UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Departament de Teoria
del Senyal i Comunicacions

Centre
Tecnològic
de Telecomunicacions
de Catalunya

Transmission strategies for wireless energy harvesting nodes

By

Maria Gregori Casas

Submitted to the Universitat Politècnica de Catalunya (UPC)
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Barcelona, June 2014

Supervised by Dr. Miquel Payaró i Llisterri
PhD program on Signal Theory and Communications

Als meus pares, Pere i Anna,
us estaré eternament agraïda
per transmetre'm els vostres valors, amor i suport.

This work has been supported by the CTTC under the 2010 PhD scholarship program and by the Catalan Government (AGAUR) under the grants:
2011FI_B 00956, 2012FI_B1 00173, and 2013FI_B2 00075.

Abstract

Over the last few decades, transistor miniaturization has enabled a tremendous increase in the processing capability of commercial electronic devices, which, combined with the reduction of production costs, has tremendously fostered the usage of the Information and Communications Technologies (ICTs) both in terms of number of users and required data rates. In turn, this has led to a tremendous increment in the energetic demand of the ICT sector, which is expected to further grow during the upcoming years, reaching unsustainable levels of greenhouse gas emissions as reported by the European Council.

Additionally, the autonomy of battery operated devices is getting reduced year after year since battery technology has not evolved fast enough to cope with the increase of energy consumption associated to the growth of the node's processing capability.

Energy harvesting, which is known as the process of collecting energy from the environment by different means (e.g., solar cells, piezoelectric generators, etc.), has become a potential technology to palliate both of these problems. However, when energy harvesting modules are placed in wireless communication devices (e.g., sensor nodes or hand-held devices), traditional transmission strategies are no longer applicable because the temporal variations of the node's energy availability must be carefully accounted for in the design.

Apart from not considering energy harvesting, traditional transmission strategies assume that the transmission radiated power is the unique energy sink in the node. This is a reasonable assumption when the transmission range is large, but it no longer holds for low consumption devices such as sensor nodes that transmit to short distances. As a result, classical transmission strategies become suboptimal in short-range communications with low consumption devices and new strategies should be investigated.

Consequently, in this dissertation we investigate and design transmission strategies for Wireless Energy Harvesting Nodes (WEHNs) by paying a special emphasis on the different sinks of energy consumption at the transmitter(s).

First, we consider a finite battery WEHN operating in a point-to-point link through a static

channel and derive the transmission strategy that minimizes the transmission completion time of a set of data packets that become available dynamically over time. The transmission strategy has to satisfy causality constrains in data transmission and energy consumption, which impose that the node cannot transmit data that is not yet available nor consume energy that has not yet been harvested.

Second, we consider a WEHN that has an infinite backlog of data to be transmitted through a point-to-point link in a time-varying linear vector Gaussian channel and study the linear precoding strategy that maximizes the mutual information given an arbitrary distribution of the input symbols while satisfying the Energy Causality Constraints (ECCs) at the transmitter.

Next, apart from the transmission radiated power, we take into account additional energy sinks in the power consumption model and analyze how these energy sinks affect to the transmission strategy that maximizes the mutual information achieved by a WEHN operating in a point-to-point link.

Finally, we consider multiple transmitter and receiver pairs sharing a common channel and investigate a distributed power allocation strategy that aims at maximizing the network sum-rate by taking into account the energy availability in the different transmitters and a generalized power consumption model.

Resum

Durant les últimes dècades, la miniaturització del transistor i la reducció dels seus costos de fabricació han provocat un augment substancial del nombre de terminals de comunicacions i del tràfic de dades requerit per aquests dispositius. Així doncs, el consum energètic del sector de les Tecnologies de la Informació i Comunicacions ha incrementat notablement. A més a més, s'espera que aquest consum segueixi creixent durant els propers anys arribant a nivells insostenibles d'emissions de gasos d'efecte hivernacle segons ha informat el Consell Europeu.

D'altra banda, la tecnologia de les bateries no ha evolucionat suficientment ràpid com per fer front a l'augment del consum energètic associat al creixement de la capacitat de processament dels dispositius. Això ha ocasionat que l'autonomia dels dispositius que operen amb bateries empitjori any rere any.

Les energies renovables (per exemple, energia solar, cinètica, etc.) s'han convertit en una solució potencial per pal·liar aquests dos problemes. No obstant això, quan els dispositius de comunicació sense fils incorporen mòduls de captació d'energies renovables, les estratègies tradicionals de transmissió deixen de ser vàlides, ja que les variacions temporals de la disponibilitat d'energia en el dispositiu han de ser considerades en el disseny.

A més a més, les estratègies de transmissió tradicionals assumeixen que la potència radiada és l'única font de consum energètic del node. Aquesta és una suposició raonable per distàncies de transmissió llargues, però deixa de ser vàlida quan es consideren dispositius de baix consum que transmeten en distàncies curtes. Com a resultat, les estratègies de transmissió clàssiques són subòptimes en comunicacions de curt abast amb dispositius de baix consum i per això, s'han d'investigar noves estratègies.

En conseqüència, en aquesta tesi doctoral s'investiguen i es dissenyen noves estratègies de transmissió per nodes sense fils que operen amb energies renovables (WEHN) posant un èmfasi especial en les diferents fonts de consum d'energia en el transmissor.

En primer lloc, la tesi investiga l'estratègia de transmissió en un enllaç punt a punt a través d'un canal estàtic que minimitza el temps de transmissió d'un conjunt de paquets de dades que

s'adquireixen al llarg del temps. L'estratègia de transmissió ha de satisfer les limitacions per causalitat en la transmissió de dades i en el consum d'energia les quals imposen que el node no pot transmetre dades que no han estat encara obtingudes o utilitzar energia que encara no ha estat adquirida.

En segon lloc, es considera un WEHN que sempre disposa de dades per a transmetre a través d'un enllaç punt a punt en un canal lineal Gaussià amb variacions temporals. En aquest escenari i, també, donada una distribució arbitrària dels símbols d'entrada, s'estudia l'estratègia de precodificació lineal que maximitza la informació mútua alhora que satisfà la causalitat d'energia en el transmissor.

A continuació, a part de la potència radiada en transmissió, s'inclouen en el model de consum energètic els costos d'activació per accés al canal i per portadora. Donat aquest model, s'analitza com aquestes fonts de consum addicionals afecten a l'estratègia de transmissió que maximitza la informació mútua d'un WEHN que opera en un enllaç punt a punt.

Finalment, la tesi considera diversos parells transmissor i receptor que comparteixen un canal comú i investiga una estratègia d'assignació de potència distribuïda la qual té com a objectiu maximitzar la suma de les taxes de transmissió dels diferents nodes tenint en compte la disponibilitat energètica en cada transmissor que està basada en un model de consum energètic generalitzat.

Acknowledgements

It was four years ago, but it seems like it was only yesterday that I entered the Centre Tecnològic de Telecomunicacions de Catalaunya (CTTC) for the first time and Laura Casaus showed me the different offices and labs introducing me to my future colleagues; simply, time flies when one enjoys every single moment of every single day. Of course, all this joy is due to the great companions and friends at CTTC. In the following lines, I will try to thank you all; I hope I do not forget anyone. :)

First and foremost, I would like to thank my PhD advisor, Dr. Miquel Payaró i Llisterri, for his valuable help, suggestions, comments, and constructive criticisms that have greatly helped me to keep learning day after day. Miquel introduced me into the exciting worlds of convex optimization and linear precoder design and taught me the preciseness and rigor required in all the technical derivations. I also want to thank Miquel for being always flexible and comprehensive allowing me to select and study the problems for which I felt more interest. In short, I extremely appreciate Miquel's support both in professional and personal terms.

I also want to thank Professor Daniel Pérez Palomar for giving me the opportunity to join his research lab in the Hong Kong University of Science and Technology, treating me as one more of his PhD students, and letting me assist to his course on Convex Optimization. During this internship, I really consolidated several concepts of convex optimization and learned many approaches to solve nonconvex optimization problems.

Additionally, I thank Jesús Gómez, Kostas Stamatiou, Antonio Pascual-Iserte, and Miquel Calvo for reviewing part of this thesis.

Next, I would like to thank the CTTC both for the received funding and for giving me the opportunity to carry out my PhD thesis in such good conditions and work environment.

This amazing work environment is created by the great people working at CTTC; I thank all of them. Specially, I would like to thank the colleagues that have partially or fully overlapped their PhD studies with mine:

- First, I thank the guys in the coo- *wait for it* -lest office ever: Jaume Ferragut for always bringing up funny debates; Ana Maria Galindo for broadcasting happiness and joy in the office, making it a fu*nn*y Q-learning place; Moises Espinosa for his interesting early morning talks; the casanova of the lab, Angelos Antonopoulos, for constantly making fun of me; Pol Blasco for the funny terrace parties; Musbah Shaat, Vahid Joroughi and Alessandro Acampora for sharing with me a bit of their culture; Laia Nadal for being such a good person; Jessica Moysen for teaching us funny Mexican words that make us laugh for weeks; and Giussepe Cocco for his tremendous sense of humor.

- Second, the guys in the coo- *wait for it* -kie office: Iñaki Estella, Andrea Bartoli, Daniel Sacristan, Javier Arribas, Anica Bukva, Lazar Berbakov, Tatjana Predojev, Biljana Bojovic, Kyriaki Niotaki, Juan Manuel Castro, and Miquel Calvo.

- Finally, I also want to thank Onur Tan, Konstantinos Ntontin, Miguel Angel Vazquez, Adrià Gusi, Javier Rubio, Aleksandar Brankovic, Rakesh Kumar, Italo Atzeni and the Greek gang leaded by Alexandra Bousia.

Apart from the PhD students, I would like to thank all the players in the PMT pàdel tournament for the funny weekly matches; specially, I thank Edu Díaz for the organization.

També vull agraïr als meus amics i companys de viatge durant tots aquests anys. En primer lloc, a Judit Aliberas, Aida Abiad, Laia Simó, Adrià Barbosa, Carmina Gil i Eric Juan ja que cada moment que aconseguim passar els set junts és sempre únic i innolvidable. En segon lloc, a la Gemma Roger i en Lluís Pons pel vostre suport i amistat durant la meva estada a Hong Kong. Tot seguit, a en Lluís Parcerisa i la Neus Roca (*la nano-people*), i a la resta del Quelf-team, Emili Nogués, Ivan Pardina i Enric Pulido. A tot el grup de Valldemia, Montse Damont, Griselda Abril, Núria Farràs, Laura Ricós, Lluís Polo, Sandra Castañé, Aurora Gil, i Carla Aleña amb qui sempre aconseguim estar al dia. A la Gemma Molina i el Christian Roldán per la vostra amistat al llarg dels anys. Als konjos etíops, Clarà Macià, Laura Gonzàlez, Jesús Berdún, Llorenç Mercer i Kaja Miko ja que quan ens reunim és impossible parar de riure.

Per últim, vull agraïr i dedicar aquesta tesi a la meva família: Pere Gregori, Ana Casas i Xavi Gregori; Angelina Gómez i Pere Casas; Pilar Isla i Pere Gregori; Jordi Gregori; Nina Casas, Jordi Figueras, Marc Figueras i Aina Figueras; Joan Carles Gregori, Marta Palacios, Oriol Gregori i Sergi Gregori. Sense el vostre amor i suport aquesta tesi no seria possible.

List of Publications

Journals

[J1] M. Gregori, M. Payaró, G. Scutari, and D. P. Palomar, "Sum-rate maximization for energy harvesting nodes with a generalized power consumption model," *in preparation*, 2014.

[J2] M. Gregori and M. Payaró, "On the optimal resource allocation for a wireless energy harvesting node considering the circuitry power consumption," *accepted in IEEE Trans. on Wireless Communications*, Jun. 2014.

[J3] M. Gregori and M. Payaró, "On the precoder design of a wireless energy harvesting node in linear vector Gaussian channels with arbitrary input distribution," *IEEE Trans. on Communications*, vol. 61, no. 5, pp. 1868–1879, May 2013.
NEWCOM# 2013 Best young researcher's paper award.
Committee:
Muriel Médard, MIT Boston, USA
Petar M. Djurić, Stony Brook University, USA
Björn Ottersten, University of Luxembourg, Luxembourg.

[J4] M. Gregori and M. Payaró, "Energy-efficient transmission for wireless energy harvesting nodes," *IEEE Trans. on Wireless Communications*, vol. 12, no. 3, pp. 1244–1254, Mar. 2013.

International Conferences

[C1] M. Gregori and M. Payaró, "Multiuser communications with energy harvesting transmitters," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2014.

[C2] M. Gregori, A. Pascual-Iserte, and M. Payaró, "Mutual information maximization for a wireless energy harvesting node considering the circuitry power consumption," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2013, pp. 4238–4243.

[C3] M. Gregori and M. Payaró, "Throughput maximization for a wireless energy harvesting node considering the circuitry power consumption," in *Proceedings of the IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2012, pp. 1–5.

[C4] M. Gregori and M. Payaró, "Optimal power allocation for a wireless multi-antenna energy harvesting node with arbitrary input distribution," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2012, pp. 5794 –5798.

[C5] M. Gregori and M. Payaró, "Efficient data transmission for an energy harvesting node with battery capacity constraint," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2011, pp. 1–6.

[C6] M. Payaró, M. Gregori, and D. P. Palomar, "Yet another entropy power inequality with an application," in *Proceedings of the IEEE International Conference on Wireless Communications and Signal Processing (WCSP)*, Nov. 2011, pp. 1–5.

Contents

List of acronyms

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CDMA	Code Division Multiple Access
CSI	Channel State Information
CWF	Classical Water-Filling
DAI	Data Arrival Information
DCC	Data Causality Constraint
DMT	Discrete Multi-Tone
DSL	Digital Subscriber Line
DWF	Directional Water-Filling
EBS	Empty Buffers Strategy
ECC	Energy Causality Constraint
EH	Energy Harvesting
EHI	Energy Harvesting Information
FSA	Forward Search Algorithm
GSM	Global System for Mobile communications
$\mathcal{H}_g$**WF**	Mercury/Water-Filling
$\mathcal{H}_g$**WFA**	Mercury/Water-Filling Algorithm

ICT Information and Communications Technology

IFFT Inverse Fast Fourier Transform

i.i.d. independent and identically distributed

ISCA Iterative Smooth and Convex approximation Algorithm

IV Indicator Variable

KKT Karush–Kuhn–Tucker

MIMO Multiple-Input Multiple-Output

MMSE Minimum Mean-Square Error

MSE Mean-Square Error

MUI MultiUser Interference

NDA Non Decreasing water level Algorithm

OFDM Orthogonal Frequency Division Multiplexing

QoS Quality of Service

RF Radio Frequency

RFID Radio Frequency IDentification

SCA Successive Convex Approximation

SINR Signal to Interference plus Noise Ratio

SISO Single-Input Single-Output

SNR Signal to Noise Ratio

SSA Successive Smooth Approximation

WEHN Wireless Energy Harvesting Node

WLAN Wireless Local Area Network

w.r.t. with respect to

Notation

$\mathbb{R}, \mathbb{C}, \mathbb{N}$	The set of real, complex, and natural numbers, respectively.
$\mathbb{R}_+$	The set of nonnegative real numbers.
$\mathbb{R}^n, \mathbb{R}_+^n$	The set of vectors of dimension n with entries in $\mathbb{R}$ and $\mathbb{R}_+$, respectively.
$\mathbb{R}^{n\times m}, \mathbb{R}_+^{n\times m}, \mathbb{C}^{n\times m}$	The set of $n \times m$ matrices with entries in $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{C}$, respectively.
x	Scalar.
$\mathbf{x}$	Column vector.
$\mathbf{X}$	Matrix.
$\mathcal{X}$	Set.
$\|\|\mathbf{x}\|\|$	Euclidian norm of the vector $\mathbf{x}$, i.e., $\sqrt{\mathbf{x}^{\mathrm{H}}\mathbf{x}}$.
$\|\mathcal{X}\|$	Cardinality of the set $\mathcal{X}$.
$[\mathbf{x}]_n$	n-th component of the vector $\mathbf{x}$.
$[\mathbf{X}]_{pq}$	Element in the p-th row and q-th column of matrix $\mathbf{X}$.
$\mathbf{I}_n$	Identity matrix of order n.
$\mathbf{1}_n$, $\mathbf{0}_n$	Column vector of n ones or zeros. The dimension n might be omitted when it can be deduced from the context.
$(\cdot)^{\mathsf{T}}$	Transpose operator.
$(\cdot)^{\mathrm{H}}$	Conjugate transpose operator.
$\mathsf{Tr}(\cdot)$	Trace operator.
$\mathsf{diag}(\mathbf{X})$	Column vector that contains the diagonal elements of the matrix $\mathbf{X}$.
$\mathsf{Diag}(\mathbf{x})$	Diagonal matrix where the diagonal entries are given by the vector $\mathbf{x}$.
$\mathsf{vec}(\mathbf{X})$	Column vector that stacks the columns of $\mathbf{X}$.
$\mathbf{x} = (\mathbf{x}_n)_{n=1}^N$	Column vector constructed by stacking vectors $\mathbf{x}_n$, i.e., $\mathbf{x} = [\mathbf{x}_1^{\mathsf{T}}, \ldots, \mathbf{x}_N^{\mathsf{T}}]^{\mathsf{T}}$.
$\frac{\mathrm{d}f(x)}{\mathrm{d}x}$	Derivative of the scalar function $f(x)$ with respect to (w.r.t.) x.

$\frac{\partial f(x,y)}{\partial x}$	Partial derivative of the scalar function $f\,(x, y)$ w.r.t. x.
$\nabla_{\mathbf{x}} f(\mathbf{x})$	Gradient of the function $f\,(\mathbf{x})$ w.r.t. $\mathbf{x}$.
$\mathsf{D}_{\mathbf{X}}\mathbf{F}$	Jacobian of the matrix function $\mathbf{F}$ w.r.t. the matrix variable $\mathbf{X}$.
$\int_a^b f(x)\mathrm{d}x$	Integral of $f\,(x)$ w.r.t. x in the interval [a,b].
$\otimes$	Kronecker product.
$\circ$	Hadamard product.
$[x]_a^b$	Projection of x in the interval $[a, b]$.
$[x]^+$	Projection of x into $\mathbb{R}_+$, i.e., $\max\{0, x\}$.
$\mathbb{E}\{\cdot\}$	Expected value.
$\mathsf{var}\{\cdot\}$	Variance.
$\leq, \geq$	Smaller and greater than or equal inequality, respectively.
$\preceq, \succeq$	Componentwise smaller and greater than or equal inequality, respectively.
$\arg$	Argument.
$\mathrm{maximize}_{\mathbf{x}}\, f(\mathbf{x})$	Maximize the objective function $f(\mathbf{x})$ w.r.t. the optimization variable $\mathbf{x}$.
$\mathrm{minimize}_{\mathbf{x}}\, f(\mathbf{x})$	Minimize the objective function $f(\mathbf{x})$ w.r.t. the optimization variable $\mathbf{x}$.
$\mathbf{x}^\star$	Optimal value of a given optimization problem.
$\max, \min$	Pointwise maximum and minimum.
$\mathcal{H}^\ell(\cdot)$	Left continuous unit step function, i.e., $\mathcal{H}^\ell(x) = 1$ if $x > 0$ and $\mathcal{H}^\ell(x) = 0$, otherwise.
$\mathcal{H}^r(\cdot)$	Right continuous unit step function, i.e., $\mathcal{H}^r(x) = 1$ if $x \geq 0$ and $\mathcal{H}^r(x) = 0$, otherwise.
$\Pi(\cdot)$	Unit pulse in the interval $[0, 1]$, i.e., $\Pi(x) = 1$ if $x \in [0, 1]$ and $\Pi(x) = 0$, otherwise.
$\mathcal{W}_0(\cdot)$	Positive branch of the Lambert function.
$\log(\cdot)$	Natural logarithm.
$\log_b(\cdot)$	Base-b logarithm.
$\cap, \cup$	Intersection and union, respectively.
$\subset, \subseteq$	Proper subset and subset, respectively.
$\sim$	Distributed according to.
$\approx$	Approximately equal to.
$\mathbb{B}\,(n, q)$	Binomial distribution of parameters n and q.
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	Complex circularly symmetric Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$.
$\mathcal{N}(m, \sigma^2)$	Gaussian distribution with mean m and variance σ^2.

Chapter 1

Introduction

"The Information and Communications Technologies (ICTs) industry is in a unique position to demonstrate leadership in reducing its footprint, through structural change and innovation as well as by leading the way in identifying and creating efficient solutions for other socio-economic sectors to follow.

