Universitá degli Studi di Padova
Dipartimento di Ingegneria dell'Informazione
Laurea Magistrale in Ingegneria delle Telecomunicazioni

Transmission policies for Wireless Sensor Networks with Energy Harvesting Devices

Strategie di trasmissione per reti di sensori wireless con dispositivi Energy Harvesting

Studente
Alessandro Biason

Supervisore
Chiar.mo Prof. Michele Zorzi

Anno Accademico 2013/2014
7 Luglio 2014

alla mia famiglia
e a Sabrina

Abstract

Recently Wireless Sensor Networks (WSNs) with energy harvesting capabilities are experiencing increasing interest with respect to traditional ones, due to their promise of extended lifetime. In this work we consider three different scenarios with one or two Energy Harvesting Devices (EHDs). In all the cases the main goal is to maximize the *long-term average importance* associated with the transmitted data. The devices, at each time instant, have data of different importance levels to be transmitted, as well as different battery energy levels.

In the first part we consider the case of a pair of devices coordinated by a Central Controller (CC). Assuming a negligible processing cost in terms of energy, our objective is to identify *low-complexity* transmission policies, that achieve good performance with respect to the optimal one. We numerically show that two policies, namely the Balanced Policy (BP) and the Heuristic Constrained Energy Independent Policy (HCEIP), despite being independent of the battery energy levels, achieve near optimal performance in most cases of interest, and can be easily found with an adaptation to the ambient energy supply. Moreover, we derive an analitycal approximation of the BP and show that this policy can be considered as a good lower bound for the performance of the Optimal Policy.

In the second part we analyse the case of two devices, an Energy Harvesting Transmitter (EHTX) that sends data to an Energy Harvesting Receiver (EHRX). We consider several scenarios: centralized case, semi-independent case and totally-independent case. In the last two, we analyze different cases as a function of the quantity of information the devices have. We find that the outage information is essential for the performance of the system.

Finally, we study the case of one EHD with a FIFO data queue. We derive an analytical sufficient condition for the queue stability and we use this information to design a Low Complexity Policy (LCP). This is a particular case of Energy Independent Policy (EIP) and it is easy to implement with the knowledge of the ambient energy supply and the packets arrival rate. We numerically verify that LCP achieves good performance with respect to the Optimal Policy.

Sommario

Recentemente le reti di sensori wireless (WSNs) con capacità di energy harvesting stanno suscitando crescente interesse rispetto a quelle tradizionali, grazie alla loro capacità di aumentare la durata operativa della rete. In questa tesi si considerano tre diversi scenari con uno o due Energy Harvesting Devices (EHDs). In tutti i casi l'obiettivo principale è massimizzare, in condizioni asintotiche, l'importanza media associata ai dati trasmessi. I dispositivi in ogni istante temporale possono trasmettere dati di diversa importanza e possiedono vari livelli d'energia della batteria.

Nella prima parte si considera il caso di una coppia di dispositivi coordinati da un controllore centrale. Assumendo un costo di elaborazione trascurabile in termini energetici, l'obiettivo è identificare strategie di trasmissione a basse complessità che raggiungano buone prestazioni rispetto a quelle ottime. Si verifica numericamente che due strategie, la Balanced Policy (BP) e la Heuristic Constrained Energy Independent Policy (HCEIP), indipendentemente dal livello energetico delle batterie, hanno prestazioni quasi ottime in molti casi d'interesse e possono essere facilmente trovate con un semplice adattamento alla disponibilità energetica dell'ambiente. Inoltre, si deriva un'approssimazione analitica di BP e si mostra che questa strategia può essere considerata un buon limite inferiore alle prestazioni della Optimal Policy.

Il secondo problema consiste nell'analisi del comportamento di una coppia di dispositivi: un Energy Harvesting Transmitter (EHTX) che trasmette dati a un Energy Harvesting Receiver (EHRX). Si considerano i diversi scenari di caso centralizzato, semi indipendente e totalmente indipendente. Negli ultimi due, si analizzano diversi casi in funzione della quantità dell'informazione che i dispositivi possiedono. Si verifica inoltre che l'informazione di outage è essenziale per le prestazioni del sistema.

Infine, si studia il caso di un EHD con una coda di dati FIFO. Si deriva una condizione analitica sufficiente per la stabilità della coda e si utilizza tale informazione per costruire una strategia di trasmissione a bassa complessità. Questa è un caso particolare di Energy Independent Policy (EIP) e può essere facilmente implementata conoscendo la probabilità d'arrivo dei quanti energetici e dei pacchetti. Si verifica numericamente che la strategia ideata ottiene buone prestazioni rispetto a quella ottima.

Contents

Acronyms

BP Balanced Policy.

C Centralized case.

CC Central Controller.

CEIP Constrained Energy Independent Policy.

EHD Energy Harvesting Device.

EHRX Energy Harvesting Receiver.

EHTX Energy Harvesting Transmitter.

EIP Energy Independent Policy.

FIFO First In First Out.

HCEIP Heuristic Constrained Energy Independent Policy.

LCP Low Complexity Policy.

MC Markov Chain.

OCEIP Optimal-CEIP.

OEIP Optimal-EIP.

OP Optimal Policy.

pdf probability density function.

PIA Policy Iteration Algorithm.

RX Receiver.

S,L Semi-Independent case with Local knowledge.

S,P Semi-Independent case with Partial knowledge.

SNR Signal to Noise Ratio.

WSN Wireless Sensor Network.

Chapter 1

Introduction

Wireless Sensor Networks (WSNs) have been designed to sense the environment using IT equipment placed in every object in our daily life. Their application fields span from area monitoring to industrial processes monitoring and health care monitoring. Moreover, the devices can be deployed really close to the phenomenon to sense, *e.g.*, in inaccessible or disaster areas [1].The reference scenario for this kind of networks is multi-hop, distributed, autonomous and battery-powered; the sensor nodes have sensing and transmitting/receiving capabilities but they are subject to energy and computational resources constraints due to their limited power source. Generally, in a WSN, nodes report data to a particular device called *sink*, who has higher computational and storage capabilities. The sink can be considered as a gateway between the WSN and other networks. A node sends a packet towards the sink when it senses something or if the sink explicitly sends a *query* message. Packets can reach the sink through the WSN, with a multi-hop mechanism. This strategy is used in order to reduce the energy expenditure, that is a common problem in this scenarios. Indeed, in WSNs, reducing the energy cost may greatly reduce the devices maintenance cost: the nodes are distributed in a wide area or in a complex environment, therefore it is difficult to replace their batteries. Thus, differently from wired networks, WSN algorithms have to focus also on power conservation, considering both the throughput and the energetic constraint. To increase the lifetime of the network, many strategies are adopted: choosing the best modulation strategy [2], exploiting a duty cycle mechanism (sleep/listen) [3], reducing the amount of data to transmit [4, 5], optimizing the routing techniques [6, 7] or/and MAC [8] or using an efficient transmission scheduling [9]. Beyond these, in the last years, the Energy Harvesting mechanism [10], that is the process of exploiting the ambient energy to power a device, is steadily gaining popularity in deployments. In nature there are several free source of energy [11, 12], *e.g.*, solar, heat, motion or electromagnetic, thus this technique was already used for many other reasons [13].

A device with harvesting capabilities is defined as an Energy Harvesting Device (EHD). For an EHD, energy conservation is no longer the main concern. Differently from traditional networks, where the main objective is to minimize a particular metric, *e.g.*, delay [14], with EHDs the goal is to efficiently manage the energy gathered from the environment. Intuitively, when a finite battery is available, an EHD should judiciously perform its assigned task based on its available energy, becoming more “conservative" as its energy supply runs low to ensure uninterrupted operation, and more “aggressive" when energy is abundant, to avoid that harvested energy is wasted due to lack of storage space.

Problems addressed in this thesis

In this work we consider several problems:

1. two EHDs with a Central Controller (CC);

2. an Energy Harvesting Transmitter (EHTX) and Energy Harvesting Receiver (EHRX) pair;

3. an EHD with a data queue.

In all these problems, we consider system models consisting of EHDs that report data of varying "importance" levels, with the aim of maximizing the long-term average importance of the reported data. Practical examples of these scenarios include: a network of temperature sensing EHDs, where a high temperature measurement can be an indicator of overheating or fire; a sensor network which routes different priority packets [15]; data transmission over a fading channel where the EHDs adjust the transmitted redundancy according to the instantaneous channel realization, in which case the importance level corresponds to the instantaneous rate [10]. As in prior works, *e.g.*, see [16], [17] and [18], we assume a slotted-time system and i.i.d. Bernoulli energy arrivals, where the importance values follow an arbitrary continuous distribution.

In the first problem we consider the centralized case of two EHDs, in which a Central Controller (CC) allows at most the transmission of one EHD at the same time. CC knows the battery energy levels and the packet importance levels of each EHD. This approach can be considered as an upper bound to the decentralized scenario in terms of overall performance and is a first step towards the study of multi-user EHD networks. It was shown in [17] that, for the class of binary (transmit/no transmit) policies, the optimal policy dictates the transmission of data with importance above a given threshold, which is a function of the joint energy levels available at the two EHDs. In this work, using the structure of the Optimal Policy (OP), we derive different suboptimal policies which are shown to perform closely to the optimal one. Firstly, we introduce the Energy Independent Policy (EIP), *i.e.*, a threshold policy which, on average, transmits with a constant probability, independent of the battery energy levels. Furthermore, in accordance to the values of the average energy harvesting rates, we define the Balanced Policy (BP) and the Heuristic Constrained Energy Independent Policy (HCEIP), that are particular cases of EIP. These are *low-complexity* policies, that do not require any optimization to compute the transmission probabilities. The main implication of these results is that near-optimal performance can be obtained *without* precise knowledge of the energy stored in the sensor batteries at any given time and with a *simple adaptation* to the ambient energy supply. Moreover, we find an analytical approximation of the BP, numerically showing that this can be considered as a good analytical lower bound for the Optimal Policy in most cases of interest. The results of this first part were published in [19].

The second problem consists of the study of a system with an EHTX and an EHRX. We consider several cases:

1. *Centralized case*: the EHDs are coordinated by a Central Controller. We suppose that EHTX transmits if and only if EHRX can receive. This is the optimal scenario because energy is never wasted.

2. *Semi-Independent case*: the EHDs are not coordinated but we assume that if there are no transmissions, then EHRX does not waste energy listening to the channel. For example, in a real scenario, EHRX can listen to the channel at the beginning of a slot and if it does not sense anything, then it goes to sleep for that slot. We make several assumptions on the amount of information the devices have:

   - *global knowledge*: each EHD knows the energy state of the other one;
   - *partial knowledge*: each EHD knows if the battery of the other device is empty or not (*outage information*);
   - *local knowledge*: the EHDs do not have any information about the energy level of the other device.

3. *Totally-Independent case*: the EHDs are not coordinated, *i.e.*, they take every action independently from the other device. Also here the device may have different degrees of knowledge of the overall energy status as in the previous case.

The last considered problem deals with an EHD with a First In First Out (FIFO) data queue. Initially we analyze the case of an infinite data queue and we suppose that the transmission of a packet requires the expenditure of one energy quantum. We show that if the average number of packet arrivals is greater than the average number of energy arrivals, then a policy that makes the queue stable may not exist. Furthermore, we find an analytical stability condition and we use this one to design a Low Complexity Policy (LCP) (with two variations), that is a particular case of Energy Independent Policy. Finally, we show that LCP achieves good performance with respect to the Optimal Policy.

In [20] the problem of maximizing the average value of the reported data was addressed, considering an energy-aware replenishable sensor in a continuous-time system, using policy iteration to derive the optimal thresholds. A network of energy-limited sensors was studied in [15], where the objective was the investigation of the relaying of packets of different priorities, but without focusing on energy harvesting capabilities. [10] studied heuristic delay-minimizing policies and sufficient stability conditions for an EHD with a data queue. In addition, [21, 22] dealt with a system model similar to ours, but had the objective of evaluating the probability of detection of a randomly occurring event, and [23] devoted the use of RF-energy harvesting to the performance improvement of passive RFID systems. In [24] and [25] energy harvesting receivers were studied, but the relation with the transmitter was not considered. Also, in [25] the energy harvest model assumed deterministic energy quanta arrivals. The Energy Harvesting Transmitter and Energy Harvesting Receiver pair was analyzed in [26], but without a Markov model and without considering several degrees of knowledge of the system. [10] dealt with the case of a data queue, but did not consider a small queue. Also in [27] an EHD with a data queue was treated, but the impact of a small battery was not examined.

This work is organized as follows. In Chapter 2 we study the two users centralized case. In Chapter 3 we analyze the EHTX and EHRX case. In Chapter 4 we present the case of one EHD with a data queue. Chapter 5 concludes the thesis.

Chapter 2

Two EHDs - Centralized case

2.1 System Model

We analyze the case of two Energy Harvesting Devices (EHDs) that report data to a common Central Controller (CC). Since our next results are based on the system model proposed in [17], we report it here for the sake of completeness.

Time is slotted and slot k corresponds to the time interval $[k, k+1),\ k \in \mathbb{N}_0$. Each device, in every slot k, has a packet to transmit, *i.e.*, its data queue is always nonempty. We assume that:

- in slot k, the nodes can transmit over the intervals $[k, k+\delta_{i,k}),\ i \in \{1, 2\}$, where $\delta_{i,k} \in (0, 1]$ are the duty cycles;
- the channel is ideal, *i.e.*, it does not introduce any errors;
- retransmissions are not considered, therefore, if a packet is not transmitted, it is lost;
- since the CC coordinates the two devices, the EHDs cannot transmit simultaneously, so as to avoid collisions.

We model the energy storage capability of each EHD as a buffer of size $e_{\max,i}+1,\ i \in \{1, 2\}$, where $e_{\max,i}$ is the maximum amount of energy that the EHD i can store in its battery. Each position in the buffers can hold one energy quantum and the transmission of one packet requires the expenditure of one energy quantum. The set of energy levels of the i-th device is $\mathcal{E}_i = \{0, 1, \ldots, e_{\max,i}\}$. The energy status at $k+1$ depends upon the energy state in slot k, the choice of transmitting or not in that slot and the arrival process. In particular, the following equation holds:[1]

$$E_{i,k+1} = \min\{[E_{i,k} - Q_{i,k}]^+ + B_{i,k}, e_{\max,i}\}, \qquad i \in \{1, 2\}, \tag{2.1}$$

where $\{B_{i,k}\}$ is the arrival process, that models the randomness in the energy that can be harvested from the environment. $\{B_{i,k}\}$ is equal to one if an energy quantum is harvested, and to zero otherwise. We assume that $\{B_{i,k}\}$ is an i.i.d. Bernoulli random arrival process with mean $\bar{b}_i \in (0, 1]$ independent over time and across EHDs. Note that we exclude the case $\bar{b}_i = 0$ because this is the single EHD scenario, already largely studied (*e.g.*, see [16]).

[1] with the notation $[a]^+$ we indicate $\max\{0, a\}$

Furthermore, an energy quantum harvested in the slot k can be used only in a time slot $> k$.
$\{Q_{i,k}\}$ is the action process, that represents the amount of energy devoted to the transmission in the slot k. $\{Q_{i,k}\}$ is equal to one if a packet is transmitted, zero otherwise. Note that we adopt a model without collisions, therefore the pair $(Q_{1,k}, Q_{2,k})$ is always different from $(1, 1)$.

Since the batteries are finite, the definition of *energy outage* and *energy overflow* can be introduced:

Definition 1 (energy outage)**.** *In slot k, for EHD i, energy outage occurs if* $E_{i,k} = 0$.

Definition 2 (energy overflow)**.** *In slot k, for EHD i, energy overflow occurs if* $(E_{i,k} = e_{\max,i}) \cap (B_{i,k} = 1) \cap (Q_{i,k} = 0)$.

An EHD is in outage when its battery is empty. In this condition, no transmissions can be performed, regardless of the importance of the current data packet. Differently, energy overflow occurs when the battery is fully charged, a device does not transmit any packets and $B_{i,k} = 1$. In this case, the quantum energy is lost, therefore this event may represent a future lost transmission opportunity.

We introduce the importance value of the current data packet for the EHD i as $V_{i,k}$. We model $V_{i,k}$ as a continuous random variable with probability density function (pdf) $f_{V_i}(v_i)$, $v_i \geq 0$. We also suppose that $\{V_{i,k}\}$ are i.i.d over time and across the two devices.

With the introduced quantities, we define the *state of the system*

$$\mathbf{S}_k = (E_{1,k}, E_{2,k}, V_{1,k}, V_{2,k}). \tag{2.2}$$

The assumption of two users can be relaxed in order to consider larger networks, but would require a higher computational cost. However, conceptually, the step from one to two users is the most interesting as it introduces access-related problems (contention and collisions) that do not exist in single-user systems. Furthermore, our model assumes that the process of harvested energy $B_{i,k}$ is statistically independent in different time slots. This is a particular case of the Generalized Markov model described in [28], where B_k depends on the values of an underlying "scenario" process, which is itself modeled as a Markov Chain with memory L. In [28], it is shown that different energy sources can be efficiently modeled by means of different values of L. In particular, piezoelectric energy is well described using $L = 0$, while solar is better characterized by $L = 1$. In this work, we use $L = 0$ to maintain the analysis simpler: the case for $L > 0$ can be modeled, though at the price of additional complexity, following an approach similar to [29], and is left for future study.

2.2 Policy Definition and General Optimization Problem

We consider a policy μ that, fixed a value of $\mathbf{S}_k$, determines which device transmits at the time slot k, *i.e.*, μ gives the pair $(Q_{1,k}, Q_{2,k}) \in \mathcal{Q} = \{(0,0), (0,1), (1,0)\}$. Formally, a policy is a probability measure on the action space $\mathcal{Q}$, parameterized by state $\mathbf{S}_k$: given $\mathbf{S}_k = (e_1, e_2, v_1, v_2) \in \mathcal{E}_1 \times \mathcal{E}_2 \times \mathbb{R}^+ \times \mathbb{R}^+$,[2]

$$\mu((i,j); \mathbf{e}, \mathbf{v}) = \mathbb{P}(\text{draw } i \text{ from EHD } 1 \cap \text{draw } j \text{ from EHD } 2 | \mathbf{S}_k), \qquad (i,j) \in \mathcal{Q}. \tag{2.3}$$

[2]in the next, we indicate a pair (a_1, a_2) with the bold notation "$\mathbf{a}$"

For notational simplicity we introduce the following three probabilities:

$$\mu_0(\mathbf{e},\mathbf{v}) \triangleq \mu((0,0);\mathbf{e},\mathbf{v}) = \mathbb{P}(\text{none transmits}|\mathbf{S}_k), \tag{2.4}$$

$$\mu_1(\mathbf{e},\mathbf{v}) \triangleq \mu((1,0);\mathbf{e},\mathbf{v}) = \mathbb{P}(\text{only EHD 1 transmits}|\mathbf{S}_k), \tag{2.5}$$

$$\mu_2(\mathbf{e},\mathbf{v}) \triangleq \mu((0,1);\mathbf{e},\mathbf{v}) = \mathbb{P}(\text{only EHD 2 transmits}|\mathbf{S}_k). \tag{2.6}$$

We define the *long-term average reward* G, using a policy μ, as:

$$G(\mu,\mathbf{S}_0) = \lim_{K\to\infty} \inf \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1}(Q_{1,k}V_{1,k} + Q_{2,k}V_{2,k}|\mathbf{S}_0)\right], \tag{2.7}$$

where:

- $\mathbf{S}_0 = (E_{1,0}, E_{2,0}, V_{1,0}, V_{2,0})$ is the initial system state;
- the expectation is taken with respect to $\{Q_{1,k}, Q_{2,k}, V_{1,k}, V_{2,k}, B_{1,k}, B_{2,k}\}$;
- the policy μ selects $Q_{1,k}$ and $Q_{2,k}$.

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

$$\mu^* = \arg\max_{\mu} \{G(\mu,\mathbf{S}_0)\}. \tag{2.8}$$

It can be proved (see [17]), that the optimal policy μ^* must have a threshold structure with respect to the importance of the current data packet: a pair of thresholds $(v_{1,\text{th}}(\mathbf{e}), v_{2,\text{th}}(\mathbf{e}))$ is associated to every pair of joint energy level $\mathbf{e}$. In particular it holds:

$$\begin{cases} \mu_0(\mathbf{e},\mathbf{v}) = 1, & \text{if } v_1 \le v_{1,\text{th}}(\mathbf{e}),\ v_2 \le v_{2,\text{th}}(\mathbf{e}), \\ \mu_1(\mathbf{e},\mathbf{v}) = 1, & \text{if } v_1 > v_{1,\text{th}}(\mathbf{e}),\ v_1 - v_{1,\text{th}}(\mathbf{e}) \ge v_2 - v_{2,\text{th}}(\mathbf{e}), \\ \mu_2(\mathbf{e},\mathbf{v}) = 1, & \text{if } v_2 > v_{2,\text{th}}(\mathbf{e}),\ v_2 - v_{2,\text{th}}(\mathbf{e}) > v_1 - v_{1,\text{th}}(\mathbf{e}). \end{cases} \tag{2.9}$$

Therefore, in the next, we consider only policies with a threshold structure.

We define the *marginal transmission probability* of the EHD i as:

$$\eta_i(\mathbf{e}) \triangleq \mathbb{E}_{\boldsymbol{V}}[\mu_i(\mathbf{e},\mathbf{v})] = \mathbb{P}(Q_{i,k} = 1|\boldsymbol{E}_k = \mathbf{e}), \qquad i \in \{1,2\}. \tag{2.10}$$

$\eta_i(\mathbf{e})$ represents the probability that the device i is the only transmitter given the global energy status. Note that $\eta_1(\mathbf{e}) + \eta_2(\mathbf{e}) \le 1$, and in particular we define the probability that no EHD transmits as:

$$\eta_0(\mathbf{e}) \triangleq 1 - \eta_1(\mathbf{e}) - \eta_2(\mathbf{e}). \tag{2.11}$$

An important point is that, due to the threshold structure of the policy, there is a one-to-one mapping between μ, $v_{i,\text{th}}$ and η_i. Therefore, even if the devices make choices based upon μ, in the next analysis we deal with $\eta_1,\ \eta_2$.

The *expected reward* can be defined as a function of the marginal transmission probabilities pair $\boldsymbol{\eta} \triangleq [\eta_1, \eta_2]$:

$$g(\boldsymbol{\eta}(\mathbf{e})) = \mathbb{E}[Q_{1,k}V_{1,k} + Q_{2,k}V_{2,k}|\boldsymbol{E}_k = \mathbf{e}]. \tag{2.12}$$

Both the policy μ and the transition probabilities of the time-homogeneous Markov Chain (MC) related to the energy states can be formulated as a function of $\boldsymbol{\eta}$. Consequently, if $\boldsymbol{\eta}$ induces an irreducible Markov Chain, (2.7) does not depend on the initial state $\mathbf{S}_0$, and the long-term average reward given a policy $\boldsymbol{\eta}$ can be rewritten as

$$G_{\boldsymbol{\eta}} = \sum_{e_1=0}^{e_{\max,1}} \sum_{e_2=0}^{e_{\max,2}} \pi_{\boldsymbol{\eta}}(\mathbf{e}) g(\boldsymbol{\eta}(\mathbf{e})), \tag{2.13}$$

where $\pi_{\boldsymbol{\eta}}(\mathbf{e})$ is the steady-state probability of being in the energy state $\mathbf{e}$, given a policy $\boldsymbol{\eta}$. Thus, the optimization problem (2.8) can be formulated as:

$$\boldsymbol{\eta}^* = \arg\max_{\boldsymbol{\eta}} \{G_{\boldsymbol{\eta}}\} \tag{2.14}$$

and can be solved via standard optimization techniques, like the Policy Iteration Algorithm (PIA) [30].

2.2.1 Maximization of the Transmission Rate

Until the meaning of $V_{i,k}$ is not specified, $G_{\boldsymbol{\eta}}$ is only a mathematical expression. A practical important case to study is when $V_{i,k}$ is the *achievable rate* by the EHD i in the slot k. With this assumption, the function $G_{\boldsymbol{\eta}}$ becomes the *long-term average transmission rate* from the two EHDs to a Receiver (RX) and the problem is to maximize this quantity.

