stochastic battery degradation model has been extrapolated from manufacturer-provided data and realistic deterministic models proposed in the literature, and has been shown to fit well the behavior of real batteries for what concerns their storage capacity degradation over time. Note, however, that different battery degradation models can be easily accommodated in the proposed framework, depending on the available experimental data and the desired accuracy. Based on the proposed model, we have formulated the policy optimization problem as the maximization of the battery lifetime, subject to a minimum guaranteed QoS in each battery degradation status. We have shown that this problem can be solved efficiently by a sequential linear programming optimization algorithm over the degradation states of the battery. The numerical evaluation gives evidence of the fact that a lifetime-aware management policy significantly improves the lifetime of the sensor node with respect to a "greedy" operation policy, while guaranteeing the QoS.

Appendix 4.A: Proof of Theorem 4.3.1

Proof of Theorem 4.3.1. For the proof of the theorem, we present a general result of Markov chains. The relationship to the specific problem considered in this paper is provided at the end of the proof. Consider a finite Markov chain $\{Z_k\} \subseteq \mathcal{Z} \equiv \{1, 2, \ldots, N_\mathrm{t}+1\}$, where the state space $\mathcal{S}$ is partitioned into a set of transient states $\mathcal{Z}_\mathrm{t} \equiv \{1, \ldots, N_\mathrm{t}\}$ forming a communicating class, and the absorbing state $\mathcal{Z}_\mathrm{a} \equiv \{N_\mathrm{t} + 1\}$, with transition matrix

$$\mathbf{P}_\epsilon = \begin{bmatrix} (\mathbf{I}_{N_t} - \epsilon \mathbf{P}_\mathrm{a})\mathbf{P}_\mathrm{t} & \epsilon \mathbf{P}_\mathrm{a} \mathbf{1}_{N_t} \\ \mathbf{0}_{N_t}^T & 1 \end{bmatrix}, \tag{4.39}$$

where $\mathbf{0}_K$ is a $K \times 1$ vector with all entries equal to zero; $\mathbf{1}_K$ is an $K \times 1$ vector with all entries equal to one; $\mathbf{I}_K$ is the $K \times K$ identity matrix; $\mathbf{P}_\mathrm{t}$ is the $N_\mathrm{t} \times N_\mathrm{t}$ transition probability matrix associated with transitions in $\mathcal{Z}_\mathrm{t}$, given that the Markov chain is not absorbed by $\mathcal{Z}_\mathrm{a}$; $\mathbf{P}_\mathrm{a}$ is an $N_\mathrm{t} \times N_\mathrm{t}$ diagonal matrix with strictly positive diagonal elements, and $\epsilon \mathbf{P}_\mathrm{a}(i,i) \in (0,1)$ is the probability of moving from state i to the absorbing state $N_\mathrm{t} + 1$, where the scaling parameter ϵ can take any value in $(0, 1/\max_i \mathbf{P}_\mathrm{a}(i,i))$ (we will be interested in $\epsilon \to 0$). In the following, $\mathbf{e}_{1,K}$ denotes the first column of $\mathbf{I}_K$. Moreover, for convenience we drop the dependence of $\mathbf{0}_K$, $\mathbf{1}_K$, $\mathbf{I}_K$ and $\mathbf{e}_{1,K}$ on K in the notation whenever the size K can be deduced from the context.

We assume that $\mathbf{P}_\mathrm{t}$ is a regular stochastic matrix (*i.e.*, the associated Markov chain is irreducible and aperiodic). Therefore, $\mathbf{X}_\epsilon = (\mathbf{I} - \epsilon \mathbf{P}_\mathrm{a})\, \mathbf{P}_\mathrm{t}$ is a primitive matrix and, from the Perron-Frobenius Theorem [75], there is a real positive eigenvalue λ_ϵ of $\mathbf{X}_\epsilon$, with algebraic multiplicity 1, such that any other eigenvalue ξ of $\mathbf{X}_\epsilon$ has $|\xi| < \lambda_\epsilon$. Since $\mathbf{X}_\epsilon$ is continuous in ϵ, λ_ϵ is also continuous. We denote the corresponding right eigenvector as $\mathbf{v}_\epsilon$, *i.e.*,

$$(\mathbf{X}_\epsilon - \lambda_\epsilon \mathbf{I})\mathbf{v}_\epsilon = \mathbf{0}. \tag{4.40}$$

We normalize the eigenvector $\mathbf{v}_\epsilon$ so that the sum of its elements equals N_t[4], *i.e.*, $\mathbf{1}^T \mathbf{v}_\epsilon = N_t$, so that $\mathbf{v}_\epsilon$ is uniquely defined for each $\epsilon > 0$, and is continuous in ϵ. Since $\mathbf{X}_0 = \mathbf{P}_\mathrm{t}$ is a regular stochastic matrix, we have $\lambda_0 = 1$ and $\lambda_\epsilon < 1$ for $\epsilon > 0$. Moreover, $\mathbf{v}_0 = \mathbf{1}$ and there exists a unique $\pi_{\mathrm{t},\infty}$ such

[4]This is always possible since the Perron-Frobenius Theorem guarantees that there always exists an eigenvector associated to the eigenvalue λ_ϵ with all positive elements [75].

that $\pi_{\mathrm{t},\infty} = \pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{t}}$. We can thus write $\mathbf{X}_0$ as

$$\mathbf{X}_0 = \mathbf{U}_0\mathbf{D}_0\mathbf{U}_0^{-1}, \tag{4.41}$$

where $\mathbf{D}_0$ is the Jordan normal form of $\mathbf{X}_0$, and $\mathbf{U}_0$ is the matrix whose columns are the corresponding generalized eigenvectors. Without loss of generality, $\mathbf{D}_0$ is given by

$$\mathbf{D}_0 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{J}_0 \end{bmatrix}, \tag{4.42}$$

where $\mathbf{J}_0$ is a block diagonal matrix, whose diagonal blocks are given by the Jordan blocks corresponding to the eigenvalues of $\mathbf{X}_0$ inside the unit circle. Therefore, $\mathbf{U}_0\mathbf{e}_1 = \mathbf{1}$ and $\mathbf{e}_1^T\mathbf{U}_0^{-1} = \pi_{\mathrm{t},\infty}$, since $\mathbf{1}$ and $\pi_{\mathrm{t},\infty}$ are, respectively, the right and left eigenvectors of $\mathbf{X}_0$ associated to the eigenvalue 1.

Recall, from standard results on absorbing Markov Chains (see [23]), that the expected time until absorption is given by

$$\mathcal{T}_\epsilon(\pi_{\mathrm{t},0}) = \pi_{\mathrm{t},0}\left(\mathbf{I} - \mathbf{X}_\epsilon\right)^{-1}\mathbf{1}, \tag{4.43}$$

where $\pi_{\mathrm{t},0}$ is an initial distribution over $\mathcal{Z}_{\mathrm{t}}$. Note that, when $\epsilon > 0$, the eigenvalues of $\mathbf{X}_\epsilon$ are all strictly inside the unit circle, so that $\mathbf{I} - \mathbf{X}_\epsilon$ is invertible and (4.43) is well defined. We prove that

$$\mathcal{T}_\epsilon(\pi_{\mathrm{t},0}) = \frac{1}{\epsilon\pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{a}}\mathbf{1}} + \mathcal{O}(1), \text{ for } \epsilon \to 0, \tag{4.44}$$

or equivalently, by definition of $\mathcal{O}(x)$,

$$\lim_{\epsilon\to 0}\left|\mathcal{T}_\epsilon(\pi_{\mathrm{t},0}) - \frac{1}{\epsilon\pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{a}}\mathbf{1}}\right| < \infty. \tag{4.45}$$

We have

$$\begin{aligned}
&\mathcal{T}_\epsilon(\pi_{\mathrm{t},0}) - \frac{1}{\epsilon\pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{a}}\mathbf{1}} \overset{(a)}{=} \pi_{\mathrm{t},0}\left(\mathbf{I} - \mathbf{X}_\epsilon\right)^{-1}\mathbf{1} - \frac{\pi_{\mathrm{t},0}\mathbf{1}}{\epsilon\pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{a}}\mathbf{1}} \\
&= \frac{1}{\epsilon\pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{a}}\mathbf{1}}\pi_{\mathrm{t},0}\left(\mathbf{I} - \mathbf{X}_\epsilon\right)^{-1}\left[\epsilon\mathbf{1}\pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{a}}\mathbf{1} - \left(\mathbf{I} - \mathbf{X}_\epsilon\right)\mathbf{1}\right] \\
&\overset{(b)}{=} \frac{1}{\pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{a}}\mathbf{1}}\pi_{\mathrm{t},0}\left(\mathbf{I} - \mathbf{X}_\epsilon\right)^{-1}\left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right)\mathbf{P}_{\mathrm{a}}\mathbf{1} = \pi_{\mathrm{t},0}\left(\mathbf{I} - \mathbf{X}_\epsilon\right)^{-1}\left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right)\mathbf{x},
\end{aligned} \tag{4.46}$$

where we have defined the vector $\mathbf{x} = (\pi_{\mathrm{t},\infty}\mathbf{P}_{\mathrm{a}}\mathbf{1})^{-1}\mathbf{P}_{\mathrm{a}}\mathbf{1}$. In step (a), we have used the fact that $1 = \pi_{\mathrm{t},0}\mathbf{1}$. In step (b), we have used the fact that $\mathbf{X}_\epsilon\mathbf{1} = (\mathbf{I} - \epsilon\mathbf{P}_{\mathrm{a}})\,\mathbf{P}_{\mathrm{t}}\mathbf{1} = (\mathbf{I} - \epsilon\mathbf{P}_{\mathrm{a}})\,\mathbf{1}$. Let

$$\mathbf{U}_\epsilon = \mathbf{U}_0 + (\mathbf{v}_\epsilon - \mathbf{1})\mathbf{e}_1^T. \tag{4.47}$$

Since $\mathbf{U}_0$ is invertible, there exists $\epsilon_{\mathrm{th}} > 0$ such that $\mathbf{U}_\epsilon$ is also invertible, for all $\epsilon \in (0, \epsilon_{\mathrm{th}})$, by continuity. For any such ϵ, we can thus write

$$\mathbf{X}_\epsilon = \mathbf{U}_\epsilon\mathbf{D}_\epsilon\mathbf{U}_\epsilon^{-1}, \text{ where } \mathbf{D}_\epsilon = \begin{bmatrix} \lambda_\epsilon & \mathbf{r}_\epsilon \\ \mathbf{0} & \mathbf{J}_\epsilon \end{bmatrix}, \tag{4.48}$$

and, using the fact that $\mathbf{e}_1^T[\mathbf{0}, \mathbf{I}]^T = \mathbf{0}^T$, hence $\mathbf{U}_\epsilon[\mathbf{0}, \mathbf{I}]^T = \mathbf{U}_0[\mathbf{0}, \mathbf{I}]^T$,

$$\begin{bmatrix} \mathbf{r}_\epsilon \\ \mathbf{J}_\epsilon \end{bmatrix} = \mathbf{U}_\epsilon^{-1}\mathbf{X}_\epsilon\mathbf{U}_\epsilon \begin{bmatrix} \mathbf{0}^T \\ \mathbf{I} \end{bmatrix} = \mathbf{U}_\epsilon^{-1}\mathbf{X}_\epsilon\mathbf{U}_0 \begin{bmatrix} \mathbf{0}^T \\ \mathbf{I} \end{bmatrix}. \tag{4.49}$$

Then, using (4.48) and the fact that $\mathbf{I} - \mathbf{X}_\epsilon = \mathbf{U}_\epsilon(\mathbf{I} - \mathbf{D}_\epsilon)\mathbf{U}_\epsilon^{-1}$, we obtain

$$\begin{aligned}(\mathbf{I} - \mathbf{X}_\epsilon)^{-1} &= \mathbf{U}_\epsilon \begin{bmatrix} 1 - \lambda_\epsilon & -\mathbf{r}_\epsilon \\ \mathbf{0} & \mathbf{I} - \mathbf{J}_\epsilon \end{bmatrix}^{-1} \mathbf{U}_\epsilon^{-1} = \mathbf{U}_\epsilon \begin{bmatrix} \frac{1}{1-\lambda_\epsilon} & \frac{1}{1-\lambda_\epsilon}\mathbf{r}_\epsilon(\mathbf{I} - \mathbf{J}_\epsilon)^{-1} \\ \mathbf{0} & (\mathbf{I} - \mathbf{J}_\epsilon)^{-1} \end{bmatrix} \mathbf{U}_\epsilon^{-1} \\ &= \frac{1}{1-\lambda_\epsilon}\mathbf{v}_\epsilon\mathbf{e}_1^T\mathbf{U}_\epsilon^{-1} + \frac{1}{1-\lambda_\epsilon}\mathbf{v}_\epsilon\mathbf{r}_\epsilon(\mathbf{I} - \mathbf{J}_\epsilon)^{-1}\,[\mathbf{0}, \mathbf{I}]\,\mathbf{U}_\epsilon^{-1} + \mathbf{U}_\epsilon[\mathbf{0}, \mathbf{I}]^T(\mathbf{I} - \mathbf{J}_\epsilon)^{-1}[\mathbf{0}, \mathbf{I}]\mathbf{U}_\epsilon^{-1}. \end{aligned} \tag{4.50}$$

In the last step, we have used the fact that $\mathbf{I} = \mathbf{e}_1\mathbf{e}_1^T + [\mathbf{0}, \mathbf{I}]^T[\mathbf{0}, \mathbf{I}]$ and $\mathbf{U}_\epsilon\mathbf{e}_1 = \mathbf{v}_\epsilon$, hence

$$\mathbf{U}_\epsilon = \mathbf{v}_\epsilon\mathbf{e}_1^T + \mathbf{U}_\epsilon[\mathbf{0}, \mathbf{I}]^T[\mathbf{0}, \mathbf{I}], \tag{4.51}$$

$$\mathbf{U}_\epsilon^{-1} = \mathbf{e}_1\mathbf{e}_1^T\mathbf{U}_\epsilon^{-1} + [\mathbf{0}, \mathbf{I}]^T[\mathbf{0}, \mathbf{I}]\mathbf{U}_\epsilon^{-1}, \tag{4.52}$$

$$\text{and } (\mathbf{I} - \mathbf{D}_\epsilon)^{-1} = \frac{1}{1-\lambda_\epsilon}\mathbf{e}_1\mathbf{e}_1^T + \frac{1}{1-\lambda_\epsilon}\mathbf{e}_1\mathbf{r}_\epsilon(\mathbf{I} - \mathbf{J}_\epsilon)^{-1}[\mathbf{0}, \mathbf{I}] + [\mathbf{0}, \mathbf{I}]^T(\mathbf{I} - \mathbf{J}_\epsilon)^{-1}[\mathbf{0}, \mathbf{I}]; \tag{4.53}$$

the result is then obtained by substituting these expressions, by expanding the products and by noting that $[\mathbf{0}, \mathbf{I}]\mathbf{e}_1 = \mathbf{0}$. Since $\mathbf{J}_0$ is the Jordan matrix corresponding to eigenvalues of $\mathbf{X}_0$ within the unit circle, $\mathbf{I} - \mathbf{J}_0$ is invertible, hence, by continuity, $\mathbf{I} - \mathbf{J}_\epsilon$ is invertible for sufficiently small ϵ. By

replacing (4.50) into (4.46), we thus get

$$\mathcal{T}_\epsilon(\pi_{\mathrm{t},0}) - \frac{1}{\epsilon \pi_{\mathrm{t},\infty} \mathbf{P}_\mathrm{a} \mathbf{1}} = A(\epsilon) + B(\epsilon) + C(\epsilon), \tag{4.54}$$

where we have defined

$$\begin{aligned} A(\epsilon) &= \tfrac{1}{1-\lambda_\epsilon} \pi_{\mathrm{t},0} \mathbf{v}_\epsilon \mathbf{e}_1^T \mathbf{U}_\epsilon^{-1} \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) \mathbf{x}, \\ B(\epsilon) &= \tfrac{1}{1-\lambda_\epsilon} \pi_{\mathrm{t},0} \mathbf{v}_\epsilon \mathbf{r}_\epsilon (\mathbf{I} - \mathbf{J}_\epsilon)^{-1} [\mathbf{0}, \mathbf{I}] \mathbf{U}_\epsilon^{-1} \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) \mathbf{x}, \\ C(\epsilon) &= \pi_{\mathrm{t},0} \mathbf{U}_\epsilon \left[\mathbf{0}, \mathbf{I}\right]^T (\mathbf{I} - \mathbf{J}_\epsilon)^{-1} \left[\mathbf{0}, \mathbf{I}\right] \mathbf{U}_\epsilon^{-1} \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) \mathbf{x}. \end{aligned} \tag{4.55}$$

We finally show that the limit of each term above exists and is finite for $\epsilon \to 0$, thus proving (4.45). Regarding the first term $A(\epsilon)$, since $\mathbf{e}_1^T \mathbf{U}_0^{-1} \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) = \pi_{\mathrm{t},\infty} \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) = \mathbf{0}^T$, we obtain

$$A(\epsilon) = \frac{1}{1-\lambda_\epsilon} \pi_{\mathrm{t},0} \mathbf{v}_\epsilon \mathbf{e}_1^T \left(\mathbf{U}_\epsilon^{-1} - \mathbf{U}_0^{-1}\right) \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) \mathbf{x}. \tag{4.56}$$

Moreover, from (4.47), we have

$$\mathbf{U}_\epsilon^{-1} - \mathbf{U}_0^{-1} = \mathbf{U}_\epsilon^{-1} (\mathbf{U}_0 - \mathbf{U}_\epsilon) \mathbf{U}_0^{-1} = -\mathbf{U}_\epsilon^{-1} (\mathbf{v}_\epsilon - \mathbf{1}) \mathbf{e}_1^T \mathbf{U}_0^{-1} = -\mathbf{U}_\epsilon^{-1} (\mathbf{v}_\epsilon - \mathbf{1}) \pi_{\mathrm{t},\infty}. \tag{4.57}$$

Substituting (4.57) in (4.56), we obtain $A(\epsilon) = 0$, since $\pi_{\mathrm{t},\infty} \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) = \mathbf{0}^T$.

For the second term $B(\epsilon)$, substituting the expression of $\mathbf{r}_\epsilon = \mathbf{e}_1^T \mathbf{U}_\epsilon^{-1} \mathbf{X}_\epsilon \mathbf{U}_0 [\mathbf{0}, \mathbf{I}]^T$ given by (4.49) into (4.55), and using the fact that $\mathbf{e}_1^T \mathbf{U}_0^{-1} \mathbf{X}_0 \mathbf{U}_0 \left[\mathbf{0}, \mathbf{I}\right]^T = \pi_{\mathrm{t},\infty} \mathbf{U}_0 \left[\mathbf{0}, \mathbf{I}\right]^T = \mathbf{0}^T$, we obtain

$$B(\epsilon) = \frac{1}{1-\lambda_\epsilon} \pi_{\mathrm{t},0} \mathbf{v}_\epsilon \mathbf{e}_1^T (\mathbf{U}_\epsilon^{-1} \mathbf{X}_\epsilon - \mathbf{U}_0^{-1} \mathbf{X}_0) \mathbf{U}_0 \left[\mathbf{0}, \mathbf{I}\right]^T (\mathbf{I} - \mathbf{J}_\epsilon)^{-1} [\mathbf{0}, \mathbf{I}] \mathbf{U}_\epsilon^{-1} \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) \mathbf{x}. \tag{4.58}$$

Moreover, using (4.47) and (4.41), $\mathbf{U}_\epsilon \mathbf{U}_0^{-1} \mathbf{X}_0 \mathbf{U}_0 = \mathbf{X}_0 \mathbf{U}_0 + (\mathbf{v}_\epsilon - \mathbf{1}) \mathbf{e}_1^T \mathbf{D}_0 = \mathbf{X}_0 \mathbf{U}_0 + (\mathbf{v}_\epsilon - \mathbf{1}) \mathbf{e}_1^T$, and therefore, since $\mathbf{X}_\epsilon = (\mathbf{I} - \epsilon \mathbf{P}_\mathrm{a}) \mathbf{P}_\mathrm{t}$,

$$(\mathbf{U}_\epsilon^{-1} \mathbf{X}_\epsilon - \mathbf{U}_0^{-1} \mathbf{X}_0) \mathbf{U}_0 = \mathbf{U}_\epsilon^{-1} (\mathbf{X}_\epsilon \mathbf{U}_0 - \mathbf{U}_\epsilon \mathbf{U}_0^{-1} \mathbf{X}_0 \mathbf{U}_0) = -\mathbf{U}_\epsilon^{-1} (\epsilon \mathbf{P}_\mathrm{a} \mathbf{X}_0 \mathbf{U}_0 + (\mathbf{v}_\epsilon - \mathbf{1}) \mathbf{e}_1^T). \tag{4.59}$$

Therefore, by substituting (4.59) into (4.58), and noting that $\mathbf{e}_1^T [\mathbf{0}, \mathbf{I}]^T = \mathbf{0}^T$, we obtain

$$B(\epsilon) = -\frac{\epsilon}{1-\lambda_\epsilon} \pi_{\mathrm{t},0} \mathbf{v}_\epsilon \mathbf{e}_1^T \mathbf{U}_\epsilon^{-1} \mathbf{P}_\mathrm{a} \mathbf{X}_0 \mathbf{U}_0 \left[\mathbf{0}, \mathbf{I}\right]^T (\mathbf{I} - \mathbf{J}_\epsilon)^{-1} [\mathbf{0}, \mathbf{I}] \mathbf{U}_\epsilon^{-1} \left(\mathbf{1}\pi_{\mathrm{t},\infty} - \mathbf{I}\right) \mathbf{x}. \tag{4.60}$$

Moreover, by left-multiplying each side of (4.40) by $\pi_{\mathrm{t},\infty}$, for $\epsilon > 0$ we obtain

$$\frac{1-\lambda_\epsilon}{\epsilon} = \frac{\pi_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{P}_\mathrm{t}\mathbf{v}_\epsilon}{\pi_{\mathrm{t},\infty}\mathbf{v}_\epsilon} \to \pi_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1} > 0, \tag{4.61}$$

where the limit holds for $\epsilon \to 0$, since $\mathbf{v}_\epsilon \to \mathbf{1}$, $\mathbf{P}_\mathrm{t}\mathbf{1} = \mathbf{1}$ and $\pi_{\mathrm{t},\infty}\mathbf{1} = 1$. Therefore, $B(\epsilon)$ for $\epsilon \to 0$ is bounded, since $\mathbf{I} - \mathbf{J}_0$ is invertible. Similarly, the limit of $C(\epsilon)$ for $\epsilon \to 0$ is bounded. (4.44) is thus proved.

Similarly, from [23], the total cost/reward accrued before the process is absorbed by $\mathcal{Z}_\mathrm{a}$ is

$$\mathcal{C}_\epsilon^{\mathrm{tot}}(\pi_{\mathrm{t},0}) = \pi_{\mathrm{t},0}\left(\mathbf{I} - \mathbf{X}_\epsilon\right)^{-1}\mathbf{c}, \tag{4.62}$$

where $\mathbf{c} = [c(s)]_{s\in\mathcal{S}}$ is the cost/reward vector. We prove that

$$\frac{\mathcal{C}_\epsilon^{\mathrm{tot}}(\pi_{\mathrm{t},0})}{\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})} = \pi_{\mathrm{t},\infty}\mathbf{c} + \mathcal{O}(\epsilon). \tag{4.63}$$

Equivalently,

$$\lim_{\epsilon\to 0}\left|\frac{\mathcal{C}_\epsilon^{\mathrm{tot}}(\pi_{\mathrm{t},0})}{\epsilon\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})} - \frac{1}{\epsilon}\pi_{\mathrm{t},\infty}\mathbf{c}\right| < \infty. \tag{4.64}$$

Using (4.62) and (4.43), we obtain

$$\frac{\mathcal{C}_\epsilon^{\mathrm{tot}}(\pi_{\mathrm{t},0})}{\epsilon\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})} - \frac{1}{\epsilon}\pi_{\mathrm{t},\infty}\mathbf{c} = \frac{\mathcal{C}_\epsilon^{\mathrm{tot}}(\pi_{\mathrm{t},0}) - \mathcal{T}_\epsilon(\pi_{\mathrm{t},0})\pi_{\mathrm{t},\infty}\mathbf{c}}{\epsilon\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})} = \frac{\pi_{\mathrm{t},0}\left(\mathbf{I} - \mathbf{X}_\epsilon\right)^{-1}\left(\mathbf{I} - \mathbf{1}\pi_{\mathrm{t},\infty}\right)\mathbf{c}}{\epsilon\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})}. \tag{4.65}$$

We now compute the limit of the numerator and denominator of (4.65) separately. For the denominator $\epsilon\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})$, from (4.44), $\epsilon\mathcal{T}_\epsilon(\pi_{\mathrm{t},0}) = \left(\pi_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1}\right)^{-1} + \mathcal{O}(\epsilon)$, hence $\lim_{\epsilon\to 0}\epsilon\mathcal{T}_\epsilon(\pi_{\mathrm{t},0}) = \left(\pi_{\mathrm{t},\infty}\mathbf{P}_\mathrm{a}\mathbf{1}\right)^{-1}$, which is positive and bounded. Therefore, (4.64) holds as long as the numerator of (4.65) is bounded. This is directly shown since the numerator of (4.65) equals the last expression of (4.46) when $\mathbf{c} = -\mathbf{x}$, which, as previously shown, is bounded for $\epsilon \to 0$, for any bounded $\mathbf{x}$.

The connection to the problem at hand is obtained as follows. In health state h, the set of transient states ($\mathcal{Z}_\mathrm{t}$ in the proof of the theorem) is $\mathcal{E}(h) \times \{h\} \times \mathcal{A}$. The absorbing state $\mathcal{Z}_\mathrm{a}$ corresponds to the set $\mathcal{E}(h-1) \times \{h-1\} \times \mathcal{A}$, so that $\mathcal{C}_\epsilon^{\mathrm{tot}}(\pi_{\mathrm{t},0})$ and $\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})$ count, respectively, the expected total cumulative reward earned and total time spent by the process $\{\boldsymbol{S}_k\}$ while in health state h, until it is absorbed by the lower health state $h-1$. The initial distribution $\pi_{\mathrm{t},0}$ corresponds to the state

distribution in the set $\mathcal{E}(h) \times \{h\} \times \mathcal{A}$, when the process $\{\boldsymbol{S}_k\}$ first hits the health state h (this event occurs at time K_h, as defined in (4.8)), as induced by policy μ, by (4.3) and by the energy harvesting process. The transition probability matrix $\mathbf{P}_{\mathrm{t}}$ is associated to transitions within the transient states $\mathcal{E}(h) \times \{h\} \times \mathcal{A}$. $\mathbf{P}_{\mathrm{t}}$ is a function of the policy μ_h employed in health state h. The probability matrix $\mathbf{P}_{\mathrm{a}}$ has diagonal components given by the degradation probabilities $p_H(h; e)$. Therefore, $\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})$ and $\frac{1}{\epsilon \pi_{\mathrm{t},\infty} \mathbf{P}_{\mathrm{a}} \mathbf{1}}$ correspond to (4.10) and (4.19); $\frac{\mathcal{C}_\epsilon^{\mathrm{tot}}(\pi_{\mathrm{t},0})}{\mathcal{T}_\epsilon(\pi_{\mathrm{t},0})}$ and $\pi_{\mathrm{t},\infty}\mathbf{c}$ correspond to (4.11) and (4.18), respectively. □

Chapter 5

Conclusions

In this thesis, we have investigated the potential offered by Cognitive Radio and Energy Harvesting to cope, respectively, with spectrum and energy scarcity in today's wireless networks. We have employed a stochastic optimization approach to optimize the utilization of the available resources, recurring, in particular, to the theory of Markov Decision Processes.