...

This can be done for instance by replacing products with on-line services (e.g. company newsletters), by moving business to the internet (e.g. costumer's support), by adopting new ways of working (tele-working and flexi-work enhanced by video-conferencing and tele-presence tools) and by exploring the viability of using green suppliers and energy from renewable resources."

The European Comission [1].

1.1 Motivation

The discovery of the transistor in 1947 revolutionized the field of electronics, becoming the fundamental component of current electronic devices. The transistor's inventors John Bardeen, Walter Brattain, and William Shockley were worldwide recognized with the Nobel Prize in Physics in 1956.

A few years later, in 1965, Gordon Moore accurately predicted that the number of transistors that can be placed in an integrated circuit would double every two years [2]. Since then, transistor miniaturization has enabled a tremendous increase in the processing capability of commercial electronic devices, which combined with the reduction of production costs has tremendously fostered the sales of electronic equipments. As a result, the usage of the ICTs has exponentially grown during last years both in terms of number of users and required data

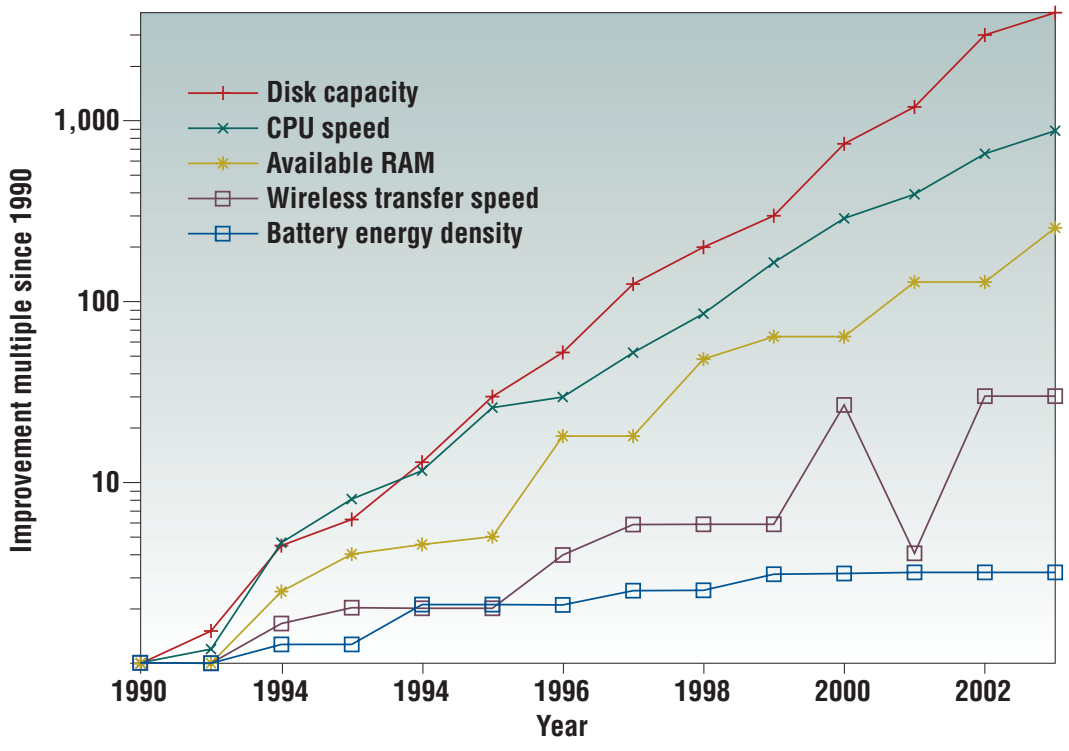

Figure 1.1: Relative improvements in laptop computing technology from 1990–2003 [4].

rates. This growth has in turn increased the energetic demand of the ICT sector; in 2008, the ICT community expended 7.15% of the global electricity bill [3] and the analysts predict that the energetic demand of the ICT sector will double by 2020 [3]. The unsustainability of this situation has yielded the European Council to target for 2020 a 20% reduction in emissions compared to 1990 levels and a 20% share of renewable energies in overall European Union energy consumption [1].

Another recent technological trend is that battery powered devices are becoming broadly used due to the high mobility provided to users and because they can be deployed in places in which there is no access to the power grid. Unfortunately, the battery autonomy of these devices drops year after year because, as depicted in Figure 1.1 [4], battery technology has not evolved fast enough to cope with the increase in energy consumption associated to the growth of the processing capability of the devices.

In this context, energy harvesting, which is known as the process of collecting energy from the environment by different means (e.g. solar cells, piezoelectric generators, etc.), has emerged as a potential technology both to expand the lifetime of battery powered devices by recharging their batteries and to reduce the carbon footprint in order to meet the 2020 targets of the European Council.

Energy harvesting has a great potential in a myriad of applications. For example, energy harvesters can be used to power devices with high mobility requirements (e.g., hand-held de-

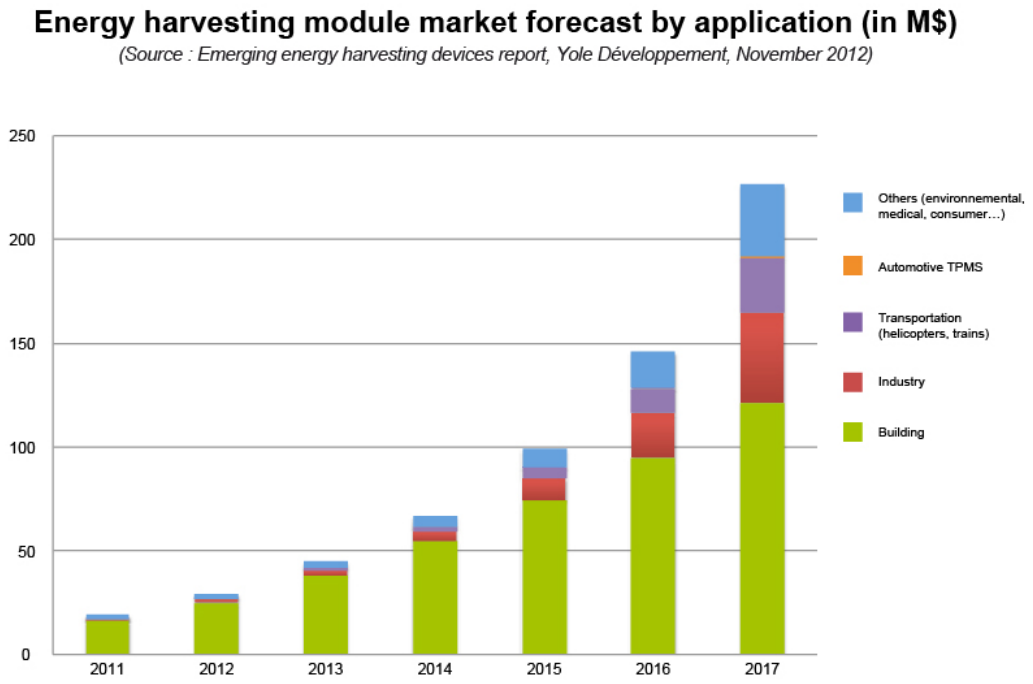

Figure 1.2: Expected sales of energy harvesting modules by application [5].

vices like cell-phones, laptops, tablets, etc.), which cannot be constantly connected to the power grid, or devices that are placed in remote locations where it is not practical to replace batteries (e.g., sensor nodes). Accordingly, energy scavenging modules will be ubiquitously deployed and used in networking applications such as wireless sensor networks, machine to machine communications, the Internet of Things, etc. This is confirmed in the market study performed by Yole Développement [5] that, as shown in Figure 1.2, predicts a tremendous increase on the sales of energy harvesting modules in the upcoming years. The main market drivers are the huge installation cost reduction (no wiring is needed to connect the devices to the power grid), and their being maintenance free (there is no need to replace the batteries) [5].

Up until very recently, many research efforts have focused on the design of communication systems and transmission strategies that are able to provide the exponentially increasing bit rate demand of the network users given some Quality of Service (QoS) requirements. As argued above, currently, users do not only require higher bit rates but also more battery autonomy and mobility, which can be obtained thanks to modern energy harvesting technologies. Nevertheless, when nodes are powered by energy harvesting modules, traditional transmission strategies, such as the well-known Classical Water-Filling (CWF) (which is presented in §2.2.1), are no longer applicable because the temporal evolution of the node's available energy must be carefully accounted for in the design. Accordingly, energy harvesting opens a new

research paradigm for the design of transmission strategies.

Additionally, classical transmission strategies generally consider that the transmission radiated power is the unique source of energy consumption of the node. This is a reasonable assumption when the transmission range is large, but it no longer holds for low consumption devices such as sensor nodes that transmit to short distances. As a result, classical transmission strategies become suboptimal in short-range communications with low consumption devices and new strategies should be investigated.

To summarize all what has been said above, the study and design of transmission strategies for energy harvesting nodes is required in order to enlarge the autonomy of battery operated devices and, at the same time, to reduce the carbon footprint of the ICT community; it is key that these transmission strategies not only account for the transmission radiated power but also for other relevant sources of energy consumption.

1.2 Outline of the dissertation and research contributions

This dissertation considers transmitter nodes equipped with energy harvesting modules, which palliate the battery autonomy problem, and investigates transmission strategies that take into account the energy availability variations in the node. More precisely, the thesis studies theoretical bounds on the best achievable performance in different scenarios as well as practical transmission strategies that can be implemented in Wireless Energy Harvesting Nodes (WEHNs). As depicted in Figure 1.3, the dissertation is structured in seven chapters and two different performance measures are considered: (i) transmission completion time minimization, which is studied in Chapter 3; and (ii) mutual information maximization, which is investigated in Chapters 4-6 for different transmitter architectures, network topologies, and sources of energy consumption.

The following lines summarize the contents of the different chapters.

Chapter 1

The current chapter has motivated the conducted research by answering the question "Why is the design of new transmission strategies for WEHNs necessary?" and now is presenting the outline and research contributions of this dissertation.

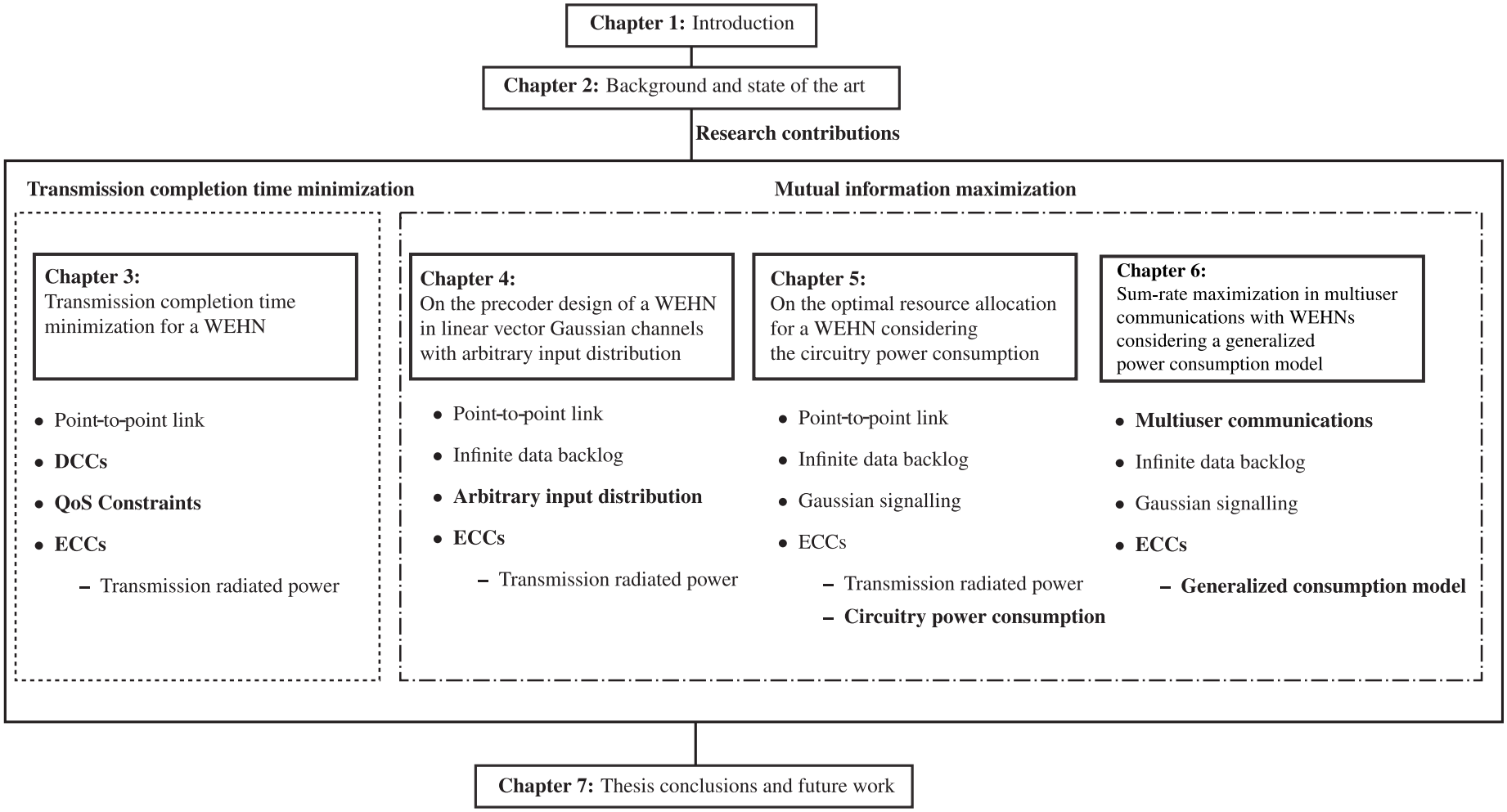

Figure 1.3: Outline of the dissertation.

Chapter 2

The second chapter introduces the main characteristics of WEHNs, presents the main facts that must be accounted for in the design of transmission strategies for WEHNs, and overviews the state of the art on known transmission strategies.

Chapter 3

The third chapter considers a point-to-point link where the transmitter is a WEHN that acquires the data along time. The chapter investigates the scheduling or power allocation strategy that minimizes the transmission completion time of all the data packets by using the harvested energy while satisfying some generic QoS constraints. Additionally, since both data and energy arrive dynamically to the node, the resource allocation strategy must satisfy a set of Data Causality Constraints (DCCs) and Energy Causality Constraints (ECCs), which are formally introduced in §2.1.5.

The contributions of this chapter were presented in the 2011 edition of the Global Telecommunications Conference [6] and published in a journal publication [7]:

- M. Gregori and M. Payaró, "Efficient data transmission for an energy harvesting node with battery capacity constraint," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2011, pp. 1–6.
- M. Gregori and M. Payaró, "Energy-efficient transmission for wireless energy harvesting nodes," *IEEE Trans. on Wireless Communications*, vol. 12, no. 3, pp. 1244–1254, Mar. 2013.

Chapter 4

The fourth chapter considers an energy harvesting node that has an infinite backlog of data to be transmitted through a point-to-point link in a linear vector Gaussian channel. Given an arbitrary distribution of the input symbols, the chapter investigates the linear precoding strategy that maximizes the mutual information while satisfying the ECCs at the transmitter. Accordingly, the linear precoding strategy must take into account that the mutual information of finite constellations asymptotically saturates.

The contributions of this chapter were presented in the 2012 edition of the International Conference on Communications [8] and published in 2013 as a journal publication [9], which obtained the 2013 best young researcher's paper award within the NEWCOM# project funded by the European Commission:

- M. Gregori and M. Payaró, "Optimal power allocation for a wireless multi-antenna energy harvesting node with arbitrary input distribution," in *Proceedings of the IEEE International Conference on Communications (ICC)*, Jun. 2012, pp. 5794 –5798.

- M. Gregori and M. Payaró, "On the precoder design of a wireless energy harvesting node in linear vector Gaussian channels with arbitrary input distribution," *IEEE Trans. on Communications*, vol. 61, no. 5, pp. 1868–1879, May 2013.
  NEWCOM# 2013 Best young researcher's paper award.
  Committee:
  Muriel Médard, MIT Boston, USA
  Petar M. Djurić, Stony Brook University, USA
  Björn Ottersten, University of Luxembourg, Luxembourg.

Chapter 5

The fifth chapter investigates the resource allocation strategy that maximizes the mutual information of a WEHN transmitting in a point-to-point link when the circuitry power consumption, e.g., the consumption of the Radio Frequency (RF) chain, is considered within the ECCs.

The contributions of this chapter were presented in two international conferences: in the 2012 edition of Vehicular Technology Conference [10], and in the 2013 edition of the Wireless Communications and Networking Conference [11]. Additionally, a journal publication has been recently accepted for publication [12].

- M. Gregori and M. Payaró, "Throughput maximization for a wireless energy harvesting node considering the circuitry power consumption," in *Proceedings of the IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2012, pp. 1–5.

- M. Gregori, A. Pascual-Iserte, and M. Payaró, "Mutual information maximization for a wireless energy harvesting node considering the circuitry power consumption," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2013, pp. 4238–4243.

- M. Gregori and M. Payaró, "On the optimal resource allocation for a wireless energy harvesting node considering the circuitry power consumption," *accepted in IEEE Trans. on Wireless Communications*, Jun. 2014.

Chapter 6

Chapter 6 considers the case where multiple transmitter and receiver pairs share a common channel and investigates a distributed transmission strategy that aims at maximizing the sum

of the achieved mutual information in the different links while satisfying the ECCs in all the transmitters. The ECCs consider a generalized power consumption model that accounts for a broad class of energy sinks such as the circuitry power consumption and the startup cost of the transmitter, which is associated to off-on transitions of the transmitter.

The contributions of this chapter will be soon submitted for journal publication:

- M. Gregori, M. Payaró, G. Scutari, and D. P. Palomar, "Sum-rate maximization for energy harvesting nodes with a generalized power consumption model," *in preparation*, 2014.

Chapter 7

The final chapter concludes the dissertation and points some possible future research directions.

Other research contributions

Some of the work performed during this PhD thesis has not been included in this dissertation; however, the results have been published in the following international conferences [13, 14]:

- M. Gregori and M. Payaró, "Multiuser communications with energy harvesting transmitters," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2014.

- M. Payaró, M. Gregori, and D. P. Palomar, "Yet another entropy power inequality with an application," in *International Conference on Wireless Communications and Signal Processing (WCSP)*, Nov. 2011, pp. 1–5.

Chapter 2

Background and state of the art

"Energy harvesting has grown from long-established concepts into devices for powering ubiquitously deployed sensor networks and mobile electronics. Systems can scavenge power from human activity or derive limited energy from ambient heat, light, radio, or vibrations."

Joseph A. Paradiso and Thad Starner [4].

This chapter is divided into three sections: the first section gives an overview of the structure of a WEHN and presents the key factors that influence on the design of transmission strategies for WEHNs; whereas, the second and third sections overview the state of the art on well-known transmission strategies for non-harvesting nodes and WEHNs, respectively.

2.1 Characteristics of wireless energy harvesting nodes

A WEHN is a battery operated device equipped with one or more energy harvesters that transmits data through a wireless medium.

The basic components of a WEHN, which is depicted in Figure 2.1, are:

- **Power transducer(s):** Different types of energy might be present in the surroundings of a WEHN (e.g., light, temperature or wind energy); the power transducer is the element in charge of converting these energy sources into usable electric power. Each energy source has associated a specific transducer and a wireless node can be equipped with different transducers, as it is later presented in §2.1.1.

- **Power storage system:** The electric energy at the output of the transducer is stored in the power storage system, typically a rechargeable battery, a supercapacitor, a solid state battery or hybrid solutions. Then, the different circuits of the node are powered with

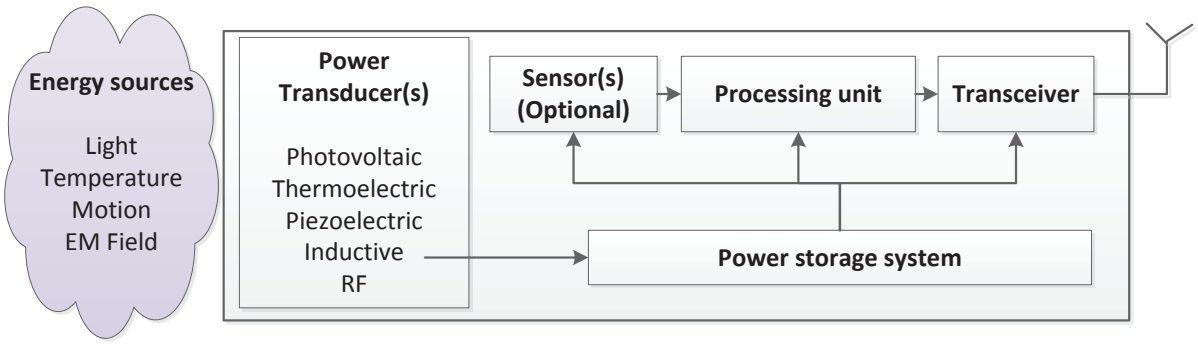

Figure 2.1: Structure of a WEHN.

the stored energy. This architecture is known as harvest-store-use [15, 16].[1] The storage system has a finite capacity, which might degrade after several charge and discharge cycles; additionally, the power storage system may suffer from energy leakage.