Through Shannon's formula, the transmission rate can be related to the normalized channel gains $H_{1,k}$ and $H_{2,k}$ and to the average Signal to Noise Ratio (SNR) of the link i Λ_i as $V_{i,k} \propto \ln(1 + \Lambda_i H_{i,k})$. In particular we assume

$$V_{i,k} = \ln(1 + \Lambda_i H_{i,k}). \tag{2.15}$$

We also suppose that $H_{1,k}$ and $H_{2,k}$ are i.i.d. across EHDs and over time and that $H_{i,k}$ has pdf $f_{H_i}(h_i) = e^{-h_i},\ h_i > 0$, that is an exponential distribution with unit mean. The total SNR enjoyed by the device i in slot k is $\Lambda_i H_{i,k}$.

Since we are dealing with threshold policies, we are interested in the threshold value of $V_{i,k}$, that corresponds to a threshold value of $H_{i,k}$:

$$v_{i,\text{th}} = \ln(1 + \Lambda_i h_{i,\text{th}}) \Leftrightarrow h_{i,\text{th}} = \frac{e^{v_{i,\text{th}}} - 1}{\Lambda_i}. \tag{2.16}$$

As in [17], in the next, we consider $\Lambda_i \ll 1$ (low SNR regime [10]), therefore the expression (2.15) becomes:

$$V_{i,k} \approx \Lambda_i H_{i,k} \tag{2.17}$$

and the marginal transmission probabilities can be written as (see [17]):

$$\eta_0(\mathbf{e}) = \left(1 - e^{-h_{1,\text{th}}}\right)\left(1 - e^{-h_{2,\text{th}}}\right), \tag{2.18}$$

$$\eta_1(\mathbf{e}) = e^{-h_{1,\text{th}}}\left(1 - \frac{\Lambda_2}{\Lambda_1 + \Lambda_2} e^{-h_{2,\text{th}}}\right), \tag{2.19}$$

$$\eta_2(\mathbf{e}) = e^{-h_{2,\text{th}}}\left(1 - \frac{\Lambda_1}{\Lambda_1 + \Lambda_2} e^{-h_{1,\text{th}}}\right). \tag{2.20}$$

The reward function (2.15) is (we neglect the dependence on the joint energy state $\mathbf{e}$ for notational convenience):

$$g(\boldsymbol{\eta}) = \sum_{i=1}^{2} \Lambda_i e^{-h_{i,\text{th}}}(h_{i,\text{th}}+1) - \frac{\Lambda_1\Lambda_2}{\Lambda_1+\Lambda_2} e^{-(h_{1,\text{th}}+h_{2,\text{th}})}(h_{1,\text{th}}+h_{2,\text{th}}+1) \tag{2.21}$$

In addition, it is possible to derive the channel thresholds $h_{1,\text{th}}$ and $h_{2,\text{th}}$ as functions of the marginal probabilities η_1 and η_2 by performing the inversion of (2.19) and (2.20):

$$\begin{aligned} h_{1,\text{th}} &= \ln\left(\frac{-(\eta_1+\eta_2)\Lambda_2 + (\Lambda_1+\Lambda_2)(\eta_1+1)+\Delta}{2(\Lambda_1+\Lambda_2)\eta_1}\right), \\ h_{2,\text{th}} &= \ln\left(\frac{2\Lambda_2}{(\eta_1+\eta_2)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1)-\Delta}\right), \end{aligned} \tag{2.22}$$

where

$$\Delta = \sqrt{[(\eta_1+\eta_2)\Lambda_2 + (\Lambda_1+\Lambda_2)(1-\eta_1)]^2 - 4\Lambda_2\eta_2(\Lambda_1+\Lambda_2)}.$$

2.2.2 Definition of the Analyzed Policies

The Optimal Policy (OP) dictates the transmission of data with importance above a given threshold, which is a function of the global system energy status $\mathbf{e}$. In the next, using the structure of the OP, we derive several suboptimal policies which are shown to perform close to the optimal one. These are low-complexity policies, that, differently from the OP, do not require any computationally demanding optimization processes to compute the transmission probabilities. Furthermore, the transmission probabilities do not require a precise knowledge of the energy stored in the sensor batteries at any given time and they can be evaluated with a simple adaptation to the ambient energy supply.

- Optimal Policy (OP): the optimal policy that, for each $\bar{\mathbf{b}}$ and $\mathbf{e}_{\max}$, identifies the values of $\boldsymbol{\eta}(\mathbf{e})$ maximizing (2.7). This optimization problem requires to determine approximately $2\cdot(e_{\max,1}+1)(e_{\max,2}+1)$ variables. Since a rough optimization would not be computationally efficient, the PIA can be used. However, when the batteries sizes are not small, this process is still costly from a computational point of view.
- Energy Independent Policy (EIP): a policy in which

$$\eta_i(\mathbf{e}) = \eta_i \chi\{e_i > 0\}, \qquad i \in \{1,2\},\ \eta_i \in [0,1], \tag{2.23}$$

$$\eta_1 + \eta_2 \leq 1, \tag{2.24}$$

where χ is the indicator function. With this policy, the values of $\eta_1(\mathbf{e})$ and $\eta_2(\mathbf{e})$ do not depend on the batteries status, provided that they are not empty. This is interesting because, generally, the devices do not know precisely the state of charge of the batteries (see [31] and [32]).
In particular, we are interested in the Optimal-EIP (OEIP), that maximizes the long-term average reward $G_{\boldsymbol{\eta}}$ under the constraints (2.23) and (2.24). Note that an optimization problem should be solved also in this case, but now the variables are only two, independently from the maximum batteries capability.
Special cases of EIP include:

– Balanced Policy (BP): a particular case of EIP where

$$\eta_i = \bar{b}_i, \qquad i \in \{1, 2\} \tag{2.25}$$

and, obviously, it is defined only for $\bar{b}_1 + \bar{b}_2 \leq 1$, otherwise the constraint (2.24) would not be satisfied. It can be shown that this policy is asymptotically optimal for large batteries (see [17]). Note that no optimization is required to find $\boldsymbol{\eta}$. BP, in each nonzero energy state, "balances" the EHD operations because it matches the energy consumption rate to the energy harvesting rate.
An approximation of the performance obtained with the BP will be discussed in Section 2.3.1.

– Constrained Energy Independent Policy (CEIP): a particular case of EIP where

$$\eta_1 + \eta_2 = 1, \tag{2.26}$$

$$\min\{\eta_1, \eta_2\} \leq \min\{\bar{b}_1, \bar{b}_2\}. \tag{2.27}$$

Even if this policy can be defined for every value of $\bar{\mathbf{b}}$, we focus on the case $\bar{b}_1 + \bar{b}_2 > 1$, because in the other case we have already defined the BP. In particular, we are interested in the Optimal-CEIP (OCEIP), that maximizes the long-term average reward $G_{\boldsymbol{\eta}}$ under the constraints (2.26) and (2.27). Note that even in this case an optimization process is required, but the variable to optimize is only one (η_1 or η_2) due to the constraint (2.26).

  ∗ Heuristic Constrained Energy Independent Policy (HCEIP): a particular case of CEIP. We defined this policy only when $e_{\max} \triangleq e_{\max,1} = e_{\max,2}$. If we suppose $\bar{b}_1 \leq \bar{b}_2$, we use the following marginal transmission probabilities

$$\eta_1 = \begin{cases} \bar{b}_1 & \text{if } \bar{b}_1 \leq \Psi_1 \\ \frac{0.5-\Psi_1}{\Psi_2-\Psi_1}(\bar{b}_1 - \Psi_1) + \Psi_1 & \text{if } \Psi_1 < \bar{b}_1 < \Psi_2 \\ 0.5 & \text{if } \bar{b}_1 \geq \Psi_2 \end{cases}, \qquad \eta_2 = 1 - \eta_1, \tag{2.28}$$

  where Ψ_1 and Ψ_2 are two thresholds that depend upon $\bar{b}_2$ and $e_{\max}$. The case $\bar{b}_2 \leq \bar{b}_1$ is obtained by symmetry. Note that, for fixed $e_{\max}$ and $\bar{\mathbf{b}}$, the probabilities are immediately determined, without any optimization. Indeed, the HCEIP was designed in order to achieve the performance of the OCEIP without a high computational cost. We discuss the HCEIP in Section 2.4.1.

In Table 2.1 we compare the previous policies, according to the value of $\bar{b}_1 + \bar{b}_2$.

Table 2.1: Available policies for different values of $\bar{b}_1 + \bar{b}_2$

EH condition	Available policies
$\bar{b}_1 + \bar{b}_2 \leq 1$	OP, EIP, BP
$\bar{b}_1 + \bar{b}_2 > 1$	OP, EIP, CEIP, HCEIP

In the next we define G_{P} as the long-term average reward $G_{\boldsymbol{\eta}}$ under a policy P. Similarly, $\eta_{1,\mathrm{P}}$, $\eta_{2,\mathrm{P}}$ and $\pi_P(\mathbf{e})$ are the marginal transmission and the steady-state probabilities under a policy P.
The following inequalities chains hold:

$$\begin{cases} G_{\mathrm{OP}} \geq G_{\mathrm{OEIP}} \geq G_{\mathrm{BP}} & \text{when } \bar{b}_1 + \bar{b}_2 \leq 1 \\ G_{\mathrm{OP}} \geq G_{\mathrm{OEIP}} \geq G_{\mathrm{OCEIP}} \geq G_{\mathrm{HCEIP}} & \text{when } \bar{b}_1 + \bar{b}_2 > 1 \end{cases} \tag{2.29}$$

and in particular we compare the performance of two policies according to the following metric.

Definition 3 (reward precision)**.** *The reward precision of two policies A and B is defined as:*

$$\mathcal{R}_{\mathrm{A}}^{\mathrm{B}} \triangleq \frac{G_{\mathrm{A}} - G_{\mathrm{B}}}{G_{\mathrm{A}}} \tag{2.30}$$

Note that:

- if $\mathcal{R}_{\mathrm{A}}^{\mathrm{B}} \geq 0$, then policy A has better performance than policy B;
- if $\mathcal{R}_{\mathrm{A}}^{\mathrm{B}}$ is close to zero, than the two policies have similar performance.

Thus, policy B is considered a good lower bound for policy A if the two previous conditions hold, *i.e.*, $\mathcal{R}_{\mathrm{A}}^{\mathrm{B}} \gtrsim 0$.

2.3 Balanced Policy

The *Balanced Policy (BP)* is a threshold policy where, on average, each device transmits in all non-zero energy levels with probability b_i, *i.e.*,

$$\eta_{i,\mathrm{BP}}(\mathbf{e}) = \bar{b}_i \chi\{e_i > 0\}, \qquad i \in \{1, 2\}. \tag{2.31}$$

Note that, in each non-zero energy state, the BP matches the energy consumption rate to the energy harvesting rate, thus balancing the EHD operation.

This policy can be considered as a special case of EIP when $\bar{b}_1 + \bar{b}_2 \leq 1$.

Except in some particular cases, an analytical formulation of G_{BP} cannot be computed. Therefore, in the next, we derive its approximation $\hat{G}_{\mathrm{BP}}$, so as to characterize the performance obtained by BP in a closed-form expression.[3] Furthermore, we will numerically show that:

$$\mathcal{R}_{\mathrm{BP}}^{\widehat{\mathrm{BP}}} \gtrsim 0, \tag{2.32}$$

i.e., the $\widehat{\mathrm{BP}}$ is a good lower bound of the BP. Since we find an analytical expression of $G_{\widehat{\mathrm{BP}}}$ and $G_{\mathrm{OP}} \geq G_{\mathrm{BP}} \geq G_{\widehat{\mathrm{BP}}}$, it is possible to find an analytical lower bound to the optimum reward G_{OP}.

The basic idea, in order to reduce the problem complexity, is to divide the set of energy states in four classes and force the steady-state probability of all states in the same class to be equal. With this hypothesis, we solve a reduced system of steady-state equations and we find explicitly the value of the *approximate steady-state distribution*, from which the *approximate reward function* $\hat{G}_{\mathrm{BP}}$ can be computed.

[3] In the next, $\widehat{\mathrm{BP}}$ refers to the approximate performance of the BP.

2.3.1 Computation of the Approximate Reward Function

To find the exact expression of G_{BP}, the steady-state probabilities should be found. This can be done solving a system of $(e_{\max,1}+1)(e_{\max,2}+1)$ equations.

In order to simplify the problem, we introduce the approximate steady-state distribution $\hat{\pi}_{\text{BP}}(e_1, e_2)$ that can be found according to the following working assumption:

$$\hat{\pi}_{\text{BP}}(e_1, e_2) = \begin{cases} \pi_0, & \text{if } e_1 = 0,\ e_2 = 0, \\ \pi_1, & \text{if } e_1 > 0,\ e_2 = 0, \\ \pi_2, & \text{if } e_1 = 0,\ e_2 > 0, \\ \pi_{12}, & \text{if } e_1 > 0,\ e_2 > 0 \end{cases} \tag{2.33}$$

and we find the previous values solving the reduced system of equations involving states $(0,1)$, $(1,0)$ and $(1,1)$ and the normalization equation $\sum_{e_1=0}^{e_{\max,1}} \sum_{e_2=0}^{e_{\max,2}} \hat{\pi}_{\text{BP}}(e_1, e_2) = 1$. In Figure 2.1 we represent the considered transmission probabilities with continuous lines.

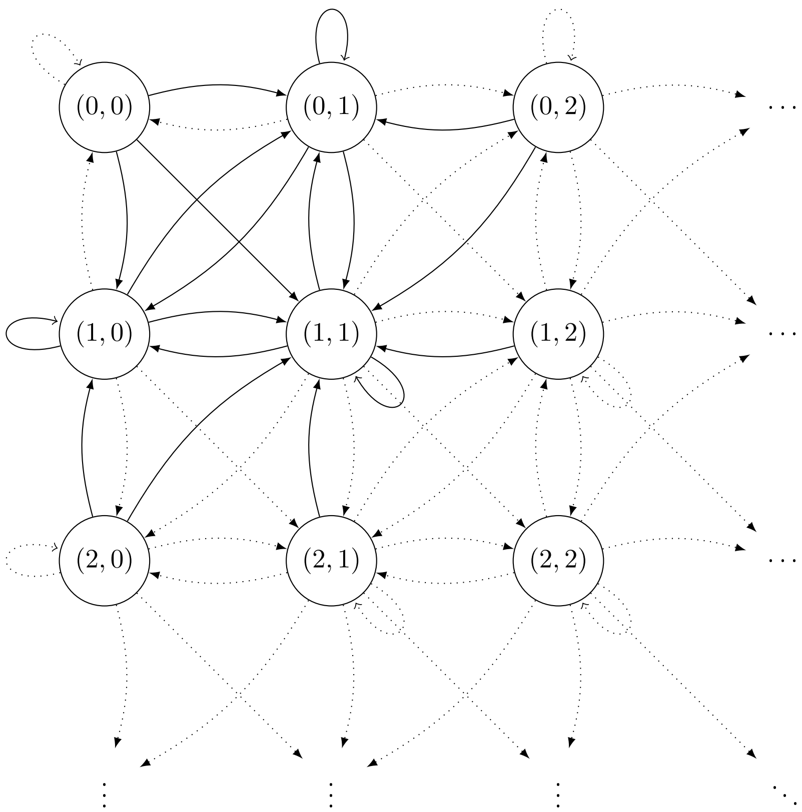


Figure 2.1: Sub-graph of the MC of the system. Continuous lines represent the transitions considered for the computation of the approximate balanced policy.

In particular, we solve the system of the four following equations:

$$\begin{aligned}\hat{\pi}_{\mathrm{BP}}(0,1) =&\hat{\pi}_{\mathrm{BP}}(0,0)p_{(0,0)\to(0,1)} + \hat{\pi}_{\mathrm{BP}}(0,1)p_{(0,1)\to(0,1)} + \hat{\pi}_{\mathrm{BP}}(0,2)p_{(0,2)\to(0,1)}+\\ &+\hat{\pi}_{\mathrm{BP}}(1,0)p_{(1,0)\to(0,1)} + \hat{\pi}_{\mathrm{BP}}(1,1)p_{(1,1)\to(0,1)},\end{aligned} \tag{2.34}$$

$$\begin{aligned}\hat{\pi}_{\mathrm{BP}}(1,0) =&\hat{\pi}_{\mathrm{BP}}(0,0)p_{(0,0)\to(1,0)} + \hat{\pi}_{\mathrm{BP}}(0,1)p_{(0,1)\to(1,0)}+\\ &+\hat{\pi}_{\mathrm{BP}}(1,0)p_{(1,0)\to(1,0)} + \hat{\pi}_{\mathrm{BP}}(1,1)p_{(1,1)\to(1,0)}+\\ &+\hat{\pi}_{\mathrm{BP}}(2,0)p_{(2,0)\to(1,0)},\end{aligned} \tag{2.35}$$

$$\begin{aligned}\hat{\pi}_{\mathrm{BP}}(1,1) =&\hat{\pi}_{\mathrm{BP}}(0,0)p_{(0,0)\to(1,1)} + \hat{\pi}_{\mathrm{BP}}(0,1)p_{(0,1)\to(1,1)} + \hat{\pi}_{\mathrm{BP}}(0,2)p_{(0,2)\to(1,1)}+\\ &+\hat{\pi}_{\mathrm{BP}}(1,0)p_{(1,0)\to(1,1)} + \hat{\pi}_{\mathrm{BP}}(1,1)p_{(1,1)\to(1,1)} + \hat{\pi}_{\mathrm{BP}}(1,2)p_{(1,2)\to(1,1)}+\\ &+\hat{\pi}_{\mathrm{BP}}(2,0)p_{(2,0)\to(1,1)} + \hat{\pi}_{\mathrm{BP}}(2,1)p_{(2,1)\to(1,1)},\end{aligned} \tag{2.36}$$

$$\sum_{e_1=0}^{e_{\max,1}} \sum_{e_2=0}^{e_{\max,2}} \hat{\pi}_{\mathrm{BP}}(e_1,e_2) = 1. \tag{2.37}$$

It can be easily proved that the solution is:

$$\begin{pmatrix}\pi_0\\ \pi_1\\ \pi_2\\ \pi_{12}\end{pmatrix} = \frac{1}{D}\begin{pmatrix}1-\bar{b}_1-\bar{b}_2+2\bar{b}_1\bar{b}_2\\ 1-\bar{b}_2\\ 1-\bar{b}_1\\ 1\end{pmatrix}, \tag{2.38}$$

where

$$D = (e_{\max,1}+1)(e_{\max,2}+1) - \bar{b}_1(e_{\max,2}+1) - \bar{b}_2(e_{\max,1}+1) + 2\bar{b}_1\bar{b}_2. \tag{2.39}$$

The *approximate reward function* $\hat{G}_{\mathrm{BP}}$ is thus defined as:

$$\hat{G}_{\mathrm{BP}} = e_{\max,1}\pi_1 g_1 + e_{\max,2}\pi_2 g_2 + e_{\max,1}e_{\max,2}\pi_{12}g_{12} \tag{2.40}$$

where we have defined $g_1 \triangleq g(\bar{b}_1,0)$, $g_2 \triangleq g(0,\bar{b}_2)$ and $g_{12} \triangleq g(\bar{b}_1,\bar{b}_2)$.[4]

2.3.2 Approximation Accuracy

We introduce the vectors $\hat{\boldsymbol{\pi}}^{\mathrm{tot}}$ and $\boldsymbol{\pi}^{\mathrm{tot}}$ as follow:

$$\hat{\boldsymbol{\pi}}^{\mathrm{tot}} = \begin{pmatrix}\pi_0\\ \hat{\pi}_1^{\mathrm{tot}}\\ \hat{\pi}_2^{\mathrm{tot}}\\ \hat{\pi}_{12}^{\mathrm{tot}}\end{pmatrix} \triangleq \begin{pmatrix}\pi_0\\ e_{\max,1}\pi_1\\ e_{\max,2}\pi_2\\ e_{\max,1}e_{\max,2}\pi_{12}\end{pmatrix}, \tag{2.41}$$

[4]We always assume that $g(\eta_1,\eta_2)$ is an increasing function of η_1 and η_2, therefore $g_i \leq g_{12},\ i=1,2$

$$\boldsymbol{\pi}^{\text{tot}} = \begin{pmatrix} \pi(0,0) \\ \pi_1^{\text{tot}} \\ \pi_2^{\text{tot}} \\ \pi_{12}^{\text{tot}} \end{pmatrix} \triangleq \begin{pmatrix} \pi(0,0) \\ \sum_{e_1=1}^{e_{\max,1}} \pi_{\text{BP}}(e_1, 0) \\ \sum_{e_2=1}^{e_{\max,2}} \pi_{\text{BP}}(0, e_2) \\ \sum_{e_1=1}^{e_{\max,1}} \sum_{e_2=1}^{e_{\max,2}} \pi_{\text{BP}}(e_1, e_2) \end{pmatrix}. \tag{2.42}$$

Note that the long-term expected reward and its approximation can be written as:

$$G_{\text{BP}} = \pi_1^{\text{tot}} g_1 + \pi_{12}^{\text{tot}} g_2 + \pi_{12}^{\text{tot}} g_{12}, \tag{2.43}$$

$$\hat{G}_{\text{BP}} = \hat{\pi}_1^{\text{tot}} g_1 + \hat{\pi}_{12}^{\text{tot}} g_2 + \hat{\pi}_{12}^{\text{tot}} g_{12}, \tag{2.44}$$

therefore the approximation $\hat{G}_{\text{BP}}$ of G_{BP} is accurate if $\hat{\boldsymbol{\pi}}^{\text{tot}}$ is close to $\boldsymbol{\pi}^{\text{tot}}$. Note that $\hat{\pi}_{\text{BP}}(e_1, e_2) \approx \pi_{\text{BP}}(e_1, e_2)$ for all e_1 and e_2 is a sufficient but not necessary condition to achieve $\hat{\boldsymbol{\pi}}^{\text{tot}} \approx \boldsymbol{\pi}^{\text{tot}}$.

We introduce two parameters to evaluate the performance of the approximation:

- *quadratic distance* $\mathcal{D}$: we compute $\mathcal{D}$ as:

$$\mathcal{D} \triangleq \|\hat{\boldsymbol{\pi}}^{\text{tot}} - \boldsymbol{\pi}^{\text{tot}}\|_2^2 = (\pi_0 - \pi(0,0))^2 + \sum_{i \in \{e_1, e_2, e_1 e_2\}} (\hat{\pi}_i^{\text{tot}} - \pi_i^{\text{tot}})^2. \tag{2.45}$$

  The approximation is as good as $\mathcal{D}$ is close to zero.

- *reward precision* $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}}$: according to the definition (2.30), $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}}$ is equal to:

$$\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}} \triangleq \frac{G_{\text{BP}} - \hat{G}_{\text{BP}}}{G_{\text{BP}}}. \tag{2.46}$$

  The approximation is as good as $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}}$ is close to zero. Furthermore, if $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}} > 0$, $\hat{\text{BP}}$ is a lower bound for BP.

In Section 2.6 we present the results of the numerical evaluation, showing that $\mathcal{D} \approx 0$ and $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}} \gtrsim 0$.