Within the Cognitive Radio framework, we have investigated a technique to exploit the Type-I Hybrid Automatic Retransmission reQuest (Type-I HARQ) protocol implemented by the licensed users. We have shown that the use of HARQ opens up opportunities for a more efficient utilization of the spectrum by unlicensed users. In particular, the proposed scheme exploits the temporal redundancy introduced by the use of HARQ by the licensed users to enable interference cancellation techniques at the receiver of the unlicensed users.

Within the Energy Harvesting (EH) paradigm, we have studied a general model where an EH Sensor (EHS) needs to report data of varying importance to a Fusion Center (FC), under a stochastic EH process. For the single EHS scenario, we investigated the interplay between the finite battery storage and the time-correlation in the EH process, demonstrating, both theoretically and numerically, that near optimal performance can be attained by a balanced policy, which solely adapts to the EH state, but not to the exact amount of energy available in the battery. We have then investigated a random multiaccess problem, and designed policies that maximize the aggregate data reporting performance of the network. Also for this scenario, we have designed low-complexity policies, which only loosely depend on the energy level in the battery. Overall, our results and analysis are encouraging for practical EHS design, as they indicate that near-optimal data reporting performance can be achieved with low-complexity policies, suitable for practical implementation.

Finally, we have proposed a stochastic framework, suitable for policy optimization, to model the degradation of the battery capacity over time, and we have formulated an optimization problem which captures the trade-off between battery lifetime and Quality of Service. We have demonstrated that a degradation-aware policy significantly improves the lifetime of the sensor compared to "greedy" policies, while guaranteeing the minimum required QoS. This study represents one step further towards a more realistic performance characterization of harvesting based systems.

Appendix A

UWB Sparse/Diffuse Channel Estimation

A.1 Introduction

Ultra Wide-Band (UWB) signaling had been originally proposed as a technology for indoor mobile and multiple-access communications [76–78]. Due to its significant bandwidth, UWB offers high precision localization [79], robustness against multipath fading [80] and immunity to narrow-band interference [81], thus representing a compelling solution for applications such as short-range, high-speed broadband access [82], Wireless Body Area Networks (WBANs) [83], covert communication links, through-wall imaging, high-resolution ground-penetrating radar and asset tracking [84–86]. However, the performance of coherent UWB transceivers relies on the availability of accurate channel estimates (*e.g.*, [87–89]). Thus, it is important to design channel estimation strategies that exploit the structural and statistical properties of UWB propagation to achieve the best estimation accuracy.

The significant transmission bandwidth of UWB systems enables a fine-grained delay resolution at the receiver, of the order of $1\,\mathrm{ns}$. In many environments, only some of the resolvable delay bins carry significant multipath energy, yielding a sparse channel structure [85, 90]. For this reason, UWB channel estimation strategies based on compressive sensing and sparse approximation techniques [91–94] have been proposed in the literature, and they have been shown to outperform conventional unstructured estimators [95, 96]. Also, localization techniques that exploit the information about the specular multipath structure of the UWB channel have been proposed (see, *e.g.*, [97, 98]).

However, recent propagation studies suggest that, for some environments, such as indoor, WBANs and vehicular scenarios, diffuse (dense) components of the impulse response arise. These are caused by propagation processes such as diffuse scattering [99], or unresolvable MultiPath Components

(MPCs). Moreover, UWB channels exhibit a significant frequency dispersion [100] due to the large transmission bandwidth employed. While irrelevant for conventional narrow-band systems, this effect results in a pulse broadening and spreading of the MPC energy over multiple resolvable delay bins. These propagation mechanisms are not properly modeled by a purely sparse channel.

Recent work explores these effects. In [101], a geometry-based stochastic UWB model is proposed, consisting of a statistical model for the diffuse component. The model developed in [102] combines a geometric approach to model the resolvable MPCs, and a stochastic approach to model the diffuse tail associated with each MPC. In [99], the spatial structure of the diffuse MPCs is investigated, and its parameters are extracted from the measurements. In [103], the impact of diffuse scattering on the characteristics of vehicular propagation channels in highway environments is evaluated, and the Doppler frequency-delay characteristics of diffuse components are analyzed. In [104], a low-complexity model of diffuse scattering is proposed for vehicular radio channels. While these prior models were targeted towards performance assessment, herein we develop a simplified UWB channel model suitable for channel estimation purposes and estimator analysis.

Exploitation of structure in channel models can lead to estimation strategies with strong performance: in [88], a Maximum Likelihood (ML) estimator is designed which exploits the clustered structure of the UWB channel. In [89], a joint channel estimation and decoding technique for Bit-Interleaved Coded Orthogonal Frequency Division Multiplexing is designed, based on a two-state Gaussian mixture prior to model the sparse/diffuse structure of the channel, and on an hidden Markov prior to model clustering among the large taps. Therein, more structure is assumed, *e.g.*, clustering of the taps, and further the scheme is semi-blind. In [105], an ML framework is developed for parameter estimation in multi-dimensional channel sounding. Therein, the channel comprises a deterministic component, resulting from specular reflection, and a stochastic component modeling diffuse scattering.

A.1.1 Contributions

In this chapter, based on the analysis of the propagation mechanisms peculiar to UWB systems, we present a novel Hybrid Sparse/Diffuse (HSD) UWB channel model. In particular, we propose statistical models for the sparse and diffuse components. We identify three physically motivated scenarios that differ in the amount of side information available at the receiver (*e.g.*, channel sparsity level, Power Delay Profile (PDP) of the diffuse or sparse component).

In Sec. A.5, for each scenario, Bayesian channel estimators are derived. In particular, we propose the *Generalized MMSE* (GMMSE) and the *Generalized Thresholding* (GThres) estimators, for the scenario where the statistics of the specular coefficients are unknown. We present a Mean-Squared Error (MSE) analysis of the GMMSE and the GThres estimators, in the asymptotic regimes of high and low Signal to Noise Ratios (SNR). We also design an Expectation-Maximization (EM) algorithm for the estimation of the PDP of the diffuse component, which exploits the structure of the PDP over the channel delay dimension to enhance the estimation accuracy. Moreover, we analyze the scenario with a non-orthogonal pilot sequence, and establish a connection between the GThres estimator and conventional sparse approximation algorithms proposed in the literature.

Finally, in Sec. A.9, we compare the proposed algorithms to unconstrained estimators, which do not exploit the structure of the UWB channel, and conventional sparse estimators, which, on the other hand, ignore the diffuse component of the channel. We also validate the simplified HSD channel model and the channel estimation strategies, based on a realistic UWB channel model developed in [102]. The numerical results show that the new channel estimation methods considerably improve the Mean-Squared Error (MSE) accuracy and the Bit Error Rate (BER) performance over conventional unstructured estimators, *e.g.*, Least Squares (LS), and purely sparse estimators, thus suggesting the importance of a proper model for the UWB channel. Specifically, a purely sparse estimator, by ignoring the diffuse component, is not able to capture important phenomena in UWB, *e.g.*, pulse distortion [106] and diffuse scattering [100], thus failing to accurately estimate the channel. In contrast, the HSD model, despite its simplicity, can effectively capture important UWB propagation mechanisms, such as fine delay resolution, scattering from rough surfaces and frequency dispersion. Moreover, due to its hybrid structure, the HSD model is robust and covers a wide range of practical scenarios, where the channel exhibits either a sparse, diffuse or hybrid nature.

A.1.2 Structure of the chapter

This chapter is organized as follows. In Sec. A.2, we introduce the notation. In Sec. A.3, we overview the UWB propagation mechanisms. In Sec. A.4, we present the system model and we introduce the HSD channel model. In Sec. A.5, we present channel estimators based on the HSD model. In Sec. A.6, we perform an asymptotic MSE analysis of these estimation schemes, and we discuss the results. In Sec. A.7, we present an EM algorithm for the PDP estimation of the diffuse component. In Sec. A.8, we analyze the case with a non-orthogonal pilot sequence. In Sec. A.9, we

present simulation results. Sec. A.10 concludes this chapter. The proofs of the theorems and lemmas are provided in the appendices at the end of the chapter.

A.2 Notation

We use lower-case and upper-case bold letters for column vectors ($\mathbf{a}$) and matrices ($\mathbf{A}$), respectively. The scalar $\mathbf{a}_k$ (or $\mathbf{a}(k)$) denotes the kth entry of vector $\mathbf{a}$, and $\mathbf{A}_{k,j}$ (or $\mathbf{A}(k,j)$) denotes the (k,j)th entry of matrix $\mathbf{A}$. A positive definite (positive semi-definite) matrix $\mathbf{A}$ is denoted by $\mathbf{A} \succ 0$ ($\mathbf{A} \succeq 0$). The transpose, complex conjugate of matrix $\mathbf{A}$ is denoted by $\mathbf{A}^*$. We define the square root of $\mathbf{A} \succeq 0$ with eigenvalue decomposition $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ as $\sqrt{\mathbf{A}} = \mathbf{U}\sqrt{\mathbf{D}}\mathbf{U}^*$. The $K \times K$ unit matrix is defined as $\mathbf{I}_K$. The trace operator is denoted by $\mathrm{tr}\,(\mathbf{A}) = \sum_k \mathbf{A}_{k,k}$. The vector $\mathbf{a} \odot \mathbf{b}$ is the component-wise (Schur) product of vectors $\mathbf{a}$ and $\mathbf{b}$. The indicator function is given by $\mathcal{I}\,(\cdot)$. We use $p(\cdot)$ to indicate a continuous or discrete probability distribution, and $\Pr\,(\cdot)$ to indicate the probability of an event. The expectation of random variable x, conditioned on y, is denoted by $\mathbb{E}\,[x|y]$. The Gaussian distribution with mean $\mathbf{m}$ and covariance $\mathbf{\Sigma}$ is written as $\mathcal{N}\,(\mathbf{m}, \mathbf{\Sigma})$, whereas the circularly symmetric complex Gaussian distribution is denoted by $\mathcal{CN}\,(\mathbf{m}, \mathbf{\Sigma})$;[1] the Bernoulli distribution with parameter q is denoted by $\mathcal{B}(q)$, and the exponential distribution with mean m by $\mathcal{E}(m)$. The indicator function is denoted by $\mathcal{I}\,(\cdot)$.

A.3 UWB channel propagation and modeling overview

In this section, we overview the state of the art of UWB channel propagation and modeling. The aim is to determine an appropriate UWB channel model, which captures the main UWB propagation mechanisms. Neglecting pulse distortion [106] for simplicity, a time-varying channel in the continuous time can be represented as [107]

$$h(\tau, t) = \sum_l a_l(t)\delta(\tau - \tau_l(t)), \tag{A.1}$$

where $\delta\,(\cdot)$ is the Kronecker delta function, t is the time dimension and τ is the channel delay. The sum is over the MPCs, with time-varying amplitude $a_l(t)$ and delay $\tau_l(t)$. If we consider a UWB system with center frequency f_0 and transmission bandwidth W, the discrete baseband time-varying

[1] For a vector $\mathbf{x} = \mathbf{x}_R + i\mathbf{x}_I \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{x}_R = \mathrm{Re}(\mathbf{x})$, $\mathbf{x}_I = \mathrm{Im}(\mathbf{x})$ and $i = \sqrt{-1}$, we define the covariance matrices of its real and imaginary parts as $\mathbb{E}[\mathbf{x}_R\mathbf{x}_R^*] = \mathbb{E}[\mathbf{x}_I\mathbf{x}_I^*] = \frac{\mathrm{Re}(\mathbf{\Sigma})}{2}$ and $\mathbb{E}[\mathbf{x}_I\mathbf{x}_R^*] = -\mathbb{E}[\mathbf{x}_R\mathbf{x}_I^*] = \frac{\mathrm{Im}(\mathbf{\Sigma})}{2}$.

impulse response of the channel is given by

$$h_{bb}(n,t) = \sum_l a_l(t) e^{-i2\pi f_0 \tau_l(t)} \mathrm{sinc}\left(n - W\tau_l(t)\right), \tag{A.2}$$

where $\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$ is the sinc function, and $n \in \mathbb{Z}$ is the discrete channel delay. Due to the large transmission bandwidth of UWB systems, MPCs arising from reflections and scattering in the environment spaced apart (in the delay domain) by more than $\frac{1}{W}$, which is typically of the order of a fraction of a ns, can be resolved at the receiver. Then, by neglecting leakage effects due to the sampling of the sinc function off its peak, (A.2) is commonly approximated by the following *sparse* discrete baseband representation:

$$h_{bb}(n,t) \simeq \sum_l a_l(t) e^{-i2\pi f_0 \tau_l(t)} \delta\left(n - \mathrm{rd}\left(W\tau_l(t)\right)\right), \tag{A.3}$$

where $\mathrm{rd}(x)$ returns the closest integer to x.

However, in many practical scenarios of interest (*e.g.*, indoor environments), diffuse components, that cannot be described by the above model, arise. These are created mainly by the following phenomena: a large number of unresolved paths, diffuse scattering [100], pulse distortion resulting from the frequency dependence of the gain and efficiency of the antennas and of the dielectric or conductive materials, and diffraction effects [106]. In [101], the following frequency response has been proposed, modeling the contribution from all these effects:

$$H_{UWB}(f) = \left(\mathcal{S}_{LOS}(f) + \sum_k \mathcal{S}_k(f) + \mathcal{D}(f)\right) \frac{f^{-m}}{F}, \tag{A.4}$$

where f is frequency. In particular, we recognize in $\mathcal{S}_{LOS}(f)$ and $\sum_k \mathcal{S}_k(f)$ the contributions from the line of sight and the resolvable MPCs, respectively, *i.e.*, the MPCs whose inter-arrival time is larger than $\frac{1}{W}$, giving rise to a *sparse* component in the time domain. The term $\mathcal{D}(f)$ represents the *diffuse* component due to multipath interference, and is associated with the non-resolvable MPCs. Finally, $\frac{f^{-m}}{F}$ models the frequency distortion of the channel, where F is a normalization factor and m is the frequency decay exponent. Note that, in this model, the diffuse component is independent of the realization of the discrete MPCs, while, in contrast, the work in [102] models the diffuse component as a diffuse tail associated with each specular component.

It is worth noting that the level of channel diffuseness or sparseness depends primarily on two

factors: the transmission bandwidth and the environment. In fact, the larger the transmission bandwidth, the finer the delay resolution at the receiver, and the sparser the channel is expected to be. On the other hand, an environment with many scatterers or rough surfaces, *e.g.*, an indoor scenario or WBANs, is more likely to give rise to a dense channel, due to the richer interaction among the MPCs. Dense channels have been observed, *e.g.*, in gas stations [101], industrial [108], office [85] and vehicular environments [103]. We thus expect a dense or hybrid channel representation to be relevant in these or similar scenarios.

Spatio-temporal scale of variation in the UWB channel

We now consider the spatio-temporal variation of the channel, due to the relative motion of the scatterers, receiver and transmitter in the environment. For ease of exposition, we consider movement of the receiver only. Ignoring Doppler effects, which are left for future investigations, the channel time-variations affect the amount of side-information available at the receiver for the purpose of channel estimation, as discussed in Sec. A.4.2.

From the discrete baseband model (A.2), the phase[2] variation of the lth MPC over a time-interval Δt is given by $\Delta\phi_l \triangleq 2\pi \frac{c_0}{\lambda_0} |\tau_l(t+\Delta t) - \tau_l(t)|$, where λ_0 is the wavelength at the center frequency, and c_0 is the free space speed of light. Therefore, a significant phase variation (*e.g.*, by more than $\frac{\pi}{2}$) occurs when $\Delta\phi_l > \frac{\pi}{2}$. This quantity corresponds, in the spatial domain, to a wavelength or a fraction of it. Therefore, phase changes are expected to occur on a very small spatio-temporal scale.

Similarly, the variation of the MPC delay, over the same time-interval Δt, is given by $\Delta\tau_l \triangleq |\tau_l(t+\Delta t) - \tau_l(t)|$. Hence, a significant variation (*e.g.*, by more than one channel delay bin, $\frac{1}{W}$) occurs when $\Delta\tau_l > \frac{1}{W}$, *i.e.*, on a spatial scale of $\frac{c_0}{W}$ or roughly a number of wavelengths in the range $[0.5, 5]$, depending on the value of the transmission bandwidth W, relative to the center frequency f_0.

Finally, significant variations of the MPC amplitude $a_l(t)$, due to shadowing effects, typically correspond to a spatial scale of several wavelengths.

Note that, due to mutual interference of the unresolvable MPCs contributing to the same tap location, changes in the amplitude of the diffuse components arise over the same spatio-temporal scale as the phase changes of the MPCs (small scale fading). On the other hand, the amplitude of the resolvable MPCs vary over a much larger spatio-temporal scale (large scale fading).

[2]Note that "phase" is a narrow-band concept and can be used only as an approximation in UWB systems, in particular when the lower band edge is at $f = 0$.

Remark A.3.1. It is worth noting that the side-lobes of the sinc function in (A.2) introduce faster time-variations of the amplitude of the resolvable MPCs than the large-scale fading, over the same spatio-temporal scale as the delay variations, and account for the leakage of the MPC energy over nearby channel taps. However, this phenomenon is limited, and can be quantified as follows. The most severe leakage occurs when the MPC arrives exactly in the middle between two sampling times, in which case most of the energy ($2\mathrm{sinc}(0.5)^2 \simeq 80\%$) is spread equally between two nearby taps (each with amplitude $1 - \mathrm{sinc}(0.5) \simeq 37\%$ smaller than in the no leakage scenario, where the MPC delay is exactly an integer number of the sampling period), and the remaining 20% is leaked among the nearby taps. Therefore, the side-lobes of the sinc function account for at most a 37% variation of the amplitude of the main MPC tap in (A.2). The problem of MPCs falling in between two sample points can be modeled as a basis mismatch [109].

In the next section, we present the observation and the channel models. In particular, in Sec. A.4.1 we present the HSD model, which represents a simplification with respect to other models presented in the literature, *e.g.*, (A.4), but at the same time it captures the main propagation phenomena of the UWB channel discussed in this section: resolvable MPCs, modeled by the sparse vector (A.3), unresolvable MPCs, diffuse scattering and frequency distortion, modeled by a random, dense vector. Also, based on the analysis of the spatio-temporal scale of variation in the UWB channel, in Sec. A.4.2 we discuss different practical scenarios, differing in the side-information available at the receiver for the purpose of channel estimation, which enables more accurate estimation techniques.

A.4 System Model and Hybrid Sparse-Diffuse channel model

We consider a single-user UWB system. The source transmits a sequence of $M = N + L - 1$ pilot symbols, $x(k), k = -(L-1), \ldots, N-1$, over a channel $h(l), l = 0, \ldots, L-1$ with known delay spread $L \geq 1$. The received, discrete time, baseband signal over the corresponding observation interval of duration N is given by

$$y(k) = \sum_{l=0}^{L-1} h(l)x(k-l) + w(k), \quad k = 0, \ldots, N-1, \tag{A.5}$$

where $w(k) \in \mathcal{CN}(0, \sigma_w^2)$ is i.i.d. noise.

If we collect the N received samples in the column vector $\mathbf{y} = [y(0), y(1), \ldots, y(N-1)]^T$, we

have the following matrix representation:

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{w}. \tag{A.6}$$

Above, $\mathbf{X} \in \mathbb{C}^{N\times L}$ is the $N \times L$ Toeplitz matrix associated with the pilot sequence, having the vector of the transmitted pilot sequence $[x(-k), x(-k+1), \ldots, x(-k+N-1)]^T, k = 0, \ldots, L-1$, as its kth column, $\mathbf{h} = [h(0), h(1), \ldots, h(L-1)]^T \in \mathbb{C}^L$ is the column vector of channel coefficients, and $\mathbf{w} = [w(0), w(1), \ldots, w(N-1)]^T \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I}_N)$ is the noise vector.

We assume $\mathbf{X}^*\mathbf{X} \succ 0$, so that the LS estimate $\mathbf{h}_{\mathrm{LS}} = (\mathbf{X}^*\mathbf{X})^{-1}\mathbf{X}^*\mathbf{y}$ is a sufficient statistic [110] for the channel. Therefore, without loss of generality for the purpose of channel estimation, we consider the observation model

$$\mathbf{h}_{\mathrm{LS}} = (\mathbf{X}^*\mathbf{X})^{-1}\mathbf{X}^*\mathbf{y} = \mathbf{h} + (\mathbf{X}^*\mathbf{X})^{-1}\mathbf{X}^*\mathbf{w} = \mathbf{h} + \sqrt{\mathbf{S}}^{-1}\mathbf{n}, \tag{A.7}$$

where we have defined the SNR matrix $\mathbf{S} = \frac{\mathbf{X}^*\mathbf{X}}{\sigma_w^2} \succ 0$, and $\mathbf{n} = \frac{1}{\sigma_w^2}\sqrt{\mathbf{S}}^{-1}\mathbf{X}^*\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_L)$. With a slight abuse of notation, we will refer to the LS estimate $\mathbf{h}_{\mathrm{LS}}$ as the "observed" sequence. Moreover, we assume that the pilot sequence is orthogonal, so that $\mathbf{S}$ is a diagonal matrix. Then, the noise vector $\sqrt{\mathbf{S}}^{-1}\mathbf{n}$ in the LS estimate has independent entries. This assumption greatly simplifies the channel estimation problem. In fact, when the channel has independent entries over the delay dimension (this is the case for the HSD model we develop), a per-tap estimation approach, rather than a joint one, is optimal. The case with non-orthogonal pilot sequences is considered in Sec. A.8.

A.4.1 HSD Channel Model

The channel $\mathbf{h}$ follows the HSD model developed in [111],

$$\mathbf{h} = \mathbf{a}_s \odot \mathbf{c}_s + \mathbf{h}_d, \tag{A.8}$$

where the terms $\mathbf{a}_s \odot \mathbf{c}_s \in \mathbb{C}^L$ and $\mathbf{h}_d \in \mathbb{C}^L$ represent the *sparse*[3] and the *diffuse components*, respectively. In particular, $\mathbf{a}_s \in \{0,1\}^L$ is the *sparsity pattern*, which is equal to one in the positions of the specular MPCs, and equal to zero otherwise; its entries are drawn i.i.d. from $\mathcal{B}(q)$, where $q \ll 1$ so as to enforce sparsity. In the sequel, we refer to the non-zero entries of $\mathbf{a}_s \odot \mathbf{c}_s \in \mathbb{C}^L$ as

[3] In the following, we use the terms *sparse*, *specular* and *resolvable MPCs* interchangeably. In fact, the physical specular components (resolvable MPCs) of the channel can be modeled and represented by a sparse vector (A.3).

active sparse components. The vector of sparse coefficients, $\mathbf{c}_s \in \mathbb{C}^L$, is drawn from the continuous probability distribution $p(\mathbf{c}_s)$, with second order moment $\mathbb{E}\left[\mathbf{c}_s \mathbf{c}_s^*\right] = \mathbf{\Lambda}_s$, where $\mathbf{\Lambda}_s$ is a diagonal matrix with entries given by the PDP $\mathbf{\Lambda}_s(k,k) = \mathcal{P}_s(k), k = 0, \ldots, L-1$.[4] Finally, we use the Rayleigh fading assumption for the diffuse component, $\mathbf{h}_d \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_d)$, where $\mathbf{\Lambda}_d$ is diagonal, with entries given by the PDP $\mathbf{\Lambda}_d(k,k) = \mathcal{P}_d(k), k = 0, \ldots, L-1$.

Remark A.4.1. The Bernoulli model for $\mathbf{a}_s$ can be interpreted as a discretized Saleh-Valenzuela model [112]. In fact, according to the latter, the inter-arrival times of the specular components have an exponential distribution, whose discrete counterpart is the geometric distribution. This in turn can be interpreted as the inter-arrival time of two consecutive "1"s in a sequence of i.i.d. Bernoulli draws.

Remark A.4.2. In general, the Rayleigh fading assumption does not hold for the distribution of the sparse coefficients $p(\mathbf{c}_s)$ (unlike the diffuse ones), since only very few propagation paths contribute to an active tap in the sparse channel, thus limiting the validity of the central limit theorem. Channel measurement campaigns have shown that the large scale fading, affecting the amplitude of the entries of $\mathbf{c}_s$, can be modeled by a log-normal distribution [101]. However, for the sake of analytical tractability, in the following we either treat $\mathbf{c}_s$ as a deterministic unknown vector, when its second order moment $\mathbf{\Lambda}_s$ is unknown, or we treat it using the Gaussian approximation, when knowledge of $\mathbf{\Lambda}_s$ is available.

Remark A.4.3. Note that in [101] the amplitudes of the diffuse coefficients are modeled by a Weibull distribution, with a delay dependent shape parameter $\sigma < 2$, and approach the Rayleigh fading distribution ($\sigma = 2$) only for large excess delays. This distribution represents a fading worse than Rayleigh. However, we adopt the Rayleigh fading approximation for simplicity and tractability. Also, the sidelobes of the sinc function in (A.2) introduce correlation in the delay domain, which is not accounted for under the Rayleigh fading model. This is a common assumption in standard cellular channel models, where measurements have well established the independence of fading on different taps [113].

Despite its simplicity, we argue that the HSD model is able to capture the main UWB propagation mechanisms discussed in Sec. A.3. In fact, the resolvable specular components and the fine delay resolution are appropriately modeled by the sparse vector $\mathbf{a}_s \odot \mathbf{c}_s$, whereas diffuse scattering, multipath interference and the frequency distortion are approximated by the diffuse component $\mathbf{h}_d$. This is confirmed by simulation results in Chapter A.9, where we validate the proposed HSD model based on a realistic channel emulator [102].

[4] It is worth noting that this is not a PDP in the traditional sense, but rather represents the power profile of the active sparse components, as a function of the delay.

A.4.2 Channel Estimation scenarios

The HSD model is described by a number of deterministic parameters, namely, the sparsity level q, the PDP of the diffuse component $\mathcal{P}_d$ and the PDP of the sparse component $\mathcal{P}_s$. Accurate knowledge about some or all of these parameters may not be available at the receiver, depending on a number of factors, most importantly the length of the interval over which the channel is observed, and the dynamics of the environment.

Let $\left\{\mathbf{h}^{(j)} = \mathbf{a}_s^{(j)} \odot \mathbf{c}_s^{(j)} + \mathbf{h}_d^{(j)}, j = 0, \ldots, N_{\text{ch}} - 1\right\}$ be a sequence of N_{ch} channel realizations, spaced apart in time by Δt, corresponding to a spatial separation by $\simeq \lambda_0$, resulting from the relative motion of the receiver with respect to the scatterers and the transmitter position. Under this assumption, the samples of the diffuse component $\left\{\mathbf{h}_d^{(j)}, j \geq 0\right\}$ can be approximated as drawn independently from $\mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_d)$, due to multipath interference (Sec. A.3).

On the other hand, the positions of the active sparse coefficients $\left\{\mathbf{a}_s^{(j)}, j = 0, \ldots, N_{\text{ch}} - 1\right\}$ exhibit correlation with each other. In fact, as pointed out in Sec. A.3, a variation of the delay associated with a specular MPC by one channel delay bin occurs over a spatial scale of the order of $\frac{c_0}{W\lambda_0} \in [0.5, 5]$ wavelengths. Therefore, the positions of the "1"s observed in subsequent realizations of the sparsity pattern $\mathbf{a}_s^{(j)}$ are bound not to vary appreciably over a large spatial scale, relative to the wavelength.