- **Sensors:** The WEHN might also contain sensors and actuators. Wireless sensors have a broad class of applications like light or motion detection, environmental monitoring, machine and process control, etc. [17, 18]. The sensed data is digitalized in an analog to digital converter and stored in the data buffer until it is transmitted.

- **Processing unit:** The processing unit regulates the operation of all the functional units of the node. Accordingly, it is the element in charge of elaborating smart energy management and transmission policies so that the WEHN achieves the best performance given some figure of merit [19]. Towards this goal, the processing unit must take into account, among others, the data arrival and energy harvesting processes, the capacities of the data buffer and energy storage systems, the communication channel and the power consumption of the different components.

- **Transceiver:** The circuitry that allows the transmission/reception of RF signals through the wireless medium.

As depicted in Figure 2.2, a point-to-point energy harvesting Multiple-Input Multiple-Output (MIMO) communication system is composed of three building blocks: the energy harvesting transmitter, the channel, and the receiver. The transmitter aims at sending a given message $\boldsymbol{\omega}$ to the receiver in an efficient and reliable manner by exploiting the available Channel State Information (CSI), Data Arrival Information (DAI), and Energy Harvesting Information (EHI). To do so, the transmitter encodes the message $\boldsymbol{\omega}$ in a sequence of symbols $\mathbf{x}_n \in \mathbb{C}^{n_T \times 1}$, $n = 1, \ldots, L$, that jointly compose the codeword $\mathbf{X}(\boldsymbol{\omega}) = [\mathbf{x}_1, \ldots, \mathbf{x}_L]$ of length

[1] There exists a different node architecture, referred in the literature as harvest-use, in which the power at the output of the transducer is directly used to supply the circuitry of the node [15, 16]; however, in this dissertation we focus on the harvest-store-use architecture.

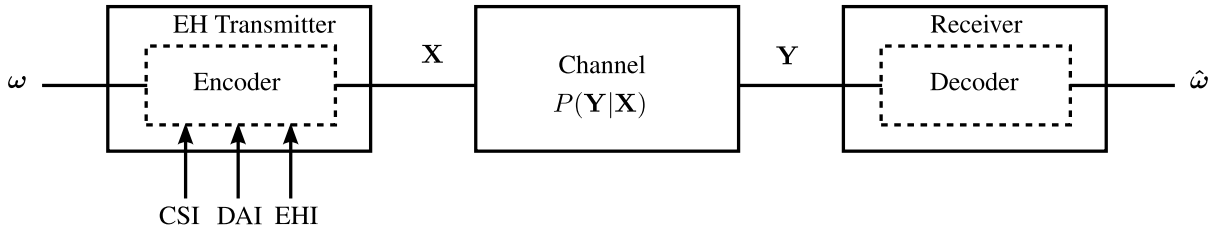

Figure 2.2: Point-to-point energy harvesting MIMO communication system.

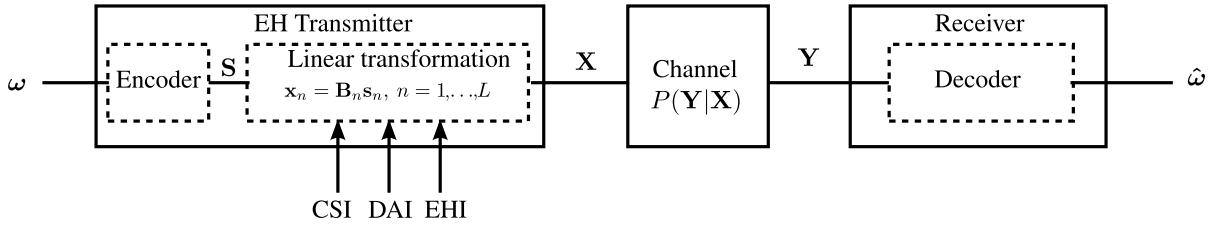

Figure 2.3: Point-to-point energy harvesting MIMO communication system with a linear transmitter.

L. These symbols are sequentially transmitted through the physical channel between the transmitter and the receiver, which adds impairments to the transmitted symbols. The receiver collects the sequence of received symbols $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L]$ with $\mathbf{y}_n \in \mathbb{C}^{n_R \times 1}$ and the decoder estimates the transmitted message given some decision rule $\hat{\boldsymbol{\omega}}(\mathbf{Y})$.

Due to the complexity of the encoder and decoder designs, it is practical in terms of simplicity and performance to impose a linear structure to the transmitter [20,21]. This means that the transmitter is divided in two blocks as depicted in Figure 2.3. The first block, the encoder, generates a sequence of data symbols $\mathbf{S}(\boldsymbol{\omega}) = [\mathbf{s}_1, \ldots, \mathbf{s}_L]$ with $\mathbf{s}_n \in \mathbb{C}^{n_S \times 1}$ from the information message independently of the CSI, DAI, and EHI, following the rules dictated by the code construction theory. The second block generates the transmitted symbols through a linear transformation of the data symbols by exploiting the available CSI, DAI, and EHI.

In this dissertation, we focus on the design of the linear transformation in the second block,

$$\mathbf{x}_n = \mathbf{B}_n \mathbf{s}_n,$$

that allows the transmitter to achieve the best performance given some figure of merit, where $\mathbf{B}_n \in \mathbb{C}^{n_T \times n_S}$ is usually referred to as linear precoding matrix.

In some scenarios, it is practical to force a diagonal structure to the precoding matrices, e.g., when the channel can be decomposed in a set of K parallel and independent subchannels. Then, the linear transmitter design reduces to a power allocation problem in which the system designer must decide the power, p_{kn}, allocated to each data stream. Thus, we have that

$$x_{kn} = \sqrt{p_{kn}} s_{kn}, \quad k = 1, \ldots, K,$$

where $K = n_T = n_R$ is the total number of independent data streams. In the following lines of this chapter, we often refer to linear precoding designs; however, the reader must recall that a power allocation problem is a particular case of a linear precoding design in which the precoding matrix is forced to have a diagonal structure.

One of the key challenges when dealing with the design of transmission strategies for WEHNs is that one must account for the temporal variations of the energy availability so that the node can operate with the best possible performance, which is not easy due to the random nature of the energy harvesting process. Among others, the following factors must be accounted for when designing transmission strategies for WEHN:

- Available energy sources and power harvesting profile of the node.
- The communication channel.
- Offline or online transmission strategies, which as explained later refer to the available knowledge of the CSI, DAI, and EHI.
- Sources of power consumption of the node.
- Chosen figure of merit and constraints.

These factors are thoroughly examined in the following subsections.

2.1.1 Energy sources and power harvesting profile

As presented in Figure 2.1, different types of energy can be present in the surroundings of wireless nodes, which can be harvested and transformed to electric power by using the appropriate conversion technologies. The characteristics of the available energy sources certainly affect the transmission strategy design. For example, an energy source can be either controllable or non-controllable; and non-controllable energy sources can be further sorted in predictable or unpredictable.

In the following lines, the characteristics of the different energy sources and the associated transducers are presented, which are summarized in Table 2.1.

2.1.1.1 Light energy

The most commonly exploited source of energy is light energy such as solar radiation or artificial light. The energy transducer is a solar panel, which is able to generate electricity through the photovoltaic effect; the amount of generated current is directly proportional to the light

Energy source	Characteristics	Harvested power
Light energy [22]		
Outdoor	Uncontrollable (predictable)	$10\,\mathrm{mW/cm^2}$
Indoor	Uncontrollable (predictable)	$10\,\mu\mathrm{W/cm^2}$
Thermal energy [22]		
Human	Uncontrollable (predictable)	$30\,\mu\mathrm{W/cm^2}$
Industrial	Uncontrollable (predictable)	$1 - 10\,\mathrm{mW/cm^2}$
Wind energy [23]	Uncontrollable (unpredictable)	$1\,\mathrm{W}$ at $2000\,\mathrm{rpm}$
Kinetic energy [16]		
Finger Motion	Controllable	$2.1\,\mathrm{mW}$
Footfalls	Uncontrollable (predictable)	$5\,\mathrm{mW}$
RF energy [22, 24]		
GSM	Uncontrollable (unpredictable)	$0.01 - 0.3\,\mu\mathrm{W/cm^2}$ at $25 - 100\,\mathrm{m}$
TV Broadcasting	Uncontrollable (unpredictable)	$0.1\,\mu\mathrm{W/cm^2}$ at $4\,\mathrm{km}$

Table 2.1: Characteristics of energy sources.

intensity, the area of the solar panel, and the efficiency of the converter that is around 10%-15% [16]. When the solar panel harvests solar energy in an outdoor environment under good weather conditions, the net electric power is around $10\,\mathrm{mW/cm^2}$; whereas, in an indoor environment, around $10\,\mu\mathrm{W/cm^2}$ can be harvested [22].

Solar energy is uncontrollable since it depends on the daily weather conditions, but it is predictable as later exposed in §2.1.1.6. Indoor light is also uncontrollable but predictable except in some specific system setups in which it is controllable, e.g., when the system designer controls the light source.

2.1.1.2 Thermal energy

Thermal energy can be harvested when the node is exposed to temperature gradients. For example, when the node has direct contact with the body, it can exploit the temperature gradient between the body and the surrounding environment [25]. Thermoelectric generators, or thermogenerators, are devices that convert temperature gradients into electrical energy by using the Seebeck effect [26]. In body applications, the temperature gradient is limited to around $15\,^{\circ}\mathrm{C}$, which limits the total harvested power to $30\,\mu\mathrm{W/cm^2}$ [22]. In opposition, higher temperature gradients are tolerated in industrial environments and the harvested power at the output of the thermoelectric generator is between $1\,\mathrm{mW/cm^2}$ and $10\,\mathrm{mW/cm^2}$ [22].

Thermal energy is in general uncontrollable but predictable; however, it depends on the specific placement of the thermogenerator.

2.1.1.3 Wind energy

A wind turbine can be used to scavenge energy from air flows. For example, the *Windlab Junior* turbine generates $1\,\mathrm{W}$ of output power when the turbine rotates at $2000\,\mathrm{rpm}$ [23,27]. A smaller

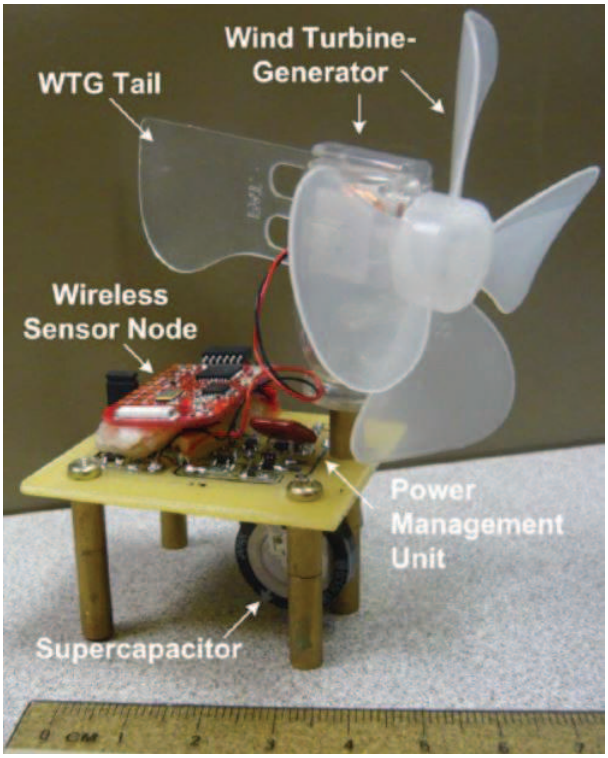

Figure 2.4: Wind turbine implemented in a wireless sensor node in [28].

wind turbine that provides an output power in the range of $7.3 - 55\,\mathrm{mW}$ was used in [28] to power the wireless sensor node depicted in Figure 2.4.

Wind energy is in general an uncontrollable and unpredictable energy source as fast variations in the air flow can easily occur. Nevertheless, if the airflow is generated by industrial applications, the harvested energy might be predictable or even controllable.

2.1.1.4 Kinetic energy

Vibrations are commonly encountered in bridges, roads, commercial buildings, automobiles, etc. Accordingly, movements or vibrations of objects are another potential source of energy for wireless nodes. The most common transducers to harvest vibrational energy are piezoelectric generators or electrostatic and electromagnetic converters [29–31].

The available energy might be due to uncontrollable environmental vibrations, human active or human passive. Human active sources require the user to perform a specific power generating motion and are generally controllable; for example, finger motion is able to produce an output power of $2.1\,\mathrm{mW}$ [16]. In opposition, human passive energy refers to the energy generated by humans in habitual gestures and movements, e.g., the heel impact in the floor while walking generates an output power of $5\,\mathrm{mW}$, which might be predictable [16].

2.1.1.5 RF energy

Ambient RF energy, e.g., from Global System for Mobile communications (GSM), Wireless Local Area Networks (WLANs), and TV broadcasting can be exploited to charge the battery of WEHNs. For example, RF energy is broadly used in active and passive Radio Frequency IDentification (RFID) technologies [32]. In [33], the authors showed that WLAN transmissions are able to power sensor nodes.

According to the Friis transmission equation, the ratio between the received power at the energy harvesting device and the radiated power in the power source is $G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2$, where G_t and G_r are the transmitter and receiver gains, λ is the wavelength of the signal, and d is the distance between the power source and the energy harvesting device. Therefore, RF harvesting is critically limited by the transmission range, d. For example, the power density at distances from $25\,\mathrm{m}$ to $100\,\mathrm{m}$ from a GSM base station ranges from $0.01\,\mu\mathrm{W/cm^2}$ to $0.3\,\mu\mathrm{W/cm^2}$ [24]; whereas the power density observed $4\,\mathrm{km}$ away from a TV broadcasting tower is $0.1\,\mu\mathrm{W/cm^2}$ [22].

RF energy is controllable in applications in which the power source is controlled by the system designer and uncontrollable (and very difficult to predict) when energy is harvested from spontaneous transmissions of other networks.

2.1.1.6 Power harvesting profile

Note that the definition of a WEHN, which is given at the beginning of this chapter, is generic and includes, among others, sensor nodes, hand-held devices (e.g., a cell-phone), or even a low-traffic base station powered with renewable energies as the one in Figure 2.5. The average data traffic and power consumption of these devices is completely different; for example, the power consumption of a wireless sensor is around $100\,\mu\mathrm{W}$ [22], a cell-phone consumes around $1\,\mathrm{W}$ [22], and a low traffic base station consumes around $150\,\mathrm{W}$ [34]. Thus, as shown in Table 2.1, current energy harvesting technologies seem specially suitable to power low power electronic circuits like sensor nodes. However, energy harvesting technologies can also be used in other applications if the energy harvesting and storage systems are dimensioned accordingly to the energetic demand of the node.

A key aspect in the design of transmission strategies for WEHNs is having knowledge of the *power harvesting profile*, i.e., the temporal evolution of the net harvested power at the output of the different transducers, which depends on the energy density available in each transducer and on its dimensions and efficiency. Having this knowledge is straightforward when all the energy sources are controllable, but it is a very difficult issue if some of the sources are uncontrollable. For predictable energy sources, it is feasible to implement prediction algorithms

Figure 2.5: Base station entirely powered with solar energy, which was used in the 2010 Mobile World Congress in Barcelona [34].

that allow the node to have an estimate of the future harvested energy; unfortunately, a little work has been done in the design of low complexity prediction algorithms that can be used in battery operated wireless nodes.

In this context, most of the efforts have focused on measuring or modeling the power harvesting profile of a solar harvester [15, 31, 35–43]. Within the Energy Harvesting Active Networked Tags (EnHANTs) project, a prototype that allows the measurement of the solar and kinetic harvested energies has been created [31, 35]; the obtained data is publicly available in [36]. This data is used in Figure 2.6 to depict the power harvesting profile obtained with a $100\,\mathrm{cm}^2$ solar panel with 10% conversion efficiency during 3 consecutive days. It is observed that the solar power harvesting profile has diurnal cycles, which can be exploited to create prediction models of the future harvested energy: In [37], the environmental energy harvesting framework was proposed to predict the harvested energy by means of an autoregressive filter, which takes into account the harvested energy in previous time instants; an exponentially weighted moving-average filter was used in [15] to exploit the diurnal and seasonal solar cycles; the authors of [44] proposed the weather conditioned moving average scheme; solar radiation models for predicting the average daily and hourly global radiation were proposed in [39–41] and references therein; finally, in [42], the authors proposed a solar (or wind) energy harvesting prediction model that uses the weather forecast.

When one aims to design transmission strategies for WEHNs, it is practical to use a discrete version of the power harvesting profile. Accordingly, the energy harvesting process is modeled as a set of energy packets arriving to the node at different time instants and with different amounts of energy: as depicted in Figure 2.7, the j-th energy packet arrives at time instant e_j seconds and a total of E_j Joules are harvested; and the term *epoch*, τ_j, is used to

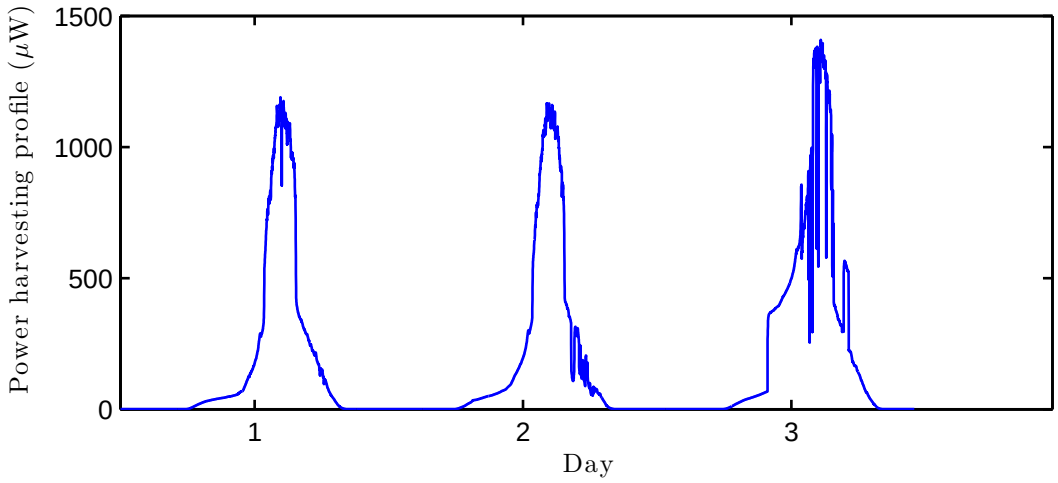

Figure 2.6: Solar power harvesting profile given a surface of $100\,\mathrm{cm}^2$ and 10% conversion efficiency. The data is taken from [36].

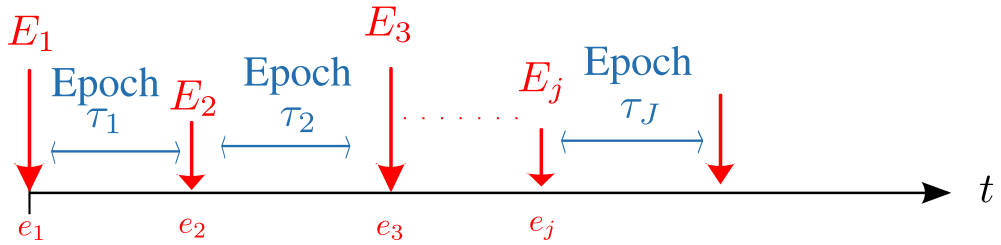

Figure 2.7: Packetized model of the energy harvesting process.

denote the period of time between two consecutive arrivals.