2.3.3 Properties

We discuss several properties of the function $\hat{G}_{\text{BP}}$ as a function of $e_{\max,i}$ and $\bar{b}_i$. The complete expression of $\hat{G}_{\text{BP}}$ is:

$$\hat{G}_{\text{BP}} = \frac{e_{\max,1}(1-\bar{b}_2)g_1 + e_{\max,2}(1-\bar{b}_1)g_2 + e_{\max,1}e_{\max,2}g_{12}}{(e_{\max,1}+1)(e_{\max,2}+1) - \bar{b}_1(e_{\max,2}+1) - \bar{b}_2(e_{\max,1}+1) + 2\bar{b}_1\bar{b}_2}. \tag{2.47}$$

Theorem 1. *$\hat{G}_{\text{BP}}$ is an increasing function of $e_{\max,1}$ or $e_{\max,2}$.*

Proof. Focusing on the first case, the goal is to prove that:

$$\hat{G}_{\text{BP}}(e_{\max,1}, e_{\max,2}) \leq \hat{G}_{\text{BP}}(e_{\max,1}+1, e_{\max,2}). \tag{2.48}$$

$e_{\max,2}$ can be considered a parameter, therefore the function $\hat{G}_{\text{BP}}$ can be rewritten as:

$$\hat{G}_{\text{BP}} = \frac{\alpha e_{\max,1} + \beta}{\gamma e_{\max,1} + \delta}, \tag{2.49}$$

where α, β, γ and δ do not depend upon $e_{\max,1}$. The previous expression is a hyperbola in $e_{\max,1}$, and it is an increasing function if the condition $\alpha\delta \geq \beta\gamma$ is satisfied:

$$\begin{aligned}
&\alpha\delta - \beta\gamma = \\
&\ ((1-\bar{b}_2)\cdot g_1 + e_{\max,2}\cdot g_{12})(e_{\max,2}+1-\bar{b}_1(e_{\max,2}+1)-\bar{b}_2+2\bar{b}_1\bar{b}_2)+ \\
&\ -(e_{\max,2}+1-\bar{b}_2)(e_{\max,2}\cdot(1-\bar{b}_1)\cdot g_2) \geq \\
&\ \geq ((1-\bar{b}_2)\cdot g_1 + e_{\max,2}\cdot g_{12})(e_{\max,2}+1-\bar{b}_1(e_{\max,2}+1)-\bar{b}_2+\bar{b}_1\bar{b}_2)+ \\
&\ -(e_{\max,2}+1-\bar{b}_2)e_{\max,2}\cdot(1-\bar{b}_1)\cdot g_2 = \\
&\ = (e_{\max,2}+1-\bar{b}_2)(1-\bar{b}_1)\Big((1-\bar{b}_2)\cdot g_1 + e_{\max,2}\cdot g_{12} - e_{\max,2}\cdot g_2\Big) \\
&\ = (e_{\max,2}+1-\bar{b}_2)(1-\bar{b}_1)\Big((1-\bar{b}_2)\cdot g_1 + e_{\max,2}\cdot(g_{12}-g_2)\Big) \geq 0,
\end{aligned} \tag{2.50}$$

where the last inequality is due to the hypothesis $g_{12} \geq g_2$.

Similarly it can be shown that $\hat{G}_{\mathrm{BP}}$ is an increasing function of $e_{\max,2}$. □

Theorem 2. *$\hat{G}_{\mathrm{BP}}$ is a convex downward function function of $e_{\max,1}$ or $e_{\max,2}$.*

Proof. It is obvious from the previous proof because of $\hat{G}_{\mathrm{BP}}$ is a hyperbola. □

Proposition 1. *When $e_{\max,1}$, $e_{\max,2}$ or both go to infinity, we obtain:*

$$\lim_{e_{\max,1}\to\infty} \hat{G}_{\mathrm{BP}} = \frac{(1-\bar{b}_2)\cdot g_1 + e_{\max,2}\cdot g_{12}}{e_{\max,2}+1-\bar{b}_2}, \tag{2.51}$$

$$\lim_{e_{\max,2}\to\infty} \hat{G}_{\mathrm{BP}} = \frac{(1-\bar{b}_1)\cdot g_2 + e_{\max,1}\cdot g_{12}}{e_{\max,1}+1-\bar{b}_1}, \tag{2.52}$$

$$\lim_{\substack{e_{\max,1}\to\infty \\ e_{\max,2}\to\infty}} \hat{G}_{\mathrm{BP}} = g_{12}. \tag{2.53}$$

From the previous equations, it is possible to derive the gap between the asymptotic reward (when both $e_{\max,i} \to \infty$) and the one achieved in a scenario where only a single device i has a large battery, as $g_{12} - \lim_{e_{\max,i}\to\infty} \hat{G}_{\mathrm{BP}}$. For example, if $e_{\max,1} \to \infty$, we have

$$g_{12} - \lim_{e_{\max,1}\to\infty} \hat{G}_{\mathrm{BP}} = (1-\bar{b}_2)\frac{g_{12}-g_1}{e_{\max,2}+1-\bar{b}_2}, \tag{2.54}$$

which allows us to justify the reward loss in the non symmetric case.

Proposition 2. *In Section 2.6, we show that the approximate reward function $\hat{G}_{\mathrm{BP}}$ numerically results to be a lower bound for the Balanced Policy. Furthermore, as $G_{\mathrm{BP}} \leq G_{\mathrm{OP}}$ for every parameter choice and, if $\bar{b}_1+\bar{b}_2 \leq 1$, $G_{\mathrm{OP}} \leq g_{12}$ (see [17]), the following inequality chain holds:*

$$\hat{G}_{\mathrm{BP}} \leq G_{\mathrm{BP}} \leq G_{\mathrm{OP}} \leq g_{12}, \tag{2.55}$$

i.e., *$\hat{G}_{\mathrm{BP}}$ can be used as an analytical lower bound for the optimal reward G_{OP}.*

2.4 Constrained Energy Independent Policy

In this Section we focus on the case $\bar{b}_1+\bar{b}_2 > 1$. With this constraint, the following policies can be defined: OP, EIP with its optimized version OEIP, CEIP with its optimized version OCEIP and, lastly, HCEIP.

Note that if $\bar{b}_1 + \bar{b}_2 > 1$ a BP where $\eta_1 = \bar{b}_1$ and $\eta_2 = \bar{b}_2$ cannot be defined, otherwise $\eta_1 + \eta_2 > 1$, which would be infeasible. Therefore, we must have:

$$\eta_1 + \eta_2 < \bar{b}_1 + \bar{b}_2 \tag{2.56}$$

and the choice of the marginal transmission probabilities is not obvious.

An EIP is a threshold policy where, on average, each device transmits in all non-zero energy levels with probability $\eta_{i,\text{EIP}}$, *i.e.*,

$$\eta_{i,\text{EIP}}(\mathbf{e}) = \eta_{i,\text{EIP}}\chi\{e_i > 0\}, \qquad i \in \{1, 2\} \tag{2.57}$$

and the CEIP is a particular case of EIP where we add two constraints:

$$\begin{aligned} &\eta_{i,\text{CEIP}}(\mathbf{e}) = \eta_{i,\text{CEIP}}\chi\{e_i > 0\}, \qquad i \in \{1, 2\} \\ &\eta_{1,\text{CEIP}} + \eta_{2,\text{CEIP}} = 1, \\ &\min\{\eta_{1,\text{CEIP}}, \eta_{2,\text{CEIP}}\} \leq \min\{\bar{b}_1, \bar{b}_2\}. \end{aligned} \tag{2.58}$$

The first one means that $\eta_{0,\text{CEIP}}$ is forced to be zero, *i.e.*, if it is possible, there is always a transmission. The second constraint will be useful to design the HCEIP described next.

Note that, with the proposed formulation, the CEIP can be defined for each pair $(\bar{b}_1, \bar{b}_2)$, but we focus on the case $\bar{b}_1+\bar{b}_2 > 1$, as, if $\bar{b}_1+\bar{b}_2 \leq 1$, the simpler Balanced Policy already achieves good performance.

It can be numerically shown (see example 2) that the Optimal-CEIP behaves differently from the Optimal-EIP, *i.e.*, $\eta_{i,\text{OCEIP}} \neq \eta_{i,\text{OEIP}}$. However, due to the first constraint $\eta_{1,\text{CEIP}} + \eta_{2,\text{CEIP}} = 1$, the CEIP has the peculiarity of allowing to reduce the number of the variables from two to one. Therefore, to compute the OCEIP, only one parameter needs to be optimized. Even if the optimization in one or two variables is not so much different from a computational point of view, we use this result to define the Heuristic Constrained Energy Independent Policy, whose marginal transmission probabilities are an approximation of the OCEIP's ones.

2.4.1 Heuristic Constrained Energy Independent Policy

The Heuristic Constrained Energy Independent Policy (HCEIP) is a particular case of CEIP. It is interesting because:

- for a given set of $\bar{b}_1$, $\bar{b}_2$, $e_{\max,1}$ and $e_{\max,2}$, it provides the values of $\eta_{1,\text{HCEIP}}$ and $\eta_{2,\text{HCEIP}}$ with no optimization needed;

- it achieves near optimal performance among the CEIPs, *i.e.*, $G_{\text{HCEIP}} \lesssim G_{\text{OCEIP}}$.

We consider only the case $e_{\max,1} = e_{\max,2} \triangleq e_{\max}$, leaving the asymmetric case as future work.

The design of the HCEIP is based on the analysis of the marginal transmission probabilities of the OCEIP. In particular we approximate $\eta_{1,\text{OCEIP}}$ and $\eta_{2,\text{OCEIP}}$ with $\eta_{1,\text{HCEIP}}$ and $\eta_{2,\text{HCEIP}}$ respectively, which are simple functions of the system parameters.

In the next we assume $\bar{b}_2 \leq \bar{b}_1$, because the other case can be found by symmetry. The behaviour of $\eta_{2,\text{OCEIP}}$ as a function of $\bar{b}_2$ can be divided in three regions:

1. a first linear zone with slope equal to one:

$$\eta_{2,\text{OCEIP}} = \bar{b}_2; \tag{2.59}$$

2. a second non-linear part:

$$\eta_{2,\text{OCEIP}} = f(\bar{b}_2); \tag{2.60}$$

3. a last constant zone:

$$\eta_{2,\text{OCEIP}} = 0.5, \tag{2.61}$$

where we recall that $\eta_{1,\text{OCEIP}} = 1 - \eta_{2,\text{OCEIP}}$.

We define the two thresholds that divide the three regions as Ψ_1 and Ψ_2 (with $\Psi_1 < \Psi_2$) (see Figure 2.2).

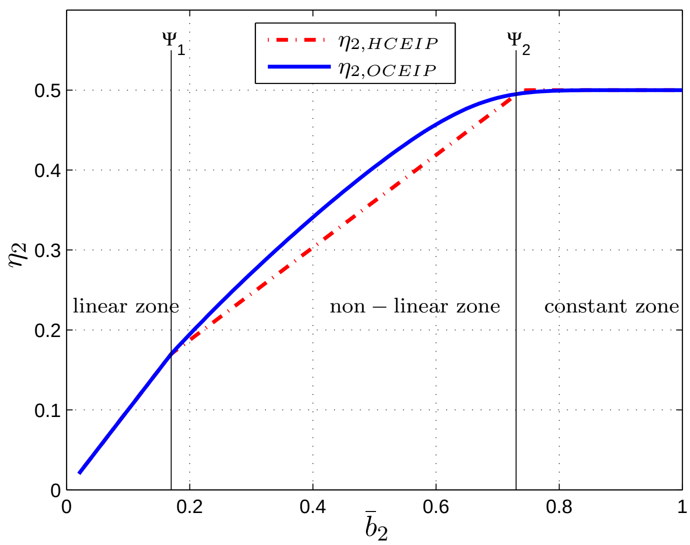


Figure 2.2: $\eta_{2,HCEIP}$ and $\eta_{2,OCEIP}$ as a function of $\bar{b}_2$ in the range $[0.02, 1]$ with $\bar{b}_1 = 0.98$ and $e_{\max} = 6$

Given Ψ_1 and Ψ_2 for each $\bar{b}_1$ and $e_{\max}$, we deduce $\eta_{2,\text{HCEIP}}$ as a function of these thresholds, approximating the central non-linear region as a straight line. With the previous assumptions, we have:

$$\eta_{2,\text{HCEIP}} = \begin{cases} \bar{b}_2, & \text{if } \bar{b}_2 \leq \Psi_1, \\ \frac{0.5-\Psi_1}{\Psi_2-\Psi_1}(\bar{b}_2 - \Psi_1) + \Psi_1, & \text{if } \Psi_1 < \bar{b}_2 \leq \Psi_2, \\ 0.5, & \text{if } \Psi_2 < \bar{b}_2. \end{cases} \tag{2.62}$$

Thresholds

In the general case, each threshold is a function of both $e_{\max}$ and $\bar{b}_1$. However, we numerically verified that:

$$\Psi_1 \triangleq \Psi_1(\bar{b}_1, e_{\max}) \approx \Psi_1(\bar{b}_1), \tag{2.63}$$

$$\Psi_2 \triangleq \Psi_2(\bar{b}_1, e_{\max}) \approx \min\{\psi_2(e_{\max}), \bar{b}_1\}, \tag{2.64}$$

that is, Ψ_1 only depends on $\bar{b}_1$, while Ψ_2 also depends on a function $\psi_2(e_{\max})$, defined in (2.66).

In Figure 2.2 we show, as an example, both $\eta_{2,\text{OCEIP}}$ and $\eta_{2,\text{HCEIP}}$ as a function of $\bar{b}_2$, with the thresholds Ψ_1 and Ψ_2 when $\bar{b}_1 = 0.98$ and $e_{\max} = 6$.

We now discuss how the derivation of Ψ_1 and Ψ_2 was accomplished. Ψ_1 was computed numerically, for different values of $e_{\max}$ and $\bar{b}_1$, as the value for which η_2 is no longer equal to $\bar{b}_2$. Since it can be observed that Ψ_1 is approximately independent of $e_{\max}$, for each value of $\bar{b}_1$ we dropped this dependence averaging Ψ_1 over all possible values of $e_{\max}$. Finally, by a linear interpolation with respect to $\bar{b}_1$, (2.65) was derived.

Ψ_2 was determined as the value for which η_2 saturates to 0.5. We found that Ψ_2 has the structure of equation (2.64), where ψ_2 was determined with a technique similar to the one used to find Ψ_1, first averaging on $\bar{b}_1$ and then interpolating in $e_{\max}$, resulting in (2.66).

The resulting heuristic expressions of the thresholds $\Psi_1(\bar{b}_1)$ and $\psi_2(e_{\max})$ are given by

$$\Psi_1(\bar{b}_1) = -0.6875\bar{b}_1 + 0.84375, \tag{2.65}$$

$$\psi_2(e_{\max}) = 0.5159e^{-0.1775 e_{\max}} + 0.5624 \tag{2.66}$$

and are valid for every possible choice of $e_{\max}$ and $\bar{b}_1$ and hence represent a general result for this scenario.

From (2.63-2.66), it is possible to compute Ψ_1 and Ψ_2, used to derive $\eta_{2,\text{HCEIP}}$ from (2.62) and $\eta_{1,\text{HCEIP}} = 1 - \eta_{2,\text{HCEIP}}$, for all values of $\bar{b}_1$, $\bar{b}_2$ and $e_{\max}$. Finally, with the marginal probabilities, G_{HCEIP} is obtained.

2.5 Examples

We now introduce two examples in the simple scenario where $e_{\max,1} = e_{\max,2} = 1$.

The transition probabilities of the Markov Chain are:

$$
\begin{aligned}
p_{(0,0)\to(k,l)} &= \begin{cases} (1-\bar{b}_1)(1-\bar{b}_2), & \text{if } k=0, l=0,\\ (1-\bar{b}_1)\bar{b}_2, & \text{if } k=0, l=1,\\ \bar{b}_1(1-\bar{b}_2), & \text{if } k=1, l=0,\\ \bar{b}_1\bar{b}_2, & \text{if } k=1, l=1; \end{cases}\\
p_{(0,1)\to(k,l)} &= \begin{cases} (1-\bar{b}_1)(1-\bar{b}_2)\eta_2(0,1), & \text{if } k=0, l=0,\\ (1-\bar{b}_1)(\eta_0(0,1)+\bar{b}_2\eta_2(0,1)), & \text{if } k=0, l=1,\\ \bar{b}_1(1-\bar{b}_2)\eta_2(0,1), & \text{if } k=1, l=0,\\ \bar{b}_1(\eta_0(0,1)+\bar{b}_2\eta_2(0,1)), & \text{if } k=1, l=1; \end{cases}\\
p_{(1,0)\to(k,l)} &= \begin{cases} (1-\bar{b}_1)(1-\bar{b}_2)\eta_1(1,0), & \text{if } k=0, l=0,\\ (1-\bar{b}_1)\bar{b}_2\eta_1(1,0), & \text{if } k=0, l=1,\\ (1-\bar{b}_2)(\eta_0(1,0)+\bar{b}_1\eta_1(1,0)), & \text{if } k=1, l=0,\\ \bar{b}_2(\eta_0(0,1)+\bar{b}_1\eta_1(1,0)), & \text{if } k=1, l=1; \end{cases}\\
p_{(1,1)\to(k,l)} &= \begin{cases} 0, & \text{if } k=0, l=0,\\ (1-\bar{b}_1)\eta_1(1,1), & \text{if } k=0, l=1,\\ (1-\bar{b}_2)\eta_2(1,1), & \text{if } k=1, l=0,\\ \bar{b}_1\eta_1(1,1)+\bar{b}_2\eta_2(1,1)+\eta_0(1,1), & \text{if } k=1, l=1. \end{cases}
\end{aligned}
\tag{2.67}
$$

For the examples we use the following notation:

$$[\eta_i] \triangleq \begin{pmatrix} \eta_i(0,0) & \eta_i(0,1) \\ \eta_i(1,0) & \eta_i(1,1) \end{pmatrix}, \tag{2.68}$$

$$[\pi] \triangleq \begin{pmatrix} \pi(0,0) & \pi(0,1) \\ \pi(1,0) & \pi(1,1) \end{pmatrix}. \tag{2.69}$$

Example 1. *we consider $\bar{b} \triangleq \bar{b}_1 = \bar{b}_2 = 0.1$ (this is the case $\bar{b}_1+\bar{b}_2 \leq 1$). Since the problem is symmetric, it follows that $\eta_1(e_1,e_2) = \eta_2(e_2,e_1)$. The EIP and BP can be defined as:*

$$[\eta_{1,\text{EIP}}] = \begin{pmatrix} 0 & 0 \\ \eta_{\text{EIP}} & \eta_{\text{EIP}} \end{pmatrix} = [\eta_{2,\text{EIP}}]^T, \tag{2.70}$$

$$[\eta_{1,\text{BP}}] = \begin{pmatrix} 0 & 0 \\ \bar{b} & \bar{b} \end{pmatrix} = [\eta_{2,\text{BP}}]^T. \tag{2.71}$$

It can be numerically found that:

$$[\eta_{1,\text{OP}}] = \begin{pmatrix} 0 & 0 \\ 0.1889 & 0.1758 \end{pmatrix} = [\eta_{2,\text{OP}}]^T, \tag{2.72}$$

$$[\eta_{1,\text{OEIP}}] = \begin{pmatrix} 0 & 0 \\ 0.1839 & 0.1839 \end{pmatrix} = [\eta_{2,\text{OEIP}}]^T. \tag{2.73}$$

and the steady-state probabilities are:

$$[\pi_{\mathrm{OP}}] = \begin{pmatrix} 0.3848 & 0.2389 \\ 0.2389 & 0.1374 \end{pmatrix}, \tag{2.74}$$

$$[\pi_{\mathrm{OEIP}}] = \begin{pmatrix} 0.3806 & 0.2427 \\ 0.2427 & 0.1339 \end{pmatrix}, \tag{2.75}$$

$$[\pi_{\mathrm{BP}}] = \begin{pmatrix} 0.2180 & 0.2557 \\ 0.2557 & 0.2706 \end{pmatrix}. \tag{2.76}$$

The long-term average rewards are:

$$G_{\mathrm{OP}} = 0.3707, \tag{2.77}$$

$$G_{\mathrm{OEIP}} = 0.3706, \tag{2.78}$$

$$G_{\mathrm{BP}} = 0.3462. \tag{2.79}$$

Note that, even if the transmission probabilities of the BP and OP are slightly different (almost a factor of two), $\mathcal{R}_{\mathrm{OP}}^{\mathrm{BP}} = 6.6\%$ *that is a small value.*

Example 2. *We consider* $\bar{b}_1 = 0.8$ *and* $\bar{b}_2 = 0.21$ *(this is one case with* $\bar{b}_1 + \bar{b}_2 > 1$*). The EIP and CEIP can be defined (according to equation* (2.58)*) as:*

$$[\eta_{1,\mathrm{EIP}}] = \begin{pmatrix} 0 & 0 \\ \eta_{1,\mathrm{EIP}} & \eta_{1,\mathrm{EIP}} \end{pmatrix}, \qquad [\eta_{2,\mathrm{EIP}}] = \begin{pmatrix} 0 & \eta_{2,\mathrm{EIP}} \\ 0 & \eta_{2,\mathrm{EIP}} \end{pmatrix}, \tag{2.80}$$

$$[\eta_{1,\mathrm{CEIP}}] = \begin{pmatrix} 0 & 0 \\ 1-\eta_{\mathrm{CEIP}} & 1-\eta_{\mathrm{CEIP}} \end{pmatrix}, \qquad [\eta_{2,\mathrm{CEIP}}] = \begin{pmatrix} 0 & \eta_{\mathrm{CEIP}} \\ 0 & \eta_{\mathrm{CEIP}} \end{pmatrix}, \tag{2.81}$$

with the constraint:

$$\eta_{\mathrm{CEIP}} \leq \bar{b}_2. \tag{2.82}$$

It can be numerically found that:

$$[\eta_{1,\mathrm{OP}}] = \begin{pmatrix} 0 & 0 \\ 0.8212 & 0.6768 \end{pmatrix}, \qquad [\eta_{2,\mathrm{OP}}] = \begin{pmatrix} 0 & 0.3830 \\ 0 & 0.2274 \end{pmatrix}, \tag{2.83}$$

$$\eta_{1,\mathrm{OEIP}} = 0.7293, \qquad \eta_{2,\mathrm{OEIP}} = 0.2340, \tag{2.84}$$

$$\eta_{\mathrm{OCEIP}} = 0.21. \tag{2.85}$$

Note that $\eta_{0,\mathrm{OEIP}} > 0$*, which proves that the condition* $\eta_{1,\mathrm{OEIP}} + \eta_{2,\mathrm{OEIP}} = 1$ *is not always satisfied by the OEIP.*

Furthermore, $\eta_{i,\mathrm{HCEIP}}$ *can be simply computed and it would be found* $\eta_{i,\mathrm{OCEIP}} = \eta_{i,\mathrm{HCEIP}}$*,* i.e.*, the approximation of the OCEIP is perfect in this case.*

The long-term average rewards are:

$$G_{\mathrm{OP}} = 1.0696, \tag{2.86}$$

$$G_{\mathrm{OEIP}} = 1.0629, \tag{2.87}$$

$$G_{\mathrm{OCEIP}} = G_{\mathrm{HCEIP}} = 1.0603. \tag{2.88}$$

it can be seen that OEIP and OCEIP are very close to the Optimal Policy.

2.6 Performance evaluation of the introduced policies

In our numerical evaluation we used the following parameters: $e_{\max,1}, e_{\max,2} \in \{1, \dots, 20\}$, $\bar{b}_1, \bar{b}_2 \in \{0.05, 0.10, \dots, 1\}$ and $\Lambda_1 = \Lambda_2$. Since there are twenty cases for each parameter, we have a total number of cases equal to $20^4 = 160000$.

In the following, we show that the reward functions $\hat{G}_{\mathrm{BP}}$ and G_{HCEIP} are good lower bounds for G_{OP} in most cases of interest.

Table 2.2: Policies comparison in the worst-case scenarios.

EH condition	Policies		$\max\{\mathcal{R}_{\mathrm{A}}^{\mathrm{B}}\}$
$\bar{b}_1 + \bar{b}_2 \gtreqless 1$	OP	OEIP	5.36%
$\bar{b}_1 + \bar{b}_2 \leq 1$	OEIP	BP	10.17%
	OP	BP	10.18%
	BP	$\hat{\mathrm{BP}}$	4.14%
$\bar{b}_1 + \bar{b}_2 > 1$	OEIP	OCEIP	1.59%
	OCEIP	HCEIP	0.45%
	OP	HCEIP	4.56%

We now comment the results shown in Table 2.2, based on the value of $\bar{b}_1 + \bar{b}_2$.

2.6.1 $\bar{b}_1 + \bar{b}_2 \lesseqgtr 1$

OP and OEIP comparison

As $\max\{\mathcal{R}_{\mathrm{OP}}^{\mathrm{OEIP}}\} = 5.36\%$, OEIP is a good lower bound for OP.