A similar consideration holds for the amplitudes of the specular components (*i.e.*, the active sparse components in the vector $\mathbf{a}_s^{(j)} \odot \mathbf{c}_s^{(j)}$), which vary according to the large scale fading, *i.e.*, over a relatively large spatial scale, compared to the rate of variation of the diffuse component (however, the side-lobes of the sinc function account for a 37% variation in the amplitude on the same spatial scale as the delay variations, as discussed in Remark A.3.1 of Sec. A.3).

This correlation structure, *i.e.*, slow amplitude and delay variations, may be exploited to enhance the estimation accuracy of the sparse component $\mathbf{a}_s^{(j)} \odot \mathbf{c}_s^{(j)}$, by tracking the position and amplitude of the resolvable MPCs over subsequent observation windows. However, in this work we consider estimation of $\mathbf{a}_s^{(j)} \odot \mathbf{c}_s^{(j)}$ based on either only one channel realization, or the statistics of the ensemble of realizations that ignores the information about the temporal sequence in which the realizations occur. We consider three different physical scenarios, dictated by the length of the observation window N_{ch}.

A.4.3 Single Snapshot of the channel

If a very short observation window is available ($N_{\mathrm{ch}} = 1$, or less than a wavelength in the spatial domain), averaging over the small scale and the large scale fading is not possible. Under this assumption, statistical information about the channel cannot be reliably collected, and the channel can reasonably be considered a deterministic and unknown vector. In this case, an LS estimate $\mathbf{h}_{\mathrm{LS}}$ may be employed. In the absence of prior information about the channel, this is a robust approach for channel estimation.

Alternatively, we may exploit further structure of the channel, *e.g.*, exponential PDP of the diffuse component, to average the fading over the delay dimension rather than over time. As shown in Sec. A.7, under this assumption, an accurate PDP estimate of $\mathbf{h}_d^{(j)}$ is possible even in the extreme case $N_{\mathrm{ch}} = 1$. We may then assume that the PDP of $\mathbf{h}_d^{(j)}$ is known at the receiver, whereas the vector $\mathbf{c}_s^{(j)}$ is modeled as deterministic and unknown.

As to the sparsity level q, letting N_{sc} be the number of resolvable scatterers, we have $q \simeq \frac{N_{\mathrm{sc}}}{L}$. This number is not expected to vary appreciably over a relatively long observation interval, and can be estimated by counting the number of resolvable MPCs which can be distinguished from the noise plus diffuse background. However, an accurate estimate of N_{sc} is obtained by averaging the small-scale fading and the noise over subsequent channel realizations. Hence, we model q as a deterministic and unknown parameter.

A.4.4 Averaging over the Small scale fading

When a larger observation window is available (corresponding, in the spatial domain, to a few wavelengths, $N_{\mathrm{ch}} > 1$), averaging over the small scale fading (amplitude and phase of the diffuse component) may be possible. In this case, the PDP of $\mathbf{h}_d^{(j)}$ can be estimated accurately by averaging over subsequent realizations of the fading process.

In this scenario, we assume that $\mathbf{\Lambda}_d$ is perfectly known at the receiver. This knowledge can be exploited by performing a Minimum MSE (MMSE) estimate of $\mathbf{h}_d^{(j)}$, which achieves a better accuracy than LS. On the other hand, due to the inability to average over the large-scale fading, which affects the variation of the amplitude of the resolvable MPCs, $\mathbf{c}_s^{(j)}$ is treated as deterministic and unknown.

Table A.1. *Estimation scenarios considered.*

	Scenario	sparsity q	PDP $\mathbf{\Lambda}_s$	PDP $\mathbf{\Lambda}_d$
S0	Single snapshot (unstructured)	unknown	unknown	unknown
S1	Single snapshot (PDP structure exploited)	unknown	unknown	known
S2	Avg. over Small scale fading	known	unknown	known
S3	Avg. over Small&Large scale fading	known	known	known

A.4.5 Averaging over the Small scale and the Large scale fading

Finally, when the observation interval spans several wavelengths ($N_{\rm ch} \gg 1$), averaging over the large scale, other than the small scale fading, is possible.

In this scenario, we assume that $\mathbf{\Lambda}_d$, $\mathbf{\Lambda}_s$ and q are known at the receiver. This information can be exploited to compute a linear-MMSE estimate of $\mathbf{c}_s^{(j)}$ and $\mathbf{h}_d^{(j)}$, thus enhancing the estimation accuracy over an unstructured estimate (*e.g.*, LS).

The main scenarios of interest, and the side information at the receiver, are listed in Table A.1. Scenario S0 will not be further considered, since the channel is estimated via LS. The next chapter is devoted to the design and analysis of channel estimators based on the HSD model.

A.5 HSD estimators

A.5.1 MMSE Estimator

When $\mathbf{\Lambda}_d$, $\mathbf{\Lambda}_s$ and q are known, we can devise an MMSE estimator. By exploiting the orthogonality of the pilot sequence, we can use a per-tap estimation approach. The MMSE estimate of the kth delay bin is given by the posterior mean of the channel, given the observed channel sample $\mathbf{h}_{\rm LS}(k)$ [110],

$$\begin{aligned}\hat{\mathbf{h}}_{\rm MMSE}(k) &= \Pr\left(\mathbf{a}_s(k)=0|\mathbf{h}_{\rm LS}(k)\right)\mathbb{E}\left[\mathbf{h}_d(k)|\mathbf{h}_{\rm LS}(k),\mathbf{a}_s(k)=0\right]\\ &+\Pr\left(\mathbf{a}_s(k)=1|\mathbf{h}_{\rm LS}(k)\right)\mathbb{E}\left[\mathbf{c}_s(k)+\mathbf{h}_d(k)|\mathbf{h}_{\rm LS}(k),\mathbf{a}_s(k)=1\right],\end{aligned} \tag{A.9}$$

where we have conditioned on the realization of the sparsity bit $\mathbf{a}_s(k)$. In particular, the sum is over the posterior mean under the two hypotheses $\mathbf{a}_s(k)=1$ and $\mathbf{a}_s(k)=0$, weighted by their posterior distribution $\Pr\left(\mathbf{a}_s(k)=1|\mathbf{h}_{\rm LS}(k)\right)$ and $\Pr\left(\mathbf{a}_s(k)=0|\mathbf{h}_{\rm LS}(k)\right)$, respectively.

In order to compute (A.9), we use the circular Gaussian approximation for $\mathbf{c}_s(k)$.[5] Under this

[5] As discussed in Remark A.4.2 in Sec. A.4, the large scale fading is commonly modeled by a log-normal prior; however, due to the difficulty in handling it, the Rayleigh fading approximation is used, thus leading to the classical linear MMSE

assumption, $\mathbf{h}_{\mathrm{LS}}(k)|\{\mathbf{a}_s(k) = a, \mathbf{h}(k)\} \sim \mathcal{CN}(\mathbf{h}(k), 1/\mathbf{S}_{k,k})$, whereas the channel sample $\mathbf{h}(k)$, conditioned on $\mathbf{a}_s(k) = a$, is distributed as $\mathbf{h}(k)|\mathbf{a}_s(k) = a \sim \mathcal{CN}(0, \mathbf{a}_s(k)\mathcal{P}_s(k) + \mathcal{P}_d(k))$. Then, $\mathbf{h}(k)|\{\mathbf{h}_{\mathrm{LS}}(k), \mathbf{a}_s(k) = a\} \sim \mathcal{CN}(\mathbf{m}(a), \Sigma)$, with posterior mean

$$\mathbf{m}(a) = \mathbb{E}\left[\mathbf{h}(k)|\mathbf{h}_{\mathrm{LS}}(k), \mathbf{a}_s(k) = a\right] = \frac{a\mathcal{P}_s(k) + \mathcal{P}_d(k)}{1/\mathbf{S}_{k,k} + a\mathcal{P}_s(k) + \mathcal{P}_d(k)}\mathbf{h}_{\mathrm{LS}}(k). \tag{A.10}$$

From (A.9), we finally obtain

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{MMSE}}(k) = &\Pr\left(\mathbf{a}_s(k) = 0|\mathbf{h}_{\mathrm{LS}}(k)\right)\frac{\mathbf{S}_{k,k}\mathcal{P}_d(k)}{1 + \mathbf{S}_{k,k}\mathcal{P}_d(k)}\mathbf{h}_{\mathrm{LS}}(k)\\ &+ \Pr\left(\mathbf{a}_s(k) = 1|\mathbf{h}_{\mathrm{LS}}(k)\right)\frac{\mathbf{S}_{k,k}\left(\mathcal{P}_s(k) + \mathcal{P}_d(k)\right)}{1 + \mathbf{S}_{k,k}\left(\mathcal{P}_s(k) + \mathcal{P}_d(k)\right)}\mathbf{h}_{\mathrm{LS}}(k),\end{aligned}$$

where, from Bayes' rule and $\mathbf{a}_s(k) \sim \mathcal{B}(q)$, letting $\mathrm{Q}_k = \frac{\mathbf{S}_{k,k}\mathcal{P}_s(k)}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}$, we have

$$\begin{aligned}\Pr\left(\mathbf{a}_s(k) = 1|\mathbf{h}_{\mathrm{LS}}(k)\right) &= \left(1 + \frac{1-q}{q}\frac{p\left(\mathbf{h}_{\mathrm{LS}}(k)|\mathbf{a}_s(k) = 0\right)}{p\left(\mathbf{h}_{\mathrm{LS}}(k)|\mathbf{a}_s(k) = 1\right)}\right)^{-1}\\ &= \frac{1}{1 + \frac{1-q}{q}\left(1 + \mathrm{Q}_k\right)\exp\left\{-\frac{\mathrm{Q}_k}{1+\mathrm{Q}_k}\frac{\mathbf{S}_{k,k}|\mathbf{h}_{\mathrm{LS}}(k)|^2}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}\right\}}.\end{aligned} \tag{A.11}$$

A.5.2 Generalized MMSE and Generalized Thresholding Estimators

In this section, we develop estimators for scenarios S1 and S2. In particular, $\mathbf{\Lambda}_d$ is assumed to be known at the receiver, whereas $\mathbf{c}_s$ is treated as a deterministic and unknown vector. The case where $\mathbf{\Lambda}_d$ is unknown and is estimated from the observed sequence is treated in Sec. A.7.

For generality, we assume that the sparsity level q is unknown, and an estimate $\tilde{q}$ of q, which might be different from the real q, is used in the estimation phase. This choice represents a generalization with respect to [111], where the true sparsity level q is used. We will show by simulation in Chapter A.9, and by analysis in Sec. A.6, that assuming a sparsity level $\tilde{q} < q$ often improves the estimation accuracy, thus implying that knowledge of this parameter is not crucial to the performance of the estimators.

We proceed as follows. $\mathbf{c}_s$ is estimated by Maximum Likelihood (ML). Then, the estimate $\hat{\mathbf{c}}_s$ is used to perform either an MMSE or a Maximum A Posteriori (MAP) estimate of the sparsity pattern

estimator. We have numerically evaluated the performance loss incurred by using the linear MMSE estimator over an MMSE estimator based on the log-normal prior, for the simple scalar model $y = c_s + n$, where $c_s = e^{\nu_s + i\theta_s}$, with $\nu_s \sim \mathcal{N}(0, 1)$ and θ_s uniform in $[0, 2\pi]$, is the channel coefficient with log-normal amplitude, $n \sim \mathcal{CN}(0, \sigma_w^2)$ is the noise; we found that the performance loss is at most $1.67\,\mathrm{dB}$, at $0\,\mathrm{dB}$ SNR level.

$\mathbf{a}_s$, denoted by $\hat{\mathbf{a}}_s$, assuming the prior $\mathbf{a}_s \sim \mathcal{B}(\tilde{q})^L$. We refer to these estimators as the GMMSE and GThres estimators, respectively. Finally, the diffuse component $\mathbf{h}_d$ is estimated via MMSE, based on the residual estimation error $\mathbf{h}_{\mathrm{LS}} - \hat{\mathbf{a}}_s \odot \hat{\mathbf{c}}_s$. The ML estimate of $\mathbf{c}_s(k)$ is given by

$$\hat{\mathbf{c}}_s(k) = \underset{\mathbf{c}_s(k)\in\mathbb{C}}{\arg\min} \left\{ -\ln p\left(\mathbf{h}_{\mathrm{LS}}(k) | \mathbf{c}_s(k), \mathbf{a}_s(k) = 1\right)\right\} = \mathbf{h}_{\mathrm{LS}}(k), \tag{A.12}$$

where we have used the fact that, when conditioned on $\mathbf{a}_s(k) = 0$, the observation $\mathbf{h}_{\mathrm{LS}}(k)$ does not depend on $\mathbf{c}_s(k)$, and $\mathbf{h}_{\mathrm{LS}}(k) | \left\{\mathbf{c}_s(k), \mathbf{a}_s(k) = 1\right\} \sim \mathcal{CN}\left(\mathbf{c}_s(k), [\mathbf{S}_{k,k}]^{-1} + \mathcal{P}_d(k)\right)$. We thus obtain $\hat{\mathbf{c}}_s = \mathbf{h}_{\mathrm{LS}}$. Using the estimate $\hat{\mathbf{c}}_s(k) = \mathbf{h}_{\mathrm{LS}}(k)$ and conditioning on $\mathbf{a}_s(k) = a,\ a \in \{0, 1\}$, the MMSE estimate of the diffuse component $\mathbf{h}_d(k)$ is given by

$$\hat{\mathbf{h}}_d^{(a)}(k) = \mathbb{E}\left[\mathbf{h}_d(k) | \mathbf{h}_{\mathrm{LS}}(k), \hat{\mathbf{c}}_s(k), \hat{\mathbf{a}}_s(k) = a\right] = \frac{\mathbf{S}_{k,k}\mathcal{P}_d(k)}{1 + \mathbf{S}_{k,k}\mathcal{P}_d(k)} (1 - a)\, \mathbf{h}_{\mathrm{LS}}(k). \tag{A.13}$$

Finally, by combining the estimates $\hat{\mathbf{a}}_s$, $\hat{\mathbf{c}}_s$ and $\hat{\mathbf{h}}_d^{(a)}$, the overall HSD estimate is given by

$$\hat{\mathbf{h}}(k) = \hat{\mathbf{a}}_s(k)\mathbf{h}_{\mathrm{LS}}(k) + (1 - \hat{\mathbf{a}}_s(k)) \frac{\mathbf{S}_{k,k}\mathcal{P}_d(k)}{1 + \mathbf{S}_{k,k}\mathcal{P}_d(k)} \mathbf{h}_{\mathrm{LS}}(k). \tag{A.14}$$

We now develop the MMSE and MAP estimates of $\mathbf{a}_s(k)$.

A.5.3 Generalized MMSE Estimator

The MMSE estimate of the sparsity bit $\mathbf{a}_s(k)$ is given by

$$\hat{\mathbf{a}}_s^{(\mathrm{GMMSE})}(k) = \mathbb{E}\left[\mathbf{a}_s(k) | \mathbf{h}_{\mathrm{LS}}(k), \hat{\mathbf{c}}_s(k)\right] = \Pr\left(\mathbf{a}_s(k) = 1 | \mathbf{h}_{\mathrm{LS}}(k), \hat{\mathbf{c}}_s(k)\right). \tag{A.15}$$

Using Bayes' rule, $\hat{\mathbf{c}}_s(k) = \mathbf{h}_{\mathrm{LS}}(k)$, and assuming $\mathbf{a}_s(k) \sim \mathcal{B}(\tilde{q})$, we have

$$\hat{\mathbf{a}}_s^{(\mathrm{GMMSE})}(k) = \frac{1}{1 + e^{\alpha} \exp\left\{-\frac{\mathbf{S}_{k,k}|\mathbf{h}_{\mathrm{LS}}(k)|^2}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}\right\}}, \tag{A.16}$$

where we have defined $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right)$.

A.5.4 Generalized Thresholding Estimator

Using Bayes' rule and the ML estimate $\hat{\mathbf{c}}_s(k) = \mathbf{h}_{\mathrm{LS}}(k)$, the MAP estimate of $\mathbf{a}_s$ is given by

$$\hat{\mathbf{a}}_s^{(\mathrm{GThres})}(k) = \arg\max_{a\in\{0,1\}} \left\{\ln \Pr\left(\mathbf{a}_s(k) = a | \mathbf{h}_{\mathrm{LS}}(k), \hat{\mathbf{c}}_s(k)\right)\right\} \tag{A.17}$$

$$= \arg\min_{a\in\{0,1\}} \left\{(1-a)\frac{\mathbf{S}_{k,k}\left|\mathbf{h}_{\mathrm{LS}}(k)\right|^2}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)} + a\ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right)\right\} = \mathcal{I}\left(|\mathbf{h}_{\mathrm{LS}}(k)|^2 \geq \alpha\left(1/\mathbf{S}_{k,k} + \mathcal{P}_d(k)\right)\right).$$

This solution consists in a thresholding of the LS estimate, hence the name *Generalized Thresholding estimator*, where the diffuse component represents noise for the estimation of the sparse coefficients. For this reason, the threshold is proportional, by a factor α, to the sum of the noise strength $1/\mathbf{S}_{k,k}$ and the power of the diffuse component $\mathcal{P}_d(k)$. It is worth noting that, if $\alpha \leq 0$ (*i.e.*, $\tilde{q} \geq \frac{1}{2}$), then $\hat{\mathbf{a}}_s^{(\mathrm{GThres})}(k) = 1$, and the GThres estimator trivially reduces to the LS solution.

A.6 MSE analysis

Let $\hat{\mathbf{h}}^{(X)}$ be any estimator, where X is an estimator label. We define the MSE of the estimator $\hat{\mathbf{h}}^{(X)}$, as a function of the SNR matrix $\mathbf{S}$, as

$$\mathrm{MSE}^{(X)}(\mathbf{S}) = \frac{1}{L}\mathbb{E}\left[\left\|\hat{\mathbf{h}}^{(X)} - \mathbf{h}\right\|_2^2\right] = \frac{1}{L}\sum_k \mathrm{MSE}_k^{(X)}\left(\mathbf{S}_{k,k}\right), \tag{A.18}$$

where, owing to the use of per-tap estimation approaches, the sum is over the MSE terms associated with the estimation of the kth channel coefficient, *i.e.*,

$$\mathrm{MSE}_k^{(X)}\left(\mathbf{S}_{k,k}\right) = \mathbb{E}\left[\left|\hat{\mathbf{h}}^{(X)}(k) - \mathbf{h}(k)\right|^2\right]. \tag{A.19}$$

The expectation is computed with respect to the joint probability distribution $p(\mathbf{a}_s)p(\mathbf{c}_s)p(\mathbf{h}_d)p(\mathbf{n})$. In this section, we study the asymptotic behavior of each term $\mathrm{MSE}_k^{(X)}\left(\mathbf{S}_{k,k}\right), k = 0, \ldots, L-1$, in the limit of high ($\mathbf{S}_{k,k} \to +\infty$) and low ($\mathbf{S}_{k,k} \to 0^+$) SNR.

For the sake of a more concise notation, we define $y = \mathbf{h}_{\mathrm{LS}}(k)$, $\hat{h}(y) = \hat{\mathbf{h}}(k)$, $a_s = \mathbf{a}_s(k)$, $c_s = \mathbf{c}_s(k)$, $h_d = \frac{1}{\sqrt{\mathcal{P}_d(k)}}\mathbf{h}_d(k)$ (normalized to have unit variance), $h = \mathbf{h}(k)$, $n = \mathbf{n}(k)$, $S = \mathbf{S}_{k,k}$ and $P_d = \mathcal{P}_d(k)$. From (A.8) and (A.7), we can then rewrite the observation model associated with the kth channel entry as $y = a_s c_s + \sqrt{P_d}h_d + \frac{1}{\sqrt{S}}n$, where $a_s \sim \mathcal{B}(q)$, $h_d \sim \mathcal{CN}(0,1)$, $n \sim \mathcal{CN}(0,1)$.

For the LS estimator, we have $\mathrm{mse}_k^{(\mathrm{LS})}(S) \triangleq S\mathrm{MSE}_k^{(\mathrm{LS})}(S) = \mathbb{E}\left[S\left|y-h\right|^2\right] = 1$. Hence,

the normalized MSE, $\mathrm{mse}_k^{(\mathrm{LS})}(S)$, is a constant, independent of the SNR. Herein, we show that the GMMSE and GThres estimators exhibit the same behavior in the asymptotic high and low SNR, *i.e.*, letting $\mathrm{mse}_k^{(X)}(S) \triangleq S\mathrm{MSE}_k^{(X)}(S)$, we have

$$\lim_{S\to 0(\infty)} \mathrm{mse}_k^{(X)}(S) = \text{const.} > 0,\ X \in \{\mathrm{GMMSE}, \mathrm{GThres}\},$$

for a proper constant, which depends on the asymptotic regime and on the estimator. To this end, let

$$f^{(X)}\left(\sqrt{S}y, n\right) = S\left|\hat{h}(y) - h\right|^2. \tag{A.20}$$

Then, we have

$$\mathrm{mse}_k^{(X)}(S) = \mathbb{E}\left[f^{(X)}\left(\sqrt{S}h + n, n\right)\right], \tag{A.21}$$

where the expectation is calculated with respect to $h = a_s c_s + \sqrt{P_d}h_d$ and $n \sim \mathcal{CN}(0,1)$, which are independent of the SNR S. From Lemma A.10.1 in Appendix A.A, we can exchange the limit operator with the expectation, yielding, for $S_{\lim} \in \{0, +\infty\}$,

$$\lim_{S\to S_{\lim}} \mathrm{mse}_k^{(X)}(S) = \mathbb{E}\left[\lim_{S\to S_{\lim}} f^{(X)}\left(\sqrt{S}h + n, n\right)\right]. \tag{A.22}$$

We evaluate (A.22) for the GMMSE and GThres estimators in Secs. A.6.1 and A.6.2, respectively.

A.6.1 Generalized MMSE estimator

Substituting the expression of the GMMSE estimator (A.14) and (A.16) in (A.20), we obtain, after some algebraic manipulation,

$$f^{(\mathrm{GMMSE})}\left(\sqrt{S}y, n\right) = \left| n - \frac{e^{\alpha}\exp\left\{\frac{-S|y|^2}{1+SP_d}\right\}\frac{\sqrt{S}y}{1+SP_d}}{1 + e^{\alpha}\exp\left\{\frac{-S|y|^2}{1+SP_d}\right\}}\right|^2. \tag{A.23}$$

We distinguish the three cases $S \to +\infty$ with $P_d = 0$, $S \to +\infty$ with $P_d > 0$, and $S \to 0$.

A.6.1.1 High SNR with no diffuse component: $S \to +\infty$, $P_d = 0$

When $P_d = 0$, we have $\sqrt{S}y = \sqrt{S}a_s c_s + n$ and

$$f^{(\mathrm{GMMSE})}\left(\sqrt{S}a_s c_s + n, n\right) = \left| n - \frac{e^{\alpha}\exp\left\{-\left|\sqrt{S}a_s c_s + n\right|^2\right\}}{1 + e^{\alpha}\exp\left\{-\left|\sqrt{S}a_s c_s + n\right|^2\right\}}\left(\sqrt{S}a_s c_s + n\right)\right|^2 .$$

In the limit of high SNR, we obtain

$$\begin{cases} \lim\limits_{S\to+\infty} f^{(\mathrm{GMMSE})}\left(\sqrt{S}c_s + n, n\right) = |n|^2, \ a_s = 1, \ \text{a.e.}, \\ \lim\limits_{S\to+\infty} f^{(\mathrm{GMMSE})}\left(n, n\right) = \frac{|n|^2}{\left(1+e^{\alpha}\exp\left\{-|n|^2\right\}\right)^2}, \ a_s = 0, \end{cases}$$

where a.e. stands for *almost everywhere*, *i.e.*, the limit holds except on a set with probability measure zero. In particular, this set is given by $\{c_s = 0\}$, which has probability measure zero since c_s is a continuous random variable. From (A.22), by averaging over $a_s \sim \mathcal{B}(q)$ and $n \sim \mathcal{CN}(0,1)$, we thus obtain

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GMMSE})}(S) = q\mathbb{E}\left[|n|^2\right] + (1-q)\mathbb{E}\left[\frac{|n|^2}{\left(1 + e^{\alpha}\exp\left\{-|n|^2\right\}\right)^2}\right] = q + (1-q)g(\alpha),$$

where we have defined $g(\alpha) = e^{-\alpha}\ln\left(1 + e^{\alpha}\right)$ and we have used Lemma A.10.2 in Appendix A.A. Therefore, in the high SNR regime (*i.e.*, letting $\sigma_w^2 \to 0$, which scales the SNR matrix $\mathbf{S}$ to infinity) with no diffuse component, $\mathcal{P}_d(k) = 0, \forall k$, using (A.18), we obtain the limiting MSE behavior

$$\mathrm{MSE}^{(\mathrm{GMMSE})}(\mathbf{S}) = \frac{1}{L}\sum_{k=0}^{L-1}\frac{\mathrm{mse}_k^{(\mathrm{GMMSE})}(\mathbf{S}_{k,k})}{\mathbf{S}_{k,k}} \simeq_{\infty} \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})\left(q + (1-q)g(\alpha)\right),$$

where we have defined $\simeq_{\infty}$ as the high SNR approximation, and we have denoted the MSE of the LS estimator as $\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S}) = \frac{1}{L}\mathrm{tr}\left(\mathbf{S}^{-1}\right)$.

A.6.1.2 High SNR with diffuse component: $S \to +\infty$, $P_d > 0$

From (A.23), we have $\lim_{S\to+\infty} f^{(\mathrm{GMMSE})}\left(\sqrt{S}h + n, n\right) = |n|^2$. Then, from (A.22),

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GMMSE})}(S) = \mathbb{E}\left[|n|^2\right] = 1. \tag{A.24}$$

From (A.18), the limiting behavior of the overall MSE in the high SNR, with $\mathcal{P}_d(k) > 0, \forall k$, is given by $\mathrm{MSE}^{(\mathrm{GMMSE})}(\mathbf{S}) \simeq_\infty \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})$.

A.6.1.3 Low SNR: $S \to 0$

From (A.23), we have

$$\lim_{S\to 0} f^{(\mathrm{GMMSE})}\left(\sqrt{S}h + n, n\right) = \left| \frac{n}{1 + e^{\alpha} \exp\left\{ - |n|^2 \right\}} \right|^2 .$$

Then, using (A.22) and Lemma A.10.2 in Appendix A.A, we obtain

$$\lim_{S\to 0} \mathrm{mse}_k^{(\mathrm{GMMSE})}(S) = \mathbb{E}\left[\left| \frac{n}{1 + e^{\alpha} \exp\left\{ - |n|^2 \right\}} \right|^2 \right] = g(\alpha).$$

Then, from (A.18), the overall MSE in the low SNR regime behaves like

$$\mathrm{MSE}^{(\mathrm{GMMSE})}(\mathbf{S}) \simeq_0 \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S}) g(\alpha), \tag{A.25}$$

where we have defined $\simeq_0$ as the low SNR approximation.

A.6.2 Generalized Thresholding estimator

Substituting the expression of the GThres estimator (A.14) and (A.17) in (A.20), we obtain, after some algebraic manipulation,

$$\begin{aligned} f^{(\mathrm{GThres})}\left(\sqrt{S}h + n, n\right) &= \mathcal{I}\left(\left| \sqrt{S}h + n \right|^2 \geq \alpha(1 + SP_d) \right) |n|^2 \\ &+ \mathcal{I}\left(\left| \sqrt{S}h + n \right|^2 < \alpha(1 + SP_d) \right) \left| \frac{\sqrt{S}h - SP_d n}{1 + SP_d} \right|^2 . \end{aligned} \tag{A.26}$$

Note that, if $\alpha \leq 0$, then we have a trivial thresholding operation, and the estimator is equivalent to LS. This case is of no interest. In the following, therefore, we study the case $\alpha > 0$.