Note that this *packetized model* of the energy harvesting process can be understood as a sampling of the continuous power profile with the proper scaling factor to convert from powers to energies; this is depicted in Figures 2.8 and 2.9. First, Figure 2.8 depicts a time sampling of the continuous power harvesting profile in Figure 2.6 from 8 am to 1 pm of day 1 given a sampling window of 30 minutes. Then, assuming that the harvested power remains constant throughout the sampling window and that all the energy is collected by the end of the window in an energy packet, Figure 2.9 depicts the amplitude and arrival times of the different energy packets. It is important to remark that this packetized model of the energy harvesting process is able to capture any power harvesting profile by making the sampling window sufficiently small. Additionally, more samples can be taken in time intervals with strong variations of the power harvesting profile (e.g., from 11 am to 1 pm in Figure 2.8).

Additionally, this packetized model is also able to capture the energy harvesting process of nodes that have an hybrid storage system composed of a battery and a supercapacitor. Under this particular storage system, the harvested energy is first temporarily stored in the supercapacitor and then, when a substantial amount of energy is stored in the supercapacitor, it is

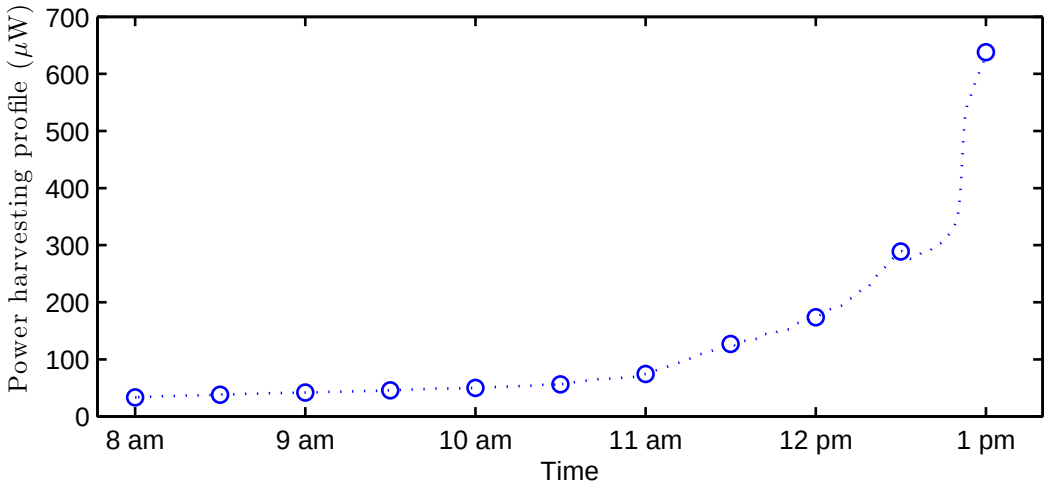

Figure 2.8: Sampling of the continuous power harvesting profile in Figure 2.6 from 8 am to 1 pm of day 1 given a sampling window of 30 minutes.

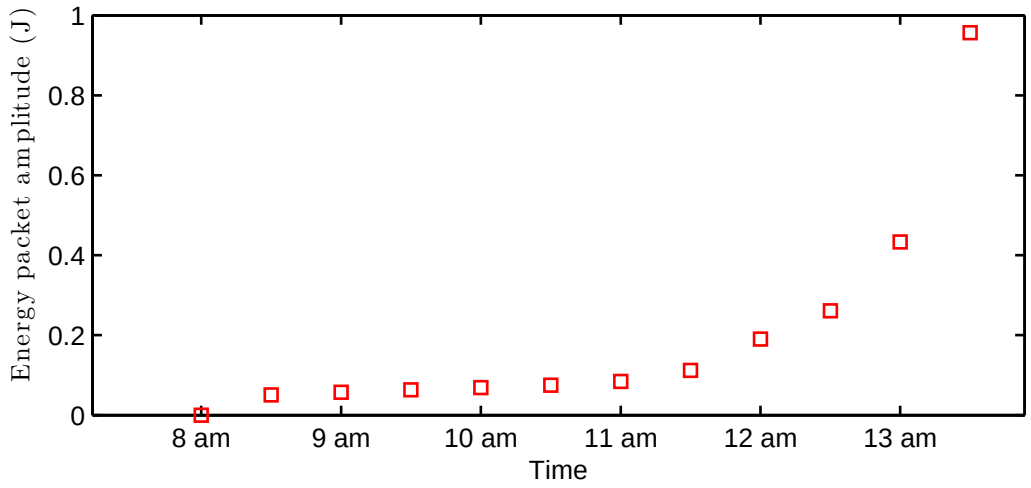

Figure 2.9: Packetized model of the power harvesting profile in Figure 2.8.

transferred to the battery as an energy packet. It has been shown that the discharge capacity and life cycle of Lithium-Ion batteries can be improved by recharging the battery in pulses [45], which motivates the use of a hybrid storage system.

Accordingly, throughout the dissertation, the energy harvesting process is characterized by a *packetized model*.

2.1.2 The communications channel

Throughout this dissertation, we have only considered memoryless channels in which the channel output probability distribution depends only on the current channel input. The following

lines present the channel models that are relevant to the following chapters.

2.1.2.1 The linear vector Gaussian channel

A linear vector Gaussian channel is such that the output of the channel, $\mathbf{y} \in \mathbb{C}^{n_R \times 1}$, is a linear function of the input, $\mathbf{x} \in \mathbb{C}^{n_T \times 1}$, i.e.,

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{w}, \tag{2.1}$$

where $\mathbf{G} \in \mathbb{C}^{n_R \times n_T}$ is the channel matrix and $\mathbf{w} \in \mathbb{C}^{n_R \times 1}$ is a circularly symmetric complex Gaussian random variable with zero mean and covariance matrix given by $\mathbf{R}_w$, $\mathbf{w} \sim \mathcal{CN}(0, \mathbf{R}_w)$, that models the thermal noise and other undesired effects in the receiving RF front-ends.

Many practical communication systems of interest can be modeled as a linear vector Gaussian channel. For example, in a multi-antenna wireless channel, n_T and n_R denote the number of antennas at the transmitter and receiver, respectively, and the element in the r-th row and t-th column of $\mathbf{G}$ denotes the channel coefficient between the t-th antenna at the transmitter and r-th antenna at the receiver. Other examples of systems that can be modeled as (2.1) are: Digital Subscriber Line (DSL); Code Division Multiple Access (CDMA); multicarrier systems like Orthogonal Frequency Division Multiplexing (OFDM) or Discrete Multi-Tone (DMT); etc. [20].

2.1.2.2 Parallel transmissions over Gaussian scalar channels

A particular case of special interest throughout the dissertation occurs when the channel matrix is diagonal, i.e., $\mathbf{G} = \mathsf{Diag}([g_1, \ldots, g_K]^{\mathsf{T}})$, since then the system model can be rewritten as a set of K parallel scalar channels:

$$y_k = g_k x_k + w_k, \quad k = 1, \ldots, K. \tag{2.2}$$

2.1.2.3 Interfering multiuser communications

In practice, it is common to encounter scenarios where several transmitter and receiver pairs must share the same channel. When this happens, the transmission of the different users generally interfere each other, which must be accounted for in the design of the transmission strategy of each user.

In Chapter 6, we consider a Gaussian interference channel composed of $\mathcal{T}$ transmitter and receiver pairs sharing the same band over SISO frequency-selective links composed of K

parallel subcarriers. In this setup, the received signal at the t-th receiver and k-th subchannel is

$$y_t(k) = g_{tt}(k)x_t(k) + \underbrace{w_t(k)}_{\text{Noise}} + \underbrace{\sum_{t' \neq t} g_{t't}(k)x_{t'}(k)}_{\text{MUI}}, \tag{2.3}$$

where $g_{tr}(k)$ is the channel value between the t-th transmitter and the r-th receiver at the k-th subcarrier; $x_t(k)$ is the transmitted symbol by user t at the k-th subcarrier; $w_t(k)$ denotes the noise at the t-th receiver and k-th subcarrier; and the term $\sum_{t' \neq t} g_{t't}(k)x_{t'}(k)$ is the MultiUser Interference (MUI) at the t-th receiver and k-th subcarrier.

Note that, in practical communication systems, several independent channel accesses are produced over time. Throughout this dissertation, we denote the temporal accesses with the index n. Accordingly, we will further index the expressions in (2.1), (2.2), and (2.3) by n.

2.1.3 Offline and online transmission strategies

There exist two well established approaches for the design of transmission strategies, which apply to both WEHNs and classical non-harvesting nodes, namely, online and offline. These approaches differ on the available knowledge at the transmitter of random parameters that influence transmission, e.g., CSI, DAI, and EHI.

In *offline transmission strategies*, the transmitter node has full knowledge (i.e., from the past, present, and future realizations) of these random parameters.

In opposition, *online transmission strategies* consider that the transmitter only has causal knowledge (i.e., from the past and present realizations) of these random parameters and maybe some statistical information regarding its future behaviour.

The study of offline transmission strategies is of key importance due to the following reasons.

(1.) The offline transmission strategy can be found independently of the specific choice of energy transducers in the node since, as introduced in §2.1.1.6, the packetized model of the energy harvesting process applies to any power harvesting profile. In opposition, online transmission strategies must account for the power harvesting profiles and statistical information of each specific energy source.

(2.) In some scenarios, it is indeed feasible to have full knowledge of these random parameters. For example, when the channel is time-static, the transmitter knows the arrival time of the data packets (e.g., a sensor that takes periodic measurements or when there is a sufficiently long backlog of data to be transmitted), and the energy source is controllable

or uncontrollable but predictable in the time window in which the transmission scheme is being designed (e.g., the solar power harvesting profile can be predicted with the models introduced in §2.1.1.6).

In other scenarios, where indeed the transmitter only has causal knowledge of these random parameters:

(3.) The optimal offline transmission strategy gives a bound on the achievable performance by any online strategy:

- When the transmitter is designed to maximize some utility function, then the optimal offline transmission strategy gives an upper bound on the achievable utility.
- In opposition, when the design objective is the minimization of a cost function, then the optimal offline transmission strategy gives a lower bound on the achievable cost.

(4.) In many cases, it provides analytical and intuitive solutions, which can be later used for the design of online transmission strategies.

Therefore, the derivation of the offline transmission strategy is a good first step to gain insight for the later design of the online transmission strategies.

2.1.4 Sources of power consumption

Traditional transmission strategies consider that the radiated power is the unique source of energy consumption at the transmitter (as it is later presented in §2.2). This is a reasonable assumption when the transmission power is large, which occurs in long-range communications, since it dominates over other sinks of energy. However, in certain applications, the transmission power might be comparable to the remaining energy sinks. For example, in energy efficient network topologies, transmission distances may be below $10\,\mathrm{m}$ and the circuitry energy consumption caused by the different components of the RF chain becomes relevant, even dominating over the radiated power [46, 47]. This often occurs to low-consumption devices such as sensor nodes that are powered by energy harvesting.

A more realistic power consumption model is given in [47–51], where the total consumed power is modeled as[2]

$$\mathcal{C}^1(p) = \begin{cases} \frac{\xi}{\eta} p + \alpha & \text{if } p > 0, \\ \delta & \text{if } p = 0, \end{cases} \tag{2.4}$$

[2]Different power consumption models are used through the dissertation. Accordingly, the i-th power consumption model is denoted as $\mathcal{C}^i$.

where p denotes the transmission radiated power; α accounts for both the power consumption of the digital to analog converter and for the consumption of the different components of the radio frequency chain, which includes the mixer, the filters, and the synthesizer; ξ and η are the power amplifier output back-off and drain efficiency, respectively [48]; and δ models the circuitry consumption when the transmitter is silent, which is much smaller than α.

For convenience, we introduce the following equivalent formulation of (2.4):

$$\mathcal{C}^1(p) = \frac{\xi}{\eta} p + (\alpha - \delta)\mathcal{H}^\ell(p) + \delta, \tag{2.5}$$

where $\mathcal{H}^\ell(x)$ is the left continuous unit step function defined as

$$\mathcal{H}^\ell(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases} \tag{2.6}$$

This model is able to capture the consumption of having the node "on", but it is still a clear simplification from reality because, among others, it ignores the power consumption of transitions between the "off" and "on" states, which are associated to the startup time of the transmitter [51].

As it is later presented in §2.2.1.3 and Chapters 5 and 6, the discontinuity of $\mathcal{C}^1$ at the origin substantially complicates the design of transmission strategies with respect to (w.r.t.) solely considering the transmission radiated power.

2.1.5 Figure of merit and constraints

There are several figures of merit or objective functions that can be investigated when designing transmission strategies. Some examples that have been considered in the literature are:

(1.) Minimization of the Mean-Square Error (MSE) [52, 53].

(2.) Minimization of the Bit Error Rate (BER) [53].

(3.) Minimization of the total energy consumption to transmit a certain amount of information by a given deadline [54, 55].

(4.) Maximization of the mean Signal to Interference plus Noise Ratio (SINR) [53].

(5.) Maximization of the mutual information [56–59].

(6.) Minimization of the total transmission time of a certain amount of information [60].

In this dissertation, we focus on the figures of merit 5 and 6 leaving the remaining ones as possible future research directions.

Given any figure of merit, the system performance is, in practice, limited by the available resources (e.g., a finite amount of energy, bandwidth, transmission time, etc.) or by some specific design requirements (e.g., a minimum Signal to Noise Ratio (SNR) at the receiver, a maximum delay constraint, etc.).

In the following lines, we summarize the common limitations that appear when designing transmission strategies for WEHNs.

- **Energy Causality Constraints (ECCs):** As it has been briefly introduced in the previous chapter, the presence of energy harvesters implies a loss of optimality of the traditional transmission policies for non-harvesting nodes because the ECCs must be taken into account, which impose that the energy cumulatively used by the node at a certain time instant must be no greater than the energy cumulatively harvested and stored on the battery. Accordingly, the ECCs mainly depend on the battery capacity (since energy lost due to battery overflows cannot be later used), the packetized model of the energy harvesting process (cf. §2.1.1.6), and the considered power consumption model (cf. §2.1.4). Other factors such as battery capacity degradation or leakages can also be considered in the ECCs (see , e.g., [61, 62]); however, this dissertation considers an ideal storage system. The ECCs can be imposed instantaneously or by averaging over the transmitted symbols; nevertheless, most of the works in the literature consider the averaged ECCs since it is commonly assumed that the dynamics of the energy harvesting process are much slower than the symbol duration.

- **Instantaneous mask constraints:** The temporal and/or spectral mask constraints limit the maximum transmitted power at a certain time instant and/or over a certain subcarrier. These constraints are generally imposed either by radio regulatory bodies or by the maximum output power at the power amplifier.

- **Data Causality Constraints (DCCs):** This set of constraints applies when the data to be transmitted is collected dynamically over time, which occurs, for instance, when the node takes periodic measurements of some event(s) through the sensor(s). Similarly to the ECCs above, the DCCs impose that the data cumulatively transmitted by the node at a certain time instant must be no greater than the data that have cumulatively arrived to the data buffer.

- **Finite data buffer constraint:** This constraint applies when the buffer to store the data to be transmitted is finite and imposes that no data is lost due to buffer overflows.

- **QoS constraints:** QoS constraints impose a given performance on the quality of the received message. For example, it might be convenient to bound the maximum transmission delay, the maximum probability of error or the minimum SNR.

2.1.6 Linear transmitter design problem formulation

The linear transmitter design problem can be, in general, mathematically formulated as the following optimization problem

$$\underset{\{\mathbf{B}_n\}_{n=1}^N}{\text{minimize}} \quad f\left(\{\mathbf{B}_n\}_{n=1}^N\right) \tag{2.7a}$$

$$\text{subject to} \quad g_i\left(\{\mathbf{B}_n\}_{n=1}^N\right) \leq 0, \quad i = 1, \ldots, m, \tag{2.7b}$$

where N denotes the number of channel accesses in which the transmitter is being designed; $\{\mathbf{B}_n\}_{n=1}^N$ is the set of precoding matrices from the channel access 1 to N, which are the optimization variables; f is the objective function; and the functions g_i denote a set of m inequality constraints. These functions depend on all the factors presented above (packetized model of the energy harvesting process, channel model, offline or online implementation, sources of power consumption, figure of merit, practical design constraints, etc.).

Note that for compactness we have formulated the transmitter design problem as a discrete-time linear precoding design; however, without loss of generality, it can also be formulated as a time continuous problem or as a power allocation problem (when the precoding matrix is forced to have a diagonal structure). In the remaining chapters of the dissertation, we deal both with continuous and discrete-time designs.

Convex optimization theory[3] dictates that if the functions f and g_i are convex, then the optimization problem in (2.7) is a convex program (or convex optimization problem) whose optimal solution can be found by standard convex optimization techniques [63].

Unfortunately, in many transmitter designs of interest, e.g., the problems explored in Chapters 3-6, the optimization problem in (2.7) is nonconvex and, accordingly, determining its solution is more involved.

Depending on the structure of the problem, one may still be able to derive a globally optimal solution to the nonconvex problem. For example, in some problems of interest, it is possible to reformulate the problem as an equivalent convex optimization problem that can be solved by convex optimization theory, e.g., by doing a change of variables. However, such a reformulation does not always exist; in these situations, a common approach in the literature

[3]The following chapters of this dissertation make use of convex optimization theory, which is not introduced here since it is currently well-known in the information and communication theory community; the interested reader is referred to [63].

is to relax some of the constraints to obtain a more tractable problem or to simply accept a feasible solution that performs close to the optimal one.

There exist several nonconvex optimization algorithms in the literature (e.g., gradient-based descend schemes [64], Successive Convex Approximation (SCA) algorithms [65–67], feasible sequential quadratic programming [68], parallel variable distribution [69], etc.), which are iterative and guarantee converge to some locally optimal solution.

When the constraints are convex, but the objective function is nonconvex, a common approach is to use gradient-based descend schemes, e.g., the Frank-Wolfe method [64], which linearizes the objective function around the current iterate and solves the resulting convex program to find the updated direction. In opposition, the algorithm in [66], which is based on SCA, convexifies only the nonconvex part of the objective function, thus exploiting any possible degree of convexity in the objective function, which results in a faster convergence speed. The algorithm in [67] generalizes the work in [66] to deal also with nonconvex constraints.

Throughout this dissertation, we explore different alternatives to deal with nonconvex transmitter design problems.

The following section presents the relevant state of the art of known transmission strategies for non-harvesting nodes.

2.2 State of the art on transmission strategies for non-harvesting nodes

2.2.1 Maximization of the mutual information

Channel capacity was originally derived by Shannon for single-user time invariant channels as the maximum mutual information between the channel input and output. Equivalently, the channel capacity is the maximum data transmission rate such that the probability of error can be made arbitrarily small by using a sufficiently long transmission block [70].

When the channel is time-varying, several definitions of channel capacity apply depending on the available CSI, e.g., the instantaneous capacity, the ergodic capacity or outage capacity (see the tutorial in [71] for more details). This dissertation considers the idealistic situation in which both the transmitter and receiver know perfectly and instantaneously the channel matrix, which happens when the channel is static or when the channel coherence time is much larger than the symbol duration so that it is constant for a sufficient number of transmissions. In such scenario, the channel can be considered as time invariant.

2.2.1.1 Classical Water-Filling (CWF)

The instantaneous capacity of the linear vector Gaussian channel in (2.1), i.e., $\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{w}$, is obtained by using a Gaussian code with transmission covariance matrix $\mathbf{Q}$, $\mathbf{x} \sim \mathcal{CN}(0, \mathbf{Q})$, as

$$\underset{\mathbf{Q}}{\text{maximize}} \quad I(\mathbf{x}; \mathbf{y}) \tag{2.8a}$$

$$\text{subject to} \quad \mathsf{Tr}(\mathbf{Q}) \leq \bar{P}, \tag{2.8b}$$

where $I(\mathbf{x}; \mathbf{y}) \triangleq \log\det(\mathbf{I}_{n_R} + \mathbf{R}_w^{-1}\mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{H}})$ is the mutual information [56] and the constraint in (2.8b) imposes a limitation on the average transmit power.