2.6.2 $\bar{b}_1 + \bar{b}_2 \leq 1$

OEIP and BP comparison

BP is not always a good lower bound for OEIP, as $\max\{\mathcal{R}_{\mathrm{OEIP}}^{\mathrm{BP}}\} = 10.17\%$. However, in general, BP is rather close to OEIP, especially if $e_{\max,1}$ and/or $e_{\max,2}$ are not too small. In particular, we noticed that $\mathcal{R}_{\mathrm{OEIP}}^{\mathrm{BP}}$ is high when both $\bar{b}_1$ and $\bar{b}_2$ are close to 0, which is not a very practical scenario. For example, if $e_{\max,1} = e_{\max,2} = 1$ and $\bar{b}_1 = 0.4$, $\bar{b}_2 = 0.2$, $\mathcal{R}_{\mathrm{OEIP}}^{\mathrm{BP}} = 1.2\%$, but if $e_{\max,1} = e_{\max,2} = 5$, this value decreases to 0.14%. We can therefore say that BP is a good lower bound for OEIP in most cases of interest.

OP and BP comparison

From Sections 2.6.1 and 2.6.2, we can state that in most cases of interest BP is a good lower bound for OP. As an example, in Figure 2.3, we plot the worst case scenario ($\bar{b}_1 = \bar{b}_2 = 0.05$) for $\mathcal{R}_{\mathrm{OP}}^{\mathrm{BP}}$, with different values of $e_{\max,i}$. It can be seen that $\mathcal{R}_{\mathrm{OP}}^{\mathrm{BP}}$ decreases quickly when $e_{\max,1}$ and/or $e_{\max,2}$ increase, *i.e.*, the Balanced Policy is better for high values of $e_{\max}$, as expected.

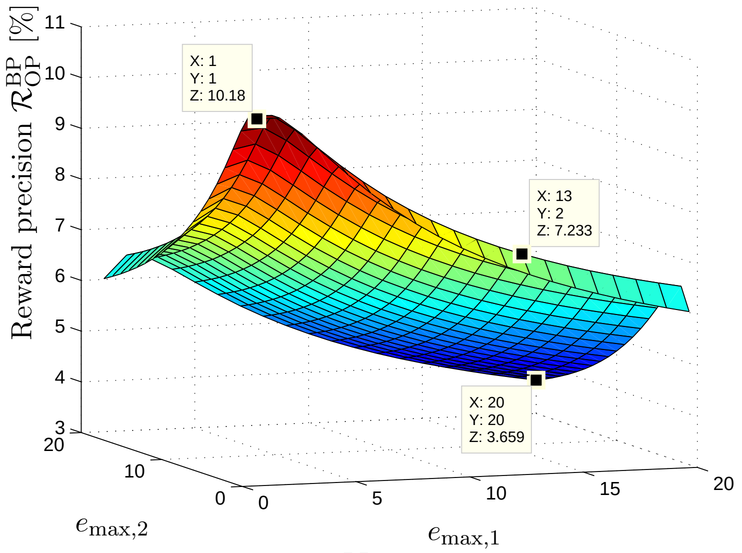


Figure 2.3: Percentage reward precision $\mathcal{R}_{\text{OP}}^{\text{BP}}$ when $\bar{b}_1 = \bar{b}_2 = 0.05$ (worst case) and $e_{\max,1}$ and $e_{\max,2}$ range between 1 and 20

$\hat{\text{BP}}$ and BP comparison

For each pair of $e_{\max,1}$ and $e_{\max,2}$, the maximum of $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}}$ is reached when $\bar{b}_1 = \bar{b}_2 = 0.5$ and, in particular, in the worst possible case $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}}$ is equal to 4.14%. Moreover, the lower bound of $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}}$ is 0, *i.e.*, $\hat{G}_{\text{BP}}$ is always lower than G_{BP}. Since $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}} \gtrsim 0$, we can state that the approximate balanced policy can be considered as a good lower bound for the BP. Note that this result is not obvious, as $\hat{G}_{\text{BP}}$ has been derived as an approximation of G_{BP}. In Figure 2.4, we depict the reward precision $\mathcal{R}_{\text{BP}}^{\hat{\text{BP}}}$ for $e_{\max,1} = e_{\max,2} = 2$, for different values of $\bar{b}_1$ and $\bar{b}_2$. It can be seen that $\bar{b}_1 = \bar{b}_2 = 0.5$ is the worst case, but for low values of $\bar{b}_1$ or $\bar{b}_2$, the two rewards are comparable.

For this comparison we introduced also the quadratic distance $\mathcal{D}$ (equation (2.45)). In Figure 2.5 we represent this quantity as a function of $e_{\max,1} = e_{\max,2}$. The results comply with the previous ones, *i.e.*, the approximation becomes better when the battery sizes increase and when the average harvesting rates decrease.

2.6.3 $\bar{b}_1 + \bar{b}_2 > 1$

OEIP and OCEIP comparison

The difference between OCEIP and OEIP is mainly due to the constraint $\eta_1 + \eta_2 = 1$. However, we verified that $\eta_{1,\text{OEIP}} + \eta_{2,\text{OEIP}} \approx 1$, even for the optimal unconstrained EIP. The results show that $\mathcal{R}_{\text{OEIP}}^{\text{OCEIP}}$ is always lower than 1.59%, and the worst case occurs when $\bar{b}_1 + \bar{b}_2 \gtrsim 1$. In fact, in this case, OEIP is more conservative than OCEIP, *i.e.*, $\eta_{1,\text{OEIP}} < \eta_{1,\text{OCEIP}}$ and $\eta_{2,\text{OEIP}} < \eta_{2,\text{OCEIP}}$, because OEIP generally attempts to avoid energy outage, whose probability increases if $\bar{b}_1$ and $\bar{b}_2$ decrease.

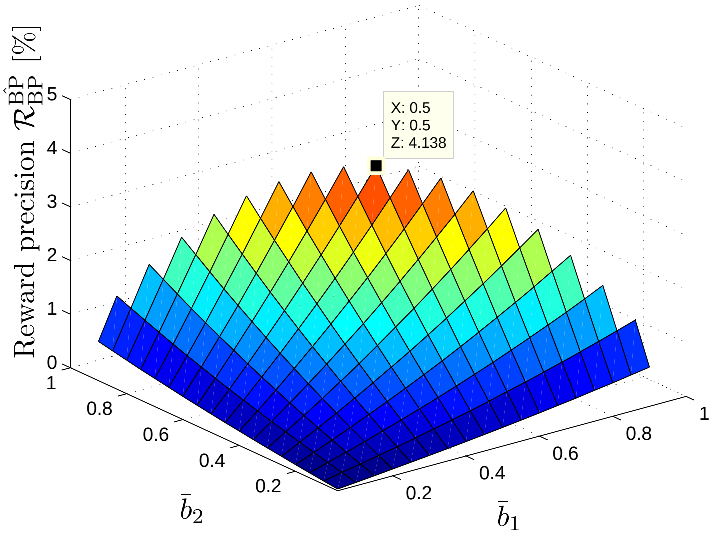


Figure 2.4: Percentage reward precision $\mathcal{R}_{\rm BP}^{\hat{\rm BP}}$ when $e_{\max,1} = e_{\max,2} = 2$ and $\bar{b}_1$ and $\bar{b}_2$ range between 0.05 and 1 with $\bar{b}_1 + \bar{b}_2 \leq 1$

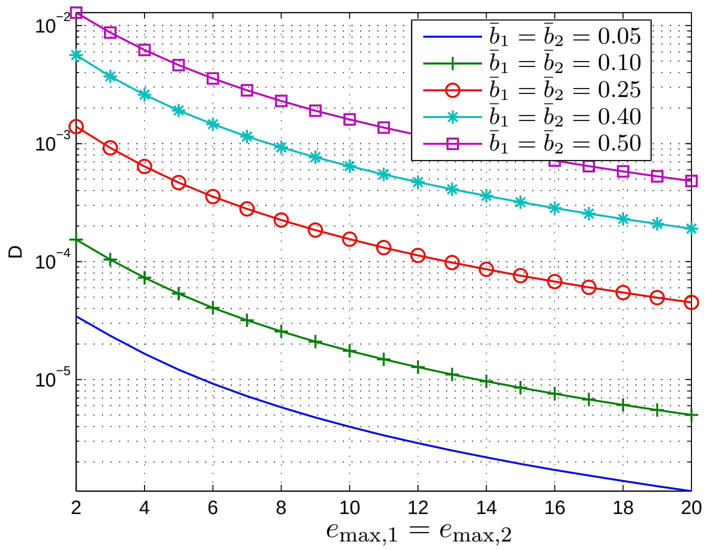


Figure 2.5: Quadratic distance $\mathcal{D}$ for several values of $\bar{b}_1$ and $\bar{b}_2$ when $e_{\max,1} = e_{\max,2} \in \{2, \ldots, 20\}$

OCEIP and HCEIP comparison

We verified that the quantity $\mathcal{R}_{\rm OCEIP}^{\rm HCEIP}$ has a maximum that is less than 0.45%, *i.e.*, with the considered parameters, the heuristic approximation can be considered as a good lower bound for OCEIP. Clearly, the approximation performs worse when $\bar{b}_2 \in (\Psi_1, \Psi_2)$, where we fitted a non-linear function with a straight line. In Figure 2.6, we show the reward precision $\mathcal{R}_{\rm OCEIP}^{\rm HCEIP}$ as a function of $\bar{b}_2$ for the same parameters of Figure 2.2. It can be seen that $\mathcal{R}_{\text{OCEIP}}^{\text{HCEIP}}$ is low even in the non-linear zone, where the approximation is worse.

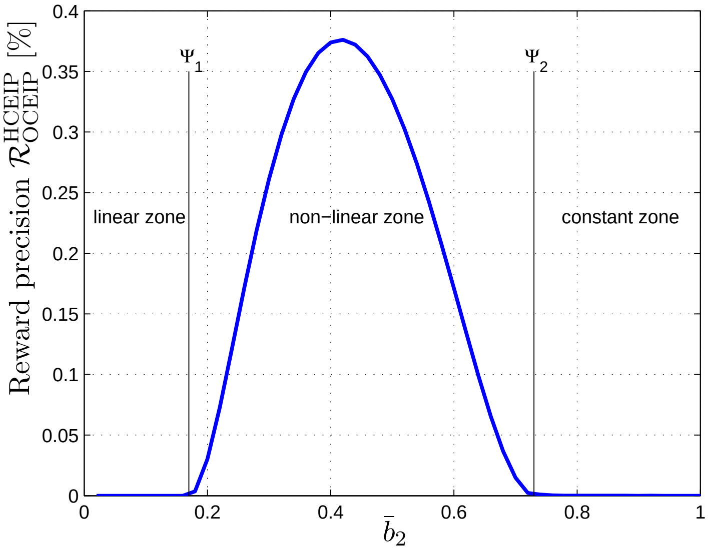


Figure 2.6: Percentage reward precision $\mathcal{R}_{\text{OCEIP}}^{\text{HCEIP}}$ as a function of $\bar{b}_2$ that range in $[0.02, 1]$ with $\bar{b}_1 = 0.98$ and $e_{\text{max}} = 6$

OP and HCEIP comparison

By the aforementioned results, it follows that

$$G_{\text{OP}} \gtrsim G_{\text{OEIP}} \gtrsim G_{\text{OCEIP}} \gtrsim G_{\text{HCEIP}}, \tag{2.89}$$

i.e., HCEIP is a good lower bound for OP. In particular, from Table 2.2, $\max\{\mathcal{R}_{\text{OP}}^{\text{HCEIP}}\} = 4.56\%$: this is an interesting result, as HCEIP, differently from OP, can be analytically formulated with a closed form expression.

Finally, in Figures 2.7 and 2.8 we compare G_{OP}, G_{OEIP}, $\hat{G}_{\text{BP}}$ and G_{HCEIP}, when $e_{\text{max},1} = e_{\text{max},2} \in \{1, \ldots, 20\}$. It can be seen that all policies approach OP for high values of $e_{\text{max},i}$ and are very close already for $e_{\text{max}} = 20$.

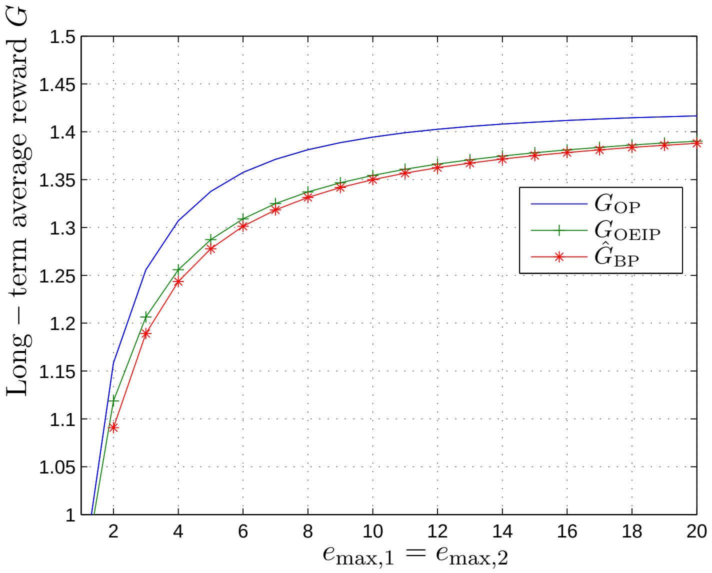


Figure 2.7: Comparison of different reward functions G when $e_{\max,1} = e_{\max,2} \in \{1, \ldots, 20\}$ and $\bar{b}_1 = \bar{b}_2 = 0.4$

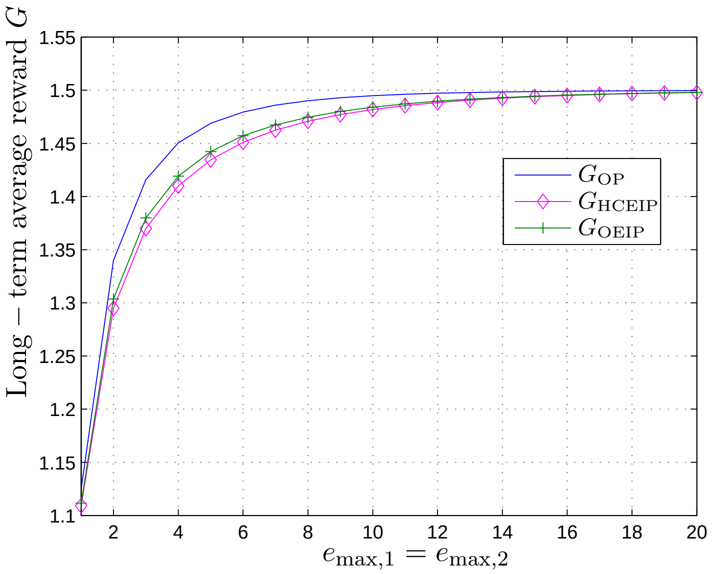


Figure 2.8: Comparison of different reward functions G when $e_{\max,1} = e_{\max,2} \in \{1, \ldots, 20\}$ and $\bar{b}_1 = \bar{b}_2 = 0.55$

2.7 Conclusions

In the first part of this work we studied the case of two Energy Harvesting Devices (EHDs) which report data of different importance to a common Receiver (RX) and are managed by a Central Controller (CC). We distinguished two cases, depending on the global energy harvesting rate of the system. In the first case, when $\bar{b}_1 + \bar{b}_2 \leq 1$, we used the Balanced Policy (BP) and derived an analytic approximation of its performance. We showed that BP can be considered as a good lower bound for the Optimal Policy (OP) when $\bar{b}_1$ and $\bar{b}_2$ are not too close to zero. Furthermore, we numerically derived an approximate reward function of BP, that is always lower than or equal to the real reward function. In this way, we found an analytical lower bound to BP and, consequently, to OP. Future work will complete the analysis of BP in the particular case $e_{\max,1} = 1$ or $e_{\max,2} = 1$. In the case $\bar{b}_1 + \bar{b}_2 > 1$, we introduced a chain of policies that leads to the Heuristic Constrained Energy Independent Policy (HCEIP), a policy computable with no optimization which nevertheless achieves good performance with respect to OP. We leave as future work the case $e_{\max,1} \neq e_{\max,2}$.

Chapter 3

EH Transmitter and EH Receiver pair

In this part we study a network composed of an Energy Harvesting Transmitter (EHTX) and an Energy Harvesting Receiver (EHRX). EHTX can receive packets from other nodes in the WSN or can generate a packet after sensing. EHTX transmits packets to EHRX, that may or may not receive the data, depending on its current energy level and on the packet importance.

3.1 System model

The model is similar to the one presented in Chapter 2, therefore we emphasize only the differences.

The transmitter sends packets to the receiver. Time is slotted and slot k is the time interval $[k, k+1), k \in \mathbb{N}_0$. In slot k, the node can transmit over the interval $[k, k+\delta_k)$, where $\delta_k \in (0, 1]$ is the duty cycle.

We model the energy storage capability of both transmitter and receiver as a buffer and each position in the buffer can hold one energy quantum. The transmission and the reception of one packet require one energy quantum each, and the maximum number of quanta that can be stored in the batteries is $e_{\max,t}$ and $e_{\max,r}$ for EHTX and EHRX, respectively. We assume that EHRX can consume an energy quantum also for listening to the channel, *i.e.*, EHRX may consume energy even if there is no transmission.

At time $k+1$, the energies in the buffers are:

$$E_{t,k+1} = \min\{(E_{t,k} - Q_{t,k})^+ + B_{t,k}, e_{\max,t}\}, \tag{3.1}$$
$$E_{r,k+1} = \min\{(E_{r,k} - Q_{r,k})^+ + B_{r,k}, e_{\max,r}\}, \tag{3.2}$$

where $\{B_{t,k}\}$ and $\{B_{r,k}\}$, defined as in (2.1), are the energy arrival processes (i.i.d. Bernoulli) with mean $\bar{b}_t$ and $\bar{b}_r$, $\{Q_{t,k}\}$ is the transmitter action process (1 if there is a transmission, 0 otherwise) and $\{Q_{r,k}\}$ is the reception action process (1 if EHRX receives or listens to the channel, 0 otherwise).

In the next we consider several scenarios:

- *Centralized case (C)*: EHTX and EHRX are coordinated by a Central Controller (CC). In this case the following implications hold:

$$Q_{t,k} = 1 \Leftrightarrow Q_{r,k} = 1, \tag{3.3}$$

  i.e., EHTX transmits a packet if and only if EHRX can receive it.

- *Semi-Independent case*: EHTX and EHRX are not coordinated but we assume:

$$Q_{t,k} = 0 \Rightarrow Q_{r,k} = 0, \tag{3.4}$$

  because we suppose that if there is no transmission, the receiver does not consume energy.

  However, in this case, the following event may happen:

$$\{Q_{t,k} = 1 \cap Q_{r,k} = 0\}. \tag{3.5}$$

  The devices can have different degrees of knowledge about the overall energy status and in particular we can divide the following cases:

  - *global knowledge*: EHTX knows $E_{r,k}$ and EHRX knows $E_{t,k}$.
  - *partial knowledge*: EHTX knows only if EHRX is in outage and vice-versa. In this case the following implications hold:

$$\begin{aligned} E_{t,k} = 0 \Rightarrow Q_{r,k} = 0, \\ E_{r,k} = 0 \Rightarrow Q_{t,k} = 0. \end{aligned} \tag{3.6}$$

  - *local knowledge*: EHTX does not have any information about $E_{r,k}$ and vice-versa.

- *Totally-Independent case*: EHTX and EHRX are not coordinated. In this case, the following events may happen:

$$\begin{aligned} \{Q_{t,k} = 1 \cap Q_{r,k} = 0\}, \\ \{Q_{t,k} = 0 \cap Q_{r,k} = 1\}. \end{aligned} \tag{3.7}$$

  Also here the devices can have different degrees of knowledge about the overall energy status and we can divide the cases as before.

We also define $V_k \in \mathbb{R}^+$ as the current packet importance. V_k has probability density function (pdf) $f_V(v),\ v \geq 0$. In all cases, we suppose that both EHTX and EHRX know the packet importance, otherwise the problem should be defined in another way.
The *state of the system* in slot k is defined as

$$\mathbf{S}_k = (E_{t,k}, E_{r,k}, V_k). \tag{3.8}$$

3.2 Policy Definition and General Optimization Problem

Given $\mathbf{S}_k$, a policy determines $(Q_{t,k}, Q_{r,k}) \in \{(0,0),(0,1),(1,0),(1,1)\}$ in slot k.

$\mu((i,j);\mathbf{e},v)$ is the probability of drawing i and j energy quanta from EHTX and EHRX respectively, given that $\mathbf{E}_k = \mathbf{e}$ and $V_k = v$.
The *long-term average reward* under a policy μ is defined as:

$$G(\mu, \mathbf{S}_0) = \lim_{K\to\infty} \inf \frac{1}{K}\mathbb{E}\left[\sum_{k=0}^{K-1} Q_{t,k}Q_{r,k}V_k \middle| S_0\right]. \tag{3.9}$$

Note that, differently from the case of Chapter 2, in the previous formula we have the product of $Q_{t,k}$ and $Q_{r,k}$, because the reward increases only if EHTX transmits and EHRX receives at the same time.
The optimization problem is to find the μ^* such that:

$$\mu^* = \arg\max_{\mu}\{G(\mu, S_0)\}. \tag{3.10}$$

Theorem 3. *The optimal policy has the following structure:*

$$\begin{cases} \mu((1,1);\mathbf{e},v) = 1, & \text{if } v > v_{\text{th}}(\mathbf{e}), \\ \mu((0,0);\mathbf{e},v) = 1, & \text{if } v \le v_{\text{th}}(\mathbf{e}). \end{cases} \tag{3.11}$$

Proof. See Appendix A.2. □

However, in the next we consider also the following sub-optimal policy:

$$\begin{cases} \mu((1,1);\mathbf{e},v) = 1, & \text{if } v > v_{t,\text{th}}(\mathbf{e}) \cap v > v_{r,\text{th}}(\mathbf{e}), \\ \mu((1,0);\mathbf{e},v) = 1, & \text{if } v > v_{t,\text{th}}(\mathbf{e}) \cap v \le v_{r,\text{th}}(\mathbf{e}), \\ \mu((0,1);\mathbf{e},v) = 1, & \text{if } v \le v_{t,\text{th}}(\mathbf{e}) \cap v > v_{r,\text{th}}(\mathbf{e}), \\ \mu((0,0);\mathbf{e},v) = 1, & \text{if } v \le v_{t,\text{th}}(\mathbf{e}) \cap v \le v_{r,\text{th}}(\mathbf{e}), \end{cases} \tag{3.12}$$

i.e., we introduce a threshold for EHTX and another threshold for EHRX. In general, $v_{t,\text{th}}(\mathbf{e})$ and $v_{r,\text{th}}(\mathbf{e})$ can be different, especially in the cases where the devices do not have global information about the system. However, if $v_{t,\text{th}}(\mathbf{e}) = v_{r,\text{th}}(\mathbf{e})$, then the sub-optimal policy degenerates in the optimal one.
We define the *marginal probabilities* as:

$$\eta_{ij}(\mathbf{e}) \triangleq \mathbb{E}_V[\mu((i,j);\mathbf{e},v)] = \mathbb{P}(Q_{t,k} = i, Q_{r,k} = j | E_{t,k} = e_t \cap E_{r,k} = e_r), \qquad i,j \in \{0,1\} \tag{3.13}$$

and we indicate as $\boldsymbol{\eta}(\mathbf{e})$ the vector $[\eta_{00}(\mathbf{e}), \eta_{01}(\mathbf{e}), \eta_{10}(\mathbf{e}), \eta_{11}(\mathbf{e})]$. Furthermore, we introduce the *marginal transmission and reception probabilities* as:

$$\eta_t(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{10}(\mathbf{e}) + \eta_{11}(\mathbf{e}), \tag{3.14}$$

$$\eta_r(\mathbf{e}) \triangleq \mathbb{P}(Q_{r,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{01}(\mathbf{e}) + \eta_{11}(\mathbf{e}), \tag{3.15}$$

respectively. Similarly to [16], the thresholds can be found as:

$$\eta_t(\mathbf{e}) = \int_{v_{t,\mathrm{th}}(\mathbf{e})}^{\infty} f_V(\nu)\mathrm{d}\nu, \tag{3.16}$$

$$\eta_r(\mathbf{e}) = \int_{v_{r,\mathrm{th}}(\mathbf{e})}^{\infty} f_V(\nu)\mathrm{d}\nu \tag{3.17}$$

and, from equations (3.12) and (3.13), we obtain:

$$\eta_{00}(\mathbf{e}) = \int_{0}^{\min\{v_{t,\mathrm{th}}(\mathbf{e}), v_{r,\mathrm{th}}(\mathbf{e})\}} f_V(\nu)\mathrm{d}\nu, \tag{3.18}$$

$$\eta_{10}(\mathbf{e}) = \begin{cases} \int_{v_{t,\mathrm{th}}(\mathbf{e})}^{v_{r,\mathrm{th}}(\mathbf{e})} f_V(\nu)\mathrm{d}\nu, & \text{if } v_{r,\mathrm{th}}(\mathbf{e}) > v_{t,\mathrm{th}}(\mathbf{e}), \\ 0, & \text{otherwise}, \end{cases} \tag{3.19}$$

$$\eta_{01}(\mathbf{e}) = \begin{cases} \int_{v_{r,\mathrm{th}}(\mathbf{e})}^{v_{t,\mathrm{th}}(\mathbf{e})} f_V(\nu)\mathrm{d}\nu, & \text{if } v_{t,\mathrm{th}}(\mathbf{e}) > v_{r,\mathrm{th}}(\mathbf{e}), \\ 0, & \text{otherwise}, \end{cases} \tag{3.20}$$

$$\eta_{11}(\mathbf{e}) = \int_{\max\{v_{t,\mathrm{th}}(\mathbf{e}), v_{r,\mathrm{th}}(\mathbf{e})\}}^{\infty} f_V(\nu)\mathrm{d}\nu. \tag{3.21}$$

The *expected reward* is:

$$g(\boldsymbol{\eta}(\mathbf{e})) \triangleq \mathbb{E}_V[Q_{t,k}Q_{r,k}V_k | E_{t,k} = e_t \cap E_{r,k} = e_r] = \int_{v_{\mathrm{th}}(\mathbf{e})}^{\infty} v f_V(v)\mathrm{d}v, \tag{3.22}$$

where we find the threshold v_{th} as:

$$\eta_{11}(\mathbf{e}) = \int_{v_{\mathrm{th}}(\mathbf{e})}^{\infty} f_V(\nu)\mathrm{d}\nu, \tag{3.23}$$

$$v_{\mathrm{th}}(\mathbf{e}) \triangleq \max\{v_{t,\mathrm{th}}(\mathbf{e}), v_{r,\mathrm{th}}(\mathbf{e})\}. \tag{3.24}$$

Note that the reward is non zero only if both EHTX and EHRX consume an energy quantum.