Similarly to the GMMSE estimator, we distinguish the three cases $S \to +\infty$ with $P_d = 0$, $S \to +\infty$ with $P_d > 0$, and $S \to 0$.

A.6.2.1 High SNR with no diffuse component: $S \to +\infty$, $P_d = 0$

When $P_d = 0$ we have $y = a_s c_s + \sqrt{S}^{-1} n$ and

$$f^{(\mathrm{GThres})}\left(\sqrt{S}a_s c_s + n, n\right) = \mathcal{I}\left(\left|\sqrt{S}a_s c_s + n\right|^2 \geq \alpha\right)|n|^2 + \mathcal{I}\left(\left|\sqrt{S}a_s c_s + n\right|^2 < \alpha\right)\left|\sqrt{S}a_s c_s\right|^2.$$

We have

$$\begin{cases} \lim\limits_{S\to+\infty} f^{(\mathrm{GThres})}\left(\sqrt{S}c_s + n, n\right) = |n|^2,\ a_s = 1,\ \text{a.e.}, \\ \lim\limits_{S\to+\infty} f^{(\mathrm{GThres})}\left(n, n\right) = \mathcal{I}\left(|n|^2 \geq \alpha\right)|n|^2,\ a_s = 0, \end{cases}$$

where the first limit holds a.e., *i.e.*, except on the set with zero probability measure $\{c_s = 0\}$. From (A.22), we then obtain

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GThres})}(S) = q\mathbb{E}\left[|n|^2\right] + (1-q)\mathbb{E}\left[\mathcal{I}\left(|n|^2 \geq \alpha\right)|n|^2\right] = q + (1-q)w(\alpha),$$

where in the last step we have used the fact that $|n|^2 \sim \mathcal{E}(1)$ to compute the second expectation term, and we have defined $w(\alpha) = e^{-\alpha}(1+\alpha)$. Then, from (A.18), the overall MSE in the high SNR regime with $\mathcal{P}_d(k) = 0, \forall k$, behaves like

$$\mathrm{MSE}^{(\mathrm{GThres})}(\mathbf{S}) \simeq_\infty \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})\left(q + (1-q)e^{-\alpha}(1+\alpha)\right).$$

A.6.2.2 High SNR with diffuse component: $S \to +\infty$, $P_d > 0$

From (A.26), we have $\lim\limits_{S\to+\infty} f^{(\mathrm{GThres})}\left(\sqrt{S}h + n, n\right) = |n|^2$. Then, from (A.22), we obtain

$$\lim_{S\to+\infty} \mathrm{mse}_k^{(\mathrm{GThres})}(S) = \mathbb{E}\left[|n|^2\right] = 1. \tag{A.27}$$

Therefore, in the high SNR regime with $\mathcal{P}_d(k) > 0, \forall k$, the GThres estimator performs like

$$\mathrm{MSE}^{(\mathrm{GThres})}(\mathbf{S}) \simeq_\infty \mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S}). \tag{A.28}$$

A.6.2.3 Low SNR: $S \to 0$

From (A.26), we have

$$\lim_{S\to 0} f^{(\mathrm{GThres})}\left(\sqrt{S}h+n,n\right)=\mathcal{I}\left(|n|^2\geq\alpha\right)|n|^2\,. \tag{A.29}$$

Then, from (A.22), we obtain

$$\lim_{S\to+\infty}\mathrm{mse}_k^{(\mathrm{GThres})}(S)=\mathbb{E}\left[\mathcal{I}\left(|n|^2\geq\alpha\right)|n|^2\right]=w(\alpha).$$

Therefore, in the low SNR regime, the GThres estimator performs like

$$\mathrm{MSE}^{(\mathrm{GThres})}\left(\mathbf{S}\right)\simeq_0\mathrm{MSE}^{(\mathrm{LS})}\left(\mathbf{S}\right)e^{-\alpha}\left(1+\alpha\right). \tag{A.30}$$

A.6.3 Discussion

The asymptotic MSE behavior of the GMMSE and GThres estimators is summarized in Table A.2. A plot is given in Fig. A.1. We compare their limiting behavior with the (unstructured) LS estimator and with the Oracle estimator, which assumes the HSD model, perfect knowledge of $\mathbf{a}_s$, and treats $\mathbf{c}_s$ as a deterministic unknown vector. The latter, by knowing $\mathbf{a}_s$, performs an LS estimate of $\mathbf{c}_s$ and an MMSE of $\mathbf{h}_d$. Its MSE as a function of the SNR matrix $\mathbf{S}$ is given by

$$\mathrm{MSE}^{(\mathrm{Oracle})}\left(\mathbf{S}\right)=q\mathrm{MSE}^{(\mathrm{LS})}\left(\mathbf{S}\right)+\frac{1-q}{L}\sum_{k=0}^{L-1}\frac{\mathcal{P}_d(k)}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}.$$

The limiting MSE behavior in the table is normalized to $\mathrm{MSE}^{(\mathrm{LS})}\left(\mathbf{S}\right)$. Then, a value smaller than 1 indicates that the estimation accuracy, in the corresponding regime, improves over LS. Moreover, the smaller the value, the better the asymptotic MSE accuracy.

Notice that, in the high SNR with diffuse component, all estimators achieve the LS MSE accuracy. In fact, in this regime the diffuse component is strong compared to the noise level, *i.e.*, $\mathcal{P}_d(k)\gg 1/\mathbf{S}_{k,k}$, hence the observed channel exhibits a dense structure, yielding the same accuracy as LS. On the other hand, in the high SNR with no diffuse component, the GMMSE and GThres estimators achieve a better estimation accuracy than LS. Their limiting behavior can be explained as follows. When $\mathbf{a}_s(k)=1$ (with probability q), the active sparse coefficients $\mathbf{c}_s(k)$, which are much stronger than the noise background in the high SNR, are always correctly detected, and are estimated

Table A.2. *Asymptotic MSE behavior of* LS, Oracle, GMMSE *and* GThres *estimators.* $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right), g(\alpha) = e^{-\alpha}\ln\left(1+e^{\alpha}\right), w(\alpha) = e^{-\alpha}\left(1+\alpha\right)$.

$\frac{\mathrm{MSE}^{(\mathbf{X})}(\mathbf{S})}{\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})}$	**High SNR, $\mathbf{\Lambda}_d = \mathbf{0}$**	**High SNR, $\mathbf{\Lambda}_d \succ 0$**	**Low SNR**
LS, GThres, $\alpha \leq 0$	1	1	1
Oracle	q	1	q
GMMSE	$q + (1-q)g(\alpha)$	1	$g(\alpha)$
GThres, $\alpha > 0$	$q + (1-q)w(\alpha)$	1	$w(\alpha)$

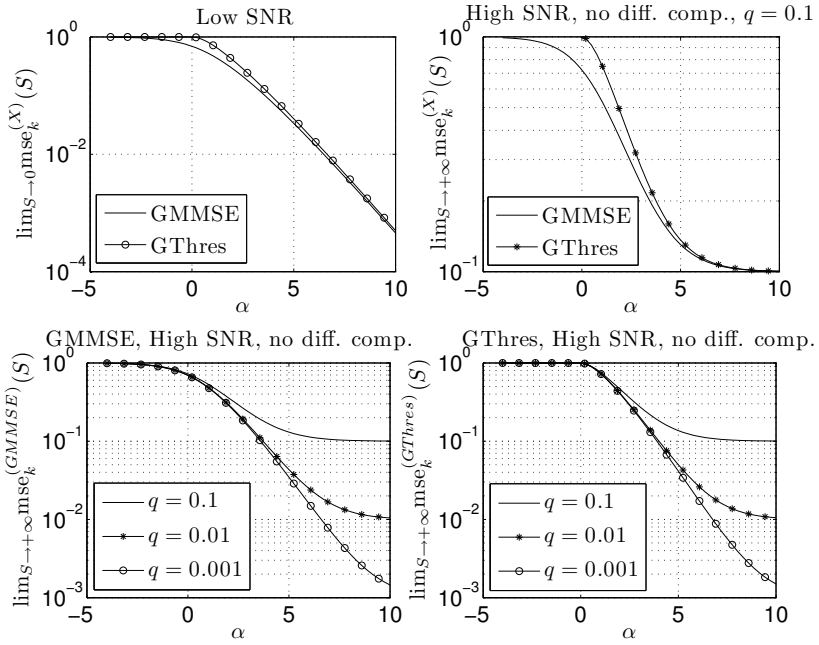


Figure A.1. *High and Low asymptotic SNR behavior of the* GMMSE *and* GThres *estimators as a function of* $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right)$.

with the same estimation accuracy as LS. On the other hand, when $\mathbf{a}_s(k) = 0$ (with probability $1-q$), the GMMSE (respectively, GThres) estimator incurs a mis-detection error $\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})g(\alpha)$ ($\mathrm{MSE}^{(\mathrm{LS})}(\mathbf{S})w(\alpha)$), due to strong noise samples which are mis-detected as active sparse components.

Moreover, since $g(\alpha)$ and $w(\alpha)$ are decreasing functions of $\alpha \in \mathbb{R}$ (*i.e.*, increasing functions of $\tilde{q} \in (0,1)$), with $\lim_{\alpha\to-\infty} g(\alpha) = w(0) = 1$ and $\lim_{\alpha\to+\infty} g(\alpha) = \lim_{\alpha\to+\infty} w(\alpha) = 0$, the MSE is a decreasing function of α (*i.e.*, an increasing function of $\tilde{q}$). In particular, for small values of α, the estimates of $\mathbf{a}_s$ in (A.16) and (A.17) approach 1 for both the GMMSE and the GThres estimators, hence the overall HSD estimate (A.14) approaches the LS solution, yielding the same LS accuracy. Conversely, for increasing values of α, the GMMSE and GThres estimators approach the MSE

accuracy of the Oracle estimator. Note that, the larger α, the larger the threshold level of the GThres estimator in (A.17), hence the fewer noise samples are mis-detected as active sparse components, and the smaller the overall mis-detection error and MSE accuracy (a similar interpretation holds for the GMMSE estimator).

Similarly, in the low SNR, the MSE of the GMMSE and GThres estimators is a decreasing function of α. In particular, a better MSE than the Oracle estimator is achieved for α sufficiently large. In fact, the main source of error is associated with the LS estimates of the sparse coefficients. On the other hand, the MMSE estimate of the diffuse component is forced to zero at small SNR values, hence the resulting MSE approaches the channel energy floor. Therefore, the larger α (alternatively, the smaller $\tilde{q}$), the smaller the weight given to the LS estimates of the sparse coefficients in (A.14) with respect to the MMSE estimates of the diffuse coefficients, and the better the estimation accuracy. In the limit $\alpha \to +\infty$ (*i.e.*, $\tilde{q} \to 0^+$), the GMMSE and GThres estimators treat the channel as being purely diffuse, hence the MMSE estimate of the channel is forced to zero and the MSE approaches the channel energy floor.

We conclude that, in the asymptotic SNR regimes, using $\alpha > \ln \frac{1-q}{q}$ (*i.e.*, $\tilde{q} < q$) improves the performance of the GMMSE and GThres estimators compared to assuming the true sparsity prior q. Hence, it is beneficial to use a conservative approach, *i.e.*, to assume the sparse component to be sparser than it actually is. However, this behavior does not always hold for medium SNR, where in fact a larger α (*i.e.*, a smaller $\tilde{q}$) may induce a larger MSE. This behavior can be seen by studying the two extreme cases $\alpha \to -\infty$ and $\alpha \to +\infty$, *i.e.*, $\tilde{q} \to 1$ and $\tilde{q} \to 0$, respectively. In the first case ($\alpha \to -\infty$, $\tilde{q} \to 1$), the two estimators are equivalent to LS, yielding the same MSE accuracy as LS. Conversely, when $\alpha \to +\infty$ (*i.e.*, $\tilde{q} \to 0^+$), the channel is treated as being diffuse only and is estimated via MMSE. The MSE in this case is given by

$$\begin{aligned} \mathrm{MSE}^{(\mathrm{Diff})}(\mathbf{S}) &= \frac{1}{L}\sum_{k=0}^{L-1} \mathbb{E}\left[\left|\frac{\mathbf{S}_{k,k}\mathcal{P}_d(k)}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}\mathbf{h}_{LS}(k) - \mathbf{h}(k)\right|^2\right] \\ &= \frac{1}{L}\sum_{k=0}^{L-1}\left(q\frac{\mathcal{P}_s(k)}{\left(1+\mathbf{S}_{k,k}\mathcal{P}_d(k)\right)^2} + \frac{\mathcal{P}_d(k)}{1+\mathbf{S}_{k,k}\mathcal{P}_d(k)}\right), \end{aligned} \tag{A.31}$$

which performs worse than LS, for any value of the SNR matrix $\mathbf{S}$, for sufficiently large values of $\mathcal{P}_s(k), k = 0, \ldots, L-1$. Hence, in medium SNR we expect a trade-off between large values of α (*i.e.*, small values of $\tilde{q}$), which induce sparsity in the estimate of the sparse component, and small values of α, which, on the other hand, induce a less sparse solution and privilege the diffuse channel

component.

It is worth noting that the MMSE estimator of the channel, which assumes perfect knowledge of q, $\mathbf{\Lambda}_s$ and $\mathbf{\Lambda}_d$, minimizes the MSE when the true sparsity level $\tilde{q} = q$ is employed. We conclude that the uncertainty about the sparse coefficients, which are treated as deterministic and unknown under the GMMSE and GThres estimators, is compensated by employing a conservative approach in the estimation of the sparse component.

Finally, for a given value of α, the GMMSE estimator achieves a better MSE accuracy than the GThres estimator, in the asymptotic regimes. In fact, the MMSE estimate of $\mathbf{a}_s(k)$ (A.16), *i.e.*, the posterior probability of $\mathbf{a}_s(k) = 1$, incorporates also the reliability associated with an active sparse component, and therefore, the closer the estimate to one, the more likely an active sparse component. On the other hand, the MAP estimate of $\mathbf{a}_s(k)$, by allowing only the two extreme values of $\hat{\mathbf{a}}_s(k) \in \{0, 1\}$, completely discards the reliability associated with these estimates, thus incurring a performance degradation.

A.7 Structured PDP Estimation of the diffuse component

In the derivation of the GMMSE and GThres estimators in the previous section, we have assumed that the PDP of the diffuse component $\mathbf{h}_d$ is perfectly known at the receiver. However, in a practical system, this is unknown, and therefore needs to be estimated.

Herein, we develop a structured estimate of the PDP $\mathcal{P}_d$, when the observation interval is too short to allow time-averaging over the small scale fading. By exploiting prior information about the structure of the PDP, we can average the small scale fading over the delay dimension, rather than over subsequent realizations of the fading process, thus enhancing the estimation accuracy.

We assume an exponential PDP model [101, 113, 114] $\mathcal{P}_d(k) = \beta e^{-\omega k}, \quad k = 0, \ldots, L-1$, where the deterministic, unknown parameters $\beta \geq 0$ and $\omega \geq 0$ represent the relative power and the decay rate of the PDP, respectively. We derive an ML estimate of these parameters, using the EM algorithm (the general EM framework is presented in, *e.g.*, [115]). For simplicity, we assume a single channel snapshot. However, the following derivation can be extended to include a sequence of channel realizations. Moreover, we treat the vector $\mathbf{c}_s$ as a deterministic unknown parameter, and we assume a sparsity level $\tilde{q}$ (possibly, $\neq q$), which is consistent with the design choice of the GMMSE and GThres estimators.

Let the HSD channel and the observed sequence be given by (A.8) and (A.7), respectively.

From (A.8), if $\mathbf{a}_s(k) = 1$, then $\mathbf{h}_{\rm LS}(k) = \mathbf{c}_s(k) + \mathbf{h}_d(k) + \sqrt{\mathbf{S}_{k,k}}^{-1}\mathbf{n}(k)$. In this case, since $\mathbf{c}_s(k)$ is a deterministic, unknown parameter, the observed sample $\mathbf{h}_{\rm LS}(k)$ does not provide statistical information to estimate the diffuse component (hence, its power). In fact, the ML estimate of $\mathbf{c}_s(k)$ is $\hat{\mathbf{c}}_s(k) = \mathbf{h}_{\rm LS}(k)$ (A.12). The estimated contribution from the noise and the diffuse component is then $\mathbf{h}_{\rm LS}(k) - \hat{\mathbf{c}}_s(k) = 0$, and the estimate of $\mathbf{h}_d(k)$, given by (A.13), is forced to zero. Therefore, the observations corresponding to the active sparse components should be neglected. Conversely, all the statistical information to estimate the PDP parameters ω and β is contained in the vector $(\mathbf{1} - \mathbf{a}_s) \odot \mathbf{h}_{\rm LS} = (\mathbf{1} - \mathbf{a}_s) \odot (\mathbf{h}_d + \sqrt{\mathbf{S}}^{-1}\mathbf{n})$, which is obtained by zeroing the contribution from the active sparse components. Unfortunately, $\mathbf{a}_s$ is unknown in advance, hence it needs to be estimated from the observed sequence.

In employing the EM algorithm to estimate the PDP parameters β and ω, we assume $\mathbf{a}_s$ and $(\mathbf{1} - \mathbf{a}_s) \odot \mathbf{h}_d$ as the hidden variables. Moreover, we discard the contribution of the active sparse components to the observed sequence, as justified above. Then, letting $\hat{\beta}$, $\hat{\omega}$ be the current estimates of the deterministic unknown parameters β and ω, respectively, in the E-step we compute

$$\begin{aligned}
\mathcal{L}(\beta,\omega;\hat{\beta},\hat{\omega}) &\triangleq -\mathbb{E}\left[\ln p\left((\mathbf{1}-\mathbf{a}_s)\odot\mathbf{h}_{\rm LS},(\mathbf{1}-\mathbf{a}_s)\odot\mathbf{h}_d,\mathbf{a}_s|\,\beta,\omega\right)|\,\mathbf{h}_{\rm LS},\hat{\beta},\hat{\omega}\right] \qquad \text{(A.32)}\\
&\overset{(a)}{=} -\mathbb{E}\left[\ln p\left((\mathbf{1}-\mathbf{a}_s)\odot\mathbf{h}_{\rm LS}|\,(\mathbf{1}-\mathbf{a}_s)\odot\mathbf{h}_d,\mathbf{a}_s\right)|\,\mathbf{h}_{\rm LS},\hat{\beta},\hat{\omega}\right] - \mathbb{E}\left[\ln p\left(\mathbf{a}_s\right)|\,\mathbf{h}_{\rm LS},\hat{\beta},\hat{\omega}\right]\\
&\quad -\mathbb{E}\left[\ln p\left((\mathbf{1}-\mathbf{a}_s)\odot\mathbf{h}_d|\,\mathbf{a}_s,\beta,\omega\right)|\,\mathbf{h}_{\rm LS},\hat{\beta},\hat{\omega}\right] \overset{(b)}{\propto} -\mathbb{E}\left[\ln p\left((\mathbf{1}-\mathbf{a}_s)\odot\mathbf{h}_d|\,\mathbf{a}_s,\beta,\omega\right)|\,\mathbf{h}_{\rm LS},\hat{\beta},\hat{\omega}\right]\\
&\overset{(c)}{=} -\sum_{\mathbf{x}\in\{0,1\}^L}\Pr\left(\mathbf{a}_s=\mathbf{x}|\,\mathbf{h}_{\rm LS},\hat{\beta},\hat{\omega},\hat{\mathbf{c}}_s=\mathbf{h}_{\rm LS}\right)\mathbb{E}\left[\ln p\left((\mathbf{1}-\mathbf{a}_s)\odot\mathbf{h}_d|\,\mathbf{a}_s=\mathbf{x},\beta,\omega\right)|\,\mathbf{h}_{\rm LS},\mathbf{a}_s=\mathbf{x},\hat{\beta},\hat{\omega}\right]\\
&= \sum_{k=0}^{L-1}(1-\hat{q}_{\rm post}(k))\left[\ln\left(\beta e^{-\omega k}\right) + \frac{\mathbb{E}\left[|\mathbf{h}_d(k)|^2\Big|\,\mathbf{h}_{\rm LS}(k),\mathbf{a}_s(k)=0,\hat{\beta},\hat{\omega}\right]}{\beta e^{-\omega k}}\right] \triangleq \mathcal{R}(\beta,\omega;\hat{\beta},\hat{\omega})
\end{aligned}$$

where, in the last step, we have defined the posterior probability of an active sparse component

$$\hat{q}_{\rm post}(k) = \Pr\left(\mathbf{a}_s(k)=1|\mathbf{h}_{\rm LS}(k),\hat{\beta},\hat{\omega},\hat{\mathbf{c}}_s(k)=\mathbf{h}_{\rm LS}(k)\right) = \frac{1}{1+\frac{1-\tilde{q}}{\tilde{q}}\exp\left\{-\frac{\mathbf{S}_{k,k}|\mathbf{h}_{\rm LS}(k)|^2}{1+\mathbf{S}_{k,k}\hat{\beta}e^{-\hat{\omega}k}}\right\}}. \quad \text{(A.33)}$$

In particular, in step (a) we have expressed the likelihood function in terms of its conditional probabilities. Moreover, we have used that fact that the term $(\mathbf{1} - \mathbf{a}_s) \odot \mathbf{h}_{\rm LS} = (\mathbf{1} - \mathbf{a}_s) \odot (\mathbf{h}_d + \sqrt{\mathbf{S}}^{-1}\mathbf{n})$ is independent of the PDP parameters β, ω, when conditioned on $(\mathbf{1} - \mathbf{a}_s) \odot \mathbf{h}_d$ and $\mathbf{a}_s$, and the prior distribution of $\mathbf{a}_s$ is independent of β, ω. In step (b), we have neglected the terms which are independent of the optimization parameters β, ω. In step (c), the expectation is first conditioned on $\mathbf{a}_s = \mathbf{x}$,

and then averaged over the posterior probability of $\mathbf{a}_s \in \{0,1\}^L$. The conditional expectation of $|\mathbf{h}_d(k)|^2$ is given by

$$\mathbb{E}\left[\left|\mathbf{h}_d(k)\right|^2 \middle| \mathbf{h}_{\mathrm{LS}}(k), \mathbf{a}_s(k)=0, \hat{\beta}, \hat{\omega}\right] = \frac{\hat{\mathcal{P}}_d(k)^2}{(\hat{\mathcal{P}}_d(k)+1/\mathbf{S}_{k,k})^2}\left|\mathbf{h}_{\mathrm{LS}}(k)\right|^2 + \frac{\hat{\mathcal{P}}_d(k)}{1+\hat{\mathcal{P}}_d(k)\mathbf{S}_{k,k}}, \quad \text{(A.34)}$$

where $\hat{\mathcal{P}}_d(k) = \hat{\beta}e^{-\hat{\omega}k}$ is the current estimate of the prior variance of $\mathbf{h}_d(k)$. In the M-step, the term $\mathcal{L}(\beta,\omega;\hat{\beta},\hat{\omega})$ is minimized with respect to the optimization parameters β,ω. We obtain

$$\left\{\tilde{\beta},\tilde{\omega}\right\} = \underset{\beta\geq 0,\omega\geq 0}{\arg\min}\mathcal{L}(\beta,\omega;\hat{\beta},\hat{\omega}) = \underset{\beta\geq 0,\omega\geq 0}{\arg\min}\mathcal{R}(\beta,\omega;\hat{\beta},\hat{\omega}) = \underset{\beta\geq 0,\omega\geq 0}{\arg\min}\sum_{k=0}^{L-1}(1-\hat{q}_{\mathrm{post}}(k))\ln\left(\beta e^{-\omega k}\right)$$
$$+\sum_{k=0}^{L-1}(1-\hat{q}_{\mathrm{post}}(k))\frac{\mathbb{E}\left[\left|\mathbf{h}_d(k)\right|^2\middle|\mathbf{h}_{\mathrm{LS}}(k),\mathbf{a}_s(k)=0,\hat{\beta},\hat{\omega}\right]}{\beta e^{-\omega k}}. \quad \text{(A.35)}$$

By defining, for $k=0,\ldots,L-1$,

$$\begin{cases} A_k = \frac{L(1-\hat{q}_{\mathrm{post}}(k))\mathbb{E}\left[\left|\mathbf{h}_d(k)\right|^2\middle|\mathbf{h}_{\mathrm{LS}}(k),\mathbf{a}_s(k)=0,\hat{\beta},\hat{\omega}\right]}{\sum_{p=0}^{L-1}(1-\hat{q}_{\mathrm{post}}(p))}, \\ Z = \frac{\sum_{p=0}^{L-1}p(1-\hat{q}_{\mathrm{post}}(p))}{\sum_{p=0}^{L-1}(1-\hat{q}_{\mathrm{post}}(p))}, \end{cases} \quad \text{(A.36)}$$

the M-step (A.35) is equivalent to

$$\left\{\tilde{\beta},\tilde{\omega}\right\} = \arg\min_{\beta\geq 0,\omega\geq 0}\ln\beta - \omega Z + \frac{1}{\beta L}\sum_{k=0}^{L-1}A_k e^{\omega k}. \quad \text{(A.37)}$$

We have the following theorem.

Theorem A.7.1. *There is a unique solution* $\left\{\tilde{\beta},\tilde{\omega}\right\}$ *to*

$$\left\{\tilde{\beta},\tilde{\omega}\right\} = \arg\min_{\beta\geq 0,\omega\geq 0}\ln\beta - \omega Z + \frac{1}{\beta L}\sum_{k=0}^{L-1}A_k e^{\omega k}. \quad \text{(A.38)}$$

If $\sum_{k=0}^{L-1}(Z-k)A_k > 0$*, then* $\tilde{\omega}$ *is the unique solution in* $(0,+\infty)$ *of*

$$\sum_{k=0}^{L-1}(Z-k)A_k e^{\tilde{\omega}k} = 0. \quad \text{(A.39)}$$

Otherwise, $\tilde{\omega}=0$*. In both cases,* $\tilde{\beta} = \frac{1}{L}\sum_{k=0}^{L-1}A_k e^{\tilde{\omega}k}$.

Proof. See Appendix A.B. □

Note that, when $\sum_{k=0}^{L-1}(Z-k)A_k > 0$, the solution is a zero of a Lth order polynomial, therefore we must recur to approximate solutions. Since the solution we seek satisfies $e^{-\tilde{\omega}} \in (0,1]$, and we have proved that it is unique, we recur to the *bisection method* [61] to determine an approximate zero $\tilde{x} = e^{-\tilde{\omega}}$ of (A.39).