As shown by Telatar in [56], capacity can be achieved by diagonalizing the whitened channel matrix, which has the following Gram matrix $\mathbf{R}_G = \mathbf{G}^{\mathrm{H}}\mathbf{R}_w^{-1}\mathbf{G}$ that can be equivalently written in terms of its eigenvalue decomposition as $\mathbf{R}_G = \mathbf{U}_{R_G}\mathbf{D}_{R_G}\mathbf{U}_{R_G}^{\mathrm{H}}$, where $\mathbf{U}_{R_G}$ contains the eigenvectors of $\mathbf{R}_G$ and $\mathbf{D}_{R_G}$ contains the associated eigenvalues h_k, $k = 1, \ldots, K$, i.e., $\mathbf{D}_{R_G} = \mathsf{Diag}([h_1, \ldots, h_K]^{\mathsf{T}})$, where $K = n_T$. Accordingly, the channel is diagonalized when the transmit covariance matrix has the following structure

$$\mathbf{Q} = \mathbf{U}_{R_G}\mathbf{D}_Q\mathbf{U}_{R_G}^{\mathrm{H}},$$

where $\mathbf{D}_Q = \mathsf{Diag}([p_1, \ldots, p_K]^{\mathsf{T}})$.

As a result of diagonalizing the channel, the original linear vector channel is equivalent to a set of K parallel scalar channels whose capacity is obtained by solving the following power allocation problem:

$$\underset{\{p_k\}_{k=1}^K}{\text{maximize}} \quad \sum_{k=1}^{K} \log(1 + h_k p_k) \tag{2.9a}$$

$$\text{subject to} \quad \sum_{k=1}^{K} p_k \leq \bar{P}, \tag{2.9b}$$

where it has been used that the capacity of a scalar Additive White Gaussian Noise (AWGN) channel is $\log(1 + h_k p_k)$ [72]. Since the objective function is concave and the constraint is affine, (2.9) is a convex optimization problem and can be easily solved by means of the Karush–Kuhn–Tucker (KKT) optimality conditions [63] from where the CWF power allocation is obtained as

$$p_k^\star = \left[W - \frac{1}{h_k}\right]^+, \tag{2.10}$$

where h_k is the channel gain of the k-th parallel stream and W is the water level that is obtained by imposing that the power constraint must be satisfied with equality, $\sum_{k=1}^{K} p_k^\star = \bar{P}$.

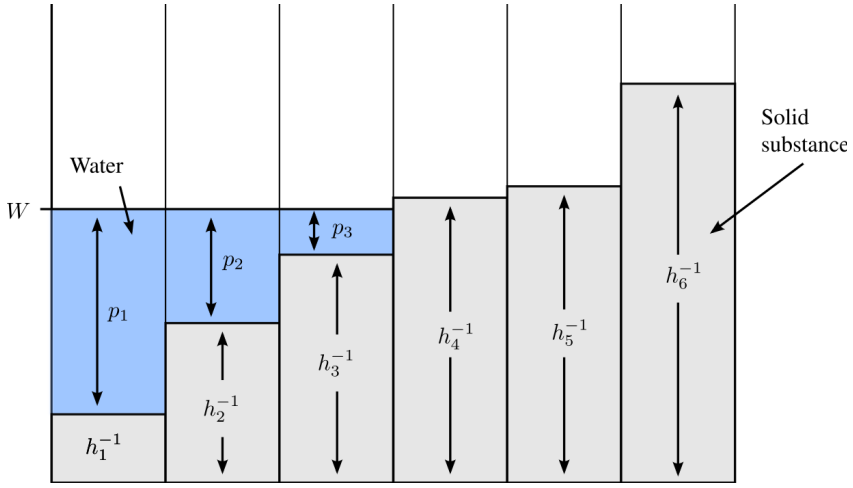

Figure 2.10: Graphical representation of CWF. In this specific representation, no power is assigned to the subchannels 4, 5, and 6.

The CWF solution accepts an intuitive graphical representation that is depicted in Figure 2.10 and proceeds as follows:

(1.) Each parallel channel is represented with a unit-base water-porous vessel.

(2.) Each vessel is filled with a solid substance up to a height equal to h_k^{-1}.

(3.) A volume of water equal to $\bar{P}$ is poured through the vessels.

(4.) Finally, the optimal power allocation in each parallel channel is the height of water in each vessel.

From the graphical representation, it is clear that higher power is assigned to the parallel channels (or eigenmodes) with higher channel gains. Additionally, zero power is assigned to those subchannels in which the water level does not reach the height $\frac{1}{h_k}$.

2.2.1.2 Mercury/Water-Filling ($\mathcal{H}_{\mathrm{g}}\mathrm{WF}$)

To derive the CWF solution in (2.10), the distribution of the input symbols is chosen to be Gaussian because it is the one that provides the highest mutual information [72]. Unfortunately, current technology cannot cope with Gaussian codes and finite constellations are used instead, e.g., $\mathcal{M}$-PAM and $\mathcal{M}$-QAM, where $\mathcal{M}$ denotes the alphabet cardinality.

In the low SNR regime, the capacity achieved with finite constellations is very close to the one achieved with Gaussian signaling. However, the mutual information asymptotically saturates when the SNR increases as no more than $\log_2 \mathcal{M}$ bit per channel use can be sent (see Figure 1 in [73]). Accordingly, in practical setups, the CWF solution in (2.10) is a theoretical limit that might be far from the real one [73].

This must be taken into account in the design of the optimal power allocation when the input symbols are constrained to belong to a finite alphabet. In opposition to the Gaussian case, where the better the channel gain, the higher the allocated power, when arbitrary constellations are used, the relation between the allocated power and the channel gain depends on the alphabet cardinality.

In [57], the authors derived the power allocation that maximizes the mutual information of a set of parallel channels, when the distribution of the input symbols is fixed and given. To do so, the authors of [57] used the relation between the mutual information and the Minimum Mean-Square Error (MMSE), which was revealed in [74] and further generalized in [75], as summarized in the following lines.

In [74], Guo et al. revealed that the derivative of the mutual information w.r.t. the SNR for a real-valued scalar Gaussian channel is proportional to the MMSE, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}h} I(x; \sqrt{h}x + w) = \frac{1}{2} mmse(h), \tag{2.11}$$

where x is the channel input, w is the observed noise with $w \sim \mathcal{N}(0,1)$ and $mmse(h) = \mathbb{E}\left\{(x - \hat{x})^2\right\}$, where $\hat{x} = \mathbb{E}\left\{x|\sqrt{h}x + w\right\}$ is the conditional mean estimator.

The mutual information in linear vector Gaussian channels was further characterized in [75], where its partial derivatives w.r.t. arbitrary system parameters were determined, e.g., the gradient w.r.t. the channel matrix, $\mathbf{G}$, was found to be

$$\nabla_{\mathbf{G}} I(\mathbf{x}; \mathbf{G}\mathbf{x} + \mathbf{w}) = \mathbf{G}\mathbf{E}, \tag{2.12}$$

where $\mathbf{E} = \mathbb{E}\left\{(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^{\mathrm{H}}\right\}$ is the MMSE matrix, and $\hat{\mathbf{x}} = \mathbb{E}\left\{\mathbf{x}|\mathbf{G}\mathbf{x} + \mathbf{w}\right\}$.

Thanks to the relationship in (2.11), the power allocation that maximizes the mutual information over a set of parallel channels (each of them denoted by a different index k) with finite alphabet inputs was derived in [57] and named Mercury/Water-Filling ($\mathcal{H}_{\mathrm{g}}\mathrm{WF}$), i.e.,

$$p_k^{\star} = \left[W - \frac{1}{h_k} G_k\left(\frac{1}{W h_k}\right)\right]^{+}, \tag{2.13}$$

where $G_k(\psi)$ is the mercury factor that depends on the input distribution and is defined as

$$G_k(\psi) = \begin{cases} \frac{1}{\psi} - mmse_k^{-1}(\psi) & \text{if } \psi \in [0,1], \\ 1 & \text{if } \psi \geq 1, \end{cases} \tag{2.14}$$

and $mmse_k^{-1}(\psi)$ is the inverse MMSE function, which returns the required SNR to achieve a certain MMSE ψ and depends on the input distribution. This result showed that the optimal

power allocation not only depends on the channel gain as in the Gaussian signaling case but also on the shape and size of the input constellation.

In particular, when the input symbols are drawn from a Gaussian distribution, $x \sim \mathcal{N}(0,1)$, then the MMSE function and the inverse MMSE function are

$$mmse(h) = \frac{1}{1+h} \tag{2.15}$$

and

$$mmse^{-1}(\psi) = \frac{1-\psi}{\psi}, \tag{2.16}$$

respectively. With this, it follows that $G_k(\psi) = 1, \forall \psi$, from where it is easy to check that the CWF solution in (2.10) is recovered after particularizing the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ in (2.13) for a Gaussian input distribution.

Interestingly, the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ solution also accepts a graphical representation that is depicted in Figure 2.11 and summarized in the following lines:

(1.) Each parallel channel is represented with a unit-base water-porous mercury-nonporous vessel.

(2.) Then, each vessel is filled with a solid substance up to a height equal to h_k^{-1}.

(3.) Given the optimal water level W, pour mercury in each vessel until the mercury level (i.e., the height of the mercury from the ground level) is $\frac{1}{h_k} G_k\left(\frac{1}{W h_k}\right)$.

(4.) Pour a volume of water equal to $\bar{P}$ (note that the resulting water level must be W).

(5.) Finally, the optimal power allocation in each parallel channel is found as the difference between the water and mercury levels.

Note that both the CWF (2.10) and the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ (2.13) are parametric solutions that depend on the water level W. In opposition to CWF, where the value of the optimal water level W is obtained directly from the graphical interpretation, in the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ solution the optimal water level W must be numerically computed in advance for the graphical interpretation to apply.

In Chapter 4, we generalize the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ solution to include energy harvesting at the transmitter and propose a graphical interpretation that does not require previous knowledge of the optimal water level.

It is important to remark that the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ presented above only applies when the channel and linear precoder matrices are diagonal. When general channel and precoder matrices are considered, finding the precoder that maximizes the mutual information is indeed quite

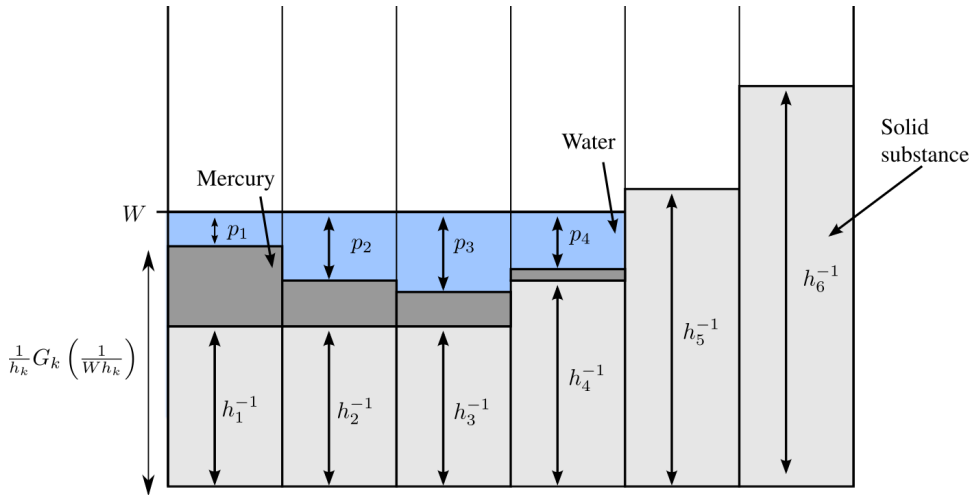

Figure 2.11: Graphical representation of $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$. In this representation, different modulation orders are used in the different streams, where the modulation order increases with the subchannel index k. In the subchannels $1-3$, it is observed that, under the same channel gain, the mercury level decreases with an increase of the modulation order and, as a result, more power is allocated to the streams with a denser constellation. Additionally, the power allocated to the fourth subchannel is greater than the power of the first subchannel in spite of having a worse channel gain.

involved as the resulting problem is nonconvex [58, 76–78]; this is argued in more detail in Chapter 4.

2.2.1.3 Glue pouring

Both CWF and $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ consider that the transmission radiated power is the unique source of energy consumption at the transmitter. As it has been argued in §2.1.4, this is a reasonable assumption for long-range communications, but it no longer holds for short-range transmissions.

In [50] and [51], the authors analyzed how CWF is modified when the circuitry consumption is taken into account. In particular, the authors considered the scalar AWGN channel,

$$y = x + w,$$

with Gaussian signaling[4]. Youssef-Massaad et al. showed that, due to the circuitry power consumption, it is preferable to transmit data during a fraction θ of the total available time.

[4]Verdú showed in [79] that pulse position modulation is optimal by transmitting arbitrarily high peaks whose position is used to carry information. This scheme has some inconveniences, e.g., the need of having the receiver "on" for the whole time; or the fact that, in practice, the instantaneous transmit power is limited by mask constraints due to the dynamic margin of the power amplifier. Accordingly, this justifies the interest of fixing the input distribution to be Gaussian [51].

Accordingly, the maximum mutual information is obtained as

$$\underset{\theta\in[0,1],p\geq 0}{\text{maximize}} \quad \theta\log(1+p) \tag{2.17a}$$

$$\text{subject to} \quad \theta(p+\alpha)\leq \bar{P}. \tag{2.17b}$$

Note that the average power constraint in (2.17b) uses the power consumption model in (2.4) where, without loss of generality, the authors considered $\frac{\xi}{\eta}=1$ and $\delta=0$.

The optimal solution to (2.17) is

$$\theta^{\star}=\min\left\{1,\frac{\bar{P}\mathcal{W}_0(e^{-1}(\alpha-1))}{(\alpha-1)(\mathcal{W}_0(e^{-1}(\alpha-1))+1)}\right\}, \tag{2.18a}$$

and

$$p^{\star}=\frac{\alpha-1}{\mathcal{W}_0(\frac{\alpha-1}{e})}-1, \tag{2.18b}$$

where $\mathcal{W}_0(\cdot)$ is the upper branch of the Lambert function [80]. Accordingly, it is shown that depending on the values of the power constraint $\bar{P}$ and the circuitry power consumption α, it is preferable to transmit information just during a fraction of the total available time and turn off the device during the remaining time. The authors named this solution as *glue pouring* because water is "glued" and does not occupy all the available degrees of freedom (in opposition to CWF where water spreads occupying all the available channel dimensions) [50].

In Chapter 5, we extend the glue pouring solution in (2.18) to cope with parallel streams at each channel access, which may have different fading levels. Additionally, we consider that the transmitter is equipped with energy harvesters. Moreover, in Chapter 6, we further generalize the power consumption model to include other sinks of energy consumption such as the startup power consumption entailed by off-on transitions of the transceiver.

2.2.2 Energy consumption minimization

The power allocation strategies derived in the previous section consider that the transmitter has an infinite amount of data to be transmitted at any given time instant. In certain applications, this does not hold since the data packets to be transmitted are acquired over time, e.g., when a sensor takes periodic measures of some process.

Consider a transmitter that collects a total of $\mathcal{D}$ data packets, where the i-th data packet arrives at the time instants $d_i\geq 0$ seconds containing $D_i>0$ bits with $i=1,\ldots,\mathcal{D}$, and let T denote the deadline constraint for the transmission of the $\sum_{i=1}^{\mathcal{D}}D_i$ bits of information.

In this setup, the authors of [54] derived the transmission strategy that minimizes the energy consumption of the node under the DCC (cf. §2.1.5). The authors exploited the fact

that energy consumption can be reduced by decreasing the transmission power or, equivalently, the rate (since for most coding and modulation schemes the achievable rate increases with the transmission power). In subsequent works, their problem was extended for different scenarios: In [81] variable length packets were considered, while [82] considered a fading channel.

The authors of [55, 83] analyzed the same problem under a different point of view, i.e., by defining the concept of *cumulative curves*, which allows the derivation of a graphical and intuitive solution that is presented next. Apart from the deadline constraint T and the DCC, Zafer and Modiano considered that the node has to satisfy some QoS constraint on the delivery of the data bits. To do so, the authors defined the following cumulative curves:

Definition 2.1 (Data Departure Curve)**.** *A data departure curve $D(t) \geq 0$, $t \geq 0$, is the total number of bits that have been cumulatively transmitted by the node in the time interval $[0, t]$.*

Definition 2.2 (Accumulated Data)**.** *The accumulated data $D_A(t)$ is the sum of data that has arrived at the node during the time interval $[0, t)$, i.e., $D_A(t) = \sum_{i=1}^{\mathcal{D}} D_i \mathcal{H}^{\ell}(t - d_i)$.*

Definition 2.3 (Minimum Data Departure)**.** *The minimum data departure, $D_{QoS}(t)$, is the smallest amount of data that the node must have cumulatively transmitted at time t such that the QoS constraint is satisfied.*

Different QoS constraints can be considered by mapping the constraint into an appropriate minimum data departure curve. Let us briefly describe two examples, introduced in [55], of how to map a certain QoS constraint to the minimum data departure curve:

- Deadline constraint: This constraint considers that the maximum permissible delay for the transmission of a certain data packet, D_k, is θ_k seconds. Then, $D_{QoS}(t)$ is a piecewise constant function that changes at instants $q_k = d_k + \theta_k$ with an increment of D_k. As a specific case, we can consider that the allowed delay for all the packets is the same, i.e., $\theta_k = \theta, \forall k$. Then, the minimum data departure is given by $D_{QoS}(t) = D_A(t - \theta)$.
- Finite data buffer constraint (cf. §2.1.5): Consider that the transmitter has a limited data queue of size $D_{\max}$. Then, in order not to lose any incoming data, the minimum data departure must be $D_{QoS}(t) = [D_A(t) - D_{\max}]^+$.

Accordingly, the energy minimization problem is to obtain the data departure curve with the least energy expenditure that satisfies the DCC, $D(t) \leq D_A(t)$, and the QoS constraint, $D_{QoS}(t) \leq D(t)$:

$$\underset{D(t)}{\text{minimize}} \quad \int_0^T g\left(\frac{\mathrm{d}D(t)}{\mathrm{d}t}\right) \mathrm{d}t \tag{2.19a}$$

$$\text{subject to} \quad D_{QoS}(t) \leq D(t) \leq D_A(t), \quad t \in [0, T], \tag{2.19b}$$

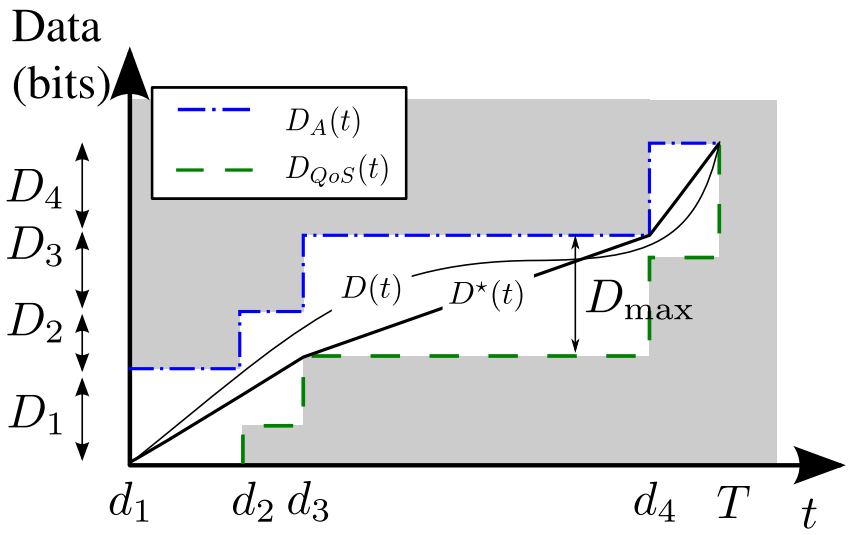

Figure 2.12: Graphical representation of the problem in (2.19). The optimal data departure curve is the tightest string tied at the origin and at $(T, D_A(T))$.

where $\frac{\mathrm{d}D(t)}{\mathrm{d}t}$ is the instantaneous transmission rate, which is the time derivative of the data departure curve, $D(t)$; $g(\cdot)$ is a time-invariant, strictly increasing, convex power-rate function; accordingly, $g(\frac{\mathrm{d}D(t)}{\mathrm{d}t})$ is the instantaneous power consumption and its integral in $[0, T]$ returns the total energy consumption. This problem is depicted in Figure 2.12 for the finite data buffer constraint. Note that a feasible data departure curve must lie within the blank region $\forall t$. Zafer and Modiano showed that independently of the specific power-rate function (as long as it is time-invariant, strictly increasing, and convex), the optimal data departure curve, $D^\star(t)$, is the one with the shortest length among all the feasible curves, which can be visualized as the tightest string tied at the origin and at $(T, D_A(T))$ when the constraints are hard boundaries.

2.3 State of the art on transmission strategies for WEHNs

When the transmitter has energy harvesting capabilities, all the strategies presented above are no longer valid because the ECCs must be taken into account in the design (cf. §2.1.5).