As in Chapter 2, the long-term average reward given a policy $\boldsymbol{\eta}$ can be rewritten as:

$$G(\boldsymbol{\eta}(\mathbf{e})) = \sum_{e_t=0}^{e_{\max,t}} \sum_{e_r=0}^{e_{\max,r}} \pi(\mathbf{e}) g(\boldsymbol{\eta}(\mathbf{e})), \tag{3.25}$$

where $\pi(\mathbf{e})$ is the steady-state probability of being in state $\mathbf{e}$.

The transmission probabilities of the Markov Chain for the generic state $\mathbf{e}$ with $e_t \in$

$\{1, \ldots, e_{\max,t} - 1\}$ and $e_r \in \{1, \ldots, e_{\max,r} - 1\}$ are:

$$p_{(e_t,e_r)\to(k,l)} = \begin{cases} (1-\bar{b}_t)(1-\bar{b}_r)\eta_{11}(\mathbf{e}), & \text{if } k = e_t - 1, l = e_r - 1, \\ (1-\bar{b}_t)\bar{b}_r\eta_{11}(\mathbf{e}) + (1-\bar{b}_t)(1-\bar{b}_r)\eta_{10}(\mathbf{e}), & \text{if } k = e_t - 1, l = e_r, \\ (1-\bar{b}_t)\bar{b}_r\eta_{10}(\mathbf{e}), & \text{if } k = e_t - 1, l = e_r + 1, \\ \bar{b}_t(1-\bar{b}_r)\eta_{11}(\mathbf{e}) + (1-\bar{b}_t)(1-\bar{b}_r)\eta_{01}(\mathbf{e}), & \text{if } \quad k = e_t, l = e_r - 1, \\ \bar{b}_t(\bar{b}_r\eta_{11}(\mathbf{e}) + (1-\bar{b}_r)\eta_{10}(\mathbf{e})) + & \\ \ +(1-\bar{b}_t)(\bar{b}_r\eta_{01}(\mathbf{e}) + (1-\bar{b}_r)\eta_{00}(\mathbf{e})), & \text{if } k = e_t, l = e_r, \\ (1-\bar{b}_t)\bar{b}_r\eta_{00}(\mathbf{e}) + \bar{b}_t\bar{b}_r\eta_{10}(\mathbf{e}), & \text{if } k = e_t, l = e_r + 1, \\ \bar{b}_t(1-\bar{b}_r)\eta_{01}(\mathbf{e}), & \text{if } k = e_t + 1, l = e_r - 1, \\ \bar{b}_t(1-\bar{b}_r)\eta_{00}(\mathbf{e}) + \bar{b}_t\bar{b}r\eta_{01}(\mathbf{e}), & \text{if } k = e_t + 1, l = e_r, \\ \bar{b}_r\bar{b}_t\eta_{00}(\mathbf{e}), & \text{if } k = e_t + 1, l = e_r + 1, \\ 0, & \text{otherwise} \end{cases} \tag{3.26}$$

and if $e_t \in \{0, e_{\max,t}\}$ and/or $e_r \in \{0, e_{\max,r}\}$ the previous probabilities must be appropriately corrected.

3.2.1 Maximization of the transmission rate

Similarly to equation (2.15), we assume:

$$V_k = \ln(1 + \Lambda H_k). \tag{3.27}$$

Λ is the Signal to Noise Ratio (SNR) and the channel gain H has pdf $f_H(h) = e^{-h},\ h > 0$. We find:

$$h_{\text{th}}(\mathbf{e}) = \ln\left(\frac{1}{\eta_{11}(\mathbf{e})}\right), \tag{3.28}$$

$$g(\boldsymbol{\eta}(\mathbf{e})) = \int_{h_{\text{th}}(\mathbf{e})}^{\infty} \ln(1 + \text{SNR}h)e^{-h}\text{d}h, \tag{3.29}$$

where h_{th} is the channel threshold associated to the importance threshold $v_{\text{th}}(\mathbf{e})$.

For a fixed policy $\boldsymbol{\eta}$, using (3.26) it is possible to find the steady state probabilities $\pi_\eta(\mathbf{e})$, therefore, with (3.29), the long-term average reward G of equation (3.25) can be found.

3.3 Centralized Case

In this case we suppose that the transmitter and the receiver are coordinated by a Central Controller, therefore EHTX transmits if only if EHRX receives the packet. This is an upper bound of the achievable performance in every scenario because energy is never wasted.

We recall that

$$Q_{t,k} = 1 \Leftrightarrow Q_{r,k} = 1. \tag{3.30}$$

The choice of the thresholds that satisfies the previous conditions is:

$$v_{t,\text{th}}(\mathbf{e}) = v_{r,\text{th}}(\mathbf{e}), \tag{3.31}$$

indeed, the events $\{Q_{t,k} = 1 \cap Q_{r,k} = 0\}$ and $\{Q_{t,k} = 0 \cap Q_{r,k} = 1\}$ cannot happen.

In terms of marginal transmission probabilities, we obtain:

$$\eta_{10}(\mathbf{e}) = \eta_{01}(\mathbf{e}) = 0, \tag{3.32}$$

$$\Rightarrow \eta_{11}(\mathbf{e}) = 1 - \eta_{00}(\mathbf{e}). \tag{3.33}$$

Note that, thanks to equation (3.32), in order to maximize (3.25), only $\eta_{11}(\mathbf{e})$ or $\eta_{00}(\mathbf{e})$ need to be optimized and this can be done via the Policy Iteration Algorithm (PIA).

From equation (3.31) and from the way G was defined (eq. (3.25)), in this case the reward function is a symmetric function of $e_{\max,t}$ and $e_{\max,r}$ if $\bar{b}_t = \bar{b}_r$. This can be seen, as an example, in Figure 3.1, where we represent the optimal reward in the Centralized case as a function of $e_{\max,t}$ and $e_{\max,r}$ when $\bar{b}_t = \bar{b}_r = 0.5$.

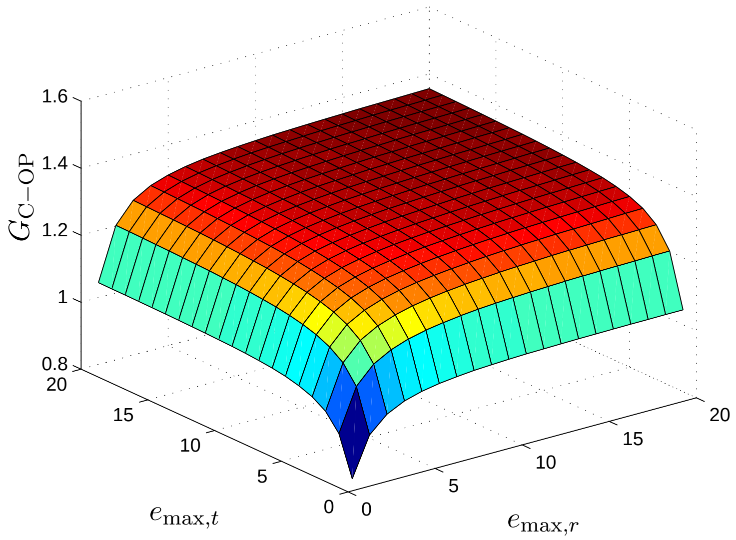


Figure 3.1: Reward function $G_{\text{C-OP}}$ when $e_{\max,t}, e_{\max,r} \in \{1, \ldots, 20\}$ and $\bar{b}_t = 0.5$, $\bar{b}_r = 0.5$

3.4 Semi-Independent case

In this scenario we suppose that if EHTX is not transmitting, then EHRX will not consume energy to listen to the channel, *i.e.*,:

$$Q_{t,k} = 0 \Rightarrow Q_{r,k} = 0. \tag{3.34}$$

The previous condition can be translated in

$$v_{r,\text{th}}(\mathbf{e}) \geq v_{t,\text{th}}(\mathbf{e}). \tag{3.35}$$

In terms of marginal transmission probabilities, we obtain:

$$\eta_{01}(\mathbf{e}) = 0. \tag{3.36}$$

Note that the transmitter can have different information about the receiver energy level and vice-versa, therefore in the next we consider several cases.

3.4.1 Global knowledge

We suppose that, in each slot, the transmitter knows the energy level of the receiver $E_{r,k}$ and vice-versa.

The optimal thresholds choice is $v_{t,\mathrm{th}}(\mathbf{e}) = v_{r,\mathrm{th}}(\mathbf{e})$. Indeed, with these thresholds, this case has the same performance as the centralized one, because the transmission probabilities are deterministic functions of $\mathbf{e}$ and v.

3.4.2 Partial knowledge

We suppose that, in each slot, EHTX knows if EHRX is in outage and vice-versa, therefore we assume:

$$\begin{aligned} E_{t,k} = 0 \Rightarrow Q_{r,k} = 0, \\ E_{r,k} = 0 \Rightarrow Q_{t,k} = 0. \end{aligned} \tag{3.37}$$

Note that the first implication is already expressed in equation (3.34) because $E_{t,k} = 0 \Rightarrow Q_{t,k} = 0$.

The marginal transmission probabilities can be written as:

$$\eta_t(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{10}(\mathbf{e}) + \eta_{11}(\mathbf{e}) = \begin{cases} \rho_t(e_t), & \text{if } e_r > 0, \\ 0, & \text{if } e_r = 0, \end{cases} \tag{3.38}$$

$$\eta_r(\mathbf{e}) \triangleq \mathbb{P}(Q_{r,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{11}(\mathbf{e}) = \begin{cases} \rho_r(e_r), & \text{if } e_t > 0, \\ 0, & \text{if } e_t = 0, \end{cases} \tag{3.39}$$

with:

$$\rho_t(e_t) \triangleq \mathbb{P}(Q_{t,k} = 1 | E_{t,k} = e_t \cap E_{r,k} > 0), \tag{3.40}$$

$$\rho_r(e_r) \triangleq \mathbb{P}(Q_{r,k} = 1 | E_{t,k} > 0 \cap E_{r,k} = e_r), \tag{3.41}$$

therefore the optimization process requires to optimize both $\rho_t(e_t)$ and $\rho_r(e_t)$ under the constraint that can be derived from equations (3.38-3.39):

$$\max_{e_r \in \{1,\dots,e_{\max,r}\}} \{\rho_r(e_r)\} < \min_{e_t \in \{1,\dots,e_{\max,t}\}} \{\rho_t(e_t)\}. \tag{3.42}$$

The previous constraint can be written in a linear form:

$$\rho_r(e_r) - \rho_t(e_t) \leq 0, \qquad \forall e_t \in \{1,\dots,e_{\max,t}\},\ \forall e_r \in \{1,\dots,e_{\max,r}\}, \tag{3.43}$$

that corresponds to a system of $e_{\max,t} \cdot e_{\max,r}$ inequalities.

Unfortunately, maximizing (3.25) considering (3.38-3.39) is not a convex problem and it can be numerically verified that there are a lot of local maxima. Since trying to solve the problem with several starting points is computational demanding, we make the following assumption:

$$\rho_t(e_t) = \begin{cases} \rho_t, & \text{if } e_t > 0, \\ 0, & \text{if } e_t = 0, \end{cases}, \qquad \rho_t \in [0,1], \tag{3.44}$$

$$\rho_r(e_r) = \begin{cases} \rho_r, & \text{if } e_r > 0, \\ 0, & \text{if } e_r = 0, \end{cases}, \qquad \rho_r \in [0,1], \tag{3.45}$$

i.e., we consider an Energy Independent Policy (EIP). With these probabilities, the maximization problem is simplified because only ρ_t and ρ_r need to be optimized.

Considering (3.44-3.45) is interesting because we show in Section 3.6 that this policy has good performance compared to the optimal ones.

3.4.3 Local knowledge

We suppose that EHTX does not have any information about the energy of EHRX and vice-versa. Similarly to the previous case, the marginal transmission probabilities can be written as:

$$\eta_t(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{10}(\mathbf{e}) + \eta_{11}(\mathbf{e}) = \rho_t(e_t), \tag{3.46}$$

$$\eta_r(\mathbf{e}) \triangleq \mathbb{P}(Q_{r,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{11}(\mathbf{e}) = \begin{cases} \rho_r(e_r), & \text{if } e_t > 0, \\ 0, & \text{if } e_t = 0, \end{cases} \tag{3.47}$$

with:

$$\rho_t(e_t) \triangleq \mathbb{P}(Q_{t,k} = 1 | E_{t,k} = e_t), \tag{3.48}$$

$$\rho_r(e_r) \triangleq \mathbb{P}(Q_{r,k} = 1 | E_{t,k} > 0 \cap E_{r,k} = e_r). \tag{3.49}$$

Note that $\eta_r(0, e_r) = 0$ because $\eta_t(0, e_r) = 0$ and we suppose that if there is no transmission, then EHRX does not consume energy. Formally: $E_{t,k} = 0 \Rightarrow Q_{t,k} = 0 \Rightarrow Q_{r,k} = 0$.

As in the previous Section, equations (3.46) and (3.47) imply the constraint (3.42).

Also in this case the maximization problem is not convex, therefore we consider again an EIP, defined as in equations (3.44) and (3.45).

3.5 Totally independent case

In this section we do not assume any hypothesis for the marginal transmission probabilities, therefore, in general, $\eta_{00}(\mathbf{e})$, $\eta_{10}(\mathbf{e})$, $\eta_{01}(\mathbf{e})$ and $\eta_{11}(\mathbf{e})$ are all greater than zero. This is the case where the devices are completely independent. As before, in general, the device may have a different amount of information about the other Energy Harvesting Device (EHD), therefore, in the next, we analyse these cases.

For this scenario we only set up the model, leaving further considerations as future work.

3.5.1 Global knowledge

As in Section 3.4.1, choosing $v_{t,\text{th}}(\mathbf{e}) = v_{r,\text{th}}(\mathbf{e})$ provides the same performance of the centralized case.

3.5.2 Partial knowledge

We suppose that, in each slot, the EHTX knows if the EHRX is in outage and vice-versa, therefore we assume:

$$\begin{aligned} E_{t,k} = 0 \Rightarrow Q_{r,k} = 0, \\ E_{r,k} = 0 \Rightarrow Q_{t,k} = 0. \end{aligned} \tag{3.50}$$

The marginal transmission probabilities can be written as:

$$\eta_t(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{10}(\mathbf{e}) + \eta_{11}(\mathbf{e}) = \begin{cases} \rho_t(e_t), & \text{if } e_r > 0, \\ 0, & \text{if } e_r = 0, \end{cases} \tag{3.51}$$

$$\eta_r(\mathbf{e}) \triangleq \mathbb{P}(Q_{r,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{01}(\mathbf{e}) + \eta_{11}(\mathbf{e}) = \begin{cases} \rho_r(e_r), & \text{if } e_t > 0, \\ 0, & \text{if } e_t = 0, \end{cases} \tag{3.52}$$

with $\rho_t(e_t)$ and $\rho_r(e_r)$ defined as in (3.40-3.41). Note that, since we consider the independent case, we obtain:

$$\eta_{11}(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 1 \cap Q_{r,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_t(\mathbf{e})\eta_r(\mathbf{e}), \tag{3.53}$$

$$\eta_{10}(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 1 \cap Q_{r,k} = 0 | \mathbf{E}_k = \mathbf{e}) = \eta_t(\mathbf{e})(1 - \eta_r(\mathbf{e})), \tag{3.54}$$

$$\eta_{01}(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 0 \cap Q_{r,k} = 1 | \mathbf{E}_k = \mathbf{e}) = (1 - \eta_t(\mathbf{e}))\eta_r(\mathbf{e}), \tag{3.55}$$

$$\eta_{00}(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 0 \cap Q_{r,k} = 0 | \mathbf{E}_k = \mathbf{e}) = (1 - \eta_t(\mathbf{e}))(1 - \eta_r(\mathbf{e})). \tag{3.56}$$

3.5.3 Local knowledge

We suppose that EHTX does not have any information about the energy of EHRX and vice-versa. Similarly to the previous case, the marginal transmission probabilities can be written as:

$$\eta_t(\mathbf{e}) \triangleq \mathbb{P}(Q_{t,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{10}(\mathbf{e}) + \eta_{11}(\mathbf{e}) = \rho_t(e_t), \tag{3.57}$$

$$\eta_r(\mathbf{e}) \triangleq \mathbb{P}(Q_{r,k} = 1 | \mathbf{E}_k = \mathbf{e}) = \eta_{01}(\mathbf{e}) + \eta_{11}(\mathbf{e}) = \rho_r(e_r), \tag{3.58}$$

with:

$$\rho_t(e_t) \triangleq \mathbb{P}(Q_{t,k} = 1 | E_{t,k} = e_t), \tag{3.59}$$

$$\rho_r(e_r) \triangleq \mathbb{P}(Q_{r,k} = 1 | E_{r,k} = e_r). \tag{3.60}$$

3.6 Performance evaluation

We focus on the Centralized and Semi-Independent cases, leaving the Totally-Independent one as future work. In our numerical evaluation we used the following parameters: $e_{\max,t}$, $e_{\max,r} \in \{1, \ldots, 20\}$, $\bar{b}_t, \bar{b}_r \in \{0.05, 0.10, 0.2, \ldots, 0.9, 0.95\}$ and $\Lambda = 10$ dB.

3.6.1 Centralized case

We notice that, when one of the two EHDs has a higher average probability of receiving energy, then the maximum energy level of this device is approximately non influential in the optimal reward. In Figure 3.2 we plot the long-term average reward (transmission rate) $G_{\mathrm{C-OP}}$ as a function of $e_{\max,t}$ and $e_{\max,r}$ when $\bar{b}_t = 0.9$ and $\bar{b}_r = 0.5$. It can be seen that, except for $e_{\max,t} = 1$, the optimal reward is independent of $e_{\max,t}$. Also, it can be noticed that $G_{\mathrm{C-OP}}$, as expected, increases with $e_{\max,r}$ and saturates when $e_{\max,r} \approx 10$ because the impact of outage and overflow decreases, therefore it is not necessary to use large batteries.

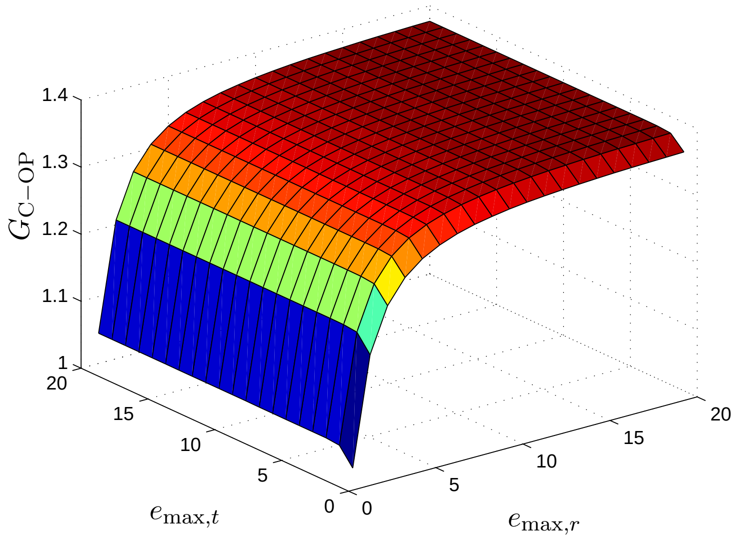


Figure 3.2: Reward function $G_{\mathrm{C-OP}}$ (Centralized case-Optimal Policy) when $e_{\max,t}, e_{\max,r} \in \{1, \dots, 20\}$ and $\bar{b}_t = 0.9$, $\bar{b}_r = 0.5$

3.6.2 Cases comparison

Table 3.1: Case acronyms.

Case acronym	Case
C	Centralized case
S,P	Semi-Independent case with Partial knowledge
S,L	Semi-Independent case with Local knowledge

In the next, with the notation "R-P" we indicate the case R with the policy P. In our work we consider the following case-policy pairs: C-OP; S,P-OP; S,P-OEIP; S,L-OP; S,L-OEIP. We recall that OP is the Optimal Policy, that is the policy that maximizes (3.25) and OEIP is the Optimal-EIP, that is, among all the EIP, the one that maximizes (3.25). EIP is a policy that does not depend on the energy levels of the batteries, provided that they are not empty.

Centralized and Semi-Independent cases

Partial knowledge In our numerical evaluation we consider an EIP because:

- it is easy to implement;
- its performance is close to the Optimal Policy (OP) in the Centralized case, that is an upper bound.

In the range of the considered parameters, we find that the reward precision (defined as in (2.30))

$$\mathcal{R}_{\mathrm{C-OP}}^{\mathrm{S,P-OEIP}} = \frac{G_{\mathrm{C-OP}} - G_{\mathrm{S,P-OEIP}}}{G_{\mathrm{C-OP}}} \tag{3.61}$$

is always lower than 0.36%. Therefore, since:

$$G_{\mathrm{S,P-OEIP}} \leq G_{\mathrm{S,P-OP}} \leq G_{\mathrm{C-OP}}, \tag{3.62}$$

it follows that:

$$\mathcal{R}_{\mathrm{C-OP}}^{\mathrm{S,P-OP}} \leq 0.36\%, \tag{3.63}$$

that is, if we suppose $\eta_{01}(\mathbf{e}) = 0$ and that EHTX knows if EHRX is in outage and vice-versa, the performance obtained is very close to the optimal one.