Finally, the overall EM algorithm consists in the iterations of the E-step (A.33), (A.36) and the M-step (A.37). The algorithm may be initialized by neglecting the noise and the sparse component, *i.e.*, assuming $\mathbf{S}_{k,k} \to +\infty$ and $\tilde{q} = 0$ in the first stage. In this case, we have $\hat{q}_{\text{post}}(k) = 0, \forall k$ in (A.33) and the parameters of the E-step (A.36) are given by

$$\begin{cases} A_k = |\mathbf{h}_{\text{LS}}(k)|^2, \quad k = 0, \ldots, L-1 \\ Z = \frac{L-1}{2}. \end{cases} \tag{A.40}$$

It is worth noting that, if we had assumed the diffuse component $\mathbf{h}_d$, rather than $(\mathbf{1} - \mathbf{a}_s) \odot \mathbf{h}_d$, as the hidden variable, and we had used all the observed sequence $\mathbf{h}_{\text{LS}}$ to estimate the unknown PDP parameters instead of $(\mathbf{1} - \mathbf{a}_s) \odot \mathbf{h}_{\text{LS}}$, then in the M-step we would have

$$\left\{\tilde{\beta}, \tilde{\omega}\right\} = \arg \min_{\beta \geq 0, \omega \geq 0} \sum_{k=0}^{L-1} (1 - \hat{q}_{\text{post}}(k)) \left[\ln\left(\beta e^{-\omega k}\right) + \frac{\mathbb{E}\left[|\mathbf{h}_d(k)|^2 \middle| \mathbf{h}_{\text{LS}}(k), \mathbf{a}_s(k) = 0, \hat{\beta}, \hat{\omega} \right]}{\beta e^{-\omega k}} \right]$$
$$+ \sum_{k=0}^{L-1} \hat{q}_{\text{post}}(k) \left[\ln\left(\beta e^{-\omega k}\right) + \frac{\hat{\beta} e^{-\hat{\omega} k}}{\beta e^{-\omega k}\left(1 + \mathbf{S}_{k,k} \hat{\beta} e^{-\hat{\omega} k}\right)} \right], \tag{A.41}$$

where we have used the fact that, since $\hat{\mathbf{c}}_s = \mathbf{h}_{\text{LS}}$, $\mathbb{E}\left[|\mathbf{h}_d(k)|^2 \middle| \mathbf{h}_{\text{LS}}(k), \hat{\mathbf{c}}_s(k), \mathbf{a}_s(k) = 1, \hat{\beta}, \hat{\omega} \right] = \frac{\hat{\beta} e^{-\hat{\omega} k}}{1 + \mathbf{S}_{k,k} \hat{\beta} e^{-\hat{\omega} k}}$. By comparing this expression with (A.35), we note one additional term. In particular, the observations associated with high probability $\hat{q}_{\text{post}}(k) \to 1$ with an active sparse component give a significant contribution to the log-likelihood function. However, these observations do not provide information about the diffuse component $\mathbf{h}_d$, since $\mathbf{c}_s$ is a deterministic, unknown vector. Conversely, in (A.35), these observations yield a negligible contribution.

Choice of the sparsity level $\tilde{q}$

We next discuss the choice of the parameter $\tilde{q}$ used to estimate the parameters β, ω. Since the EM algorithm solves the ML problem [115], we consider the general problem of maximizing the

likelihood function. Assuming the sparsity level $\tilde{q}$, the ML estimate of β, ω and $\mathbf{c}_s$ is defined as

$$\{\hat{\beta},\hat{\omega},\hat{\mathbf{c}}_s\} = \underset{\beta\geq 0,\omega\geq 0,\mathbf{c}_s}{\arg\max}\, p(\mathbf{h}_{\rm LS}|\beta,\omega,\mathbf{c}_s) = \underset{\beta\geq 0,\omega\geq 0,\mathbf{c}_s}{\arg\max} -\sum_{k=0}^{L-1}\ln\left(1/\mathbf{S}_{k,k}+\mathcal{P}_d(k)\right) + \sum_{k=0}^{L-1}\ln\left(\tilde{q}\exp\left\{-\frac{|\mathbf{h}_{\rm LS}(k)-\mathbf{c}_s(k)|^2}{1/\mathbf{S}_{k,k}+\mathcal{P}_d(k)}\right\}+(1-\tilde{q})\exp\left\{-\frac{|\mathbf{h}_{\rm LS}(k)|^2}{1/\mathbf{S}_{k,k}+\mathcal{P}_d(k)}\right\}\right),$$

where we have used the fact that $\mathbf{h}_{\rm LS}(k)|\mathbf{a}_s(k) = a \sim \mathcal{CN}(a\mathbf{c}_s(k), \mathcal{P}_d(k)+1/\mathbf{S}_{k,k})$ and $\mathcal{P}_d(k) = \beta e^{-\omega k}$. By maximizing over $\mathbf{c}_s$, we obtain $\hat{\mathbf{c}}_s = \mathbf{h}_{\rm LS}$. Then, letting $t_k(\mathcal{P}_d(k)) = \frac{|\mathbf{h}_{\rm LS}(k)|^2}{1/\mathbf{S}_{k,k}+\mathcal{P}_d(k)}$, $s(\tilde{q},t) = \ln\left(t+\frac{1-\tilde{q}}{\tilde{q}}te^{-t}\right)$ and $\mathcal{F}(\tilde{q},\beta,\omega) = \sum_{k=0}^{L-1} s(\tilde{q},t_k(\mathcal{P}_d(k)))$, we obtain

$$\begin{aligned}\{\hat{\beta},\hat{\omega}\} &= \underset{\beta\geq 0,\omega\geq 0}{\arg\max}\sum_{k=0}^{L-1}\left[\ln t_k(\mathcal{P}_d(k)) + \ln\left(1+\frac{1-\tilde{q}}{\tilde{q}}e^{-t_k(\mathcal{P}_d(k))}\right)\right]\\ &= \underset{\beta\geq 0,\omega\geq 0}{\arg\max}\sum_{k=0}^{L-1} s(\tilde{q},t_k(\mathcal{P}_d(k))) = \underset{\beta\geq 0,\omega\geq 0}{\arg\max}\mathcal{F}(\tilde{q},\beta,\omega),\end{aligned}$$

where we have added the term $\sum_{k=0}^{L-1}\ln(|\mathbf{h}_{\rm LS}(k)|^2) - L\ln\tilde{q}$, which does not affect the maximization.

Consider a given pair of parameters (β,ω), and let

$$\begin{aligned}s'(\tilde{q},t) &\triangleq \frac{{\rm d}s(\tilde{q},t)}{{\rm d}t} = \frac{\tilde{q}-(1-\tilde{q})e^{-t}(t-1)}{\tilde{q}t+(1-\tilde{q})te^{-t}},\\ \mathcal{F}'_\beta(\tilde{q},\beta,\omega) &\triangleq \frac{{\rm d}\mathcal{F}(\tilde{q},\beta,\omega)}{{\rm d}\beta} = \sum_{k=0}^{L-1} s'(\tilde{q},t_k(\mathcal{P}_d(k)))\frac{{\rm d}t_k(\mathcal{P}_d(k))}{{\rm d}\beta}.\end{aligned} \tag{A.42}$$

Similarly, we define $\mathcal{F}'_\omega(\tilde{q},\beta,\omega)$ as the derivative with respect to ω. Note that, if $\mathcal{F}'_\beta(\tilde{q},\beta,\omega) > 0$ (< 0), then there is an incentive to augment (diminish) β so as to increase the log-likelihood function $\mathcal{F}(\tilde{q},\beta,\omega)$ (the same consideration holds for $\mathcal{F}'_\omega(\tilde{q},\beta,\omega)$). We now prove that this derivative is a decreasing function of $\tilde{q}$, so that, the larger $\tilde{q}$, the smaller the incentive to increase β (and, possibly, the larger the incentive to decrease it, if the derivative becomes negative). In fact,

$$\frac{{\rm d}s'(\tilde{q},t)}{{\rm d}\tilde{q}} = \frac{1}{\tilde{q}^2}\exp\{-2s(\tilde{q},t)\}t^2e^{-t} > 0,\ \frac{{\rm d}t_k(\mathcal{P}_d(k))}{{\rm d}\beta} = -\frac{1}{\beta}t_k(\mathcal{P}_d(k))\frac{\mathcal{P}_d(k)}{1/\mathbf{S}_{k,k}+\mathcal{P}_d(k)} < 0,$$

and therefore

$$\frac{{\rm d}\mathcal{F}'_\beta(\tilde{q},\beta,\omega)}{{\rm d}\tilde{q}} = \sum_{k=0}^{L-1}\frac{{\rm d}s'(\tilde{q},t_k(\mathcal{P}_d(k)))}{{\rm d}\tilde{q}}\frac{{\rm d}t_k(\mathcal{P}_d(k))}{{\rm d}\beta} < 0.$$

Similarly, we can prove that $\mathcal{F}'_{\omega}(\tilde{q},\beta,\omega)$ is an increasing function of $\tilde{q}$, so that, the larger $\tilde{q}$, the smaller the incentive to decrease ω (and, possibly, the larger the incentive to increase it, if the derivative becomes negative).

Moreover, note that, if $\tilde{q} \geq \frac{1}{1+e^2} \simeq 0.12$, then we have $e^{-t}(t-1) \leq e^{-2} \leq \frac{\tilde{q}}{1-\tilde{q}}$ (since the left hand side is maximized for $t=2$), which implies $s'(\tilde{q},t) \geq 0, \forall t$. We conclude that, when $\tilde{q} \geq \frac{1}{1+e^2}$, the derivatives $F'_{\beta}(\tilde{q},\beta,\omega) < 0, \forall \beta \geq 0, \omega \geq 0$ and $F'_{\omega}(\tilde{q},\beta,\omega) > 0, \forall \beta \geq 0, \omega \geq 0$. Therefore, the ML estimate of β,ω gives $\hat{\beta}=0, \hat{\omega} \to +\infty$, and the PDP estimate is forced to zero.

Conversely, if we let $\tilde{q} \to 0^+$, then the contribution of the sparse component $\mathbf{a}_s \odot \mathbf{c}_s$ is neglected, and the channel is treated as being purely diffuse.

This analysis proves that the prior sparsity level $\tilde{q} \geq 0.12$ should never be used, and suggests the existence of a trade-off in the optimal algorithm parameter $\tilde{q}$, which is confirmed by simulation in Chapter A.9: in order not to force the PDP estimate to zero, $\tilde{q}$ should be "small"; however, in order to take into account the presence of the sparse component in the observations, $\tilde{q}$ should not be "too small". A further investigation on the optimal value of $\tilde{q}$ is left for future work.

A.8 Orthogonality vs non-Orthogonality of the pilot sequence

Thus far, we have assumed an orthogonal pilot sequence, which results in the optimality of per-tap estimation approaches versus joint estimation methods. In this section, we consider the non-orthogonal pilot scenario. We follow two approaches. In Sec. A.8.1, we examine the impact of using an estimator designed under the assumption of an orthogonal pilot sequence on received signals where the pilots are in fact non-orthogonal. We show that, from an MSE perspective, the effect of this mismatch can be characterized via an effective SNR loss. In Sec. A.8.2, we establish a connection between the GThres estimator and the classical sparse approximation algorithms [91–94].

A.8.1 GMMSE and GThres estimators with non-orthogonal pilot sequence

Note that in the non-orthogonal case the SNR matrix $\mathbf{S}$ is non-diagonal. In this case, the observation model associated with the kth delay bin is given by $\mathbf{h}_{\mathrm{LS}}(k) = \mathbf{h}(k) + \left[\sqrt{\mathbf{S}}^{-1}\mathbf{n}\right]_k$, where the noise term $\left[\sqrt{\mathbf{S}}^{-1}\mathbf{n}\right]_k \sim \mathcal{CN}\left(0, \left[\mathbf{S}^{-1}\right]_{k,k}\right)$. Since the GMMSE and GThres estimators, designed under the assumption of orthogonal pilot sequence, operate on a per-tap basis, the non-orthogonal case is obtained by replacing $\mathbf{S}_{k,k}$ with $1/\left[\mathbf{S}^{-1}\right]_{k,k}$ in (A.14), (A.16) and (A.17).

We now evaluate the MSE performance loss induced by a non-orthogonal pilot sequence. Let

$\mathbf{X}$ be the corresponding Toeplitz matrix. Then, the SNR matrix $\mathbf{S} = \frac{\mathbf{X}^*\mathbf{X}}{\sigma_w^2}$ has some non-zero off-diagonal elements. The effective SNR at the kth delay bin is $S_k^{(\mathrm{NO})} \triangleq 1/\left[\mathbf{S}^{-1}\right]_{k,k}$. Therefore, using (A.18) and (A.19), in the non-orthogonal case we have, for $X \in \{\mathrm{GMMSE}, \mathrm{GThres}\}$,

$$\mathrm{MSE}^{(X)}(\mathbf{S}) = \sum_{k=0}^{L-1} \mathrm{MSE}_k^{(X)}\left(1/\left[\mathbf{S}^{-1}\right]_{k,k}\right). \tag{A.43}$$

Now, consider a second scenario where the pilot sequence is orthogonal. Letting $\tilde{\mathbf{X}}$ be the associated Toeplitz matrix, and assuming that the pilot sequence has the same energy budget as in the non-orthogonal case, we have the SNR matrix $\tilde{\mathbf{S}} = \mathrm{diag}(\mathbf{S})$, where $\mathrm{diag}(\mathbf{B})$ is a diagonal matrix with the same diagonal elements as $\mathbf{B}$. The SNR at the kth delay bin is $S_k^{(\mathrm{O})} \triangleq \tilde{\mathbf{S}}_{k,k} = \mathbf{S}_{k,k}$, and the resulting MSE is given by

$$\mathrm{MSE}^{(X)}\left(\tilde{\mathbf{S}}\right) = \sum_{k=0}^{L-1} \mathrm{MSE}_k^{(X)}(\mathbf{S}_{k,k}). \tag{A.44}$$

We now prove that the effective SNRs in the non-orthogonal and orthogonal cases satisfy $S_k^{(\mathrm{O})} \geq S_k^{(\mathrm{NO})}$, $\forall k$. We can rewrite $\mathbf{S}$ as

$$\mathbf{S} = \mathbf{U}\begin{bmatrix} S_k^{(\mathrm{O})} & \mathbf{b} \\ \mathbf{b}^* & \Delta \end{bmatrix}\mathbf{U}^*, \tag{A.45}$$

for a proper $\Delta \succ 0$, row vector $\mathbf{b}$, and permutation matrix $\mathbf{U}$, where we have used the fact that $\mathbf{S}_{k,k} = S_k^{(\mathrm{O})}$. Then, from the inversion formula for 2×2 block-matrices, we have

$$S_k^{(\mathrm{NO})} = \frac{1}{[\mathbf{S}^{-1}]_{k,k}} = \left[\mathbf{U}^*\mathbf{S}^{-1}\mathbf{U}\right]_{1,1}^{-1} = S_k^{(\mathrm{O})} - \mathbf{b}\Delta^{-1}\mathbf{b}^*.$$

Finally, since $\Delta \succ 0$, we obtain $\mathbf{b}\Delta^{-1}\mathbf{b}^* \geq 0$ (with equality if and only if $\mathbf{b} = \mathbf{0}$), which proves the inequality $S_k^{(\mathrm{O})} \geq S_k^{(\mathrm{NO})}$, $\forall k$. Therefore, imperfect orthogonality of the pilot sequence yields a decrease of the effective SNR experienced on each channel delay bin, thus impairing the estimation performance.

We can quantify the loss in the estimation accuracy in the high and low SNR regimes where, as shown in Sec. A.6, for the GMMSE and GThres estimators we have $\lim_{S\to 0(+\infty)} S\mathrm{MSE}_k^{(X)}(S) = \mathrm{constant} > 0$, for a proper constant, as given in Table A.2. To this end, we define the *orthogonality*

coefficient of the pilot sequence associated with the kth delay bin as the ratio between the effective SNR experienced in the non-orthogonal case and the SNR experienced in the orthogonal case, under the same pilot energy budget, *i.e.*,

$$\eta_k = \frac{S_k^{(\mathrm{NO})}}{S_k^{(\mathrm{O})}} = \frac{1}{[\mathbf{S}^{-1}]_{k,k}\,\mathbf{S}_{k,k}} \leq 1. \tag{A.46}$$

Then, in the high and low SNR regimes, the ratio between the MSE in the orthogonal case and the MSE in the non-orthogonal case, in the kth channel bin, is given by

$$\frac{\mathrm{MSE}_k^{(X)}\left(S_k^{(\mathrm{O})}\right)}{\mathrm{MSE}_k^{(X)}\left(S_k^{(\mathrm{NO})}\right)} = \frac{S_k^{(\mathrm{NO})}}{S_k^{(\mathrm{O})}} \times \frac{S_k^{(\mathrm{O})}\mathrm{MSE}_k^{(X)}\left(S_k^{(\mathrm{O})}\right)}{S_k^{(\mathrm{NO})}\mathrm{MSE}_k^{(X)}\left(S_k^{(\mathrm{NO})}\right)} \simeq \eta_k,$$

where we have used the fact that $\lim_{S\to 0(+\infty)} S\mathrm{MSE}_k^{(X)}(S) = \text{constant}$ and the definition (A.46).

A.8.2 Exploiting the non-orthogonality of the pilot sequence

We next investigate estimators designed for the non-orthogonal case, by establishing a connection between the GThres estimator and classical sparse approximation algorithms [91–93]. In particular, we show that the GThres estimator solves

$$\left\{\hat{\mathbf{c}}_s, \hat{\mathbf{a}}_s, \hat{\mathbf{h}}_d\right\} = \arg\max_{\mathbf{c}_s,\mathbf{a}_s,\mathbf{h}_d} p\left(\mathbf{h}_{\mathrm{LS}}, \mathbf{a}_s, \mathbf{h}_d | \mathbf{c}_s\right). \tag{A.47}$$

We have $p\left(\mathbf{h}_{\mathrm{LS}}, \mathbf{a}_s, \mathbf{h}_d|\mathbf{c}_s\right) = p\left(\mathbf{h}_{\mathrm{LS}}|\mathbf{a}_s, \mathbf{h}_d, \mathbf{c}_s\right) p\left(\mathbf{a}_s\right) p\left(\mathbf{h}_d\right)$, where

$$\begin{aligned} &\mathbf{h}_{\mathrm{LS}}|\left\{\mathbf{a}_s, \mathbf{h}_d, \mathbf{c}_s\right\} \sim \mathcal{CN}\left(\mathbf{a}_s \odot \mathbf{c}_s + \mathbf{h}_d, \mathbf{S}^{-1}\right),\ \mathbf{h}_d \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Lambda}_d\right), \\ &p\left(\mathbf{a}_s\right) = \left(\frac{q}{1-q}\right)^{\|\mathbf{a}_s\|_0} (1-q)^L = \left(\frac{q}{1-q}\right)^{\|\mathbf{h}_s\|_0} (1-q)^L, \end{aligned} \tag{A.48}$$

where $\|\mathbf{x}\|_0$ is the $\mathcal{L}_0$-norm of vector $\mathbf{x}$, and $\mathbf{h}_s = \mathbf{a}_s \odot \mathbf{c}_s$ is the sparse component.

Then, from (A.47) and (A.48), we have

$$\begin{aligned} &\left\{\hat{\mathbf{c}}_s, \hat{\mathbf{a}}_s, \hat{\mathbf{h}}_d\right\} = \arg\max_{\mathbf{c}_s,\mathbf{a}_s,\mathbf{h}_d} \ln p\left(\mathbf{h}_{\mathrm{LS}}, \mathbf{a}_s, \mathbf{h}_d|\mathbf{c}_s\right) \\ &= \arg\min_{\mathbf{h}_s=\mathbf{a}_s\odot\mathbf{c}_s,\mathbf{h}_d} \left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s - \mathbf{h}_d\right)^* \mathbf{S}\left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s - \mathbf{h}_s\right) + \alpha\|\mathbf{h}_s\|_0 + \mathbf{h}_d^* \mathbf{\Lambda}_d^{-1} \mathbf{h}_d, \end{aligned} \tag{A.49}$$

where $\alpha = \ln\left(\frac{1-q}{q}\right)$. This can be viewed as an LS regression problem, with a $\mathcal{L}_0$ regularization term associated with $\mathbf{h}_s$, enforcing sparseness of the solution, and a $\mathcal{L}_2$ regularization term associated with $\mathbf{h}_d$, enforcing its Gaussian nature.

Solving with respect to $\mathbf{h}_d$ first, as a function of $\mathbf{h}_s$, we have

$$\hat{\mathbf{h}}_d\left(\mathbf{h}_s\right) = \mathbf{\Lambda}_d\left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-1}\left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right), \tag{A.50}$$

and substituting this solution into the cost function, we obtain the following optimization problem for the sparse component:

$$\hat{\mathbf{h}}_s = \hat{\mathbf{a}}_s \odot \hat{\mathbf{c}}_s = \arg\min_{\mathbf{h}_s} \alpha \left\|\mathbf{h}_s\right\|_0 + \left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right)^* \left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-1}\left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right). \tag{A.51}$$

In the orthogonal pilot case, the SNR matrix $\mathbf{S}$ is diagonal and the optimization problem (A.51) factorizes into L separate problems, one for each channel delay bin, yielding the same solution as the GThres estimator (A.17). Conversely, in the non-orthogonal case, the optimal solution requires a combinatorial search over the 2^L realizations of $\mathbf{a}_s$. This is circumvented by the use of sparse approximation algorithms [91, 116].

An equivalent problem has been addressed in [91], namely

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}\in\mathbb{C}^L} \left\|\mathbf{w} - \Phi\mathbf{z}\right\|_2^2 + \lambda\left\|\mathbf{z}\right\|_0, \tag{A.52}$$

where $\mathbf{w}$ is a noisy version of $\Phi\mathbf{z}$, and Φ is known, with $\mathbf{I}_L - \Phi^*\Phi \succ 0$. Eq. (A.51) is equivalent to (A.52) by letting $\mathbf{w} = \sqrt{\rho}\left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-\frac{1}{2}}\mathbf{h}_{\mathrm{LS}}$, $\Phi = \sqrt{\rho}\left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-\frac{1}{2}}$, $\lambda = \rho\alpha$, and $\mathbf{z} = \mathbf{h}_s$, where $\rho > 0$ is chosen so as to guarantee $\mathbf{I}_L - \Phi^*\Phi \succ 0$. The *Iterative Thresholding Algorithm* proposed in [91] may then be used to estimate $\mathbf{h}_s$, and equation (A.50) to estimate the diffuse component $\mathbf{h}_d$.

Alternatively, in [92, 93] the $\mathcal{L}_0$ cost associated with $\mathbf{h}_s$ is relaxed and the $\mathcal{L}_1$ regularization norm is used instead, thus yielding the convex problem

$$\hat{\mathbf{h}}_s = \arg\min_{\mathbf{h}_s} \left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right)^* \left(\mathbf{\Lambda}_d + \mathbf{S}^{-1}\right)^{-1}\left(\mathbf{h}_{\mathrm{LS}} - \mathbf{h}_s\right) + \alpha\left\|\mathbf{h}_s\right\|_1,$$

where we define the $\mathcal{L}_1$-norm $\left\|\mathbf{h}_s\right\|_1 = \sum_k \left|\mathbf{h}_s(k)\right|$.

As justified by the MSE analysis (Sec. A.6), a conservative $\tilde{q} < q$ may be assumed in the estima-

tion of the sparse component, by using $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right) > \ln\left(\frac{1-q}{q}\right)$.

The next chapter is devoted to the evaluation and validation of the proposed HSD channel model and channel estimation schemes.

A.9 Simulation results

A.9.1 Hybrid Sparse/Diffuse channel model

In this section, we evaluate the performance of the GMMSE and GThres estimators in a system whose channel perfectly follows the HSD model, and compare it with the asymptotic MSE behavior derived in Sec. A.6. In particular, the HSD model allows us to control the parameters (*e.g.*, sparsity level $\tilde{q}$, PDP profiles $\mathcal{P}_d$, $\mathcal{P}_s$) and to evaluate the performance of the proposed estimators in an ideal setting, *i.e.*, where the channel realizations follow exactly the HSD model, based on which the estimators have been designed. Moreover, we evaluate the performance of the estimators under a non-orthogonal pilot sequence, as discussed in Sec. A.8.

For the simulation results, we generate a channel $\mathbf{h} \in \mathbb{C}^L$ with delay spread $L = 100$. The sparsity pattern $\mathbf{a}_s \sim \mathcal{B}(q)^L$, with parameter $q = 0.1$. The vector $\mathbf{c}_s \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_s)$, where the covariance matrix $\mathbf{\Lambda}_s$ is diagonal, with exponential PDP $\mathbf{\Lambda}_s(k,k) = \mathcal{P}_s(k) = P_s e^{-\omega k}$, and $\omega = 0.05$. The diffuse component $\mathbf{h}_d \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_d)$, where the covariance matrix $\mathbf{\Lambda}_d$ is diagonal, with exponential PDP $\mathbf{\Lambda}_d(k,k) = \mathcal{P}_d(k) = \beta P_s e^{-\omega k}$. The parameter $P_s > 0$ is a normalization factor, and is chosen so that the average channel energy is L, *i.e.*, $\sum_{k=0}^{L-1} \mathbb{E}\left[|\mathbf{h}(k)|^2\right] = P_s \sum_{k=0}^{L-1} (\beta + q) e^{-\omega k} = L$. Unless otherwise stated, we use $\beta = 0.01$, hence the ratio between the energy of the sparse and diffuse components is given by $\left[\mathbb{E}[\mathbf{h}_s^* \mathbf{h}_s] / \mathbb{E}[\mathbf{h}_d^* \mathbf{h}_d]\right]_{\mathrm{dB}} = 10\,\mathrm{dB}$, where $\mathbf{h}_s = \mathbf{a}_s \odot \mathbf{c}_s$ denotes the sparse component. Unless otherwise stated, we assume an orthogonal pilot sequence, so that $\mathbf{S}$ is diagonal. For simplicity, we assume that $\mathbf{S} = S \cdot \mathbf{I}_L$, for some $S > 0$, so that we can rewrite the observation model (A.7) as

$$\mathbf{h}_{\mathrm{LS}} = \mathbf{h} + \sqrt{S}^{-1} \mathbf{n}. \tag{A.53}$$

Moreover, we define the *estimation SNR* as the average estimation SNR per channel entry, $S\mathbb{E}[\mathbf{h}^* \mathbf{h}]/L$.

We consider the following estimators:

- GMMSE and GThres estimators, for different values of the assumed sparsity level $\tilde{q} \in \{0.1, 0.01, 0.001\}$ (*i.e.*, $\alpha = \ln\left(\frac{1-\tilde{q}}{\tilde{q}}\right) \in \{2.2, 4.6, 6.9\}$);
- unstructured LS estimator;
- MMSE estimator, which assumes perfect knowledge of q, $\mathbf{\Lambda}_d$ and $\mathbf{\Lambda}_s$, and thus performs an MMSE estimate of the channel. It provides a lower bound to the estimation accuracy;
- purely sparse estimator, which ignores the diffuse component. Since a per-tap approach is optimal under an orthogonal pilot sequence, we choose a variation of the GThres estimator which assumes no diffuse component ($\mathbf{h}_d = \mathbf{0}$);
- purely diffuse estimator, which ignores the sparse component (*i.e.*, GMMSE or GThres estimators with $\tilde{q} = 0$).

In Sec. A.9.2 we compare the MSE (define in (A.18)) attained by these estimators with the asymptotic MSE behavior derived in Sec. A.6, assuming perfect knowledge of $\mathbf{\Lambda}_d$. In Sec. A.9.3 we evaluate the impact on the performance when the PDP profile $\mathbf{\Lambda}_d$ is unknown and is estimated using the PDP estimator developed in Sec. A.7. In Sec. A.9.4 we evaluate the performance under a non-orthogonal pilot sequence. Finally, in Sec. A.9.5, we evaluate the BER performance induced by channel estimation errors, when the aforementioned estimators are employed for coherent detection.