During the last years, several works have explored transmission strategies for energy harvesting nodes as summarized in the following lines.

2.3.1 Directional Water-Filling (DWF)

Given Gaussian distributed input symbols, the power allocation strategy that maximizes the mutual information of a WEHN operating in a point-to-point link during N consecutive channel

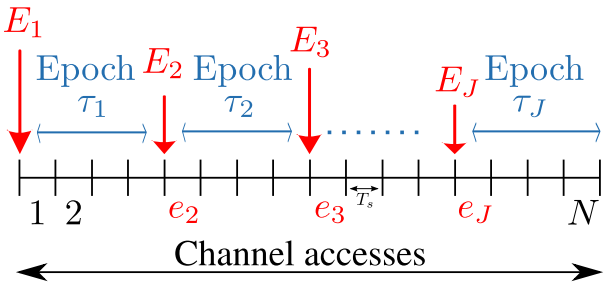

Figure 2.13: Temporal representation of the energy arrivals along channel accesses.

accesses of duration T_s is obtained as the solution to

$$\underset{\{p_n\}_{n=1}^N}{\text{maximize}} \quad \sum_{n=1}^{N} \log\left(1 + h_n p_n\right) \tag{2.20a}$$

$$\text{subject to} \quad T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} p_n \leq \sum_{j=1}^{\ell} E_j, \quad \ell = 1, \ldots, J, \tag{2.20b}$$

where the ECCs in (2.20b) assume that the energy packet arrival times are aligned at the beginning of a channel use[5] (see Figure 2.13) and impose that the total energy consumed by the end of the ℓ-th epoch is no greater than the total energy harvested at the beginning of the ℓ-th epoch.

The optimal power allocation of the problem in (2.20), which is named DWF and was initially derived in [84] for a time continuous channel, is

$$p_n^\star = \left[W_j - \frac{1}{h_n}\right]^+, \quad n \in \tau_j, \tag{2.21}$$

where n is the channel use index, and W_j is the water level of the j-th epoch, which can be found by means of the KKT optimality conditions [63]. The main difference between DWF (2.21) and CWF (2.10) is that in the former the water level is allowed to change across the different epochs according to the node energy availability.

The DWF solution also accepts a graphical representation, which is shown in Figure 2.14:

(1.) Each channel access is represented with a water-porous vessel with base equal to T_s.[6]

(2.) Each vessel is filled by a solid substance up to a height equal to h_n^{-1}.

(3.) A water *right-permeable* material is used to separate the different epochs.

[5] If an energy packet arrives in the middle of a channel access, one can always assume that the packet becomes available for the transmitter at the beginning of the following channel access since the transmission strategy can only be changed in a channel access basis.

[6] The vessel boundaries are not depicted in Figure 2.14 for the sake of simplicity.

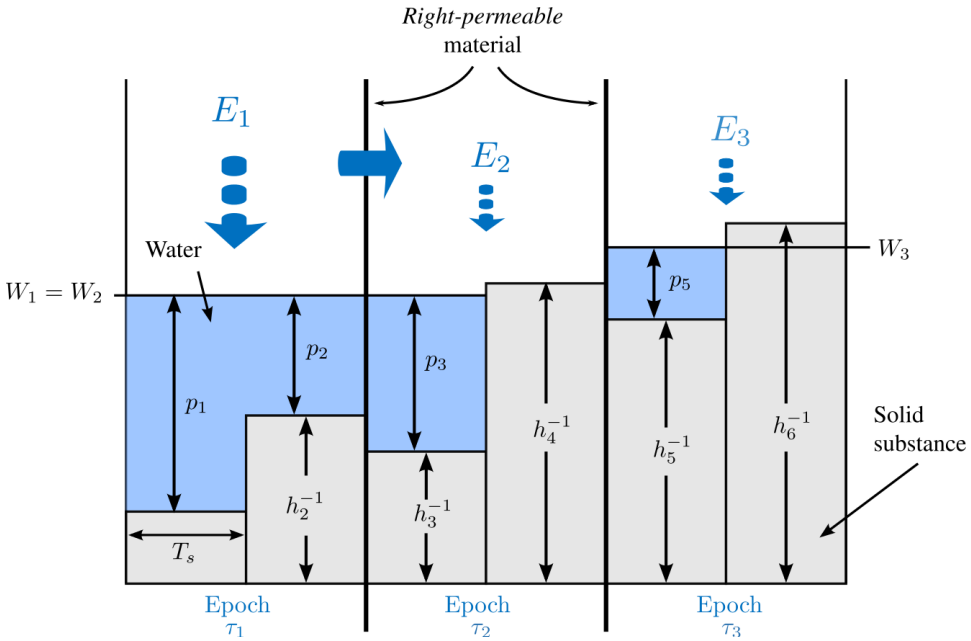

Figure 2.14: Graphical representation of the DWF. The figure represents the optimal power allocation when $E_1 > E_2 = E_3$. It is observed that some water flows from the first epoch to the second one.

(4.) Given that water volume represents energy and water height corresponds to power, the water level is progressively increased to all epochs at the same time by adding the necessary amount of water to each epoch. The maximum amount of water that can be externally added at some epoch is given by the epoch's harvested energy (depicted with the top-down arrows in Figure 2.14). When some epoch runs out of water, it uses water that flows from previous epochs (if any is available) in order to continue increasing the water level simultaneously.

(5.) When all the available water has been poured, the optimal power allocation is found as the height of the water in each vessel, i.e., $p_n^\star = [W_j - h_n^{-1}]^+$.

2.3.2 Other transmission strategies for WEHNs

Apart from the DWF [84], several works have recently considered the design of transmission strategies for WEHNs both for point-to-point links and multiuser scenarios. In the following lines, we briefly summarize some of these works:

Point-to-point links

In [85], the authors derived online and offline optimal policies by using dynamic programming and convex optimization techniques. The authors of [86] studied the coding problem from

an information theoretic perspective. The transmission policy that maximizes the throughput in a finite horizon for a finite battery capacity was derived in [87] assuming an infinite backlog of data at the transmitter. In [88], different resource allocation strategies were evaluated and compared by using real energy traces, which were collected in [35]. A learning theoretic approach was taken in [89], where data and energy arrivals were modeled as finite-state Markov processes. Markov models were used in [90] to model the battery state of the node. Finally, the stability and delay of the data queue were considered in [91] to derive the optimal transmission policies.

The concept of cumulative curves, cf. §2.2.2, has been broadly used to derive optimal transmission strategies for WEHNs. In this context, the authors of [92] considered dynamic data packet arrivals and found the transmission strategy that minimizes the delivery time of all data packets under the assumption of having infinite battery capacity. In [87], a node with finite battery capacity was studied; however, it was considered that all the data packets are available from the beginning of the transmission.

The works mentioned above consider that the transmission radiated power is the only energy sink of the node. In opposition, the works in [49, 93] derived optimal power allocation strategies by using the power consumption model $\mathcal{C}^1$ in (2.5) for a time-static and fading continuous channel, respectively.

Multiuser communication systems

Several works have explored multiuser scenarios:

- **Broadcast channel:** The minimization of the transmission completion time for a WEHN operating in a broadcast link was considered in [94–97]. References [94] and [95] assumed infinite battery capacity, whereas, the authors of [96,97] found the rate scheduling policy of the finite battery capacity case.

- **Multiple access channel:** The capacity region of a Gaussian multiple access channel was considered in [98]. Under an offline approach, [99] proposed the power allocation strategy that minimizes the transmission completion time. The stability of the packet queues was studied in [100]. The authors of [101] proposed an online algorithm to maximize the long-term average channel throughput. Low-complexity scheduling policies to maximize the sum throughput were derived in [102] without the knowledge of the power harvesting profile.

- **Relay and multi-hop channels:** Two-hop communications were considered in [103–109]. The optimal offline transmission scheme was proposed in [103] for the full-duplex

tow-hop relay channel. The optimal offline energy management policies for the source and the relay that maximize the end-to-end throughput were studied in [104] by allowing energy cooperation between the source and the relay. The authors of [109] proposed the power allocation strategy that maximizes the system throughput over a finite number of transmission time slots under conventional decode-and-forward relay (the relay retransmits the signal in the consecutive time slots) and under the buffer-aided link adaptive relaying (where the relay can retransmit the signal in any of the consecutive time slots). The relay channel (with direct link between transmitter and destination) was studied in [110–112].

- **Cooperative transmissions:** The authors of [45, 113] considered multiple energy harvesting sensors that transmit a shared common message to a distant base station and studied the power allocation strategy in the different nodes that maximizes the total data throughput.

- **Interference channel:** The power allocation that maximizes the sum throughput with a given deadline for the two-user Gaussian interference channel with energy harvesting transmitters was investigated in [114]. The two-user Gaussian interference channel was studied in [115] by considering the cost of having the transmitter "on", i.e., by using the power consumption model $\mathcal{C}^1$ in (2.5).

In the following chapter, we have generalized the concept of cumulative curves introduced by Zafer and Modiano in [83] to deal with energy harvesting at the transmitter, which has allowed us to derive the scheduling strategy that minimizes the transmission completion time while satisfying DCCs, ECCs, finite battery capacity constraint, and QoS constraints.

Chapter 3

Transmission completion time minimization for a WEHN

3.1 Introduction

This chapter considers a WEHN that acquires some data measurements over time, e.g., a sensor that eventually or periodically senses a given process like temperature, humidity, etc. Each measurement is digitalized in a sequence of bits that conform a data packet, which is stored in the data buffer until it is transmitted. In practice, the data and energy buffers can only store a finite amount of data and energy, respectively. In this context, this chapter investigates the transmission strategy that minimizes the transmission completion time of a finite amount of incoming data packets (the time by which all the data packets are transmitted), which arrive dynamically over time as generally occurs in wireless sensor nodes.

As it is presented throughout this chapter, the transmission completion time minimization problem for WEHNs is closely related to the energy minimization problem for non-harvesting nodes. In particular, the concept of cumulative curves, which has been presented in §2.2.2, is a useful tool to derive optimal transmission strategies for WEHNs.

As it has been presented in the previous chapter, given that energy harvesters are used to power the transmitter(s), the minimization of the transmission completion time was considered in [84, 87, 92] for a point-to-point link and in [94, 96, 99] for multiuser scenarios, as it is explained in more detail in the following lines.

The authors of [92] considered dynamic data packet arrivals and found the transmission strategy that minimizes the transmission completion time by assuming an infinite battery capacity.

The optimal transmission policy for a finite battery capacity node was derived in [87] for

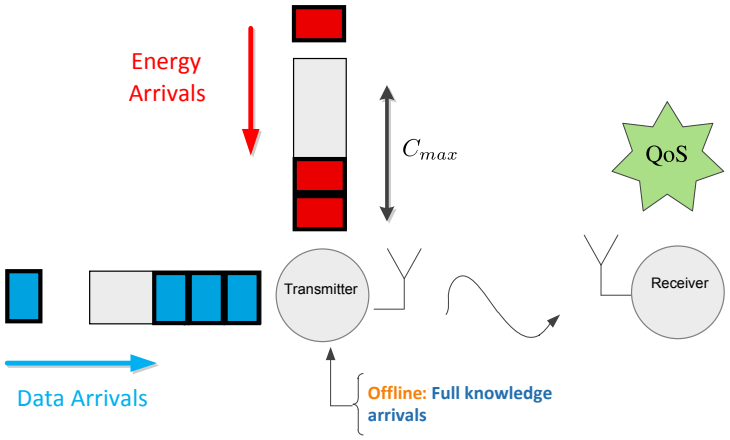

Figure 3.1: System model.

a time-static channel and in [84] for a time-varying channel. Both of these works assume that the bits to be transmitted are available from the beginning, which significantly simplifies the setup as losing energy due to battery overflows is clearly suboptimal because there is always data to be transmitted.

Again assuming that all the data bits are available from the beginning of the transmission, the optimal transmission strategy was derived in [99] for a two-user multiple access communication system, and in [94, 96] for an AWGN broadcast channel.

None of these works jointly considered dynamic data arrivals and a finite battery capacity at the transmitter because, then, there exists an inherent coupling between the data and energy domains, which substantially complicates the problem. In this context, this chapter bridges this gap by determining the transmission strategy that minimizes the transmission completion time of a WEHN with finite battery capacity and dynamical data and energy arrivals that additionally has to satisfy some generic QoS constraints (see Figure 3.1). These generic QoS constraints have the form introduced in §2.2.2 and are able to impose, among others, a finite size of the data buffer or a maximum delay on the delivery of the data packets.

The main contributions of this chapter are:

- Proposing a framework to map the constraints of the energy domain to the data domain that allows us to adapt the calculus approach proposed in [55] (cf. §2.2.2), which did not take into account energy harvesting at the transmitter, to the energy harvesting scenario.
- Studying the impact of the QoS constraint in the transmission strategy that minimizes the transmission completion time, which, to the best of our knowledge, had not been previously studied in the literature by the time of this research.
- Showing that, due to the QoS constraints, there may be situations in which no feasible transmission strategy exists. Such situations are analytically characterized in Lemma 3.4.

- Showing that, if a feasible transmission strategy exists, the optimal cumulative data departure curve is a piecewise linear function where battery overflows are only produced when the data buffer is empty.
- Developing an algorithm that either computes the optimal transmission strategy or concludes that there is no feasible solution and analytically showing its optimality.

The remaining of the chapter is structured as follows. In §3.2, the problem is mathematically formulated by using the concept of cumulative curves, which is introduced in [55], that allows an appealing visualization of the solution. The solution is characterized in §3.3. §3.4 presents the developed iterative algorithm that is able to compute the optimal solution, which is numerically evaluated in §3.5. Finally, §3.6 concludes the chapter.

3.2 Problem formulation

We consider a node with a finite battery capacity, C_{max}, that has to transmit $\mathcal{D}$ data packets by using at the most the J energy packets that it harvests over time while satisfying some QoS requirements (see Figure 3.1). We want to find the power allocation/rate scheduling strategy[1] that minimizes the transmission completion time, T.

We assume that the time instants at which the data and energy packets arrive to the node and their size (bits or Joules) are known from beforehand; accordingly, we focus on the *offline* problem (cf. §2.1.3). Hence, it is known that at the time instant $d_i \geq 0$ seconds the i-th data packet arrives containing D_i bits, with $i = 1, \ldots, \mathcal{D}$. Similarly, the j-th energy packet arrives at the instant $e_j \geq 0$ seconds and a total of E_j Joules are harvested, with $j = 1, \ldots, J$ (see Figure 3.2). Without loss of generality, the first energy arrival is produced at $e_1 = 0$ and contains the initial battery of the node E_1.

To describe our model we present the following definitions, which are based on the concept of cumulative curves introduced in §2.2.2, that are summarized in Table 3.1.

Definition 3.1 (Data Departure Curve)**.** *A data departure curve $D(t) \geq 0$, $t \geq 0$, is the total number of bits that have been cumulatively transmitted by the node in the time interval* $[0, t]$.

Definition 3.2 (Energy Expenditure Curve)**.** *An energy expenditure curve $E(t) \geq 0$, $t \geq 0$, is the energy in Joules that has been cumulatively consumed by the node in the time interval* $[0, t]$.

Let us consider a static channel with a power-rate function denoted by $g(\cdot)$, i.e., the function that, at any given time instant t, relates the transmitted power, $p(t)$, with the rate, $r(t)$,

[1] Observe that fixing the transmission power or the rate is equivalent as it will be shown next.

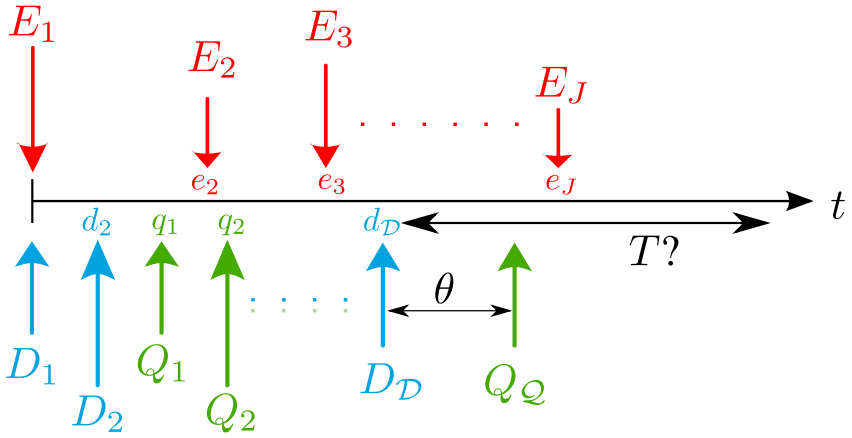

Figure 3.2: Summary of the different events considered in the time domain, namely, energy arrivals, data arrivals, and quality of service requirements. The figure represents the deadline QoS constraint, where θ is the maximum allowed delay for all the packets.

Definition	General notation	Notation at the m-th algorithm iteration
Instantaneous power	$p(t)$	
Instantaneous rate	$r(t)$	
Power-rate function	$g(\cdot)$	
Transmission completion time	T	$T^{(m)} = T - i_m$
Data departure curve	$D(t)$	$D^{(m)}(t) = D(t+i_m) - D^\star(i_m)$
Optimal data departure curve	$D^\star(t)$	$D^{\star^{(m)}}(t) = D^\star(t+i_m) - D^\star(i_m)$
Accumulated Data	$D_A(t)$	$D_A^{(m)}(t) = D_A(t+i_m) - D^\star(i_m)$
Minimum Data Departure	$D_{QoS}(t)$	$D_{QoS}^{(m)}(t) = \{D_{QoS}(t+i_m) - D^\star(i_m)\}^+$
Energy expenditure curve	$E(t)$	$E^{(m)}(t) = E(t+i_m) - E^\star(i_m)$
Optimal energy expenditure curve	$E^\star(t)$	$E^{\star^{(m)}}(t) = E^\star(t+i_m) - E^\star(i_m)$
Accumulated Battery	$B_A(t;t_x)$	$B_A^{(m)}(t) = B_A(t+i_m;i_m) - E^\star(i_m)$
Minimum Energy Expenditure	$E_{min}(t;t_x)$	$E_{min}^{(m)}(t) = \{E_{min}(t+i_m;i_m) - E^\star(i_m)\}^+$
Actual mapping of $B_A(t;i_m)$ to data domain		$\bar{D}_{B_A}^{(m)}(t)$
Effective mapping of $B_A(t;i_m)$ to data domain		$D_{B_A}^{(m)}(t)$
Effective mapping of $E_{min}(t;i_m)$ to data domain		$D_{E_{min}}^{(m)}(t)$
Equivalent upper bound on the data domain		$D_{max}^{(m)}(t) = \min\{D_A^{(m)}(t), D_{B_A}^{(m)}(t)\}$
Equivalent lower bound on the data domain		$D_{min}^{(m)}(t) = \max\{D_{QoS}^{(m)}(t), D_{E_{min}}^{(m)}(t)\}$
Discontinuities of the upper bound		$\mathcal{Z}_{max}^{(m)} = \{t \mid D_{max}^{(m)}(t^-) \neq D_{max}^{(m)}(t^+)\}$
Discontinuities of the lower bound		$\mathcal{Z}_{min}^{(m)} = \{t \mid D_{min}^{(m)}(t^-) \neq D_{min}^{(m)}(t^+)\}$
m-th pool rate		r_m
m-th pool length		ℓ_m
Beginning m-th pool		i_m
Data arrival time	d_i	$d_i^{(m)} = d_i - i_m$
Amount of data in the packet	D_i	$D_0^{(m)} = D_A(i_m) - D^\star(i_m)$ $D_i^{(m)}$ is a relabeling of D_i for $d_i > i_m$
Energy arrival time	e_j	$e_j^{(m)} = e_j - i_m$
Amount of energy in the packet	E_j	$E_1^{(m)} = B_A(i_m;i_m) - E^\star(i_m)$ $E_j^{(m)}$ is a relabeling of E_j for $e_j > i_m$
QoS requirement arrival time	q_k	$q_k^{(m)} = q_k - i_m$
Amount of data in the QoS requirement	Q_k	$Q_0^{(m)} = 0$ $Q_k^{(m)}$ is a relabeling of Q_k for $q_k > i_m$

Table 3.1: Summary of the chapter's notation.

according to $p(t) = g(r(t))$. As in [55] and [92], we make the common assumption that the function $g(\cdot)$ is time-invariant, convex, strictly increasing, and $g(0) = 0$. Note that the instantaneous rate, $r(t)$, can be expressed as the derivative w.r.t. t of the data departure curve, i.e., $r(t) = \frac{\mathrm{d}D(t)}{\mathrm{d}t}$. Similarly, the transmitted power is $p(t) = \frac{\mathrm{d}E(t)}{\mathrm{d}t}$. Then, the energy expenditure curve can be obtained from the data departure curve as follows[2]:

$$E(D(t)) = \int_0^t g\left(\frac{\mathrm{d}D(\tau)}{\mathrm{d}\tau}\right) \mathrm{d}\tau. \tag{3.1}$$

Observe that the magnitudes $D(t)$, $E(t)$, $r(t)$, and $p(t)$ are unambiguously related by (3.1) and $g(\cdot)$. Therefore, given the initial states $E(0) = 0$ and $D(0) = 0$, the design of the system to be optimized can be described by any of these magnitudes.