Local knowledge In this case we evaluate an energy independent policy. We find that, in general, $G_{\mathrm{C-OP}}$ and $G_{\mathrm{S,L-OEIP}}$ are quite distant:

$$\max\left\{\mathcal{R}_{\mathrm{C-OP}}^{\mathrm{S,L-OEIP}}\right\} \approx 30\%. \tag{3.64}$$

The following chain of inequalities holds:

$$G_{\mathrm{S,L-OEIP}} \leq G_{\mathrm{S,L-OP}} \leq G_{\mathrm{C-OP}} \tag{3.65}$$

and, since $G_{\mathrm{C-OP}} \gg G_{\mathrm{S,L-OEIP}}$, we cannot say which values $G_{\mathrm{S,L-OP}}$ assumes. However, we computed $G_{\mathrm{S,L-OP}}$ for some cases and we find $G_{\mathrm{S,L-OP}} \gtrsim G_{\mathrm{S,L-OEIP}}$. We leave as future work further investigation about $G_{\mathrm{S,L-OP}}$.

Since $\mathcal{R}_{\mathrm{C-OP}}^{\mathrm{S,P-OEIP}} \gtrsim 0$ and $\mathcal{R}_{\mathrm{C-OP}}^{\mathrm{S,L-OEIP}} \gg 0$, we can say that the outage information is essential for the devices, if an energy independent policy is used. In particular, we noticed that $\mathcal{R}_{\mathrm{C-OP}}^{\mathrm{S,L-OEIP}}$ is high when $e_{\max,r}$ and $\bar{b}_r$ are low. Indeed, in these cases, the outage probability of the receiver is high, therefore the outage information becomes important. As examples, in Figures 3.3 and 3.4 we compare $G_{\mathrm{C-OP}}$, $G_{\mathrm{S,P-OEIP}}$ and $G_{\mathrm{S,L-OEIP}}$ when $\bar{b}_t = \bar{b}_r = 0.05$ (small value). In the first figure we consider $e_{\max,r} = 1$ and in the second one $e_{\max,r} = 20$. Since in the first case the EHRX outage probability is high, $G_{\mathrm{S,L-OEIP}}$ is quite distant from the optimal reward. Instead, in the second figure, $G_{\mathrm{S,L-OEIP}}$ is close to $G_{\mathrm{S,P-OEIP}}$ because the EHRX outage probability is low. It can be also noticed that the distance between $G_{\mathrm{C-OP}}$ and $G_{\mathrm{S,P-OEIP}}$ is lower in the case $e_{\max,r} = 1$. This happens because, when $e_{\max,r} = 1$, the partial knowledge is equivalent to the global knowledge from the EHTX point of view, *i.e.*, the transmitter always knows the energy state of the receiver. Instead, in the case $e_{\max,r} = 20$, the outage knowledge is not equivalent to the global knowledge, therefore the performance with S,P-OEIP is worse than with C-OP.

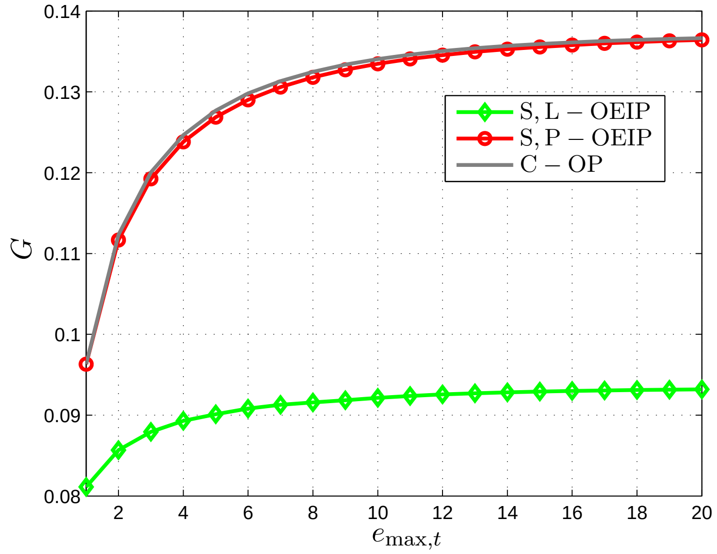


Figure 3.3: Comparison of different reward functions G when $e_{\max,t} \in \{1, \ldots, 20\}$, $e_{\max,r} = 1$ and $\bar{b}_t = \bar{b}_r = 0.05$

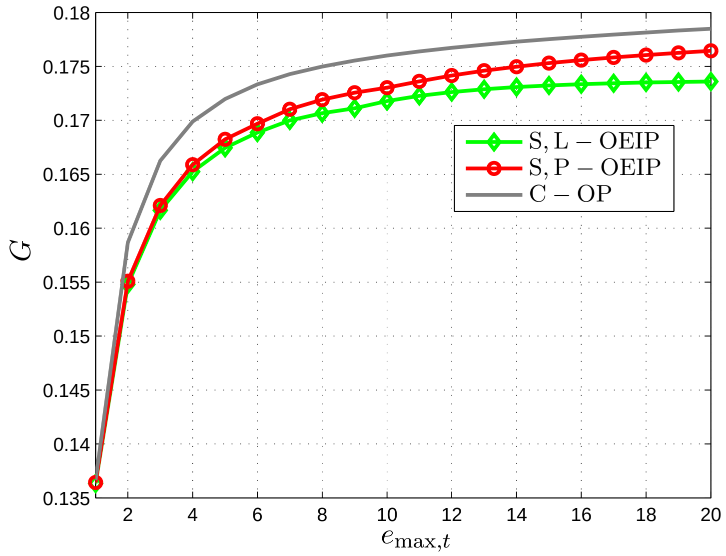


Figure 3.4: Comparison of different reward functions G when $e_{\max,t} \in \{1, \ldots, 20\}$, $e_{\max,r} = 20$ and $\bar{b}_t = \bar{b}_r = 0.05$

3.7 Conclusions

As a second problem, we studied an Energy Harvesting Transmitter (EHTX) and Energy Harvesting Receiver (EHRX) pair. We adapted the model of Chapter 2 to this case, specializing it according to the degree of knowledge the devices have. We noticed that the optimal performance in the Centralized case (C) is achieved also in the Semi-Independent or Totally-Independent cases with global knowledge. Then, we studied the cases of partial or local knowledge about the state of charge of the batteries. The former presumed that each EHD knows if the other one is in outage or not, while the second case supposed that a device has information only about its battery status. We computed the Optimal Policy (OP) and described some of its properties. Finally, we showed numerically that the outage information is essential to achieve high transmission rates. As future work we will complete the study of the Totally-Independent case and we will further investigate the optimization problem regarding the Local-knowledge cases.

Chapter 4

One EHD with a Data Queue

In this Chapter, we study a single Energy Harvesting Device (EHD) with a data queue. The device stores the packets in a FIFO queue and when the queue is non empty it can transmit the first packet to a receiver.

Initially, we study an infinite data queue and we find a sufficient condition for the queue stability. In particular, we find that if the node transmits with a probability greater than a fixed threshold, then the data queue is stable. Then, we use this information to understand how the optimal transmission policy works in the finite data queue scenario and, consequently, how to design a low complexity policy that performs close to the optimal one. This is a simple piece-wise linear transmission policy, that is easy to implement and is energy independent, therefore it can also be used in the cases of imperfect knowledge of the battery level of charge.

4.1 System model

As in Chapter 3, the model is similar to the one presented in Chapter 2, therefore we emphasize only the differences.

We consider a device that reports data to a central node. Time is slotted and slot k is the time interval $[k, k+1), k \in \mathbb{N}_0$. In slot k, the node can transmit over the interval $[k, k+\delta_k)$, where $\delta_k \in (0,1]$ is the duty cycle.

The battery of the EHD is modeled as an energy buffer. Each position in the buffer can hold one energy quantum and the transmission of one packet requires the expenditure of one energy quantum. The maximum amount of energy that can be stored in the energy buffer is $e_{\max}$ and the set of its possible energy status is $\mathcal{E} = \{0, 1, \ldots, e_{\max}\}$.

The queue policy is FIFO and its length at time k is F_k. The queue size is $f_{\max} \in \mathbb{N} \cup \{\infty\} \backslash \{0\}$. In slot k, only the header packet, if any, can be transmitted.

At time $k+1$, the energy in the buffer and the queue size are updated according to:

$$E_{k+1} = \min\{(E_k - Q_k)^+ + B_k, e_{\max}\}, \tag{4.1}$$

$$F_{k+1} = \min\{(F_k - Q_k)^+ + C_k, f_{\max}\}, \tag{4.2}$$

where $\{B_k\}$ and $\{Q_k\}$, defined as in (2.1), are the energy arrival process and the action process respectively. Similarly, $\{C_k\}$ is the packet arrival process. We assume that $\{C_k\}$ is a Bernoulli i.i.d. random arrival process with mean $\bar{c} \in (0,1]$. Furthermore, an energy quantum harvested in slot k can be used only in a time slot $> k$. Similarly, a packet generated or received in slot k can be transmitted only in a time slot $> k$.

In addition to the energy outage and overflow, in this scenario also the *queue outage* and *queue overflow* may happen:

Definition 4 (queue outage). *In slot k, queue outage occurs if $F_k = 0$.*

Definition 5 (queue overflow). *In slot k, queue overflow occurs if $(F_k = f_{\max}) \cap (C_k = 1) \cap (Q_k = 0)$.*

According to the metric to optimize, queue outage may be a situation to avoid or not. If the metric is, *e.g.*, the average delay of the packets, then the aim is to empty the queue and queue outage is a positive situation. However, in the next, supposing that the channel gain changes temporally, we consider as metric the average transmission rate. In this case queue outage should be avoided because it may induce an energy overflow situation: if the data queue is empty and the battery is fully charged, without any packet to send, any new quantum arrival is wasted. This is not convenient because a lost quantum corresponds to the loss of the possibility of transmitting a future packet, therefore it lowers the possible amount of sent data and, as a consequence, the average transmission rate. Furthermore, queue overflow should always be avoided because it corresponds to a packet loss.

Differently from the previous Chapters, we suppose that the importance level of every packet is equal. Indeed, if the packets could have different importance levels, then a FIFO queue would not be a suitable structure to manage the data and a different policy should be used. However, we analyze a system where the importance V_k associated with slot k is related to the channel gain: if the channel rate is low (high), then V_k is low (high). We model V_k as a continuous random variable with probability density function (pdf) $f_V(v)$, $v \geq 0$. Note that, in Chapters 2 and 3, a new packet is considered in every slot, therefore associating the importance level to the packets or to the channel is the same.

With the introduced quantities, we define the *state of the system*

$$\mathbf{S}_k = (E_k, F_k, V_k). \tag{4.3}$$

4.2 Policy Definition and Optimization Problem

The policy μ is defined as:

$$\mu(i; (e,f), v) = \mathbb{P}(\text{draw } i \text{ from the EHD} | (e,f), v), \qquad i \in \{0,1\}, \tag{4.4}$$

and the *long-term average reward* using a policy μ is:

$$G(\mu, \mathbf{S}_0) = \lim_{K \to \infty} \inf \frac{1}{K} \mathbb{E}\left[\sum_{k=0}^{K-1} Q_k V_k \middle| \mathbf{S}_0 \right]. \tag{4.5}$$

The optimal policy μ^* is the one that maximizes G.

Theorem 4. *The optimal policy has the following structure:*

$$\begin{cases} \mu(1; (e,f), v) = 1, & \text{if } v > v_{\text{th}}(e,f), \\ \mu(0; (e,f), v) = 1, & \text{if } v \leq v_{\text{th}}(e,f). \end{cases} \tag{4.6}$$

Proof. Similar to the proof of Theorem 3. □

The *marginal transmission probability* in state (e, f) is:

$$\eta(e, f) = \mathbb{E}_V[\mu(1; (e, f), v)] = \mathbb{P}(Q_k = 1 | E_k = e \cap F_k = f). \tag{4.7}$$

The *expected reward* is:

$$g(\eta(e, f)) = \mathbb{E}[Q_k V_k | E_k = e \cap F_k = f] \tag{4.8}$$

and, as in Chapter 2, the long-term average reward given a policy η can be rewritten as:

$$G_\eta = \sum_{e=0}^{e_{\max}} \sum_{f=0}^{f_{\max}} \pi_\eta(e, f) g(\eta(e, f)), \tag{4.9}$$

where $\pi_\eta(e, f)$ is the steady-state probability of being in state (e, f) given the policy η (in the next, for notational convenience, we neglect the dependence on η).

The transmission probabilities of the Markov Chain (MC) for the generic state (e, f) with $e \in \{1, \dots, e_{\max} - 1\}$ and $f \in \{1, \dots, f_{\max} - 1\}$ are:

$$p_{(e,f)\to(k,l)} = \begin{cases} (1-\bar{b})(1-\bar{c})\eta(e,f), & \text{if } k = e-1, l = f-1, \\ (1-\bar{b})\bar{c}\eta(e,f), & \text{if } k = e-1, l = f, \\ 0, & \text{if } k = e-1, l = f+1, \\ \bar{b}(1-\bar{c})\eta(e,f), & \text{if } k = e, l = f-1, \\ \bar{b}\bar{c}\eta(e,f) + (1-\bar{b})(1-\bar{c})(1-\eta(e,f)), & \text{if } k = e, l = f, \\ (1-\bar{b})\bar{c}(1-\eta(e,f)), & \text{if } k = e, l = f+1, \\ 0, & \text{if } k = e+1, l = f-1, \\ \bar{b}(1-\bar{c})(1-\eta(e,f)), & \text{if } k = e+1, l = f, \\ \bar{b}\bar{c}(1-\eta(e,f)), & \text{if } k = e+1, l = f+1, \\ 0, & \text{otherwise} \end{cases} \tag{4.10}$$

and if $e \in \{0, e_{\max}\}$ and/or $f \in \{0, f_{\max}\}$ the previous probabilities must be appropriately corrected.

4.2.1 Maximization of the transmission rate

Similarly to equation (2.15), we assume:

$$V_k = \ln(1 + \Lambda H_k). \tag{4.11}$$

Λ is the SNR and the channel gain H has pdf $f_H(h) = e^{-h},\ h > 0$. We find (as in [16]):

$$h_{\text{th}}(e, f) = \ln\left(\frac{1}{\eta(e, f)}\right), \tag{4.12}$$

$$g(\eta(e, f)) = \int_{h_{\text{th}}(e,f)}^{\infty} \ln(1 + \Lambda h) e^{-h} \mathrm{d}h, \tag{4.13}$$

where h_{th} is the channel threshold associated to the importance threshold $v_{\text{th}}(e, f)$.

For a fixed policy η, using (4.10) it is possible to find the steady state probabilities $\pi_\eta(e, f)$, therefore, with (4.13), the long-term average reward G_η of equation (4.9) can be found. In this case G_η is the long-term average transmission rate (in the next we call this *transmission rate* for brevity).

4.3 Analysis

We now consider $e_{\max} = 1$ and we relax later this hypothesis. In the next proposition we introduce an important relation between the steady state probabilities.

Proposition 3. *We consider a MC with $e_{\max} = 1$. The following relation for the steady-state probabilities holds:*

$$\pi(1, f) = A(f)\pi(1, f-1) + B(f)\pi(1, f+1), \qquad f \geq 1 \tag{4.14}$$

with:

$$A(f) \triangleq \frac{(1-\bar{b})\bar{c}(1-\eta(1, f-1))}{\bar{c}+\bar{b}-\bar{b}\bar{c}+\eta(1, f)(1-\bar{c}-2\bar{b}+\bar{b}\bar{c})}, \tag{4.15}$$

$$B(f) \triangleq \frac{\bar{b}\frac{1-\bar{c}}{\bar{c}}\eta(1, f+1)}{\bar{c}+\bar{b}-\bar{b}\bar{c}+\eta(1, f)(1-\bar{c}-2\bar{b}+\bar{b}\bar{c})}. \tag{4.16}$$

Proof. The relation between the steady-state probabilities is ($f > 0$):

$$\begin{aligned}
\pi(0, f) &= \pi(0, f)(1-\bar{b})(1-\bar{c}) + \pi(1, f)\eta(1, f)(1-\bar{b})\bar{c} \\
&\quad + \pi(1, f+1)\eta(1, f+1)(1-\bar{b})(1-\bar{c}) + \pi(0, f-1)(1-\bar{b})\bar{c}, \\
\pi(1, f) &= \pi(1, f)\Big(\eta(1, f)\bar{b}\bar{c} + (1-\eta(1, f))(1-\bar{c})\Big) + \pi(1, f-1)(1-\eta(1, f-1))\bar{c} \\
&\quad + \pi(1, f+1)\eta(1, f+1)(1-\bar{c})\bar{b} + \pi(0, f-1)\bar{b}\bar{c} + \pi(0, f)\bar{b}(1-\bar{c})
\end{aligned}$$

and putting together the two previous ones we obtain:

$$\pi(1, f) = \alpha(f)\pi(1, f-1) + \beta(f)\pi(1, f+1) + \delta(f)\pi(0, f-1), \tag{4.17}$$

with:

$$\alpha(f) = \frac{(1-\eta(1, f-1))\bar{c}}{1 - \Big(\eta(1, f)\bar{b}\bar{c} + (1-\eta(1, f))(1-\bar{c})\Big) - \eta(1, f)\bar{b}(1-\bar{c})\frac{(1-\bar{b})\bar{c}}{1-(1-\bar{b})(1-\bar{c})}},$$

$$\beta(f) = \frac{\eta(1, f+1)\frac{\bar{b}(1-\bar{c})}{1-(1-\bar{b})(1-\bar{c})}}{1 - \Big(\eta(1, f)\bar{b}\bar{c} + (1-\eta(1, f))(1-\bar{c})\Big) - \eta(1, f)\bar{b}(1-\bar{c})\frac{(1-\bar{b})\bar{c}}{1-(1-\bar{b})(1-\bar{c})}},$$

$$\delta(f) = \frac{\frac{\bar{b}\bar{c}}{1-(1-\bar{b})(1-\bar{c})}}{1 - \Big(\eta(1, f)\bar{b}\bar{c} + (1-\eta(1, f))(1-\bar{c})\Big) - \eta(1, f)\bar{b}(1-\bar{c})\frac{(1-\bar{b})\bar{c}}{1-(1-\bar{b})(1-\bar{c})}}.$$

We now use the fact that in the long-term, the frequency of transitions from queue level $f-1$ to f and from f to $f-1$ must be the same:

$$\pi(0, f)\bar{c} + \pi(1, f)(1-\eta(1, f))\bar{c} = \pi(1, f+1)\eta(1, f+1)(1-\bar{c}), \tag{4.18}$$

therefore, equation (4.17) can be rewritten as:

$$\pi(1, f) = A(f)\pi(1, f-1) + B(f)\pi(1, f+1),$$

with $A(f)$ and $B(f)$ as specified in (4.15-4.16). □

4.3.1 Infinite Queue

We now consider $f_{\max} = \infty$. The MC can be positive recurrent, null recurrent or transient. Its behaviour depends on the energy and packet arrival rates and on the transmission probabilities $\eta(1, f)$. Intuitively, if $\bar{b} \le \bar{c}$, the MC cannot be positive recurrent because too many packets arrive with respect to the energy arrivals. Formally this is stated in the following theorem.

Theorem 5. *We consider an irreducible and aperiodic MC and we set $f_{\max} = \infty$. If we have:*

$$\bar{c} > \bar{b}, \tag{4.19}$$

then the MC is unstable.

Proof. We consider an interval of n slots and we choose $\eta(e, f) = 1,\ \forall e > 0,\ \forall f > 0$. If the MC is not positive recurrent in this case, it cannot be positive recurrent in any case, because this choice of the marginal transmission probabilities is the most aggressive among all.

In n slots, on average, $n \cdot \bar{b}$ energy quanta and $n \cdot \bar{c}$ packets arrive. At the end of the considered slots, the queue collects, on average, at least $n \cdot (\bar{c} - \bar{b}) > 0$ packets, thus, when $n \to \infty$, the queue diverges.

More formally, we prove the theorem for the case $e_{\max} = 1$ with the drift analysis. The drift when there are $f > f_0 > 1$ packets in the queue is:

$$\begin{aligned} D(f) &= \mathbb{E}[F_{k+1} - F_k | F_k = f] = \\ &= \frac{\pi(0, f)}{\pi(0, f) + \pi(1, f)}\Big(\bar{c}(f+1) + (1-\bar{c})f\Big) + \\ &+ \frac{\pi(1, f)}{\pi(0, f) + \pi(1, f)}\Big(\bar{c}f + (1-\bar{c})(f-1)\Big) - f \\ &= (\bar{c} - \bar{b} + f) - f > 0 \end{aligned} \tag{4.20}$$

where we used the equations (4.16) and (4.18). Note that $\frac{\pi(i,f)}{\pi(0,f)+\pi(1,f)}$ is the probability of being in the state (i, f) given that $F_k = f$. Since the drift is greater than zero for each $f > f_0$ and $\mathbb{P}(f \to j) = 0,\ \forall f, j:\ j < f - 1$, then the MC is unstable (Kaplan's Theorem).

To prove the theorem in the case $e_{\max} > 1$, it is sufficient to find the relations between the steady-state probabilities (as in equations (4.15), (4.16)) and apply the drift analysis. □

Note that the previous theorem is true for every $e_{\max}$.

Theorem 6. *We consider an irreducible and aperiodic MC and we set* $f_{\max} = \infty$*. If we have:*

$$\bar{c} = \bar{b}, \tag{4.21}$$

then the MC is null recurrent.

Proof. Similar to the proof of Theorem 5. □

Vice-versa, if $\bar{b} > \bar{c}$, there always exists a choice of the marginal transmission probabilities that makes the MC positive recurrent.

Theorem 7. *We consider an irreducible and aperiodic MC and we set* $f_{\max} = \infty$*. If we have:*

$$\bar{c} < \bar{b}, \tag{4.22}$$

then there exists a choice of $\eta(1, f)$ *such that the MC is positive recurrent.*

Proof. Similar to the proof of Theorem 5. □

We now consider the steady-state probabilities whose tail decreases at least geometrically. If the MC is well defined, this is a sufficient condition for positive recurrence. Formally:

Theorem 8. *We consider an irreducible and aperiodic MC and we set* $e_{\max} = 1$ *and* $f_{\max} = \infty$*. If* $\exists f_0$ *and* $q \in (0, 1)$ *such that*

$$\frac{\pi(1, f+1)}{\pi(1, f)} \leq q < 1, \qquad \forall f \geq f_0, \tag{4.23}$$

then the MC is positive recurrent.

Proof. We want to show that the following series are convergent:

$$\sum_{f=0}^{\infty} \pi(0, f) + \sum_{f=0}^{\infty} \pi(1, f).$$

Since

$$\pi(0, f)\bar{c} = \pi(1, f+1)\eta(1, f+1)(1 - \bar{c}) - \pi(1, f)(1 - \eta(1, f))\bar{c},$$

we have that $\sum_{f=0}^{\infty} \pi(0, f)$ is proportional to $\sum_{f=0}^{\infty} \pi(1, f)$, therefore we study only the convergence of the second sum.

Exploiting the hypothesis, we have:

$$\lim_{f \to \infty} \sup \frac{\pi(1, f+1)}{\pi(1, f)} \leq q < 1,$$

therefore, by D'Alembert's criterion, the thesis is proved. □

One of our main results is stated in the next theorem, that gives a sufficient condition to have a positive recurrent MC.

Theorem 9. *We consider an irreducible and aperiodic MC and we set* $e_{\max} = 1$ *and* $f_{\max} = \infty$*. If* $\exists f_0$ *and* $q \in (0,1)$ *such that*

$$\frac{\pi(1, f_0+1)}{\pi(1, f_0)} \leq q < 1, \tag{4.24}$$

$$\frac{1-A(f)}{B(f)} \leq q < 1, \qquad \forall f > f_0, \tag{4.25}$$

then

$$\frac{\pi(1, f+1)}{\pi(1, f)} \leq q < 1, \qquad \forall f \geq f_0. \tag{4.26}$$

Proof. Proof by induction on f:

- base case: $f = f_0$

  $$\frac{\pi(1, f_0+1)}{\pi(1, f_0)} \leq q < 1$$

  is true by hypothesis.