A.9.2 Validation of the MSE analysis

In Fig. A.2, we plot the MSE of the estimators as a function of the estimation SNR, and their asymptotic MSE behavior (bold lines, with the corresponding markers for the different values of α), assuming perfect knowledge of $\mathbf{\Lambda}_d$. We note that there is a perfect match between the MSE in the high and low SNR regimes, and the asymptotic analysis developed in Sec. A.6. In particular, from an MSE perspective, it is confirmed that it is beneficial to use a conservative approach in the estimation process, *i.e.*, by assuming the sparse component to be sparser than it actually is. In fact, the optimal threshold for the GThres estimator represents a balance between the probability of mis-detecting an active sparse component as diffuse contribution and the probability of false alarm (detecting a diffuse contribution as active sparse component). A conservative approach, by employing a small threshold, reduces the false alarm probability (a similar consideration holds for the GMMSE estimator). This trend can also be observed in the medium SNR ranges. However, this property does not hold in

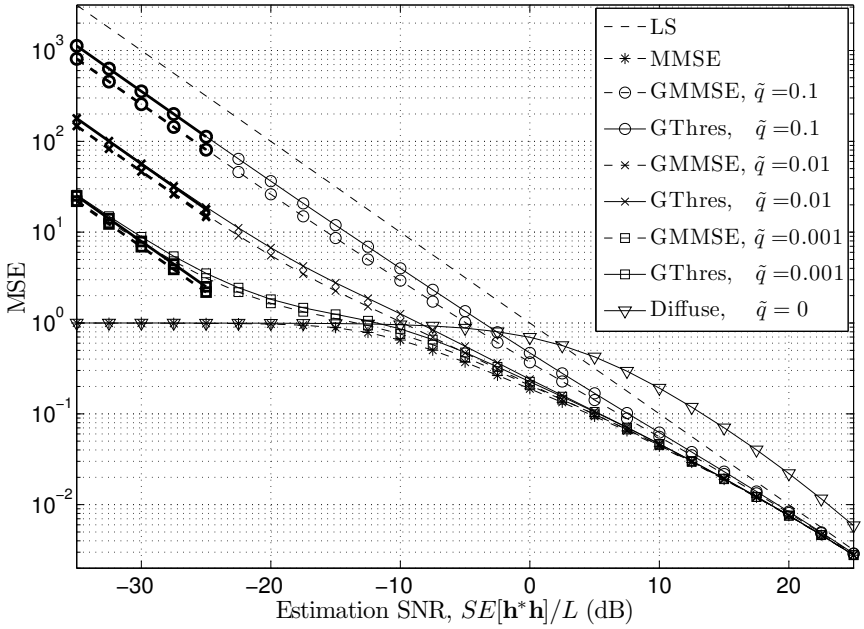


Figure A.2. *MSE of the* GMMSE *and* GThres *estimators, for the HSD channel model, with perfect knowledge of the PDP $\mathcal{P}_d(k)$. The bold lines with the corresponding markers represent the low SNR MSE behavior. The high SNR behavior is given by the* LS *estimate.* $\beta = 0.01$, $q = 0.1$.

general, as we have discussed in Sec. A.6. To see that, we also plot the accuracy of the *diffuse* estimator $\hat{\mathbf{h}}^{(\mathrm{Diff})}(k) = \frac{S\mathcal{P}_d(k)}{1+S\mathcal{P}_d(k)}\mathbf{h}_{\mathrm{LS}}(k)$, which ignores the sparse component $\mathbf{a}_s \odot \mathbf{c}_s$. This can be interpreted as a limit case of the GMMSE and GThres estimators, for $\tilde{q} \to 0$, or equivalently $\alpha \to +\infty$. Also, as predicted by the MSE analysis, for a given value of $\tilde{q}$ the GMMSE estimator outperforms the GThres estimator, in the asymptotic regimes. This is a consequence of the fact that GThres allows only the extreme values $\hat{\mathbf{a}}_s^{(\mathrm{GThres})}(k) \in \{0, 1\}$, whereas GMMSE allows a smoother transition between these two extremes.

In Fig. A.3, we plot the MSE of the estimators as a function of the SNR S, for the case with no diffuse component, $\beta = 0$. Even in this case, we notice a perfect match between the MSE in the high and low SNR regimes, and the asymptotic analysis in Sec. A.6. In particular, the larger the factor α used (the smaller $\tilde{q}$), the better the estimation accuracy. Unlike Fig. A.2, where the MSE approaches the LS estimate for high SNR, in this case we note a performance improvement. In fact, when $\beta = 0$, the estimate of $\mathbf{h}_d$ is forced to zero. Therefore, whenever the GThres estimator correctly detects $\hat{\mathbf{a}}_s(k) = \mathbf{a}_s(k) = 0$, the channel component $\mathbf{h}(k)$ is estimated with no error. On the other hand, when $\beta > 0$, a residual MMSE estimation error is incurred.

In Fig. A.4, we let vary the ratio between the energies of the sparse and diffuse components, $\mathbb{E}[\mathbf{h}_s^*\mathbf{h}_s]/\mathbb{E}[\mathbf{h}_d^*\mathbf{h}_d] = q/\beta$. The estimation SNR is $[S\mathbb{E}[\mathbf{h}^*\mathbf{h}]/L]_{\mathrm{dB}} = 10\,\mathrm{dB}$. The MSE of the

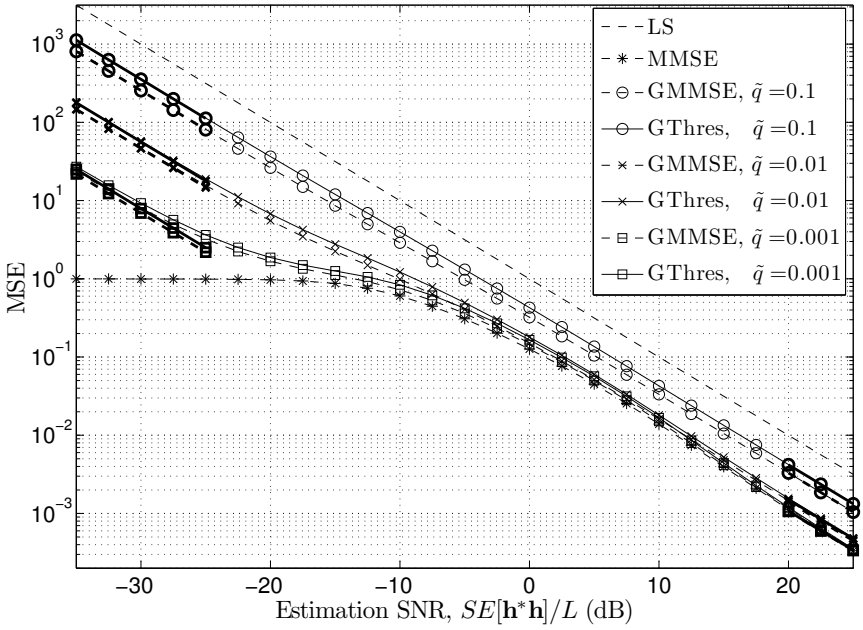


Figure A.3. *MSE of the* GMMSE *and* GThres *estimators, for the HSD channel model. The bold lines with the corresponding markers represent the high/low SNR MSE behavior.* $\beta = 0$ *(no diffuse component),* $q = 0.1$.

purely sparse estimator is also plotted in this case. Similarly to Figs. A.2 and A.3, we note that a conservative approach is beneficial from an MSE perspective. As expected, the sparse estimator performs worse than the GThres estimator, due to its inability to exploit the diffuse component of the channel. In particular, it performs closely to the GThres estimator for small values of β (*i.e.*, large values of $\mathbb{E}[\mathbf{h}_s^*\mathbf{h}_s]/\mathbb{E}[\mathbf{h}_d^*\mathbf{h}_d]$), where the diffuse component is negligible with respect to the sparse one, and incurs a performance degradation for large values of β, where the diffuse component becomes significant. Moreover, as expected, the only diffuse estimator achieves good performance for large values of β. However, it performs poorly for small values of β, where the sparse component yields a significant contribution. Note that, excluding the MMSE estimator, the GThres estimator with $\tilde{q} = 0.001$ achieves the best performance over the entire range of values considered, very close to the MMSE lower bound. This proves that the proposed methods are robust, and adapt to a wide range of estimation scenarios, where the channel exhibits either a sparse, diffuse or hybrid nature (corresponding to large, small and moderate values of $\mathbb{E}[\mathbf{h}_s^*\mathbf{h}_s]/\mathbb{E}[\mathbf{h}_d^*\mathbf{h}_d]$, respectively).

A.9.3 Evaluation of the PDP estimator

Fig. A.5 compares the MSE of the GMMSE estimator, for the two cases where $\mathbf{\Lambda}_d$ is perfectly known at the receiver, and where it is estimated from the observed sequence using the EM algorithm

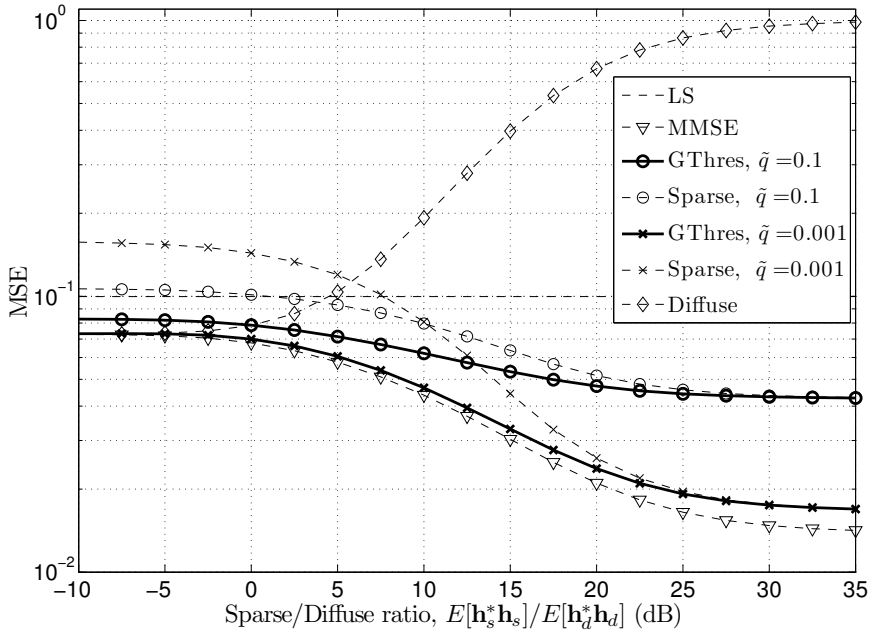


Figure A.4. *MSE of the channel estimators as a function of β, assuming perfect knowledge of the PDP of the diffuse component $\mathcal{P}_d(k)$.* $[S\mathbb{E}[\mathbf{h}^*\mathbf{h}]/L]_{\,\mathrm{dB}} = 10\,\mathrm{dB}$, $q = 0.1$

(Sec. A.7), based on only one realization of the channel. We notice that, in general, there is a small performance loss due to the unknown $\mathbf{\Lambda}_d$, mainly in the low SNR range and for small values of $\tilde{q}$ (however, no performance degradation is observed for $\tilde{q} = 0.1$). This behavior is explained by the fact that the MMSE estimate of $\mathbf{h}_d$ in (A.14) is more sensitive to errors in the estimation of $\mathbf{\Lambda}_d$ in the low SNR than in the high SNR regime. In fact, for high SNR values, it approaches the LS solution. On the other hand, for small values of $\tilde{q}$ we have the following. The posterior probability of the entries of the sparsity pattern $\mathbf{a}_s$, as a function of the factor $\alpha = \left(\frac{1-\tilde{q}}{\tilde{q}}\right)$, is given by (A.16) with $\mathbf{S}_{k,k} = S$. This is a decreasing function of α (*i.e.*, increasing function of $\tilde{q}$). As a consequence, the smaller $\tilde{q}$ the more the weight given to the right-hand term of (A.14), associated with the MMSE estimate of $\mathbf{h}_d(k)$, which is sensitive to errors in the estimate of $\mathcal{P}_d(k)$, compared to the left-hand term, associated with the LS estimate of $\mathbf{c}_s(k)$, which is independent of the PDP estimate. As a consequence, a smaller value of $\tilde{q}$ results in an overall estimate that is more sensitive to errors in the PDP estimate of $\mathbf{h}_d$. Similar considerations hold for the GThres estimator.

Fig. A.6 plots the MSE of PDP estimator of the diffuse component developed in Sec. A.7, for different values of $\tilde{q}$ and of the number of iterations of the EM algorithm, based on only one channel realization, as a function of the SNR per diffuse channel entry $S\mathbb{E}[\mathbf{h}_d^*\mathbf{h}_d]/L$. In particular, letting

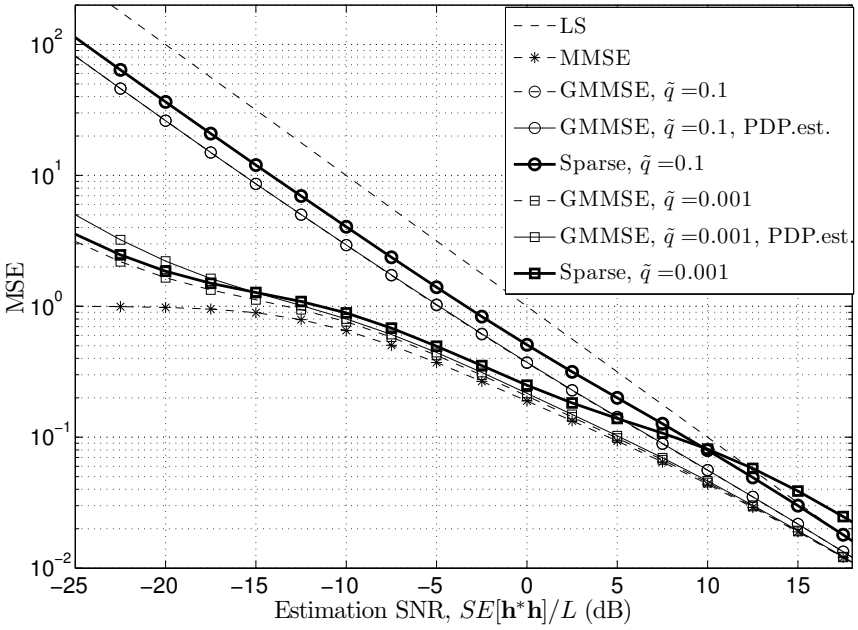


Figure A.5. *MSE of the* GMMSE *estimators, comparison between the cases where the PDP of the diffuse component is known and estimated from the data, respectively.* $\beta = 0.01$*,* $q = 0.1$*. The two curves of the* GMMSE *estimator with* $\tilde{q} = 0.1$ *where the PDP is known and estimated overlap.*

$\hat{\mathcal{P}}_d(k), k = 0, \ldots, L-1$ be an estimate of $\mathcal{P}_d(k) = \beta e^{-\omega k}$, we compute the following MSE metric:

$$\mathrm{MSE}_{\mathrm{PDP}} = \frac{1}{L}\sum_{k=0}^{L-1} \mathbb{E}\left[\left(\ln \hat{\mathcal{P}}_d(k) - \ln \mathcal{P}_d(k)\right)^2\right]. \tag{A.54}$$

The performance is compared also with an oracle estimator, which assumes perfect knowledge of $\mathbf{a}_s \odot \mathbf{c}_s$, thus being able to perfectly remove the interference from the sparse component (in particular, we use the EM estimator with $\tilde{q} = 0$). In the figure, the MSE floor refers to the ML estimator of β, ω in the noiseless scenario with no sparse component. It can be shown that, in this case, the ML estimator is obtained by setting $A_k = |\mathbf{h}_d(k)|^2$ and $Z = \frac{L-1}{2}$ in the E-step (A.36), and by solving (A.38) using the results of Theorem A.7.1. As expected, the Oracle estimator achieves the best performance, and approaches the MSE floor in the high SNR. Remarkably, the EM estimator with $\tilde{q} = 0.001$ and 300 iterations approaches the performance of the Oracle estimator, although it cannot take advantage of prior knowledge of $\mathbf{a}_s \odot \mathbf{c}_s$. This proves that the proposed method effectively removes the interference from the sparse component, by discarding the observations associated, with high probability, to the active sparse components. Interestingly, the case $\tilde{q} = 0.001$ with 20 iterations incurs a small performance degradation compared to the MSE achievable after 300 iterations, which becomes negligible for moderate and large SNR values. On the other hand, when $\tilde{q} = 0$ is used, the

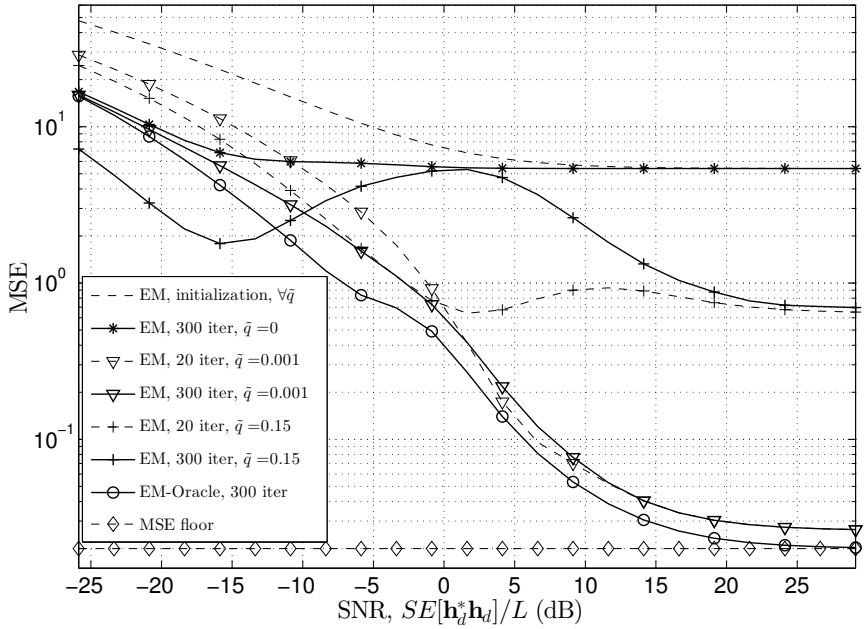


Figure A.6. *MSE of the PDP estimator of* $\mathbf{h}_d$. $\beta = 0.01$, $q = 0.1$.

presence of the sparse component is neglected and the channel is treated as being purely diffuse. In this case, a significant performance degradation is incurred. Finally, we notice that the case $\tilde{q} = 0.15$ incurs a performance degradation, compared to the case $\tilde{q} = 0.001$, which confirms our analysis in Sec. A.7. In fact, we have verified that the estimate of the PDP parameter ω diverges to $+\infty$ as the EM algorithm is iterated, so that the PDP estimate is forced to zero and the overall MSE diverges to $+\infty$.

A.9.4 Non-orthogonal pilot sequence

In Fig. A.7, we compare the MSE of the GThres estimator for the non-orthogonal and orthogonal pilot sequence cases, under the same pilot energy budget, as discussed in Sec. A.8.1. Moreover, we plot the curves associated with the modified *Iterative Thresholding Algorithm* (ITH), designed in Sec. A.8.2 based on a variation of [91] which takes into account the presence of the diffuse component. The non-orthogonal pilot sequence is generated from a CAZAC sequence of length $M = 50 = L/2$ [117]. As expected, we observe a performance loss in the non-orthogonal case, compared to the orthogonal pilot scenario with the same pilot energy budget. In fact, the GThres estimator, by employing a per-tap estimation approach, neglects any correlation among the channel taps, thus incurring a performance degradation. We measured that the *orthogonality coefficient* (A.46) ranges in the interval $\eta_k \in [0.625, 0.765]$ (note that this is a function of the delay $k \in \{0, \ldots, L-1\}$),

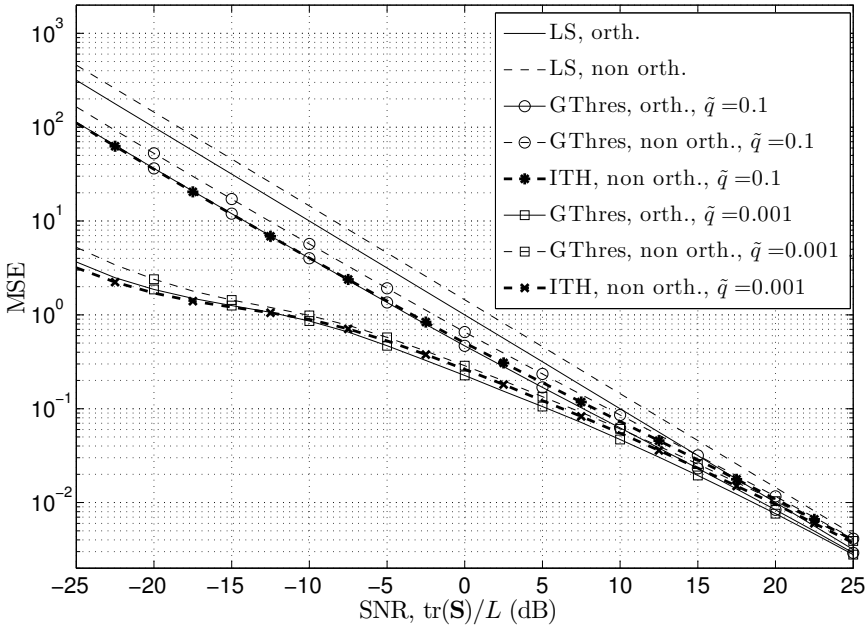


Figure A.7. *Comparison between the non-orthogonal and orthogonal pilot sequence cases.* $\beta = 0.01$, $q = 0.1$.

corresponding to an SNR loss in the range $[1.16, 2.05]$ dB. These values are confirmed by simulation, where the SNR loss induced by GThres under a non-orthogonal pilot sequence (by averaging over all channel delay taps, as in (A.18)) is approximately $[1.5, 2]$ dB. Interestingly, the performance degradation incurred by the GThres estimator is partially recovered (fully, in the low SNR regime) by the ITH algorithm, which exploits the correlation introduced by the non-orthogonal pilot sequence by estimating the channel taps jointly.

A.9.5 BER performance

Finally, in Fig. A.8 we plot the BER induced by channel estimation errors, for the case where the PDP of $\mathbf{h}_d$ is known. To this end, we define an OFDM-UWB system, employing $N_{\rm dft} = 512$ sub-carriers and a 4-QAM constellation with Gray mapping, and the bit sequence is uncoded. In the estimation phase, we use an orthogonal pilot sequence. This may be achieved, for example, by allocating an OFDM symbol with a constant modulus pilot sequence. Our observation for channel estimation has noise; in contrast, we assume no noise when evaluating the BER. As a result, the BER curves reflect the errors induced by channel estimation versus additive channel noise. In particular, let $X(n)$ be the 4-QAM symbol transmitted on the nth sub-carrier, and $H(n) = \sum_{l=0}^{L-1} \mathbf{h}(n) e^{-i2\pi \frac{ln}{N_{dft}}}$ be the DFT of the channel. Then, the received symbol is $Y(n) = H(n)X(n)$. This is equalized by

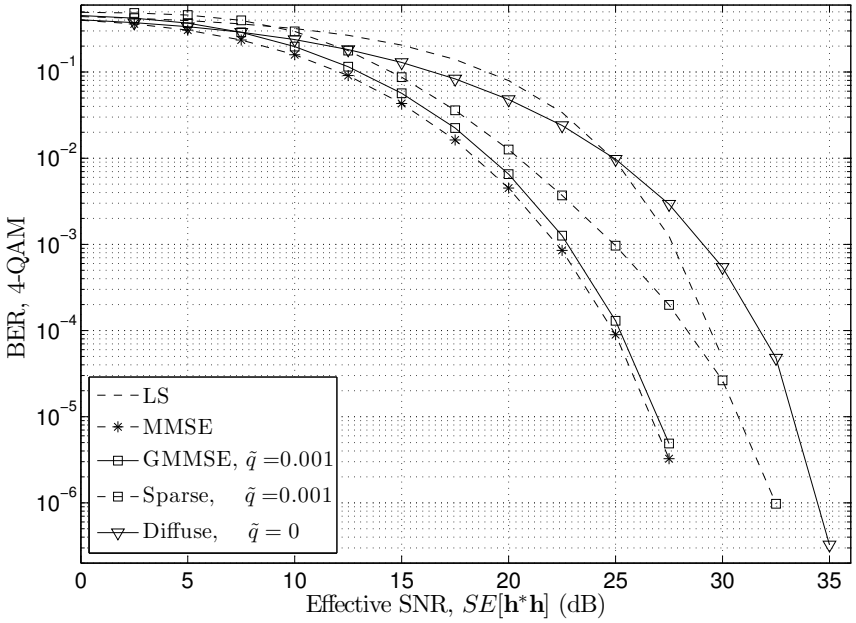


Figure A.8. *BER induced by channel estimation errors, with known PDP of* $\mathbf{h}_d$. $\beta = 0.01$, $q = 0.1$.

using the estimate $\hat{H}(n)$ of $H(n)$, *i.e.*, $\tilde{X}(n) = \frac{H(n)}{\hat{H}(n)} X(n)$, and the decision is based on a minimum distance criterion, *i.e.*, $\hat{X}(n) = \min_{x \in 4-\mathrm{QAM}} |\tilde{X}(n) - x|^2$. Moreover, the BER is averaged over the "good" sub-carriers only, which are chosen based on the heuristic carrier selection scheme

$$\left\{ k : |H(k)|^2 \geq \lambda \max_n |H(n)|^2 \right\}, \tag{A.55}$$

where $\lambda \in (0, 1)$ is a threshold value. In particular, λ is chosen so that 30% of the sub-carriers are classified as "good". The rationale behind this choice is that, in a practical system, the "bad" sub-carriers would never be used, since they are not suitable to carry information. The SNR is referred to the output of an ideal Rake receiver with perfect channel knowledge, where the estimation noise is treated as additive Gaussian noise at the receiver. This is defined as $\mathrm{SNR}_{\mathrm{rake}} = S\mathbf{h}^*\mathbf{h}$.

We notice that GMMSE estimator with $\tilde{q} = 0.001$ performs very closely to the lower bound, represented by the BER induced by the MMSE estimator, defined in Sec. A.5.1. On the other hand, both the diffuse and the purely sparse estimators perform poorly, due to their inability to exploit both the sparse and the diffuse components jointly.

A.9.6 Realistic UWB channel model

In this section, we evaluate the BER and MSE performance of the proposed estimators in a more realistic UWB channel emulator developed in [102], which we refer to as *K&P model* in the following. This approach is important as a validation of the HSD model, of the GMMSE and GThres estimators and of the analysis we have developed. We argue that the K&P model is more suitable than the model in [101] to evaluate the robustness and sensitivity of the proposed HSD channel estimation strategies to deviations from the HSD model. In fact, as explained in more detail in Sec. A.9.7, K&P models the diffuse component as a diffuse tail associated with each specular component, whereas in the HSD model the diffuse and sparse components are assumed to be independent. Therefore, it represents a deviation from the HSD model. In contrast, the model developed in [101] exhibits a better fit to the HSD model, since the diffuse component is generated independently of the specular MPCs arrivals.

A.9.7 K&P model

The K&P model combines both a geometric approach for the resolvable individual specular components (echoes), arising from reflections from the scatterers in the environment, and a statistical approach for the dense multipath clusters associated with each echo. The model also includes a frequency dependent gain decay, so that the overall channel transfer function is expressed as

$$H(f) = \sum_{l} A_l(\tau_l)\,(1 + \mathcal{D}_l(f))\,e^{-i2\pi f \tau_l}\left(1 + \frac{f}{f_0}\right)^{-\nu} \mathcal{I}\left(|f| \leq \frac{B}{2}\right), \tag{A.56}$$

The sum is over the individual echoes, with the lth echo having amplitude $A_l(\tau_l)$ and delay τ_l. $\mathcal{D}_l(f)$ is the multipath cluster associated with the lth echo, with exponential PDP and circularly symmetric Gaussian distribution in the time-domain, ν is the frequency domain decay exponent, f_0 is the center frequency, and $B < R$ is the transmission bandwidth.

The time-domain baseband representation of the channel is obtained by performing an inverse Fourier transform of (A.56), and by sampling at rate R samples per ns. We further clip the channel in the delay domain, so that only the channel window carrying most of the energy is kept. This step determines the delay spread of the channel ($L = 600$). The channel snapshot is finally normalized to have energy L, *i.e.*, $\sum_{l=0}^{L-1} |\mathbf{h}(l)|^2 = L$.