Definition 3.3 (Battery)**.** *The battery of the node, $B(t)$, is the amount of energy that the node has available at a given time instant t. We consider a battery with finite capacity C_{max}. Thus, $B(t)$ must satisfy that $0 \leq B(t) \leq C_{max}$, $\forall t \geq 0$.*

Due to the limited battery capacity, at the j-th energy arrival, some part of the harvested energy E_j may be lost. This lost energy is denoted as the j-th battery overflow, i.e., $O_j = [E_j - C_{max} + B(e_j^-)]^+$. Observe that battery overflows depend on the chosen energy expenditure curve $E(t)$ and guarantee that the battery level will never be above the battery capacity, i.e., $B(e_j^+) = B(e_j^-) + E_j - O_j \leq C_{max}$.

As battery overflows depend on the chosen $E(t)$, their value O_j cannot be computed until the energy expenditure curve $E(t)$ is fixed $\forall t \leq e_j$. Loosely speaking, one can expect that the optimal solution uses efficiently the harvested energy to transmit the available data and, at the same time, tends to minimize the total overflow of the battery and, thus, maximizes the accumulated energy stored in the battery. In other words, if overflows are minimized, the node will be able to use more energy, as, otherwise, the energy of the overflows is lost.

In the following lines, we define the accumulated battery, a concept introduced in this work that allows us to characterize the optimal solution when having finite battery capacity constraint. Let t_x denote the last time instant up to which $D(t)$ (or, equivalently, $E(t)$) is known.

Definition 3.4 (Accumulated Battery)**.** *The accumulated battery $B_A(t; t_x)$ is a real measure of the accumulated energy stored in the battery for $t \in [0, t_x)$ and it is the maximum possible accumulated energy in the battery for $t \in [t_x, \infty)$ (assuming that no overflows are produced*

[2]Without loss of generality, we can assume that $r(t)$ is right-continuous.

for $t \in (t_x, \infty)$*), i.e.,* $B_A(t;t_x) = \sum_{j=1}^{J} \left(-O_j \mathcal{H}^{\ell}(t_x - e_j) + E_j\right) \mathcal{H}^{\ell}(t - e_j)$*, where* $\mathcal{H}^{\ell}(\cdot)$ *denotes the left continuous unit step function in* (2.6)*.*

Observe that, for $t \in [0, t_x)$, $B_A(t;t_x)$, represents the real measure of the energy accumulated in the battery during $(0,t)$, because, for these time instants, overflows are known and taken into account. Alternatively, for $t \in (t_x, \infty)$, battery overflows are unknown after t_x and $B_A(t;t_x)$ models the best case scenario where the node is able to store in the battery all the harvested energy in the interval (t_x, ∞). Moreover, observe that $B_A(t;t_x) = B_A(t;t_y)$ for any $t_x, t_y \geq t$.

At every time instant, the energy stored in the battery is $B(t) = B_A(t;t) - E(t)$. Note that $B_A(t;t)$ takes into account the actual net incoming energy in the battery, whereas, $E(t)$ is the net outgoing energy. Thus, their difference results in the energy stored in the battery.

Definition 3.5 (Minimum Energy Expenditure)**.** *The minimum energy expenditure,* $E_{min}(t;t_x)$*, is the smallest amount of energy that the node must have cumulatively spent at time* $t > t_x$ *such that no overflow of the battery is produced in the interval* $(t_x, t]$*, i.e.,* $E_{min}(t;t_x) = \left[\sum_{j=1}^{J} \left(-O_j \mathcal{H}^{r}(t_x - e_j) + E_j\right) \mathcal{H}^{r}(t - e_j) - C_{max}\right]^{+}$*, where* $\mathcal{H}^{r}(\cdot)$ *denotes the right continuous unit step function.*

Note that $E_{min}(t;t_x)$ and $B_A(t;t_x)$ are right continuous and left continuous w.r.t. t, respectively. Indeed in the points in which both functions are continuous, we have that $E_{min}(t;t_x) = [B_A(t;t_x) - C_{max}]^{+}, \forall t \neq e_j, j = 1, \ldots, J$.

Definition 3.6 (Accumulated Data)**.** *The accumulated data* $D_A(t)$ *is the sum of data that has arrived at the node during the time interval* $[0,t)$*, i.e.,* $D_A(t) = \sum_{i=1}^{\mathcal{D}} D_i \mathcal{H}^{\ell}(t - d_i)$*.*

Different QoS constraints can be considered by mapping the constraint into an appropriate minimum data departure curve, which was introduced in [55], and is defined as follows:

Definition 3.7 (Minimum Data Departure)**.** *The minimum data departure,* $D_{QoS}(t)$*, is the smallest amount of data that the node must have cumulatively transmitted at time* t *such that the QoS constraint is satisfied.*

The rest of the chapter considers a general $D_{QoS}(t)$ that is a non-negative staircase function where changes are produced at time instants q_k with increments of Q_k bits for $k = 1, \ldots, \mathcal{Q}$, i.e., $D_{QoS}(t) = \sum_{k=1}^{\mathcal{Q}} Q_k \mathcal{H}^{r}(t - q_k)$. From now on, the instants q_k are called *quality requirement events*. Thus, three kind of events are considered, namely, data arrival, energy arrival, and quality requirement events, as summarized in Figure 3.2. As it has been presented in §2.2.2, the

minimum data departure curve allows to impose, among others, the following constraints: (i) no data can be lost due to overflows of the data buffer; and (ii) a maximum permissible delay for the transmission of a certain data packet.

Observe that depending on the chosen QoS constraint, it is likely that the instants q_k are equal to d_i for some values of k and i, e.g., for the finite data buffer constraint (cf. §2.2.2). Hence, we consider that two different types of events can be produced simultaneously at the same time instant.

Our goal is to find the data departure curve, $D(t)$, that minimizes the transmission completion time T of the $\mathcal{D}$ data packets, i.e., $D(T) = \sum_{i=1}^{\mathcal{D}} D_i$, while satisfying the following conditions: (i) ECC: energy must be harvested before it is used by the node or, which is the same, the battery level in the node must be greater or equal than zero. (ii) DCC: it is not possible to transmit more bits than the ones that have arrived to the node. (iii) QoS constraint: at time t, a minimum amount of data $D_{QoS}(t)$ has to be transmitted in order to preserve the link quality of service. Moreover, given two data departure curves with the same completion time, the one that requires less energy is always preferred. From all that has been said above, the problem can be expressed as follows:

$$\begin{aligned} \underset{D(t)}{\text{minimize}} \quad & T \\ \text{subject to} \quad & E(t) \leq B_A(t;t), && \forall t \in [0,T], \\ & D_{QoS}(t) \leq D(t) \leq D_A(t), && \forall t \in [0,T], \\ & D(T) = \sum_{i=1}^{\mathcal{D}} D_i. \end{aligned} \tag{3.2}$$

We want to remark the two main difficulties of the problem presented in (3.2). First, the integral relation among the data and energy domains through (3.1). Second, the fact that neither T nor $B_A(t;t)$ are known from beforehand, due to their dependence on $D(t)$. Consequently, both T and $B_A(t;t)$ will be found along with the solution to the problem.

This problem is graphically represented in Figure 3.3, where the figures at the top and bottom stand for the energy and data domains, respectively. The ECC is represented by the solid line at the top figure, whereas DCC and QoS constraints are depicted by the dot-dashed and dashed lines in the figure at the bottom, respectively. The dotted line in the energy domain represents the minimum energy expenditure curve; however, as it can be seen from (3.2), it is not a constraint of the problem. Hence, $D(t)$ and its associated $E(t)$ must lie within the blank region of the data and energy domains, respectively, in order to be a feasible solution. Three different data departure curves (A, B, and C) and their associated energy expenditure curves are shown. The curve A is not feasible since it breaks the ECC. The curve B is feasible in spite

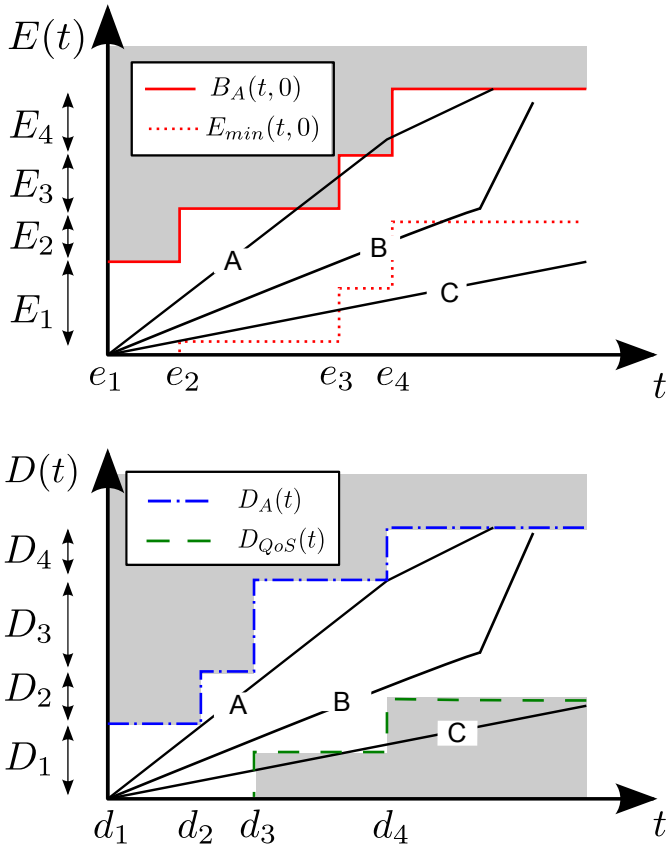

Figure 3.3: Illustrative representation of the problem presented in (3.2).

of having an overflow of the battery at e_3, which is, in general, a suboptimal strategy as we will show later. Finally, the curve C is not feasible because it does not satisfy the QoS constraint.

We want to remark that Figure 3.3 is just an illustrative representation of the problem. As mentioned before, both the accumulated battery and the minimum energy expenditure depend on the selected transmission strategy. Note that the figure shows $B_A(t;0)$ as ECC instead of the real ECC, i.e., $B_A(t;t)$. In some manner, with $B_A(t;0)$, we are showing the battery accumulated that is obtained when $D(t)$ does not produce any battery overflow. In case that overflows cannot be avoided, it would be necessary to subtract the amount of the overflow to the shown accumulated battery and minimum energy expenditure from that time instant onward. Thus, we can only show the real graphical representation of the problem once we have fixed the solution.

Note that (3.2) is not a convex optimization problem and that its conversion into a convex problem is not straightforward. Thus, we cannot directly solve (3.2). Alternatively, in next section, we model the properties that the optimal solution must satisfy, which will allow us to construct the optimal data departure curve.

3.3 Properties of the optimal solution

As pointed out previously, in this section we will characterize the optimal data departure curve, $D^\star(t)$, and its associated energy expenditure curve, $E^\star(t)$, for the problem (3.2):

Problem 3.1 (Transmission Without Events)**.** *Consider that the optimal departure curve is known up to t_1 and that we want to characterize its optimal behaviour in the time interval (t_1, t_2) where there are no changes in $D_A(t)$, $B_A(t; t_1)$, and $D_{QoS}(t)$. We also consider that the data departure curve at the boundary of the intervals is $D(t_1)$ and $D(t_2)$, respectively, and that these two points satisfy the data, energy, and QoS constraints.*

Lemma 3.1. *In Problem 3.1, $D^\star(t)$ is a straight line where the slope, or, equivalently, the transmission rate, is constant and equal to $r(t) = \frac{D(t_2)-D(t_1)}{t_2-t_1}$, $\forall t \in (t_1, t_2)$.*

Proof: The proof follows from the integral version of Jensen's inequality in a similar way to the BT-problem in [55]. ■

Corollary 3.1. *Lemma 3.1 implies that $D^\star(t)$ is a piecewise linear function such that its slope, which is equivalent to the transmission rate, can only change either at e_j, d_i or q_k.*

From the previous lemma and corollary, it follows that constant rate transmission saves energy due to the convexity of the power-rate function $g(\cdot)$. However, constant rate transmission is not optimal when a battery overflow is produced because the energy saved due to constant rate transmission is lower than the energy lost in the overflow. Consequently, the optimal solution increases the rate before the overflow until either there is no overflow or the data buffer is empty, as shown in the following lemma and its subsequent proof:

Lemma 3.2. *Under the optimal policy, battery overflows may only be produced when there is no data to be transmitted.*

Proof: See appendix 3.A.1. ■

The following lemma states that by the end of the transmission the battery must be empty, otherwise, transmission could have been finished earlier by transmitting at a higher rate.

Lemma 3.3. *The optimal solution must satisfy that, at the instant T at which all the data has been transmitted, the energy expenditure is equal to the accumulated battery, i.e., $E^\star(T) = B_A(T; T)$.*

Proof: The proof follows in a similar way to the proof of Lemma 5 in [87]. ■

In the remainder of the chapter, the term *pool* denotes each of the time intervals of $D^\star(t)$ at which transmission is done at constant rate. We define M as the total number of pools of the optimal solution, i.e., the number of linear pieces of $D^\star(t)$. Note that M is unknown a priori. Consequently, solving the problem in (3.2) is equivalent to determining the rate r_m and length ℓ_m of each pool, i.e., $\{r_m, \ell_m\}_{m=1}^{M}$. To do so, we have developed an iterative algorithm that, at the m-th iteration, determines the rate and duration of the m-th pool, i.e., $\{r_m, \ell_m\}$. We denote i_m as the instant at which the m-th pool begins, i.e., $i_m = \sum_{p=1}^{m-1} \ell_p$, where by definition $i_1 = 0$. The algorithm ends when all data has been efficiently transmitted. With this, M, T, $B_A(t;T)$, and $D^\star(t)$ are found.

To simplify the complexity of our algorithm, at the beginning of the m-th iteration[3], the origin of coordinates is moved to the point $(i_m, D^\star(i_m))$. To be coherent with the vertical and horizontal displacement of the origin of coordinates, the data and energy constraints in (3.2) must be vertically rescaled by $D^\star(i_m)$ and $E^\star(i_m)$, respectively, and temporally displaced by i_m. In the remainder of the chapter, a super-index (m) above a variable ($B_A^{(m)}(t)$, $D_A^{(m)}(t)$, $D_{QoS}^{(m)}(t)$, and $E_{min}^{(m)}(t)$) denotes that it is the rescaled version at the m-th iteration[4], e.g., $B_A^{(m)}(t) = B_A(t + i_m; i_m) - E^\star(i_m)$[5].

From the structure of the problem in (3.2), it is expected that it may not have a feasible solution whenever the node has to fulfill very tight QoS requirements, while, at the same time, it does not harvest enough energy to transmit all the required data. The following lemma is checked at every iteration of our proposed algorithm to determine whether the problem in (3.2) has a feasible solution or not.

Lemma 3.4. *The problem* (3.2) *does not have a feasible solution whenever*

$$D_{QoS}^{(m)}(q_k^{(m)}) > q_k^{(m)} g^{-1}\left(B_A^{(m)}(q_k^{(m)})/q_k^{(m)}\right), \tag{3.3}$$

for some quality requirement event $q_k^{(m)} \in (0, T^{(m)})$[6].

Proof: Whenever we encounter that the problem does not have a feasible solution is because, at some quality requirement event $q_k^{(m)}$, it is not possible to fulfill all the constraints. Let $\bar{D}(t)$ be the data departure curve that transmits the maximum amount of data in the interval $[0, q_k^{(m)}]$, i.e., $\bar{D}(t) = g^{-1}\left(B_A^{(m)}(q_k^{(m)})/q_k^{(m)}\right) t$. Note that this curve has constant rate,

[3]Note that $D^\star(t)$ is known in $[0, i_m]$.

[4]The relations among the iteration specific and general versions of the variables are given in Table 3.1.

[5]We have dropped the second argument in $B_A^{(m)}(t)$ and $E_{min}^{(m)}(t)$ since within the m-th iteration, the second argument, which denotes the last instant at which the solution is known, is always i_m.

[6]Where $X^{(m)} = X - i_m$ is the rescaled version of some temporal variable X at the m-th iteration (see Table 3.1).

empties the battery at $q_k^{(m)}$, and that the constrains are not necessarily satisfied. If, at some $q_k^{(m)} \in (0, T^{(m)})$, the QoS constraint requires more than $g^{-1}\left(B_A^{(m)}(q_k^{(m)})/q_k^{(m)}\right) q_k^{(m)}$ bits to be transmitted, then the problem does not have a feasible solution. ■

If there is no $q_k^{(m)} \in (0, T^{(m)})$ that satisfies Lemma 3.4, then the problem still may have a feasible solution and at least another pool, i.e., $\{r_m, \ell_m\}$, can be determined. In this context in the next subsection, we model how the rate changes must be produced in order to be optimal.

3.3.1 Constraints mapping into the data domain for a given pool

Within an algorithm iteration, the ECC can be mapped to the data domain, hence, allowing us to merge both constraints to the most restrictive constraint. Let us consider that the algorithm is at the beginning of the m-th iteration, i.e., the optimal solution is known up to i_m, where the rate r_m and length ℓ_m of the m-th pool must be determined. Given that transmission must be done at constant rate/power, the maximum amount of data that can be transmitted at a certain time instant t_y due to the ECC is $g^{-1}(p_y)t_y$, where $p_y = B_A^{(m)}(t_y)/t_y$.[7] With this, at the instant t_y, the upper bound on the energy expenditure curve has been mapped to an upper bound on the data departure curve, as shown in Figure 3.4 for the instants t_1 and t_2. By applying this procedure at all time instants $t_y \in (0, T^{(m)})$, we can map the whole upper bound in the energy domain to an upper bound in the data domain, which we denote by $\bar{D}_{B_A}^{(m)}(t) = g^{-1}(B_A^{(m)}(t)/t)t$ and call *actual mapping*. However, doing this computation for each time instant has a high computational cost. In summary, if a data departure curve that transmits at constant rate, i.e., $D^{(m)}(t) = r_m t$, satisfies that $D^{(m)}(t) \leq \bar{D}_{B_A}^{(m)}(t), \forall t$, then it also satisfies the ECCs.

The cost associated with the computation of $\bar{D}_{B_A}^{(m)}(t)$ can be reduced by noting that it is suboptimal that $D^{(m)}(t)$ reaches the *actual mapping* at any time instant t_1 that is not an event, i.e., $t_1 \neq d_i^{(m)}$, $t_1 \neq e_j^{(m)}$, and $t_1 \neq q_k^{(m)}$, $\forall i, j, k$. This is clearly seen in Figure 3.4. Observe that if $D^{(m)}(t_1) = \bar{D}_{B_A}^{(m)}(t_1)$, then the battery is empty at t_1, i.e., $E^{(m)}(t_1) = B_A^{(m)}(t_1)$, which follows from the definition of the actual mapping. Consequently, in order to satisfy the ECCs, the rate at t_1^+ must be zero as no energy arrival is produced at t_1. From Corollary 3.1, we know that this rate change is suboptimal and, therefore, the data departure curve can only reach the *actual mapping* in some event. Thus, to reduce the computational complexity of the *actual mapping*, we can compute the value of the mapping only at the aforementioned events and assign a constant value in the interval between events. In the rest of the chapter, we refer to this mapping as *effective mapping*, i.e., $D_{B_A}^{(m)}(t)$. Note that the *effective mapping* is an upper bound of the *actual mapping*. However, we want to remark that by using the *effective mapping*, we are

[7] Note that for the constraints mapping it is not necessary that p_y satisfies the constraints. The constraints fulfillment is enforced by the algorithm that computes the optimal solution, which is explained in Section 3.4.