- inductive hypothesis: $f = x - 1 > f_0$

  $$\frac{\pi(1, x)}{\pi(1, x-1)} \leq q < 1.$$

- inductive step: $f = x > f_0$

  $$\pi(1,x) = A(x)\pi(1,x-1) + B(x)\pi(1,x+1) > A(x)\pi(1,x) + B(x)\pi(1,x+1)$$
  $$\Leftrightarrow \frac{\pi(1,x+1)}{\pi(1,x)} < \frac{1-A(x)}{B(x)} \leq q < 1.$$

□

Remark 1. *The converse in general is not true. However, if we assume* (4.24)*, then the conditions* (4.25) *and* (4.26) *are equivalent.*

A possible choice of $\eta(1,f)$ is the queue independent one, *i.e.*, $\eta(1,f) = \eta\chi\{f > 0\}$, $\eta \in [0,1]$. With these marginal transmission probabilities, condition (4.25) becomes:

$$\eta > \eta_{\text{th}} \triangleq \frac{1}{1 + \frac{1}{\bar{c}} - \frac{1}{b}}. \tag{4.27}$$

Therefore, if we consider a policy such that

$$\eta(1,f) = \eta + \theta(1,f), \qquad \forall f > f_0, \tag{4.28}$$

with $\eta > \eta_{\text{th}}$ and $\theta(1,f) \in [0, 1-\eta]$, then condition (4.25) is satisfied.

4.3.2 Finite Queue

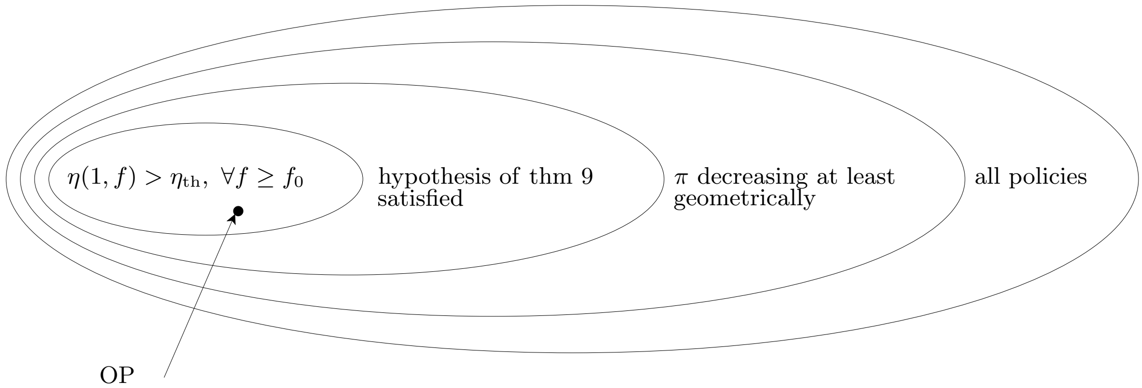


Figure 4.1: Venn diagram of the possible policies

We now consider $f_{\max} < \infty$, $e_{\max} = 1$ and $\bar{b} - \bar{c} \geq 0.05$ (we discuss later the case $\bar{b} \approx \bar{c}$). We computed numerically the Optimal Policy (OP) $\eta_{\text{OP}}(1, f)$ in several cases. With the OP, the probabilities of being in the final data queue states are low and decrease with f in order to avoid data buffer overflow. Furthermore, it can be numerically verified that the OP satisfies the hypothesis of Theorem 9 and, in particular, it always verifies the condition 4.28 (see Figure 4.1). However, note that differently from the OP, in the general case the steady-state distribution tail may not decrease. As an example, in Figure 4.2 we represent the steady state probabilities when $\bar{b} = 0.6$, $\bar{c} = 0.4$ and $f_{\max} = 50$ for a test policy $\eta_{\text{TEST}}(e, f) = 0.3\chi\{e > 0\}\chi\{f > 0\}$. It can be seen that the $\pi(e, f)$ increases with f.

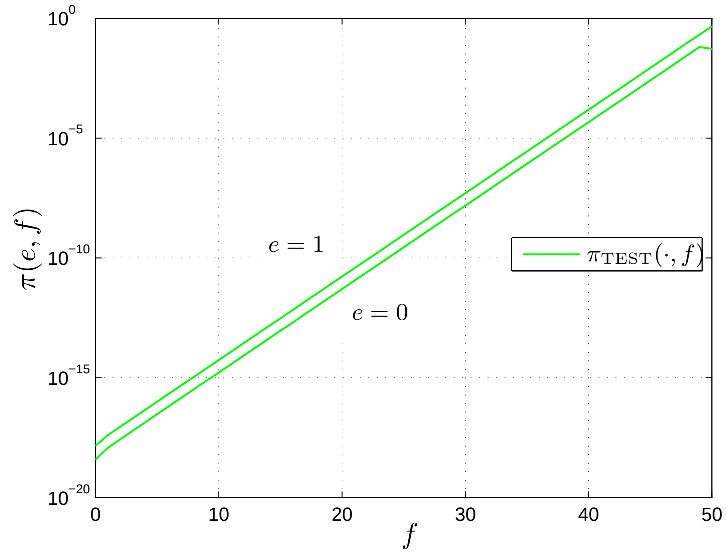


Figure 4.2: Steady-state probabilities $\pi_{\text{TEST}}(\cdot, f)$ as a function of f when $f_{\max} = 50$, $e_{\max} = 1$, $\bar{b} = 0.6$ and $\bar{c} = 0.4$.

In Figure 4.3 we represent the shape of $\eta_{\text{OP}}(1, f)$ when $\bar{b} = 0.6$, $\bar{c} = 0.4$ and $f_{\max} = 50$.

It can be seen that $\eta_{\text{OP}}(1, f)$ is low when the queue is almost empty in order to avoid queue outage, whereas it is high when f is high to avoid queue overflow (we recall that our metric is the transmission rate). The middle value is $\eta_{\text{OP}}(25) \approx \eta_{\text{th}} = 0.55$.

We verified that the shape of $\eta_{\text{OP}}(1, f)$ does not change when $\bar{b}$ and $\bar{c}$ range in $(0, 1]$. Therefore, we introduce a Low Complexity Policy (LCP) that approximates the optimal one:

$$\eta_{\text{LCP}}(f) = \begin{cases} \frac{\eta_{\text{th}}}{5} f & \text{if } f \leq 5, \\ \eta_{\text{th}} & \text{if } 5 < f \leq f_{\max} - 3, \\ \frac{\eta_{\text{end}} - \eta_{\text{th}}}{3}(f - (f_{\max} - 3)) + \eta_{\text{th}} & \text{if } f_{\max} - 3 < f, \end{cases} \tag{4.29}$$

where we define the following point:

$$\eta_{\text{end}} = \frac{1.16\bar{c} + 0.01}{\bar{c} + 0.17}. \tag{4.30}$$

To find η_{end} we computed the OP for several values of $\bar{b}$ and $\bar{c}$ and, since we had noticed that the end point η_{end} is approximately independent of $\bar{b}$, we interpolated η_{end} as a function of $\bar{c}$, finding the hyperbole of equation (4.30).

As an example, in Figure 4.3 we compare the transmission probabilities $\eta(1, f)$ for the Optimal Policy and the Low Complexity Policy when $\bar{b} = 0.6$, $\bar{c} = 0.4$, $f_{\max} = 50$ and $e_{\max} = 1$. Furthermore, with the same parameters, in Figure 4.3 we represent the steady-state probabilities $\pi(0, f)$ and $\pi(1, f)$ for the two policies. It can be seen that $\pi_{\text{OP}}(0, f)$ and $\pi_{\text{OP}}(1, f)$ decrease quickly for high values of f (at least geometrically).

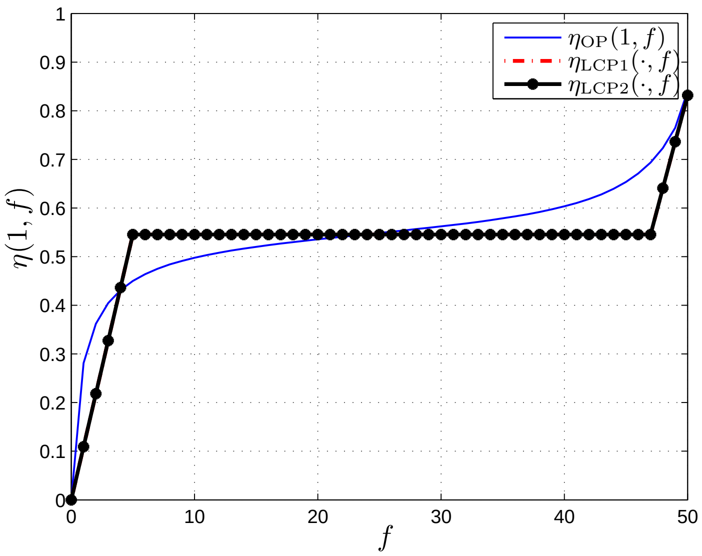
Case $\bar{b} \approx \bar{c}$

In the previous Section we neglected the case $\bar{b} \approx \bar{c}$ and we supposed a difference of at least 0.05 between the $\bar{b}$ and $\bar{c}$. Indeed, when the two values are very close, OP has a different behaviour: the tail of the steady state probabilities is not necessarily decreasing. Moreover, equation (4.28) is not satisfied, therefore LCP should be defined in another way. A possible technique to find a Low Complexity Policy for this case is, for example, to compute OP for a wide range of the parameters and to interpolate the results in order to understand the trend of the policy.

Analytically, the behaviour of the OP can be justified from the fact that, when $\bar{c}$ approaches $\bar{b}$, the system is going from the stability region to the instability one, therefore, since when we compute the policy we do not impose any stability condition, OP can easily lead to an unstable system. Indeed, when $\bar{b} \approx \bar{c}$, in order to keep the system stable, the transmission probabilities should be, generally, close to one, but this condition is not guaranteed in the computation of the OP.

4.3.3 Extension

The previous analysis assumes $e_{\max} = 1$. In the general case when $e_{\max} > 1$ a relation similar to that in equation (4.14) can be found and the previous reasoning can be followed in order to obtain a new threshold for the marginal transmission probabilities. Intuitively, when $e_{\max}$ increases, the threshold decreases. This also can be verified numerically.



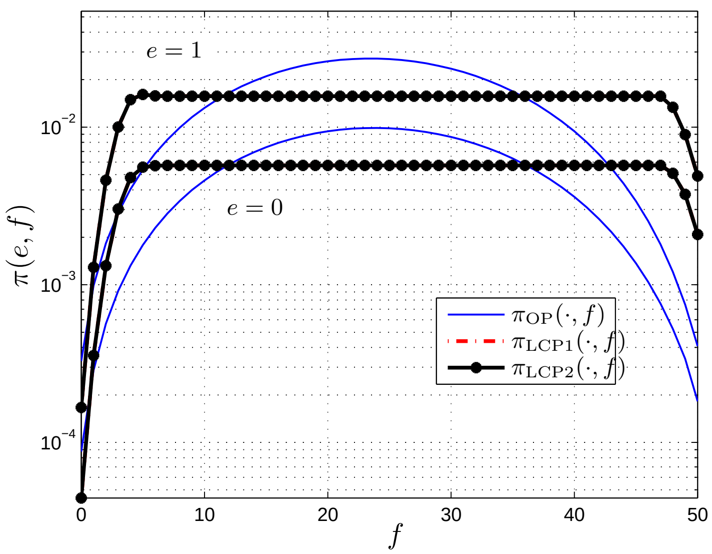


Figure 4.3: Marginal transmission probabilities $\eta_{\text{OP}}(1, f)$, $\eta_{\text{LCP}}(1, f)$ and steady-state probabilities $\pi_{\text{OP}}(\cdot, f)$, $\pi_{\text{LCP}}(\cdot, f)$ as a function of f when $f_{\max} = 50$, $e_{\max} = 1$, $\bar{b} = 0.6$ and $\bar{c} = 0.4$. Note that LCP1 is equal to LCP2 in this case.

We numerically verified that the Optimal Policy gives:

$$\eta(e_1, f) \approx \eta(e_2, f), \tag{4.31}$$

for e_1 and e_2 not too small, *i.e.*, the marginal transmission probabilities are approximately independent of the current energy state (see Figure 4.4). In order to approximate the trend of the OP, we use an Energy Independent Policy (EIP), *i.e.*, we impose $\eta(e_1, f) \approx \eta(e_2, f)$ for all e_1 and e_2. Note that, even if this approximation is rough for the low energy states, we use it anyway because it provides good results as shown in Section 4.4.

For the EIP we use the same transmission probabilities of the case $e_{\max} = 1$:

$$\eta(e, f) \triangleq \eta_{\mathrm{LCP}}(f), \qquad \forall e \in \{1, \ldots, e_{\max}\}, \tag{4.32}$$

with $\eta_{\mathrm{LCP}}(f)$ defined in (4.29) but instead of using η_{th} as defined in (4.27), we use $\eta'_{\mathrm{th}}(e_{\max})$ or $\eta''_{\mathrm{th}}(e_{\max})$ because we noticed that these are a better fit than η_{th} if $e_{\max} > 1$. The first one is defined as:

$$\eta'_{\mathrm{th}}(e_{\max}) = \begin{cases} \eta_{\mathrm{th}}, & \text{if } e_{\max} = 1, \\ \eta_{\mathrm{th}} - 0.1\chi\{\eta_{\mathrm{th}} > 0.5\}, & \text{if } e_{\max} > 1 \end{cases} \tag{4.33}$$

The offset 0.1 is used because we noticed that it gives good performance for a wide range of values of $\bar{b}$ and $\bar{c}$. However, a better approach is the following:

$$\eta''_{\mathrm{th}}(e_{\max}) \triangleq \frac{\eta_{\mathrm{th}} - \bar{c}}{e^{-\beta}} e^{-\beta e_{\max}} + \bar{c} \tag{4.34}$$

To find the previous function, we:

- fixed $\bar{b}$ and $\bar{c}$;
- found $\eta_{\mathrm{OP}}(e_{\max}, f_0)$ as a function of $e_{\max}$, with $f_0 = \lfloor \frac{f_{\max}}{2} \rfloor$;
- imposed $\eta''_{\mathrm{th}}(e_{\max})$ as the interpolation of $\eta_{\mathrm{OP}}(e_{\max}, f_0)$ (following the reasoning of the case $e_{\max} = 1$);
- noticed that a good interpolation of $\eta_{\mathrm{OP}}(e_{\max}, f_0)$ is the exponential one:

$$\eta''_{\mathrm{th}}(e_{\max}) = w_1 e^{-\beta e_{\max}} + w_2; \tag{4.35}$$

- imposed the boundary conditions $\eta''_{\mathrm{th}}(1) = \eta_{\mathrm{th}}$ and $\lim_{e_{\max}\to\infty} \eta''_{\mathrm{th}}(e_{\max}) = \bar{c}$:

$$w_1 = \frac{\eta_{\mathrm{th}} - \bar{c}}{e^{-\beta}}, \tag{4.36}$$

$$w_2 = \bar{c}; \tag{4.37}$$

  Indeed, when the battery is sufficiently large, the threshold value becomes $\bar{c}$ because transmitting with probability at least $\bar{c}$ is sufficient to keep the queue stable.

- changed $\bar{b}$ and $\bar{c}$ in order to find β. In particular, through an interpolation process, we found:

$$\beta \approx 3\bar{b} + 1.038e^{-5\bar{c}} + 0.1857 - 3(\bar{c} + 0.05) \tag{4.38}$$

  that is a straight line in $\bar{b}$ whose y-intercept is a function of $\bar{c}$.

This second policy is asymptotically optimal. In the following we refer to the two variations of the LCP with the labels LCP1 and LCP2, respectively.

We show in the next Section that these sub-optimal policies have performance close to the optimal one.

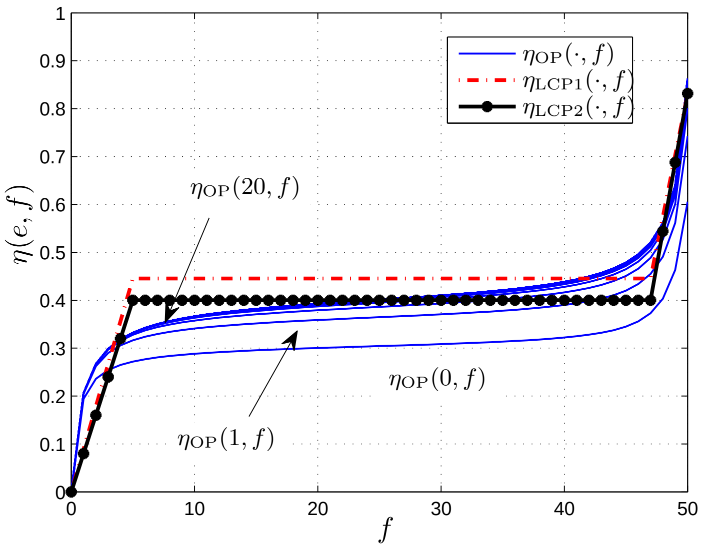


Figure 4.4: Marginal transmission probabilities $\eta_{\rm OP}(\cdot, f)$, $\eta_{\rm LCP1}(\cdot, f)$ and $\eta_{\rm LCP2}(\cdot, f)$ as a function of f when $f_{\max} = 50$, $e_{\max} = 20$, $\bar{b} = 0.6$ and $\bar{c} = 0.4$.

4.4 Performance Evaluation

We performed a numerical evaluation in order to characterize the Optimal Policy and to compare the performance of OP with LCP. We considered $e_{\max} \in \{1, \dots, 20\}$, $\bar{b}, \bar{c} \in \{0.05, 0.10, \dots, 1\}$ with $\bar{b} > \bar{c}$ and $\Lambda = 10$ dB. Furthermore, we do not consider $f_{\max}$ too small (< 10) because in a real device the buffer size is very large [10].

4.4.1 Optimal Policy

In Figure 4.5, we plot the long-term average reward (transmission rate) for different values of $\bar{c}$ when $\bar{b} = 0.5$. We notice that the transmission rate keeps increasing in the capacity of the battery until $e_{\max} \approx 6$ (this value depends on $\bar{c}$). In general, when $e_{\max} \approx 10$, the saturation region is already reached. This is because, the larger the battery, the smaller the impact of energy overflow and outage, *i.e.*, when the battery becomes sufficiently large, the improvement due to the decreased outage and overflow probabilities becomes negligible. Furthermore, the reward increases with the average packet arrival rate $\bar{c}$. Also in Figure 4.6, we represent the transmission rate as a function of $e_{\max}$, but in this case we keep $\bar{c}$ fixed and we range $\bar{b}$. It can be seen that the reward increases with $\bar{b}$ only when $e_{\max}$ is low, indeed, when $e_{\max} \approx 10$, all curves saturate to the same value. This happens because the bottleneck is due to the low average packet arrival rate, therefore, if $\bar{c}$ does not increase, the transmission rate cannot grow.

Moreover, with the considered range of $f_{\max} \geq 10$, the transmission rate is approximately independent of the queue size, *e.g.*, with $\bar{b} = 0.6$, $\bar{c} = 0.4$ and $e_{\max} = 1$ the transmission rate is equal to 1.053 if $f_{\max} = 10$ and to 1.081 if $f_{\max} = 100$. Note that this holds because we suppose $\bar{b} - \bar{c} \geq 0.05$.

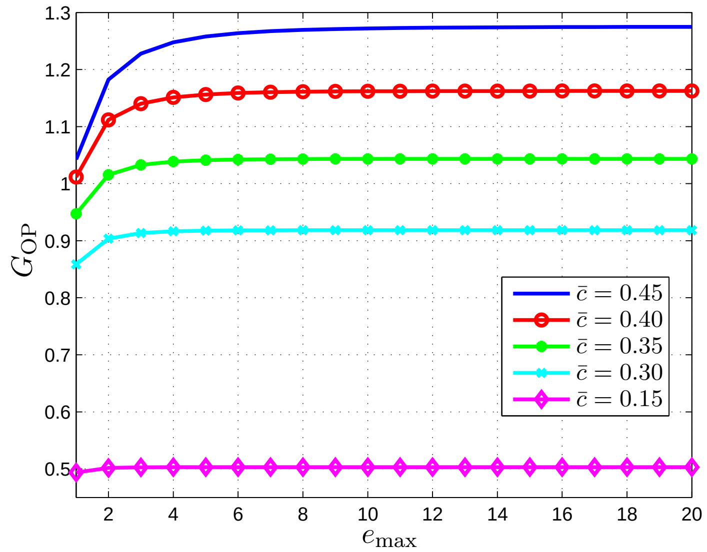


Figure 4.5: Reward functions G_{OP} for several values of $\bar{c}$ when e_{max} ranges in $\{1, \ldots, 20\}$ and $\bar{b} = 0.5$

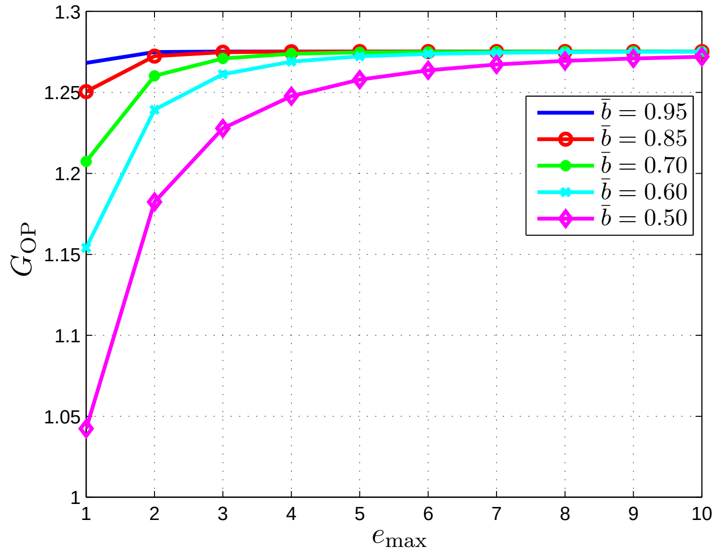


Figure 4.6: Reward functions G_{OP} for several values of $\bar{b}$ when e_{max} ranges in $\{1, \ldots, 10\}$ and $\bar{c} = 0.45$

4.4.2 Policies comparison

We computed the Optimal Policy and the two Low Complexity Policy versions in the considered range of the parameters. We verified that, with these parameters, the reward precision (defined as in (2.30))

$$\mathcal{R}_{\text{OP}}^{\text{LCP1}} = \frac{G_{\text{OP}} - G_{\text{LCP1}}}{G_{\text{OP}}} \tag{4.39}$$

is always lower than 6% and also smaller for LCP2.

In Figure 4.7 we compare G_{OP} and G_{LCP} when $e_{\max} \in \{1, \ldots, 20\}$ in a worst case scenario $\bar{b} = 0.5$ and $\bar{c} = 0.45$. Note that, even if $e_{\max}$ increases, G_{LCP1} does not approach G_{OP}, differently from G_{LCP2}.

In Figure 4.8 we represent $\mathcal{R}_{\text{OP}}^{\text{LCP1}}$ and $\mathcal{R}_{\text{OP}}^{\text{LCP2}}$ as a function of $e_{\max}$ for several values of $\bar{b}$ and $\bar{c}$. It can be seen that $\mathcal{R}_{\text{OP}}^{\text{LCP1}}$ is maximum when $\bar{b} = 0.5$ and $\bar{c} = 0.45$. In particular, we verified that this is one of the worst cases. Obviously, when $e_{\max} = 1$, $\mathcal{R}_{\text{OP}}^{\text{LCP1}}$ and $\mathcal{R}_{\text{OP}}^{\text{LCP2}}$ are low, because we designed LCP for this particular case. When $e_{\max}$ increases, $\mathcal{R}_{\text{OP}}^{\text{LCP1}}$ saturates to a constant value. This happens because $\eta_{\text{OP}}(e, f)$ is approximately independent of e (except for the first energy levels) as can be seen in Figure 4.4. Therefore, even if the battery size grows, the approximation does not get worse. Instead, for LCP2, the reward precision decreases with $e_{\max}$ (beyond a certain threshold) and it goes to zero when $e_{\max} \to \infty$, *i.e.*, it is asymptotically optimal. Note that $\mathcal{R}_{\text{OP}}^{\text{LCP1}}$ is not a decreasing function of $\bar{c}$, *e.g.*, the curve $\bar{c} = 0.3$ is lower than the one with $\bar{c} = 0.4$ but higher than the one with $\bar{c} = 0.35$.