It is worth noting that τ_l is quantized to discrete values, and equals an integer number of the sampling interval $R^{-1}\,\mathrm{ns}$. This is a simplification, which guarantees that the MPC arrival matches

Table A.3. *Main parameters for the* Office LOS *scenario in [102]*

N_{dft}	2048	Number of channel samples in the delay domain
R	$12.8\,\text{ns}^{-1}$	Sampling rate in the delay domain
B	10 GHz	Bandwidth of the UWB system
f_0	6 GHz	Center frequency
d_0	0.8 m	Reference distance for individual echo power law
δ	3	Path loss exponent for individual echo power law
G_{MP}	-20 dB	Cluster gain with respect to associated individual echo
G_{MP-LOS}	-13 dB	Additional cluster gain for LOS individual echo
γ	10 ns	Multipath cluster exponential decay parameter
ν	1.1	Frequency domain decay exponent
(x_t, y_t, z_t)	$(1.78, 4, 1.5)$ m	Coordinates of transmitter position
(x_r, y_r, z_r)	from $(3.0, 1.5, 1.5)$ m to $(4.5, 1.8, 1.5)$ m	Coordinates of mobile receiver position

exactly the sampling period. Therefore, in general, the K&P model [102] does not cope with the side-lobes of the sinc pulse, which arise when the MPCs arrive in the middle of two sampling times. However, the bandwidth limitation $B < R$ introduces a sinc filtering of the channel, so that the side-lobes of the sinc pulse affect the channel impulse response.

We choose the *Office LOS* scenario in [102] for our simulations. A summary of the main parameters are given in Table A.3. In particular, 10000 channel snapshot are generated, each corresponding to a different position of the mobile receiver along the line connecting the points $(3.0, 1.5, 1.5)$ m and $(4.5, 1.8, 1.5)$ m (where (x, y, z) m represents a point in the three dimensional space with coordinates x, y and z, measured in m relative to the origin). For each position of the mobile receiver, the arrival pattern of the resolvable MPCs, *i.e.*, their gain and delay, is determined by the relative positions of the mobile receiver, transmitter and scatterers (these are positioned on a grid in the three dimensional space). Moreover, for each channel snapshot, we generate an independent realization of the diffuse component (Rayleigh fading) and of the additive noise. We refer the interested reader to [102] for further details.

It is worth noting that the sparsity level q of the HSD model is not defined for the K&P model. This parameter may be roughly estimated as the ratio between the number of active scatterers and the delay spread L. For the *Office LOS* scenario defined in [102], we have 6 (virtual) scatterers and $L = 600$, which gives $\hat{q} \simeq 0.01$. Moreover, the PDP estimator developed in Sec. A.7 for the HSD model assumes an exponential PDP for the diffuse component, which is not defined for the K&P model. In Fig. A.9, we plot the PDP of a channel snapshot as an example, and the exponential PDP fitting, estimated using the EM algorithm developed in Sec. A.7. We note a good fitting of the exponential PDP model to the PDP of the channel realization. Remarkably, although the K&P model

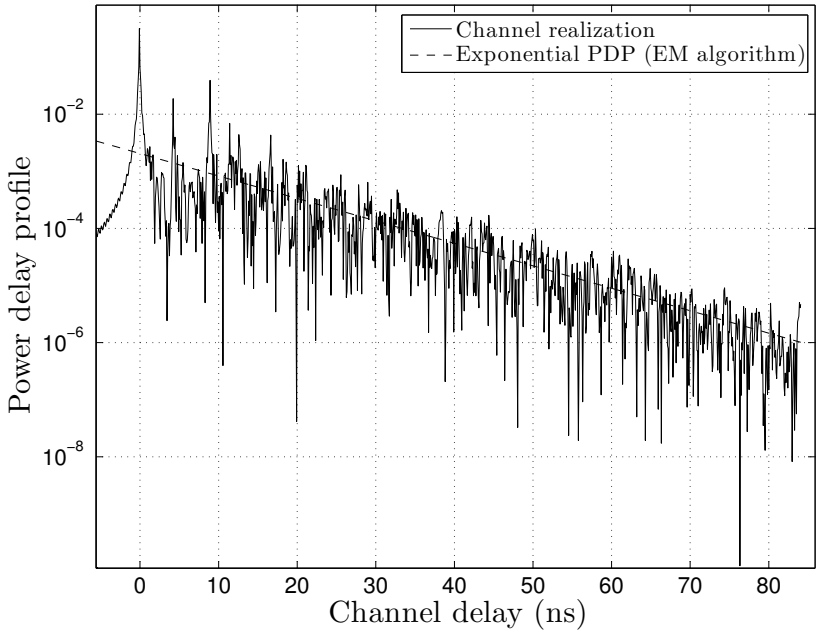


Figure A.9. *PDP of one realization of the K&P model (with parameters given in Table A.3) and exponential PDP, estimated using the EM algorithm developed in Sec. A.7.*

defines the diffuse component as a diffuse tail associated with each specular component, the overall effect, by summing the contribution from all MPCs, is that of a unique PDP tail, which fits well the exponential shape.

The channel and the PDP of the diffuse component are estimated based on a single snapshot of the channel. In particular, the PDP of the diffuse component is estimated using the EM algorithm developed in Sec. A.7. Hence, the MSE and BER results are not affected by the structure of the spatio-temporal correlation of the channel.

A.9.8 MSE performance

Figs. A.10 and A.11 plot the MSE of the GMMSE, GThres and purely sparse and diffuse estimators, for different values of the assumed sparsity level $\tilde{q}$. Since a per-tap approach is optimal in this case, for the sparse estimator we choose a variation of the GThres estimator, which assumes no diffuse component ($\mathbf{h}_d = \mathbf{0}$). The diffuse estimator assumes a purely diffuse channel, and performs a linear MMSE estimate based on the estimated PDP of the diffuse component.

In Fig. A.10, we observe that, the smaller $\tilde{q}$ (*i.e.*, the larger α), the better the estimation accuracy of the GMMSE and GThres estimators. Moreover, the GMMSE estimator outperforms the GThres estimator, for a given value of $\tilde{q}$. This is the same behavior, predicted by the MSE analysis in Sec. A.6,

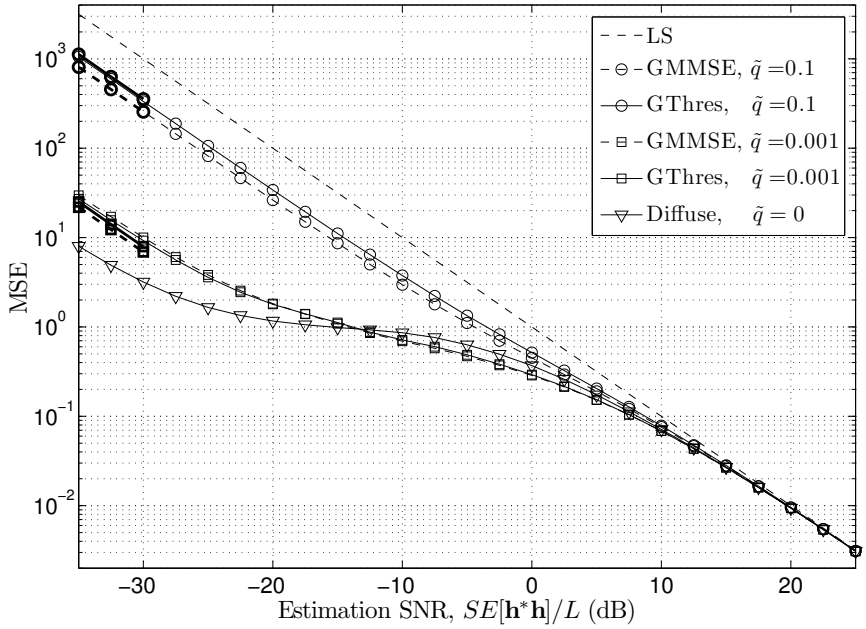


Figure A.10. *MSE of the* GMMSE *and* GThres *estimators, for the K&P channel model, LOS-Office scenario. The bold lines with the corresponding markers represent the low SNR MSE behavior. The high SNR behavior is given by the* LS *estimate.*

that we have observed in the case where the channel follows the HSD model (Fig. A.2). Remarkably, we notice a perfect match between the simulation results and the low/high SNR asymptotic behavior of the estimators (bold lines). This is a surprising result, if we consider that the K&P channel emulator deviates from the HSD model, and the PDP of the diffuse component is unknown and estimated from the data. However, note that the value of the channel delay spread, $L = 600$, allows sufficient averaging over the small scale fading in the delay dimension, so that the PDP is accurately estimated.

Moreover, we notice that the diffuse estimator outperforms the HSD estimators in the low SNR ($< -12.5\,\mathrm{dB}$). This is an expected result, which is coherent with the simulation results based on the HSD model (Fig. A.2) and with the asymptotic analysis in Sec. A.6, where we have proved that, in the low SNR, the smaller $\tilde{q}$, the better the estimation accuracy (note that the diffuse estimator corresponds to the limit case $\tilde{q} \to 0$). In fact, the diffuse estimator forces the channel estimate to zero in the low SNR, thus approaching the channel energy floor. Conversely, a performance degradation is observed for higher SNR values, with respect to the HSD estimators with $\tilde{q} = 0.001$, which achieve the best performance.

In Fig. A.11, we notice that the GMMSE estimator achieves better performance than the sparse estimator, for the same values of $\tilde{q}$. In fact, the sparse estimator does not effectively capture the diffuse component of the channel, thus incurring a performance degradation, mainly in medium and

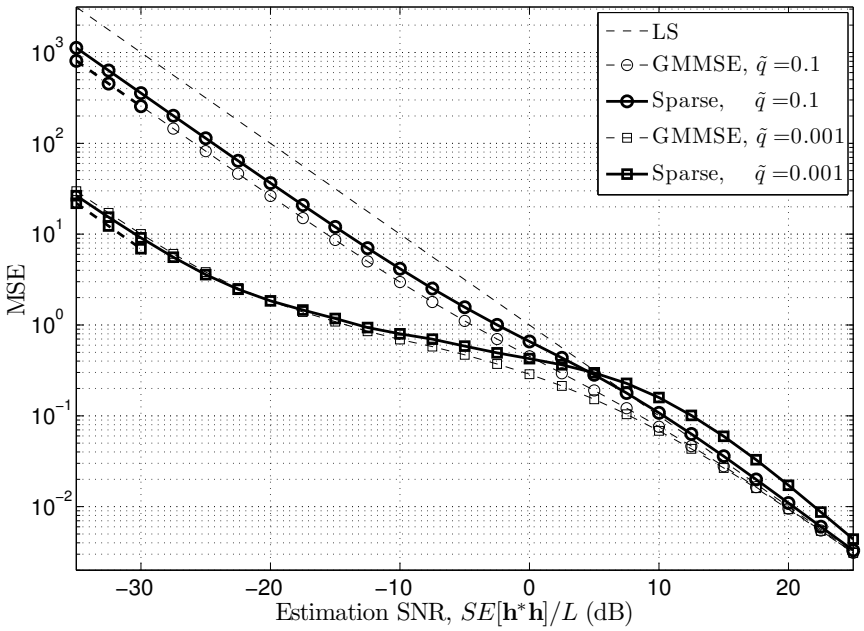


Figure A.11. *MSE of the* GMMSE *and Sparse estimators, for the K&P channel model, LOS-Office scenario.*

high SNR (in particular, in the high SNR range ($> 5\,\mathrm{dB}$), it performs even worse than LS).

A.9.9 BER performance

Similarly to Sec. A.9.5, we evaluate the BER performance induced by channel estimation errors in an OFDM-UWB system, with $N_{\mathrm{dft}} = 2048$ sub-carriers, 4-QAM constellation and transmission bandwidth $B = 10\,\mathrm{GHz}$. Since we want to evaluate the impact of channel estimation errors on the BER performance, we consider a noise-free setting, *i.e.*, no noise is added to the information symbols, whereas noise is added in the estimation phase, so as to induce channel estimation errors. The SNR is referred to the output of an ideal Rake receiver. Moreover, the BER is averaged over only the "good" sub-carriers, chosen according to the heuristic carrier selection scheme (A.55).

Fig. A.12 plots the BER associated with the GMMSE, LS, purely sparse and diffuse estimators, for different values of $\tilde{q}$. Generally, we observe that the better the MSE estimation accuracy, the smaller the BER. In particular, the best performance is achieved by the GMMSE estimator with $\tilde{q} = 0.001$. Moreover, similarly to the MSE, also the BER benefits from a conservative approach in the estimation of the sparse component, *i.e.*, it is beneficial to use small values of $\tilde{q}$. We notice that a poor BER performance is incurred by the purely sparse estimator which, in the high SNR, performs even worse than LS. Similarly, the diffuse estimator performs worse than GMMSE with $\tilde{q} = 0.001$

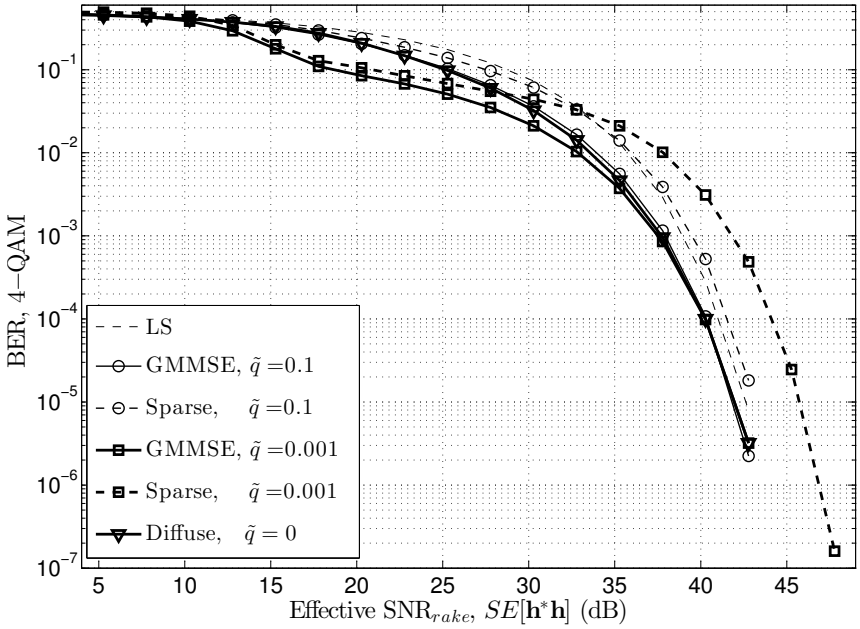


Figure A.12. *BER induced by channel estimation errors, for the K&P channel model, LOS-Office scenario. The BER is averaged over the "good" sub-carriers only (A.55).*

in the medium SNR range. As in the MSE case, the purely sparse and diffuse estimators are unable to exploit both the diffuse and sparse components of the channel jointly, thus incurring a performance degradation. Finally, we observe an irregular behavior of the GMMSE and sparse estimators with $\tilde{q} = 0.001$ around $18\,\mathrm{dB}$ SNR. We argue that this is a consequence of the fact that we do not average over independent realizations of the surrounding environment, *i.e.*, we use the particular *Office LOS* in [102], which specifies the relative positions of the scatterers, and of the transmitter/receiver pair as well.

These results show that the GMMSE and GThres estimators effectively capture the main UWB propagation phenomena, *e.g.*, the resolvable MPCs of the channel, modeled by a sparse component, unresolvable MPCs, scattering from rough surfaces and frequency dispersion, which are better modeled by a diffuse component. Also, we observe that a small performance degradation is incurred by the diffuse estimator. However, we argue that one of the strengths of the proposed HSD model and channel estimation strategies relies in their robustness and adaptability to different scenarios of interest, where the channel exhibits a sparse, diffuse or hybrid nature. Conversely, a diffuse (respectively, sparse) estimator is expected to perform poorly in sparse (diffuse) channels.

A.10 Conclusions

In this chapter, we have investigated the issue of channel estimation for UWB systems. In particular, we have proposed a novel hybrid sparse/diffuse model for the UWB channel, which is able to capture the main UWB propagation mechanisms: fine delay resolution capability, scattering from rough surfaces, frequency dispersion. We have then identified four scenarios of interest in practical systems, differing in the amount of side information available at the receiver for the purpose of channel estimation, and we have proposed channel estimators exploiting the channel structure and the side-information to enhance the estimation accuracy.

Of particular interest is the scenario where the PDP of the diffuse component is known at the receiver, and the statistics of the specular component are unknown. This is relevant when the observation interval is large enough to allow averaging over the small scale fading, but not over the large scale fading. For this scenario, we have proposed the *Generalized MMSE* and *Generalized Thresholding Estimators*. We have carried out an MSE analysis of these estimators, in the asymptotic regimes of high and low SNR. This analysis suggests that it is beneficial, from an MSE perspective, to use a conservative approach in the estimation of the sparse component, *i.e.*, to assume the sparse component to be sparser than it actually is. While this result cannot be extended to medium SNR, simulation results show that a similar behavior often holds in this regime. We have proposed an EM algorithm for the PDP estimation of the diffuse component, which exploits the exponential structure of the PDP to average the fading over the channel delay dimension, rather than over subsequent independent realizations of the fading process. Moreover, we have analyzed the case with a non-orthogonal pilot sequence, and shown that the GThres estimator can be recast as a modification of a sparse approximation algorithm proposed in the literature.

Finally, we have evaluated these estimation schemes based on a more realistic geometry-based stochastic UWB channel emulator, developed in [102]. Simulation results for this case show that the GMMSE and GThres estimators achieve better performance, in terms of both MSE and BER, than conventional unstructured (Least Squares) and purely sparse or diffuse estimators, thus suggesting that, although simplified (*e.g.*, compared to [89]), the HSD model is able to capture key UWB propagation mechanisms, such as resolvable MPCs, diffuse scattering from rough surfaces, unresolvable MPCs, and frequency dispersion.

Appendix A.A: Proofs for Section A.6

Lemma A.10.1 (Exchanging the limit with the expectation)**.** *For the* GMMSE *and* GThres *estimators of the kth channel delay bin* $\hat{h}\,(y)$*, where* $y = h + \sqrt{S}^{-1} n$ *is the observation,* $h = a_s c_s + \sqrt{P_d} h_d$ *is the HSD channel bin,* n *is the noise in the kth delay bin, we have, for* $X \in \{\mathrm{GMMSE}, \mathrm{GThres}\}$*,*

$$\lim_{S \to S_{\lim}} \mathrm{mse}_k^{(X)}(S) = \mathbb{E}\left[\lim_{S \to S_{\lim}} f^{(X)}\left(\sqrt{S}y, n\right)\right],$$

where $S_{\lim} \in \{0, +\infty\}$*, and* $\mathrm{mse}_k^{(X)}(S)$ *and* $f^{(X)}\left(\sqrt{S}y, n\right)$ *are defined in (A.21) and (A.20), respectively.*

Proof. Note from (A.14) that, for $X \in \{\mathrm{GMMSE}, \mathrm{GThres}\}$, we can write

$$\hat{h}(y) = r^{(X)}\left(\sqrt{S}|y|\right) y, \tag{A.57}$$

where $r^{(X)}\,(z)$, for $z \geq 0$, is given by

$$r^{(X)}\,(z) = \phi^{(X)}\,(z) + \left(1 - \phi^{(X)}\,(z)\right) \frac{SP_d}{1 + SP_d}. \tag{A.58}$$

The function $\phi^{(X)}\,(z)$ is the estimate of the sparsity bit a_s conditioned on $|y| = \sqrt{S}^{-1} z$, and its expression depends on the chosen estimator $X \in \{\mathrm{GMMSE}, \mathrm{GThres}\}$, specifically, from (A.16) and (A.17),

$$\phi^{(X)}\,(z) = \begin{cases} \frac{1}{1+e^{\alpha} \exp\left\{-\frac{z^2}{1+SP_d}\right\}}, & X = \mathrm{GMMSE}, \\ \mathcal{I}\left(z^2 \geq \alpha\,(1 + SP_d)\right), & X = \mathrm{GThres}. \end{cases} \tag{A.59}$$

Let $\{S_j > 0, j = 0, \ldots, +\infty\}$ be a generic SNR sequence, indexed by j, such that $\lim_{j \to +\infty} S_j = S_{\lim}$. From Lebesgue's Dominated Convergence Theorem [118], if there exists a function $g^{(X)}(h, n)$ such that

$$\begin{cases} \left|f^{(X)}\left(\sqrt{S_j}h + n, n\right)\right| \leq g^{(X)}(h, n) \quad \text{a.e.}, \quad \forall j \\ \mathbb{E}\left[g^{(X)}(h, n)\right] < +\infty, \end{cases} \tag{A.60}$$

where a.e. stands for *almost everywhere*, *i.e.*, the inequality holds except on a set with probability measure zero (with respect to the random variables $h_d \sim \mathcal{CN}(0, 1)$, $n \sim \mathcal{CN}(0, 1)$, $a_s \sim \mathcal{B}(q)$ and

c_s), then

$$\lim_{j\to+\infty} \mathrm{mse}_k^{(X)}(S_j) = \mathbb{E}\left[\lim_{j\to+\infty} f^{(X)}(\sqrt{S_j}h + n, n)\right].$$

If this property holds for any SNR sequence such that $\lim_{j\to+\infty} S_j = S_{\lim}$, then

$$\lim_{S\to S_{\lim}} \mathrm{mse}_k^{(X)}(S) = \mathbb{E}\left[\lim_{S\to S_{\lim}} f^{(X)}(\sqrt{S}y, n)\right],$$

and the Lemma is proved.

We now prove the existence of such a function $g^{(X)}(\cdot)$. Let $x = \sqrt{S}y$. Then, from (A.20) and (A.57), we have

$$f^{(X)}(x,n) = \left|r^{(X)}(|x|)\, x - \sqrt{S}h\right|^2 = \left|\left(1 - r^{(X)}(|x|)\right) x - n\right|^2,$$

where in the last step we used the fact that $\sqrt{S}h = x - n$. Using the inequality $|A+B|^2 \leq 2|A|^2 + 2|B|^2$, we have

$$f^{(X)}(x,n) \leq 2\left(1 - r^{(X)}(|x|)\right)^2 |x|^2 + 2\,|n|^2\,. \tag{A.61}$$

Moreover, from (A.58), we have, $\forall x \in \mathbb{C}$,

$$1 - r^{(X)}(|x|) = \left(1 - \phi^{(X)}(|x|)\right)\frac{1}{1+SP_d} \leq 1 - \phi^{(X)}(|x|)\,.$$

Letting $m^{(X)}(|x|) = \left(1 - \phi^{(X)}(|x|)\right)|x|$, we finally obtain $f^{(X)}(x,n) \leq 2m^{(X)}(|x|)^2 + 2\,|n|^2$.

In order to proceed, we distinguish between the estimators.

A.10.1 Generalized MMSE Estimator

For the GMMSE estimator, using the expression of $\phi^{(\mathrm{GMMSE})}(|x|)$ in (A.59), we have

$$m^{(\mathrm{GMMSE})}(|x|) = \frac{e^{\alpha}\exp\left\{-|x|^2\right\}|x|}{1 + e^{\alpha}\exp\left\{-|x|^2\right\}} \leq e^{\alpha}\exp\left\{-|x|^2\right\}|x|.$$

The right hand side is maximized at $|x| = \frac{1}{\sqrt{2}}$, and therefore we obtain the bound $m^{(\mathrm{GMMSE})}(|x|) \leq e^{\alpha}\frac{1}{\sqrt{2e}}$. Then, we have the following bound on $f^{(\mathrm{GMMSE})}(x,n)$:

$$f^{(\mathrm{GMMSE})}(x,n) \leq e^{2\alpha-1} + 2\,|n|^2 \triangleq g^{(\mathrm{GMMSE})}(h,n). \tag{A.62}$$

$g^{(\mathrm{GMMSE})}(h,n)$ is an integrable function, in fact $\mathbb{E}\left[g^{(\mathrm{GMMSE})}(h,n)\right] = e^{2\alpha-1} + 2 < +\infty$.

A.10.2 Generalized Thresholding Estimator

For the GThres estimator, using the expression of $\phi^{(\mathrm{GThres})}(|x|)$ in (A.59), we have $m^{(\mathrm{GThres})}(|x|) = \mathcal{I}\left(|x|^2 < \alpha\right)|x|$. For $|x| \geq \sqrt{\alpha}$, we have $m^{(\mathrm{GThres})}(|x|) = 0$. On the other hand, for $|x| < \sqrt{\alpha}$, we have $m^{(\mathrm{GThres})}(|x|) = |x| \leq \sqrt{\alpha}$. In general, $m^{(\mathrm{GThres})}(|x|) \leq \sqrt{\alpha}, \quad \forall |x| \geq 0$, and therefore

$$f^{(\mathrm{GThres})}(x,n) \leq 2\alpha + 2\,|n|^2 \triangleq g^{(\mathrm{GThres})}(h,n).$$

$g^{(\mathrm{GThres})}(h,n)$ is an integrable function, in fact we have

$$\mathbb{E}\left[g^{(\mathrm{GThres})}(h,n)\right] = 2\alpha + 2 < +\infty. \tag{A.63}$$

The Lemma is thus proved. □

Lemma A.10.2. *We have, for $n \in \mathcal{CN}(0,1)$,*

$$\mathbb{E}\left[\frac{|n|^2}{\left(1 + e^{\alpha}\exp\{-|n|^2\}\right)^2}\right] = e^{-\alpha}\ln\left(1 + e^{\alpha}\right). \tag{A.64}$$

Proof. We have

$$\mathbb{E}\left[\frac{|n|^2}{\left(1 + e^{\alpha}\exp\{-|n|^2\}\right)^2}\right] = \int_0^{+\infty} \frac{x}{\left(1 + e^{\alpha - x}\right)^2} e^{-x}\mathrm{d}x = \lim_{t\to+\infty}\int_0^t \frac{x}{\left(1 + e^{\alpha-x}\right)^2} e^{-x}\mathrm{d}x, \tag{A.65}$$

where we have used the substitution $x = |n|^2$, and the fact that, since $n \sim \mathcal{CN}(0,1)$, $x \sim \mathcal{E}(1)$.