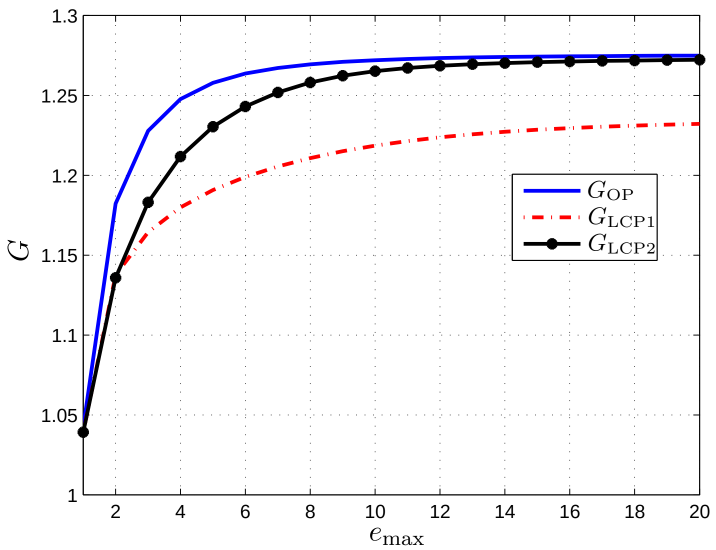


Figure 4.7: Comparison of the reward functions G_{OP}, G_{LCP1} and G_{LCP2} when $e_{\max} \in \{1, \ldots, 20\}$, $\bar{b} = 0.5$ and $\bar{c} = 0.45$

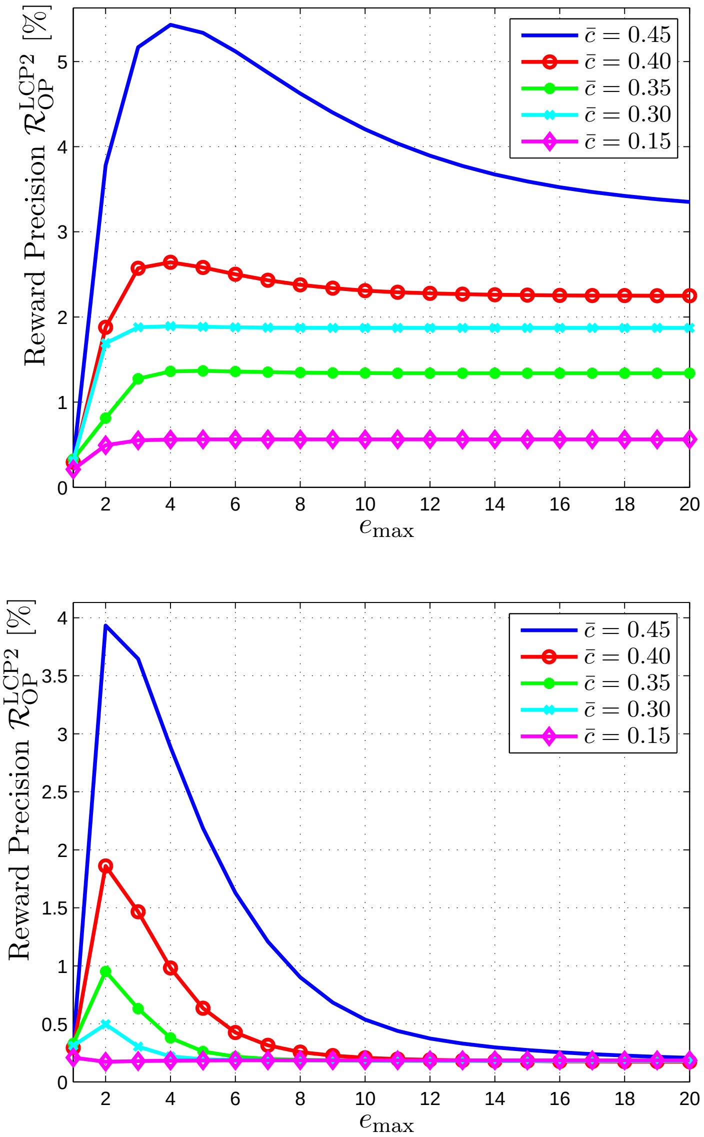


Figure 4.8: Percentage reward precisions $\mathcal{R}_{\mathrm{OP}}^{\mathrm{LPC1}}$ and $\mathcal{R}_{\mathrm{OP}}^{\mathrm{LPC2}}$ for several values of $\bar{c}$ when $e_{\max}$ ranges in $\{1, \dots, 20\}$ and $\bar{b} = 0.5$

4.5 Conclusions

In the third part of this work we studied the case of one Energy Harvesting Device (EHD) with a FIFO data queue which transmits data of different importance. Firstly we supposed $e_{\max} = 1$, $f_{\max} = \infty$ and, starting from the steady-state probabilities relation, we derived an analytical sufficient condition such that the Markov Chain is positive recurrent. Then, we considered $f_{\max} < \infty$ and verified numerically that the Optimal Policy (OP), under certain constraints, satisfies that sufficient condition. With this information, we designed a Low Complexity Policy (LCP) for the case $e_{\max} = 1$ that is a policy computable without any optimization. Furthermore, we extended LCP to the case $e_{\max} > 1$ with two variations and showed that they achieve good performance with respect to OP. We leave as future work the case $\bar{c} \gtrsim \bar{b}$. Also, we will study other metrics, *e.g.* the delay and seek a deeper understanding of LCP when the battery or the queue sizes grow so as to characterize the asymptotic behaviour of the system.

Chapter 5

Conclusions

This work studied several cases about the use of Energy Harvesting Devices (EHDs) in Wireless Sensor Networks (WSNs). In our numerical evaluations we considered the case of the transmission of packets over a channel with time-varying gain and we computed the long-term average transmission rates. Generally our main results consist in the formulation of low-complexity sub-optimal policies, that exhibit good performance, and in the definition of several models for different scenarios.

In Chapter 2 we analyzed the case of two EHDs with a Central Controller (CC). Depending on the global average harvesting rate, we introduced different sub-optimal low-complexity policies that approach the performance of the optimal one. We computed an analytical approximation for one of these policies and we showed that this is to be a good lower bound of the Optimal Policy (OP). As future work we will complete our analysis in order to cover a wider range of parameters: for the Balanced Policy (BP) we will consider the case of very small batteries and for the Heuristic Constrained Energy Independent Policy (HCEIP) we will study a context where the batteries of the devices are not equal.

A scenario with an Energy Harvesting Transmitter (EHTX) and Energy Harvesting Receiver (EHRX) pair was studied in Chapter 3. We analyzed several cases as a function of the amount of information the devices have about the system. We showed that for an Energy Independent Policy (EIP) the outage information is important to achieve high transmission rates. Future work will investigate optimal and low-complexity policies for the considered cases.

The last problem, in Chapter 4, consisted of the study of an EHD with a FIFO data queue. In the case of a very small battery and an infinite data queue, we derived several analytical results in order to characterize the queue stability. Then, we considered a finite queue, we computed the OP for several cases of interest and we introduced a Low Complexity Policy (LCP), with two variations, that approximates the trend of the OP. Numerical evaluation showed that the introduced policies are close to the optimal one. As future work, we will modify the hypothesis of the system, *e.g.* the amount of packets that can be sent in a slot, in order to consider a more realistic scenario and to characterize other metrics.

In all the problems, a future extension will be a system with a higher number of nodes. This can be useful to analyze also other problems related to the interaction among the devices, such as channel contention or routing issues. Finally, a non-ideal channel can be studied and retransmission of packets can be considered.

Appendix A

Optimization

A.1 Lagrangian relaxation

The results of this Section can be found in [33].

Consider following problem (P):

$$\min f(x), \tag{A.1}$$
$$x \in \mathbb{R}^n, \tag{A.2}$$
$$g_i(x) \leq 0, \qquad i \in I = \{1, \ldots, m\}, \tag{A.3}$$
$$h_j(x) = 0, \qquad j \in E = \{1, \ldots, p\}, \tag{A.4}$$

where, in general, $g_i(x)$ and $h_j(x)$ are non-linear continuous functions. We can define the *lagrangian multipliers*

$$u_i \in \mathbb{R}_+, \qquad \forall i \in I, \tag{A.5}$$
$$v_j \in \mathbb{R}, \qquad \forall j \in E \tag{A.6}$$

and the *auxiliary problem* (R):

$$w(u, v) = \min \mathcal{L}(x; u, v), \tag{A.7}$$
$$\mathcal{L}(x; u, v) = f(x) + \sum_{i \in I} u_i g_i(x) + \sum_{j \in E} v_j h_j(x), \tag{A.8}$$

where $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}$ is called *lagrangian function.*

Definition 6 (saddle point)**.** *The triple* $(\bar{x}, \bar{u}, \bar{v})$, $\bar{x} \in \mathbb{R}^n$, $\bar{u} \in \mathbb{R}^m_+$, $\bar{v} \in \mathbb{R}^p$ *is a saddle point if*

$$\mathcal{L}(\bar{x}, u, v) \leq \mathcal{L}(\bar{x}, \bar{u}, \bar{v}) \leq \mathcal{L}(x, \bar{u}, \bar{v}), \qquad \forall x \in \mathbb{R}^n,\ u \in \mathbb{R}^m_+,\ v \in \mathbb{R}^p. \tag{A.9}$$

Theorem 10. *Let* f, g_i $i \in I, h_j$ $j \in E$ *be continuous functions. If* $(\bar{x}, \bar{u}, \bar{v})$ *is a saddle point, then* $\bar{x}$ *is a global minimum for the problem (P).*

Remark 2. *The previous theorem gives some sufficient conditions to say that an admissible point* $\bar{x}$ *is a global minimum for (P),* i.e., *there exist some lagrangian multipliers* $\bar{u}$ *and* $\bar{v}$ *such that* $(\bar{x}, \bar{u}, \bar{v})$ *is a saddle point. In the general case this condition is not necessary.*

A.2 Proofs

Proof of Theorem 3

Proof. This proof follows the reasoning of [29].

We define $\mathcal{R}_\mu$ as the set of stationary randomized policies such that:

$$\mathbb{E}_V[\tilde{\mu}((i,j),\mathbf{e},V)] = \mathbb{E}_V[\mu((i,j),\mathbf{e},V)], \qquad \forall (i,j) \in \{(0,0),(1,1)\},\ \forall e_t, e_r,\ \forall \tilde{\mu} \in \mathcal{R}_\mu. \tag{A.10}$$

Since $\mu \in \mathcal{R}_\mu$, we have:

$$G(\mu,\mathbf{S}_0) \le \max_{\tilde{\mu}\in\mathcal{R}_\mu} G(\tilde{\mu},\mathbf{S}_0). \tag{A.11}$$

The function $G(\mu,\mathbf{S}_0)$ can be also written as in equation (3.25):

$$G(\tilde{\mu},\mathbf{S}_0) = \sum_{e_t=0}^{e_{\max,t}} \sum_{e_r=0}^{e_{\max,r}} \pi_{\tilde{\mu}}(\mathbf{e};\mathbf{S}_0)\mathbb{E}_V[\tilde{\mu}((1,1),\mathbf{e},V)V], \tag{A.12}$$

where:

$$\pi_{\tilde{\mu}}(\mathbf{e};\mathbf{S}_0) \triangleq \liminf_{K\to\infty} \frac{1}{K}\sum_{k=0}^{K-1} \mathbb{P}_{\tilde{\mu}}(\mathbf{E}_k = \mathbf{e}|\mathbf{S}_0). \tag{A.13}$$

It can be proved by induction on k that $\mathbb{P}_{\tilde{\mu}}(\mathbf{E}_k = \mathbf{e}|\mathbf{S}_0)$ depends on $\tilde{\mu}$ only through its expectation $\tilde{\eta}(\mathbf{e})$ (defined as in equation (3.13)).

From equation (A.10), we have $\tilde{\eta}(\mathbf{e}) = \eta(\mathbf{e}),\ \forall \tilde{\mu} \in \mathcal{R}_\mu$, therefore:

$$\pi_\eta(\mathbf{e};\mathbf{S}_0) \triangleq \pi_{\tilde{\mu}}(\mathbf{e};\mathbf{S}_0) = \pi_\mu(\mathbf{e};\mathbf{S}_0), \qquad \forall \tilde{\mu} \in \mathcal{R}_\mu. \tag{A.14}$$

Equation (A.11) can be rewritten as:

$$G(\mu,\mathbf{S}_0) \le \sum_{e_t=0}^{e_{\max,t}} \sum_{e_r=0}^{e_{\max,r}} \pi_\eta((1,1),\mathbf{e};\mathbf{S}_0)\mathbb{E}_V[\mu^*(\mathbf{e},V)V], \tag{A.15}$$

where $\mu^*(\mathbf{e},V)$ is the solution of the following problem P:

$$\begin{aligned} \mu^*(\mathbf{e},\cdot) = \arg\min_{\tilde{\mu}((1,1);\mathbf{e},\cdot)} & \ -\mathbb{E}_V[\tilde{\mu}((1,1);\mathbf{e},V)V], \\ \mathbb{E}_V[\tilde{\mu}((1,1);\mathbf{e},V)] - \eta_{11}(\mathbf{e}) = 0. \end{aligned} \tag{A.16}$$

We now want to find the structure of μ^* using the lagrangian relaxation. The auxiliary problem R is:

$$\begin{aligned} &\min_{\tilde{\mu}} \mathcal{L}(\tilde{\mu},v) \\ &\mathcal{L}(\tilde{\mu},v) = -\mathbb{E}_V[\tilde{\mu}((1,1);\mathbf{e},V)V] + v\Big(\mathbb{E}_V[\tilde{\mu}((1,1);\mathbf{e},V)] - \eta_{11}(\mathbf{e})\Big). \end{aligned} \tag{A.17}$$

Note that:

$$\mathcal{L}(\mu^*, v_{\text{th}}(\mathbf{e})) = -\mathbb{E}_V[\mu^*((1,1);\mathbf{e},V)V]. \tag{A.18}$$

In the following we prove that $(\mu^*((1,1);\mathbf{e},V), v_{\text{th}}(\mathbf{e}))$ is a saddle point.

1. left inequality:

$$\begin{aligned}\mathcal{L}(\mu^*, v) = &- \mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)V] + v\Big(\mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)] - \eta_{11}(\mathbf{e})\Big) = \\ &- \mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)V] = \mathcal{L}(\mu^*, v_{\text{th}}(\mathbf{e}));\end{aligned} \tag{A.19}$$

   where we used that fact that $\mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)] = \eta_{11}(\mathbf{e})$ because of the constraint of equation (A.16).

2. right inequality:

$$\begin{aligned}&\mathcal{L}(\mu^*, v_{\text{th}}(\mathbf{e})) \overset{?}{\leq} \mathcal{L}(\tilde{\mu}, v_{\text{th}}(\mathbf{e})) \\ \Leftrightarrow\ &\mathcal{L}(\mu^*, v_{\text{th}}(\mathbf{e})) = -\mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)V] \overset{?}{\leq} \\ &\qquad \overset{?}{\leq} -\mathbb{E}_V[\tilde{\mu}((1,1); \mathbf{e}, V)V] + v_{\text{th}}(\mathbf{e})\Big(\mathbb{E}_V[\tilde{\mu}((1,1); \mathbf{e}, V)] - \eta_{11}(\mathbf{e})\Big) = \\ &\qquad = -\mathbb{E}_V[\tilde{\mu}((1,1); \mathbf{e}, V)(V - v_{\text{th}}(\mathbf{e}))] - v_{\text{th}}(\mathbf{e})\eta_{11}(\mathbf{e}) \\ \Leftrightarrow\ &\mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)V] \overset{?}{\geq} \mathbb{E}_V[\tilde{\mu}((1,1); \mathbf{e}, V)(V - v_{\text{th}}(\mathbf{e}))] + v_{\text{th}}(\mathbf{e})\eta_{11}(\mathbf{e}).\end{aligned} \tag{A.20}$$

   $(\mu^*((1,1); \mathbf{e}, V), v_{\text{th}}(\mathbf{e}))$ is a saddle point if the previous inequality is true for every $\tilde{\mu}$. In particular, if we choose

$$\mu^*((1,1); \mathbf{e}, \cdot) = \underset{\tilde{\mu}((1,1); \mathbf{e}, \cdot)}{\arg\max}\ \mathbb{E}_V[\tilde{\mu}((1,1); \mathbf{e}, V)(V - v_{\text{th}}(\mathbf{e}))], \tag{A.21}$$

   then the right term is maximized when $\tilde{\mu} = \mu^*$, and the second inequality can be proved:

$$\begin{aligned}\mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)V] &\overset{?}{\geq} \mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)(V - v_{\text{th}}(\mathbf{e}))] + v_{\text{th}}(\mathbf{e})\eta_{11}(\mathbf{e}) \\ \Leftrightarrow 0 &\overset{?}{\geq} v_{\text{th}}(\mathbf{e})\Big(\eta_{11}(\mathbf{e}) - \mathbb{E}_V[\mu^*((1,1); \mathbf{e}, V)]\Big) = 0.\end{aligned} \tag{A.22}$$

Since $(\mu^*((1,1); \mathbf{e}, V), v_{\text{th}}(\mathbf{e}))$ is a saddle point, $\mu^*((1,1); \mathbf{e}, V)$ is a global minimum for the problem P (see Theorem 10).

Since $\mu^*((1,1); \mathbf{e}, \cdot)$ is defined as in equation (A.21), the structure in (3.11) is proved. □

Bibliography

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” vol. 40, no. 8, pp. 102–114, 2002. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1024422

[2] S. Cui, A. Goldsmith, and A. Bahai, “Energy-constrained modulation optimization,” vol. 4, no. 5, pp. 2349–2360, 2005. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1532220

[3] A. Sinha and A. Chandrakasan, “Dynamic power management in wireless sensor networks,” *IEEE Des Test Comput*, vol. 18, no. 2, pp. 62–74, 2001. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=914626

[4] S. J. Baek, G. de Veciana, and X. Su, “Minimizing energy consumption in large-scale sensor networks through distributed data compression and hierarchical aggregation,” vol. 22, no. 6, pp. 1130–1140, 2004. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1321225

[5] S. Pradhan, J. Kusuma, and K. Ramchandran, “Distributed compression in a dense microsensor network,” vol. 19, no. 2, pp. 51–60, 2002. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=985684

[6] S.-C. Woo and S. Singh, “Longest life routing protocol (LLRp) for ad hoc networks with highly mobile nodes,” in *Wireless Communications and Networking Confernce, 2000. WCNC. 2000 IEEE*, vol. 3, 2000, pp. 1306–1310. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=904821

[7] S. Ratnaraj, S. Jagannathan, and V. Rao, “Oedsr: Optimized energy-delay sub-network routing in wireless sensor network,” in *Networking, Sensing and Control, 2006. ICNSC ’06. Proceedings of the 2006 IEEE International Conference on*, 2006, pp. 330–335. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1673167

[8] S. Bandyopadhyay and E. Coyle, “An energy efficient hierarchical clustering algorithm for wireless sensor networks,” in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3, 2003, pp. 1713–1723. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1209194

[9] P. Nuggehalli, V. Srinivasan, and R. Rao, "Delay constrained energy efficient transmission strategies for wireless devices," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, 2002, pp. 1765–1772. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1019430

[10] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," vol. 9, no. 4, pp. 1326–1336, 2010. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5441354

[11] Z. Wan, Y. Tan, and C. Yuen, "Review on energy harvesting and energy management for sustainable wireless sensor networks," in *Communication Technology (ICCT), 2011 IEEE 13th International Conference on*, 2011, pp. 362–367. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6157897

[12] J. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," vol. 4, no. 1, pp. 18–27, 2005. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1401839

[13] E. H. Journal. Glossary: Energy harvesting. [Online]. Available: http://www.energyharvestingjournal.com/glossary/energy-harvesting-332.asp

[14] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," vol. 48, no. 5, pp. 1135–1149, 2002. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=995554

[15] R. Arroyo-Valles, A. Marques, and J. Cid-Sueiro, "Optimal selective forwarding for energy saving in wireless sensor networks," vol. 10, no. 1, pp. 164–175, 2011. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5618890

[16] N. Michelusi, K. Stamatiou, and M. Zorzi, "On optimal transmission policies for energy harvesting devices," in *Information Theory and Applications Workshop (ITA), 2012*, 2012, pp. 249–254. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6181793

[17] D. Del Testa, N. Michelusi, and M. Zorzi, "On optimal transmission policies for energy harvesting devices: the case of two users," in *Wireless Communication Systems (ISWCS 2013), Proceedings of the Tenth International Symposium on*, 2013, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6629759

[18] D. Del Testa and M. Zorzi, "Optimal policies for two-user energy harvesting device networks with imperfect state-of-charge knowledge," in *Information Theory and Applications Workshop (ITA), 2014*, 2014, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6804264

[19] A. Biason, D. Del Testa, and M. Zorzi, "Low-complexity policies for wireless sensor networks with two energy harvesting devices," in *13th IEEE IFIP Annual Mediterranean Ad Hoc Networking Workshop*, 2014, pp. 1–8.

[20] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," vol. 8, no. 2, pp. 547–551, 2009. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4786405

[21] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops, 2007. WiOpt 2007. 5th International Symposium on*, 2007, pp. 1–10. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4480065

[22] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," vol. 58, no. 7, pp. 2116–2126, 2010. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5504612

[23] F. Iannello, O. Simeone, and U. Spagnolini, "Energy management policies for passive RFID sensors with RF-energy harvesting," in *Communications (ICC), 2010 IEEE International Conference on*, 2010, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5502035

[24] R. Yates and H. Mahdavi-Doost, "Energy harvesting receivers: Optimal sampling and decoding policies," in *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE*, 2013, pp. 367–370. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6736891

[25] H. Mahdavi-Doost and R. Yates, "Energy harvesting receivers: Finite battery capacity," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, 2013, pp. 1799–1803. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6620537

[26] K. Tutuncuoglu and A. Yener, "Communicating with energy harvesting transmitters and receivers," in *Information Theory and Applications Workshop (ITA), 2012*, 2012, pp. 240–245. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6181795

[27] R. Srivastava and C. Koksal, "Basic performance limits and tradeoffs in energy-harvesting sensor nodes with finite data and energy storage," vol. 21, no. 4, pp. 1049–1062, 2013. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6334454

[28] C. K. Ho, P. D. Khoa, and P. C. Ming, "Markovian models for harvested energy in wireless communications," in *Communication Systems (ICCS), 2010 IEEE International Conference on*, 2010, pp. 311–315. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5686445

[29] N. Michelusi, K. Stamatiou, and M. Zorzi, "Transmission policies for energy harvesting sensors with time-correlated energy supply," vol. 61, no. 7, pp. 2988–3001, 2013. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6522422

[30] D. Bertsekas, *Dynamic programming and optimal control.* Athena Scientific, Belmont, Massachusetts, 2005.

[31] N. Michelusi, K. Stamatiou, L. Badia, and M. Zorzi, "Operation policies for energy harvesting devices with imperfect state-of-charge knowledge," in *Communications (ICC), 2012 IEEE International Conference on*, 2012, pp. 5782–5787. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6364958

[32] N. Michelusi, L. Badia, R. Carli, K. Stamatiou, and M. Zorzi, "Correlated energy generation and imperfect state-of-charge knowledge in energy harvesting devices," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2012 8th International*, 2012, pp. 401–406. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6314238

[33] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2009.