Let $B(x) = \frac{e^{-\alpha}}{1+e^{\alpha-x}}$ and $B'(x) \triangleq \frac{\mathrm{d}B(x)}{\mathrm{x}} = \frac{e^{-x}}{(1+e^{\alpha-x})^2}$. Then, from (A.65) we have

$$\mathbb{E}\left[\frac{|n|^2}{\left(1 + e^{\alpha}\exp\{-|n|^2\}\right)^2}\right] = \lim_{t\to+\infty}\int_0^t xB'(x)\mathrm{d}x. \tag{A.66}$$

By solving the integral in the limit by parts, we have

$$\int_0^t xB'(x)\mathrm{d}x = tB(t) - \int_0^t B(x)\mathrm{d}x = tB(t) - e^{-\alpha}\ln\left(e^t + e^{\alpha}\right) + e^{-\alpha}\ln\left(1 + e^{\alpha}\right), \quad \text{(A.67)}$$

where in the last step we used the fact that $B(x) = e^{-\alpha}\frac{\mathrm{d}}{\mathrm{d}x}\ln\left(e^x + e^{\alpha}\right)$. Finally, the result is straightforwardly obtained by substituting the expression above in (A.66), and by letting $t \to +\infty$. □

Appendix A.B: Proof of Theorem A.7.1

Proof of Theorem A.7.1. Let $f(x,\beta) = \ln\beta + Z\ln x + \frac{1}{\beta L}\sum_{k=0}^{L-1} A_k x^{-k}$, where we have defined $x = e^{-\omega} \in (0,1]$ in the argument of the minimization in (A.38). By minimizing with respect to $\beta \geq 0$, for a fixed x, we have

$$\tilde{\beta}(x) = \arg\min_{\beta\geq 0}\left\{\ln\beta + \frac{1}{\beta L}\sum_{k=0}^{L-1} A_k x^{-k}\right\} = \frac{1}{L}\sum_{k=0}^{L-1} A_k x^{-k}.$$

Substituting into $f(x,\beta)$, we obtain $f(x,\tilde{\beta}(x)) = 1 + \ln\tilde{\beta}(x) + Z\ln x$. We now minimize $f(x,\tilde{\beta}(x))$ with respect to $x \in (0,1]$. $f(x,\tilde{\beta}(x))$ is an increasing function of $x \in (0,1]$ if and only if

$$f'(x,\tilde{\beta}(x)) = \frac{\mathrm{d}f(x,\tilde{\beta}(x))}{\mathrm{d}x} = \frac{\tilde{\beta}'(x)}{\tilde{\beta}(x)} + \frac{Z}{x} > 0, \quad \text{(A.68)}$$

where $\tilde{\beta}'(x) = \frac{\mathrm{d}\tilde{\beta}(x)}{\mathrm{d}x} = -\frac{1}{L}\sum_{k=0}^{L-1} kA_k x^{-(k+1)}$. Equivalently, multiplying both sides by $x^{Z+1}\tilde{\beta}(x) > 0$, $f(x,\tilde{\beta}(x))$ is an increasing function of $x \in (0,1]$ if and only if

$$g(x) \triangleq x^{Z+1}\tilde{\beta}(x)f'(x,\tilde{\beta}(x)) = \frac{1}{L}\sum_{k=0}^{L-1} A_k x^{Z-k}(Z-k) > 0. \quad \text{(A.69)}$$

Note that $g'(x) = \frac{\mathrm{d}g(x)}{\mathrm{d}x} = \frac{1}{L}\sum_{k=0}^{L-1} A_k x^{Z-k-1}(Z-k)^2 > 0$, $\forall x \in (0,1]$. Therefore, $g(x)$ is a continue monotone increasing function of x. Moreover, since $Z < L-1$ from (A.36) and $\lim_{x\to 0^+} x^m = +\infty$ when $m < 0$, we have $\lim_{x\to 0^+} g(x) = -\infty$. Therefore, if $g(1) > 0$, *i.e.*, $\sum_{k=0}^{L-1}(Z-k)A_k > 0$, then there exists a unique $\tilde{x} \in (0,1)$ solution of $g(\tilde{x}) = 0$ such that

$$\begin{cases} g(x) > 0, & \forall x > \tilde{x} \\ g(x) < 0, & \forall x < \tilde{x}. \end{cases} \quad \text{(A.70)}$$

Equivalently, $\tilde{x} \in (0,1)$ is the unique solution of $f'(x, \tilde{\beta}(x)) = 0$ such that

$$\begin{cases} f'(x, \tilde{\beta}(x)) > 0, & \forall x > \tilde{x} \\ f'(x, \tilde{\beta}(x)) < 0, & \forall x < \tilde{x}. \end{cases} \tag{A.71}$$

As a consequence, $\tilde{x}$ is the unique minimizer of $f(x),\ x \in (0,1]$, and $\left\{\tilde{\beta}(\tilde{x}), \tilde{\omega} = -\ln \tilde{x}\right\}$ uniquely minimizes (A.38).

Conversely, if $g(1) \leq 0$, *i.e.*, $\sum_{k=0}^{L-1} (Z-k) A_k \leq 0$, then $g(x) \leq 0, \quad \forall x \in (0,1]$. This is equivalent to $f'(x, \tilde{\beta}(x)) \leq 0, \quad \forall x \in (0,1]$. As a consequence, 1 is the unique minimizer of $f(x, \tilde{\beta}(x))$, and $\left\{\tilde{\beta}(1), \tilde{\omega} = 0\right\}$ uniquely minimizes (A.38). □

List of Publications

The work presented in this thesis has in part appeared in the articles reported below.

Journal papers

Cognitive Radio Networks

[J1] **N. Michelusi**, P. Popovski, O. Simeone, M. Levorato, M. Zorzi, "Cognitive Access Policies under a Primary ARQ process via Forward-Backward Interference Cancellation" *IEEE Journal on Selected Areas in Communications*, to appear

Energy Harvesting

[J2] **N. Michelusi**, K. Stamatiou, M. Zorzi, "Transmission policies for energy harvesting sensors with time-correlated energy supply" submitted to *IEEE Transactions on Communications*

[J3] **N. Michelusi**, L. Badia, R. Carli, L. Corradini, M. Zorzi, "Energy Management Policies for Harvesting-based Wireless Sensor Devices with Battery Degradation" submitted to *IEEE Transactions on Communications*

UWB Channel Estimation

[J4] **N. Michelusi**, U. Mitra, A.F. Molisch, M. Zorzi, "UWB Sparse/Diffuse Channels, Part I: Channel Models and Bayesian Estimators" *IEEE Transactions on Signal Processing*, vol.60, no.10, pp.5307-5319, Oct. 2012

[J5] **N. Michelusi**, U. Mitra, A.F. Molisch, M. Zorzi, "UWB Sparse/Diffuse Channels, Part II: Estimator Analysis and Practical Channels" *IEEE Transactions on Signal Processing*, vol.60, no.10, pp.5320-5333, Oct. 2012

Conference/Workshop papers

Cognitive Radio Networks

[C1] **N. Michelusi**, O. Simeone, M. Levorato, P. Popovski, M. Zorzi, "Optimal cognitive transmission exploiting redundancy in the primary ARQ process", in *Information Theory and Applications Workshop (ITA)*, pp.1-10, 6-11 Feb. 2011

[C2] **N. Michelusi**, P. Popovski, M. Levorato, O. Simeone, M. Zorzi, "Cognitive transmissions under a primary ARQ process via backward interference cancellation", in *49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp.727-735, 28-30 Sept. 2011

Energy Harvesting

[C3] **N. Michelusi**, K. Stamatiou, M. Zorzi, "On optimal transmission policies for energy harvesting devices", in *Information Theory and Applications Workshop (ITA)*, pp.249-254, 5-10 Feb. 2012

[C4] **N. Michelusi**, K. Stamatiou, M. Zorzi, "Performance Analysis of Energy Harvesting Sensors with Time-Correlated Energy Supply", in *50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 1-5 October 2012, Monticello, USA

[C5] **N. Michelusi**, M. Zorzi, "Optimal Random Multiaccess in Energy Harvesting Wireless Sensor Networks", submitted to *IEEE ICC 2013 - 2nd IEEE International Workshop on Energy Harvesting, Measurement and Evaluation for Wireless Communications*

[C6] **N. Michelusi**, L. Badia, R. Carli, L. Corradini, M. Zorzi, "Impact of Battery Degradation on Optimal Management Policies of Harvesting-Based Wireless Sensor Devices", in *IEEE INFOCOM 2013*, to appear

[C7] **N. Michelusi**, K. Stamatiou, L. Badia, M. Zorzi, "Operation Policies for Energy Harvesting Devices with Imperfect State-of-Charge Knowledge", in *IEEE International Conference on Communications (ICC)*, pp.5782-5787, 10-15 June 2012

[C8] **N. Michelusi**, L. Badia, R. Carli, K. Stamatiou, M. Zorzi, "Correlated energy generation and imperfect State-of-Charge knowledge in energy harvesting devices", in *8th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp.401-406, 27-31 Aug. 2012

UWB Channel Estimation

[C9] **N. Michelusi**, U. Mitra, M. Zorzi, "Hybrid sparse/diffuse UWB channel estimation", in *IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp.201-205, 26-29 June 2011

[C10] **N. Michelusi**, B. Tomasi, U. Mitra, J. Preisig, M. Zorzi, "An evaluation of the hybrid sparse/diffuse algorithm for underwater acoustic channel estimation", in *OCEANS 2011*, pp.1-10, 19-22 Sept. 2011

[C11] **N. Michelusi**, U. Mitra, A.F. Molisch, M. Zorzi, "Hybrid sparse/diffuse channels: A new model and estimators for wideband channels", in *49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp.477-484, 28-30 Sept. 2011

Bibliography

[1] D. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific, Belmont, Massachusetts, 2005.

[2] FCC Spectrum Policy Task Force, "Report of the Spectrum Efficiency Working Group," FCC, Tech. Rep., Nov. 2002, available on http://www.fcc.gov/sptf/reports.html.

[3] J. Peha, "Approaches to spectrum sharing," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 10–12, Feb. 2005.

[4] ——, "Sharing Spectrum Through Spectrum Policy Reform and Cognitive Radio," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 708–719, Apr. 2009.

[5] J. Mitola and G. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[6] I. Akyildiz, W.-Y. Lee, M. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 40–48, Apr. 2008.

[7] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking Spectrum Gridlock With Cognitive Radios: An Information Theoretic Perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[8] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May 2007.

[9] F. F. Digham, M.-S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 21–24, Jan. 2007.

[10] L. B. Le and E. Hossain, "Resource allocation for spectrum underlay in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5306–5315, Dec. 2008.

[11] I. Maric, A. Goldsmith, G. Kramer, and S. Shamai, "On the capacity of interference channels with a partially-cognitive transmitter," in *IEEE International Symposium on Information Theory, ISIT*, June 2007, pp. 2156–2160.

[12] A. Jovicic and P. Viswanath, "Cognitive Radio: An Information-Theoretic Perspective," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3945–3958, Sep. 2009.

[13] S. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of Myopic Sensing in Multichannel Opportunistic Access," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4040–4050, Sept. 2009.

[14] R. Comroe and D. Costello, "ARQ Schemes for Data Transmission in Mobile Radio Systems," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 4, pp. 472–481, July 1984.

[15] R. Tannious and A. Nosratinia, "Cognitive Radio Protocols Based on Exploiting Hybrid ARQ Retransmissions," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2833–2841, Sep. 2010.

[16] J. Li, W. Zhang, A. Nosratinia, and J. Yuan, "Opportunistic Spectrum Sharing Based on Exploiting ARQ Retransmission in Cognitive Radio Networks," in *IEEE Global Telecommunications Conference (Globecom)*, Dec. 2010, pp. 1–5.

[17] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[18] D. J. White, *Markov Decision Processes*. Wiley, 1993.

[19] M. Levorato, U. Mitra, and M. Zorzi, "Cognitive Interference Management in Retransmission-Based Wireless Networks," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3023–3046, May 2012.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., New York, 2006.

[21] M. Levorato, S. Firouzabadi, and A. Goldsmith, "A Learning Framework for Cognitive Interference Networks with Partial and Noisy Observations," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3101–3111, Sep. 2012.

[22] K. W. Ross, "Randomized and Past-Dependent Policies for Markov Decision Processes with Multiple Constraints," *Operations Research*, vol. 37, no. 3, pp. 474–477, 1989.

[23] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Springer, 1960.

[24] N. Michelusi, P. Popovski, M. Levorato, O. Simeone, and M. Zorzi, "Cognitive transmissions under a primary ARQ process via Backward Interference Cancellation," in *49th Allerton Conference on Communication, Control, and Computing*, Sep. 2011, pp. 727–735.

[25] R. Narasimhan, "Individual Outage Rate Regions for Fading Multiple Access Channels," in *IEEE International Symposium on Information Theory, ISIT*, June 2007, pp. 24–29.

[26] R. Gallager, *Discrete Stochastic Processes*. Kluwer, Boston, 1996.

[27] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[28] D. Anthony, P. Bennett, M. C. Vuran, M. B. Dwyer, S. Elbaum, A. Lacy, M. Engels, and W. Wehtje, "Sensing through the continent: towards monitoring migratory birds using cellular sensor networks," in *Proceedings of the 11th international conference on Information Processing in Sensor Networks (ISPN)*, vol. 12, Apr. 2012, pp. 329–340.

[29] I. Dietrich and F. Dressler, "On the lifetime of wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 5, no. 1, pp. 5:1–5:39, Feb. 2009. [Online]. Available: http://doi.acm.org/10.1145/1464420.1464425

[30] Y. Chen and Q. Zhao, "On the lifetime of wireless sensor networks," *IEEE Communications Letters*, vol. 9, no. 11, pp. 976–978, Nov. 2005.

[31] M. Bhardwaj and A. Chandrakasan, "Bounding the lifetime of sensor networks via optimal role assignments," in *IEEE INFOCOM*, vol. 3, 2002, pp. 1587–1596.

[32] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Computing*, vol. 4, pp. 18–27, Jan. 2005.

[33] F. Ongaro, S. Saggini, S. Giro, and P. Mattavelli, "Two-dimensional MPPT for photovoltaic energy harvesting systems," in *Proc. 12th IEEE Workshop on Control and Modeling for Power Electronics*, June 2010.

[34] C. Luo and H. F. Hofmann, "Wideband Energy Harvesting for Piezoelectric Devices with Linear Resonant Behaviour," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 58, no. 7, pp. 1294–1301, July 2011.

[35] Y. K. Ramadass and A. P. Chandrakasan, "A Battery-less Thermoelectric Energy Harvesting Interface Circuit with 35 mV Startup Voltage," *IEEE Journal on Solid-State Circuits*, vol. 46, no. 1, pp. 333–341, Jan. 2011.

[36] T. Paing, J. Shin, R. Zane, and Z. Popovic, "Resistor emulation approach to low-power RF energy harvesting," *IEEE Transactions on Power Electronics*, vol. 23, no. 3, pp. 1494–1501, May 2008.

[37] D. Niyato, E. Hossain, M. Rashid, and V. Bhargava, "Wireless sensor networks with energy harvesting technologies: a game-theoretic approach to optimal energy management," *IEEE Wireless Communications*, vol. 14, no. 4, pp. 90–96, Aug. 2007.

[38] R. A. Valles, A. G. Marques, and J. G. Sueiro, "Optimal selective forwarding for energy saving in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 164–175, Jan. 2011.

[39] D. Niyato, E. Hossain, and A. Fallahi, "Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: performance analysis and optimization," *IEEE Transactions on Mobile Computing*, vol. 6, pp. 221–236, Feb. 2007.

[40] F. Iannello, O. Simeone, and U. Spagnolini, "Energy management policies for passive RFID sensors with RF-energy harvesting," in *IEEE International Conference on Communications (ICC)*, May 2010, pp. 1–6.

[41] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable Sensor Activation under Temporally Correlated Events," *Springer Wireless Networks (WINET)*, vol. 15, pp. 619–635, July 2009.

[42] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," *IEEE Transactions on Communications*, vol. 58, pp. 2116–2126, July 2010.

[43] M. Kashef and A. Ephremides, "Optimal scheduling for energy harvesting sources on time varying wireless channels," in *49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2011, pp. 712–718.

[44] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Control of Wireless Networks with rechargeable batteries," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 581–593, Feb. 2010.

[45] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1326–1336, Apr. 2010.

[46] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: optimal policies," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1732–1743, Sep. 2011.

[47] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Transactions on Wireless Communications*, vol. 11, pp. 1180–1189, 2012.

[48] B. Devillers and D. Gunduz, "A general framework for the optimization of energy harvesting communication systems with battery imperfections," *Journal of Communications and Networks*, vol. 14, no. 2, pp. 130–139, Apr. 2012.

[49] K. Kar, A. Krishnamurthy, and N. Jaggi, "Dynamic Node Activation in Networks of Rechargeable Sensors," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 15–26, Feb. 2006.

[50] S. Reddy and C. Murthy, "Dual-stage power management algorithms for energy harvesting sensors," *IEEE Transactions on Wireless Communications*, vol. 11, pp. 1434–1445, Apr. 2012.

[51] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 547–551, Feb. 2009.

[52] F. Iannello, O. Simeone, and U. Spagnolini, "Medium Access Control Protocols for Wireless Sensor Networks with Energy Harvesting," *IEEE Transactions on Communications*, vol. 60, no. 5, pp. 1381–1389, May 2012.

[53] F. Iannello, O. Simeone, P. Popovski, and U. Spagnolini, "Energy group-based dynamic framed ALOHA for wireless networks with energy harvesting," in *46th Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2012, pp. 1–6.

[54] V. Sharma, U. Mukherji, and V. Joseph, "Efficient energy management policies for networks with energy harvesting sensor nodes," in *46th Annual Allerton Conference on Communication, Control, and Computing*, Sep. 2008, pp. 375–383.

[55] Z. A. Eu and H.-P. Tan, "Probabilistic polling for multi-hop energy harvesting wireless sensor networks," in *IEEE International Conference on Communications (ICC)*, June 2012, pp. 271–275.

[56] D. Salerno, "Ultralow voltage energy harvester uses thermoelectric generator for battery-free wireless sensors," *Linear Technology Journal of Analog Innovation*, vol. 20, pp. 1–11, Oct. 2010.

[57] N. Michelusi, K. Stamatiou, and M. Zorzi, "On optimal transmission policies for energy harvesting devices," in *Information Theory and Applications Workshop (ITA)*, Feb. 2012, pp. 249–254.

[58] R. Howard, *Dynamic programming and Markov processes*, 1st ed. The MIT Press, 1960.

[59] C. Renner and V. Turau, "CapLibrate: self-calibration of an energy harvesting power supply with supercapacitors," in *International Conference on Architecture of Computing Systems (ARCS)*, Hannover, Germany, Feb. 2010, pp. 1–10.

[60] N. Michelusi, K. Stamatiou, L. Badia, and M. Zorzi, "Operation policies for Energy Harvesting Devices with imperfect State-of-Charge knowledge," in *IEEE International Conference on Communications (ICC)*, June 2012, pp. 5782–5787.

[61] R. L. Burden and J. D. Faires, *Numerical Analysis, 9th Edition*. Cengage Learning, 2011.

[62] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions*. New York: Dover Publications, 1965.

[63] X. Deng and A. Haimovich, "Achievable Rates Over Time-Varying Rayleigh Fading Channels," *IEEE Transactions on Communications*, vol. 55, no. 7, pp. 1397–1406, July 2007.

[64] W. Lee, "Estimate of channel capacity in Rayleigh fading environment," *IEEE Transactions on Vehicular Technology*, vol. 39, no. 3, pp. 187–189, Aug. 1990.

[65] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[66] D. Linden and T. B. Reddy, *Handbook of Batteries*, 3rd ed. McGraw Hill, 2002.

[67] Seiko Instruments Inc., Micro Battery Product Catalogue, 2011-2012. [Online]. Available: www.sii.co.jp/compo/catalog/battery_en.pdf

[68] H. N. Seiger, "Effects of depth of discharge on cycle life of near term batteries," in *Proc. 16th Intersociety Energy Conversion Engineering Conference*, Atlanta, GA, 1981, p. 102.

[69] L. H. Thaller, "Expected Cycle Life vs. Depth of Discharge Relationships of Well Behaved Single Cells and Cell Strings," NASA, Tech. Rep., 1982. [Online]. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19830002288_1983002288.pdf

[70] ——, "A Prediction Model of the Depth-of-Discharge Effect on the Cycle Life of a Storage Cell," NASA, Tech. Rep., 1987. [Online]. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19870012878_1987012878.pdf

[71] S. Drouilhet and B. L. Johnson, "A Battery Life Prediction Method for Hybrid Power Applications," 1997. [Online]. Available: http://www.nrel.gov/docs/legosti/fy97/21978.pdf

[72] B. Devillers and D. Gunduz, "Energy harvesting communication system with battery constraint and leakage," in *IEEE GLOBECOM Workshops*, Dec. 2011, pp. 383–388.

[73] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: Measurements and Algorithms," in *Proceedings of IEEE INFOCOM*, Apr. 2011, pp. 1602–1610.

[74] C.-F. Chiasserini and R. Rao, "Energy efficient battery management," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 7, pp. 1235–1245, July 2001.

[75] C. D. Meyer, Ed., *Matrix analysis and applied linear algebra*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2000.

[76] M. Win and R. Scholtz, "Impulse radio: how it works," *IEEE Communications Letters*, vol. 2, no. 2, pp. 36–38, Feb. 1998.

[77] R. Scholtz, "Multiple access with time-hopping impulse modulation," in *IEEE Military Communications Conference*, vol. 2, Oct. 1993, pp. 447–450.

[78] R. Qiu, H. Liu, and X. Shen, "Ultra-wideband for multiple access communications," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 80–87, Feb. 2005.

[79] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization via Ultra-Wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, July 2005.

[80] M. Win and R. Scholtz, "On the robustness of ultra-wide bandwidth signals in dense multipath environments," *IEEE Communications Letters*, vol. 2, no. 2, pp. 51–53, Feb. 1998.

[81] M. Chiani and A. Giorgetti, "Coexistence Between UWB and Narrow-Band Wireless Communication Systems," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 231–254, Feb. 2009.

[82] A. Batra, J. Balakrishnan, G. Aiello, J. Foerster, and A. Dabak, "Design of a multiband OFDM system for realistic UWB channel environments," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 9, Sep. 2004.

[83] T. Zasowski, G. Meyer, F. Althaus, and A. Wittneben, "Propagation effects in UWB body area networks," in *IEEE International Conference on Ultra-Wideband (ICU)*, Sep. 2005, pp. 16–21.

[84] L. Yang and G. Giannakis, "Ultra-Wideband Communications: an idea whose time has come," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 26–54, Nov. 2004.

[85] A. Molisch, D. Cassioli, C.-C. Chong, S. Emami, A. Fort, B. Kannan, J. Karedal, J. Kunisch, H. Schantz, K. Siwiak, and M. Win, "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 11, pp. 3151–3166, Nov. 2006.

[86] G. Aiello and G. Rogerson, "Ultra-Wideband Wireless Systems," *IEEE Microwave Magazine*, vol. 4, no. 2, June 2003.

[87] C. Carbonelli and U. Mitra, "Clustered Channel Estimation for UWB Multiple Antenna Systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 970–981, Mar. 2007.

[88] ——, "Clustered ML Channel Estimation for Ultra-Wideband Signals," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2412–2416, July 2007.

[89] P. Schniter, "A Message-Passing Receiver for BICM-OFDM Over Unknown Clustered-Sparse Channels," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 8, pp. 1462–1474, Dec. 2011.

[90] A. Molisch, "Ultra-Wide-Band Propagation Channels," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 353–371, Feb. 2009.

[91] T. Blumensath and M. Davies, "Iterative Thresholding for Sparse Approximations," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 629–654, 2008.

[92] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[93] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[94] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[95] C. Carbonelli, S. Vedantam, and U. Mitra, "Sparse Channel Estimation with Zero Tap Detection," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1743–1763, May 2007.

[96] J. Paredes, G. Arce, and Z. Wang, "Ultra-Wideband Compressed Sensing: Channel Estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 3, pp. 383–395, Oct. 2007.

[97] P. Meissner, T. Gigl, and K. Witrisal, "UWB sequential Monte Carlo positioning using virtual anchors," in *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sep. 2010, pp. 1–10.

[98] Y. Shen and M. Win, "Fundamental Limits of Wideband Localization-Part I: A General Framework," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4956–4980, Oct. 2010.

[99] F. Quitin, C. Oestges, F. Horlin, and P. De Doncker, "Diffuse multipath component characterization for indoor MIMO channels," in *Proceedings of the Fourth European Conference on Antennas and Propagation (EuCAP)*, Apr. 2010, pp. 1–5.

[100] A. Molisch, "Ultrawideband propagation channels-Theory, Measurement, and Modeling," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1528–1545, Sep. 2005.

[101] T. Santos, F. Tufvesson, and A. Molisch, "Modeling the Ultra-Wideband Outdoor Channel: Model Specification and Validation," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 1987–97, June 2010.

[102] J. Kunisch and J. Pamp, "An ultra-wideband space-variant multipath indoor radio channel model," in *IEEE Conference on Ultra Wideband Systems and Technologies*, Nov. 2003, pp. 290–294.

[103] Y. Zhou, X. Yin, N. Czink, T. Zemen, A. Guo, and F. Liu, "Evaluation of Doppler-Delay Properties of Diffuse Components in Vehicular Propagation Channels," in *2nd IEEE International Conference on Wireless Access in Vehicular Environments*, Dec. 2009.

[104] N. Czink, F. Kaltenberger, Y. Zhou, L. Bernado, T. Zemen, and X. Yin, "Low-Complexity Geometry-Based Modeling of Diffuse Scattering," in *Proceedings of the Fourth European Conference on Antennas and Propagation (EuCAP)*, Apr. 2010.

[105] R. Thoma, M. Landmann, and A. Richter, "RIMAX-a Maximum Likelihood Framework for Parameter Estimation in Multidimensional Channel Sounding," in *International Symposium on Antennas and Propagation (ISAP)*, Aug. 2004.

[106] R. Qiu, "A study of the ultra-wideband wireless propagation channel and optimum UWB receiver design," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1628–1637, Dec. 2002.

[107] A. Molisch, *Wireless Communications, Second Edition*, ser. Wiley-IEEE. John Wiley & Sons, 2011.

[108] J. Karedal, S. Wyne, P. Almers, F. Tufvesson, and A. Molisch, "Statistical analysis of the UWB channel in an industrial environment," in *IEEE 60th Vehicular Technology Conference*, vol. 1, Sep. 2004, pp. 81–85.

[109] Y. Chi, L. Scharf, A. Pezeshki, and A. Calderbank, "Sensitivity to Basis Mismatch in Compressed Sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, May 2011.

[110] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed. Springer, Aug. 1998.

[111] N. Michelusi, U. Mitra, and M. Zorzi, "Hybrid Sparse/Diffuse UWB channel estimation," in *IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2011, pp. 201–205.

[112] A. Saleh and R. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[113] D. Cassioli, M. Win, and A. Molisch, "The ultra-wide bandwidth indoor channel: from statistical model to simulations," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1247–1257, Aug. 2002.

[114] J. Hansen, "An analytical calculation of power delay profile and delay spread with experimental verification," *IEEE Communications Letters*, vol. 7, no. 6, pp. 257–259, June 2003.

[115] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, 1977.

[116] K. Herrity, A. Gilbert, and J. Tropp, "Sparse Approximation Via Iterative Thresholding," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, May 2006.

[117] Y. Wen, W. Huang, and Z. Zhang, "CAZAC sequence and its application in LTE random access," in *IEEE Information Theory Workshop, ITW*, Oct. 2006, pp. 544–547.

[118] R. G. Bartle, *The Elements of Integration and Lebesgue Measure*, 1st ed. Wiley-Interscience, Jan. 1995.

Acknowledgments

I would like to thank all the people I have collaborated with during these years, for their precious contributions and suggestions they gave me to improve my research activity, among these: Dr. Marco Levorato, Prof. Osvaldo Simeone, Prof. Petar Popovski, Dr. Kostas Stamatiou, Dr. Leonardo Badia, Dr. Luca Corradini, Dr. Ruggero Carli, Prof. Urbashi Mitra, Prof. Andy Molisch, Dr. James Preisig and Dr. Beatrice Tomasi.

I am especially grateful to Prof. Michele Zorzi, for being a careful advisor during these years. The work presented in this thesis would not be possible without his guidance.

I would also like to thank Prof. Urbashi Mitra, for being an inspiring supervisor during my visit at University of Southern California, USA. The period I spent there has been an exceptional chance for professional and human growth.

Moreover, I am grateful to Prof. Petar Popovski, for being an excellent supervisor during my visit at Aalborg University, Denmark. His valuable contributions have helped shaping my research activity.

Finally, I am particularly grateful to all the people who have been part of my daily life, thus becoming an irreplaceable source of moral support and authentic friendship.