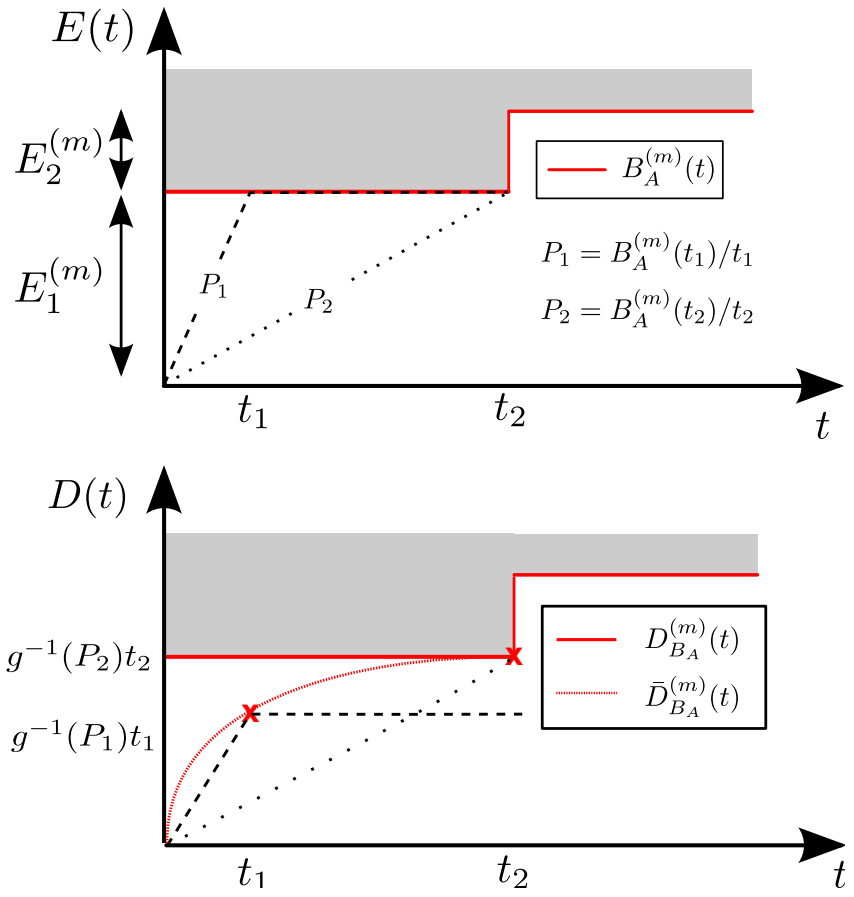

Figure 3.4: Mapping of the ECC to the data domain. The plot in the bottom shows the suboptimality to reach $\bar{D}_{B_A}^{(m)}(t)$ at any time instant that is not an event. This implies that in practice the *effective mapping* $D_{B_A}^{(m)}(t)$ can be used as the mapping of the energy constraint to the data domain.

not relaxing the constraints of the problem since both mappings are equal at the time instants where the optimal solution for the data departure curve coincides with the *actual mapping*.

A similar approach can be done to map the minimum energy expenditure curve $E_{min}^{(m)}(t)$ to the data domain. In this case, since overflows may only be produced at energy arrival events, it is only necessary to map the lower bound in the expended energy to the data domain at these time instants. For the rest of time instants, the time intervals between energy arrivals, a constant value is assigned without loss of generality, hence, obtaining $D_{E_{min}}^{(m)}(t)$.

Figure 3.5 shows a representation of the problem once the constraints in the energy domain are mapped to the data domain. Now the problem is simplified, since data and energy constraints can be merged in a single constraint that, at every time instant, is the most restrictive of the two constraints, i.e.,

$$D_{max}^{(m)}(t) = \min\{D_A^{(m)}(t), D_{B_A}^{(m)}(t)\}. \tag{3.4}$$

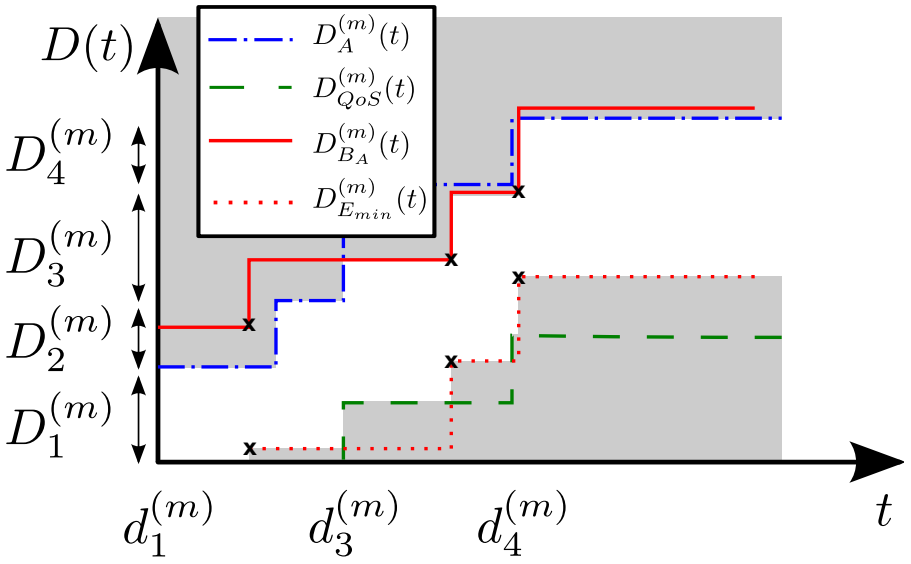

Figure 3.5: Mapping of the ECC and minimum energy expenditure to the data domain.

Similarly, the lower constraint is

$$D_{min}^{(m)}(t) = \max\{D_{QoS}^{(m)}(t), D_{E_{min}}^{(m)}(t)\}. \tag{3.5}$$

Note that $D_{max}^{(m)}(t)$ and $D_{min}^{(m)}(t)$ are only valid within an algorithm iteration and that it can occur that $D_{max}^{(m)}(t) < D_{min}^{(m)}(t)$. This happens in the following situations: (i) The node has to transmit a certain amount of data in order not to overflow the battery; however, this data is still not available, i.e, $D_{E_{min}}^{(m)}(t) > D_A^{(m)}(t)$. (ii) The node has to transmit a certain amount of data in order to satisfy the QoS constraint; however, it does not have enough energy available to do so, i.e., $D_{B_A}^{(m)}(t) < D_{QoS}^{(m)}(t)$. Note that the situation (ii) occurs when the problem does not have a feasible solution. As mentioned before, the aim of this section is to model how rate changes are produced when the problem indeed has a solution (at least up to the current algorithm iteration) and, hence, we will focus on situation (i) where an overflow of the battery is produced.

Let us define the sets of time instants at which $D_{max}^{(m)}(t)$ and $D_{min}^{(m)}(t)$ have discontinuities as $\mathcal{Z}_{max}^{(m)} = \{t \mid D_{max}^{(m)}(t^-) \neq D_{max}^{(m)}(t^+)\}$ and $\mathcal{Z}_{min}^{(m)} = \{t \mid D_{min}^{(m)}(t^-) \neq D_{min}^{(m)}(t^+)\}$, respectively. Remember that, due to Corollary 3.1, we know that $D^\star(t)$ is constant between events defined according to $\mathcal{Z}_{max}^{(m)}$ and $\mathcal{Z}_{min}^{(m)}$. In Lemmas 3.5, 3.6, and 3.7, we model the behavior of the optimal solution when the rate changes at a time instant where a single event is produced. Similarly, Lemmas 3.8 and 3.9 describe the behavior of the optimal solution when the rate changes at a time instant where two events are produced. The proofs of these lemmas are given in Appendix 3.A.2.

Lemma 3.5. *If a rate change is produced at a certain time instant ℓ_m such that $\ell_m \in \mathcal{Z}_{max}^{(m)}$ and*

$\ell_m \notin \mathcal{Z}_{min}^{(m)}$, *then* $D^{\star^{(m)}}(\ell_m) = D_{max}^{(m)}(\ell_m^-)$ *and the rate increases,* $r_m < r_{m+1}$.

Lemma 3.6. *If a rate change is produced at the time instant* ℓ_m *such that* $\ell_m \notin \mathcal{Z}_{max}^{(m)}$, $\ell_m \in \mathcal{Z}_{min}^{(m)}$ *and* $D_{max}^{(m)}(\ell_m) \geq D_{min}^{(m)}(\ell_m^+)$, *then* $D^{\star^{(m)}}(\ell_m) = D_{min}^{(m)}(\ell_m^+)$ *and the rate decreases,* $r_m > r_{m+1}$.

Lemma 3.7. *If a rate change is produced at a certain time instant* ℓ_m *such that* $\ell_m \notin \mathcal{Z}_{max}^{(m)}$, $\ell_m \in \mathcal{Z}_{min}^{(m)}$ *and* $D_{max}^{(m)}(\ell_m) < D_{min}^{(m)}(\ell_m^+)$, *then an overflow of the battery is produced at* ℓ_m, $D^{\star^{(m)}}(\ell_m) = D_{max}^{(m)}(\ell_m)$ *and the rate is zero until the next data arrival event,* $r_{m+1} = 0$.

Lemma 3.8. *If a rate change is produced at a certain time instant* ℓ_m *such that* $\ell_m \in \mathcal{Z}_{max}^{(m)}$, $\ell_m \in \mathcal{Z}_{min}^{(m)}$, *and* $D_{max}^{(m)}(\ell_m^-) \geq D_{min}^{(m)}(\ell_m^+)$, *then either* $D^{\star^{(m)}}(\ell_m) = D_{max}^{(m)}(\ell_m^-)$ *and the rate increases, or* $D^{\star^{(m)}}(\ell_m) = D_{min}^{(m)}(\ell_m^+)$ *and the rate decreases.*

Lemma 3.9. *If a rate change is produced at a certain time instant* ℓ_m *such that* $\ell_m \in \mathcal{Z}_{max}^{(m)}$, $\ell_m \in \mathcal{Z}_{min}^{(m)}$, *and* $D_{max}^{(m)}(\ell_m^-) < D_{min}^{(m)}(\ell_m^+)$, *then an overflow of the battery is produced at* ℓ_m, $D^{\star^{(m)}}(\ell_m) = D_{max}^{(m)}(\ell_m^-)$ *and the rate can either increase or decrease.*

By using these lemmas, we are able to construct an algorithm, which is presented in the next section, that iteratively finds the optimal solution, or concludes that there is no feasible solution.

3.4 Optimal data departure curve construction

In this section, we describe the developed algorithm that is able to either construct $D^\star(t)$ or, alternatively, conclude that the problem in (3.2) does not have a feasible solution. As stated in Corollary 3.1, the optimal data departure curve is a piecewise linear function. As previously explained, the developed algorithm follows an iterative process where, at the m-th iteration, the duration, ℓ_m, and rate, r_m, of a pool are determined. We will focus on the explanation of the m-th iteration since all the other iterations follow the same approach.

As shown in Figure 3.6, the algorithm is composed by three main blocks. The first block, named $checkSolution$, determines the existence of solution in the current iteration by checking the condition in Lemma 3.4. If the problem does not have a solution, the algorithm ends. Otherwise, the algorithm proceeds to the subsequent blocks to determine the rate and length of the pool.

The second block, named *checkFinish*, checks whether it is possible to transmit all the remaining data by using all the available energy at constant rate. This block is necessary

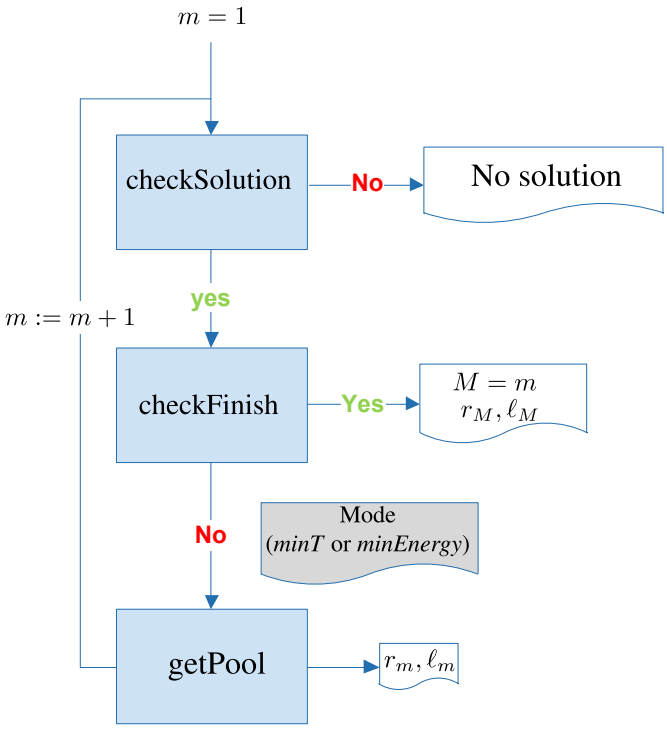

Figure 3.6: Block diagram of the iterative algorithm.

because $T^{(m)}$ is unknown. In case it is possible to finish in a single pool, it is obtained that $T^{(m)} = \ell_m$ and $M = m$. Then, the solution to problem (3.2) has been found and can be computed as

$$D^{\star}(t) = \sum_{m=1}^{M} r_m(t - i_m)\,\Pi\left(\tfrac{t-i_m}{\ell_m}\right) + r_m\ell_m\mathcal{H}^r(t - i_m - \ell_m), \tag{3.6}$$

where $\Pi(\cdot)$ is the unit pulse in the interval $[0, 1]$. The domain of $D^{\star}(t)$ is $[0, T = \sum_{m=1}^{M} \ell_m]$.

Otherwise, if it is not possible to finish in a single pool, the block *checkFinish* returns the mode ($minT$ or $minEnergy$) to be used by the third block, which we name *getPool*, to find the pool rate and length (r_m and ℓ_m) that fulfill Lemmas 3.5-3.9. The mode $minT$ is used when the node already has enough energy to finish transmission, whereas the mode $minEnergy$ is used when the node is still not able to finish transmission at any rate and, hence, the objective is to save as much energy as possible for the end of the transmission. An extended explanation of the inner behavior of each of these blocks is given in Appendix 3.A.3.

Once r_m and ℓ_m are determined, the origin of coordinates is moved to the point $(\ell_m, r_m\ell_m)$ and the variables are prepared for the new iteration. In the data domain, the iteration transmitted bits $D^{\star(m)}(\ell_m) = r_m\ell_m$ are subtracted from $D_A^{(m)}(t)$ and $D_{QoS}^{(m)}(t)$; for instance, $D_A^{(m+1)}(t) = D_A^{(m)}(t + \ell_m) - D^{\star(m)}(\ell_m)$. Similarly, in the energy domain, the expended energy $E^{\star(m)}(\ell_m) = g(r_m)\ell_m$ is subtracted from $B_A^{(m)}(t)$ and $E_{min}^{(m)}(t)$. Moreover, in case that transmitting at r_m produces a battery overflow at time instant ℓ_m, the amount of energy lost due to the overflow

is also subtracted from these variables. Finally, the mapping to the data domain is recalculated for the iteration $m+1$ and the whole procedure starts again to determine r_{m+1} and l_{m+1}.

There are two possible reasons for which the algorithm ends: (i) At some iteration, Lemma 3.4 is satisfied and, hence, the problem does not have solution. (ii) All the data has been transmitted and the optimal data departure curve has been obtained as given in (3.6).

The algorithm optimality is summarized in the following theorem and its subsequent proof:

Theorem 3.1. *The algorithm presented in this section constructs the optimal data departure curve, $D^{\star}(t)$, for the problem* (3.2)*.*

Proof: See appendix 3.A.3.4. ■

3.5 Results

To the best of our knowledge there is no other algorithm in the literature that considers altogether the ECC, DCC, the QoS constraint and the finite battery capacity. Therefore, in order to get some insights on the gain obtained with our proposed solution, we have developed a suboptimal ad-hoc strategy, namely, the Empty Buffers Strategy (EBS), that tries to empty the buffers as soon as possible, i.e., it looks for the time instant at which the next arrival (energy or data) is produced and tries to transmit at a constant rate so that the corresponding buffer is emptied by the time of the corresponding arrival.

The left y-axis of Figure 3.7 compares the normalized mean minimum T along a total of 1000 iterations, where at each iteration the data and energy arrivals are randomly generated following a uniform distribution. The amount of energy in each of the packets is normalized according to the total harvested energy which varies along the x-axis. The right y-axis shows the percentage of iterations in which there exists a feasible solution to (3.2). As shown in Figure 3.7, our proposed optimal algorithm substantially reduces the mean minimum T. If feasible solutions exist, the optimal algorithm finds the one that minimizes T; however, for some arrival profiles, the EBS is not able to find any feasible solution in spite of its existence.

Note that the EBS performs worse than the optimal strategy because it changes the rate at every packet arrival without checking whether constant power transmission is feasible between two arrivals, which would consume less energy. As a result of this extra energy consumption, the mean minimum T is higher and the probability of finding a feasible solution decreases.

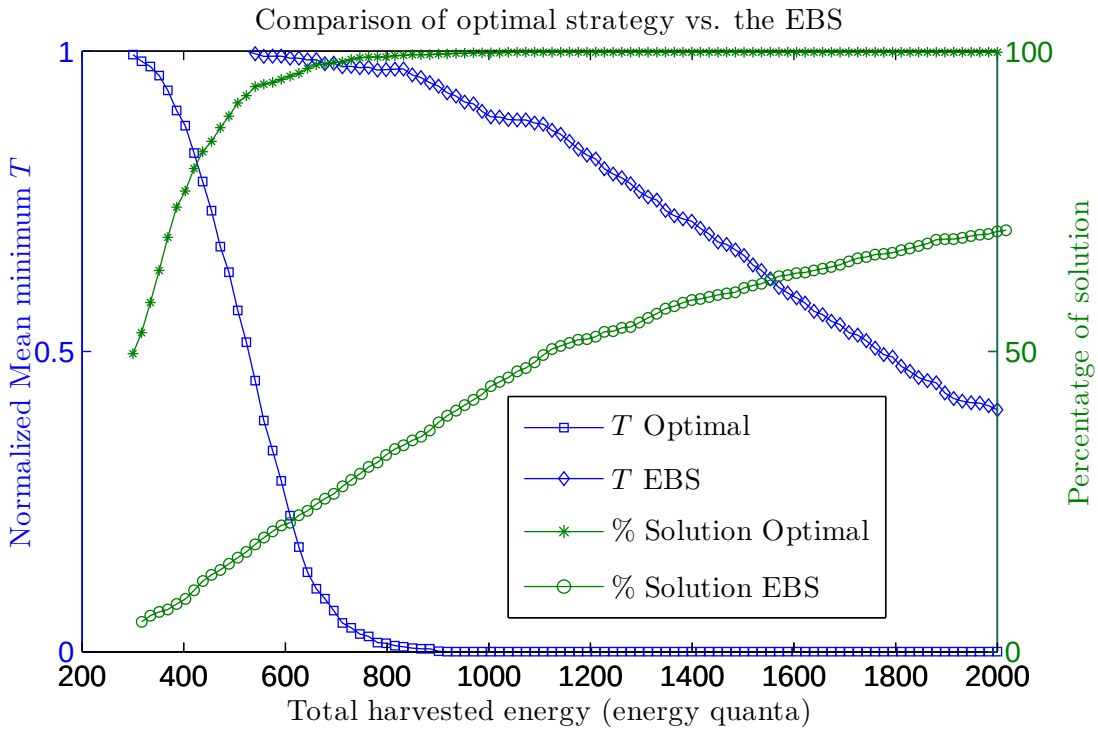

Figure 3.7: Performance comparison of our proposed optimal algorithm w.r.t. the suboptimal EBS in terms of mean minimum completion time and percentage of solutions. The lines marked with rectangles and diamonds refer to the left axis, whereas the lines marked with circles and asterisks refer to the rights axis.

3.6 Conclusions

In this chapter, the optimal transmission strategy has been obtained for WEHNs with finite battery capacity that, additionally, have to fulfill some QoS constraint. Hence, we have contributed to decrease the transmission completion time and, thus, we have increased the overall efficiency in the use of the harvested energy. Moreover, in more technical terms, we have seen that, as far as the battery does not overflow, constant rate transmission is the strategy that requires less energy to transmit a certain amount of data. However, if indeed the battery overflows, transmitting at constant rate is not optimal anymore, but the optimal strategy increases the rate before the overflow until either there is no overflow or the DCC is reached (i.e., there is no more data to transmit). We have seen that the existence of the optimal solution depends both on the dynamics of the harvesting process and on the required QoS. According to this, we have developed an algorithm that is able to determine whether the problem has a solution or not and, in case of having a solution, determines the optimal data transmission strategy.

3.A Appendix

3.A.1 Proof of Lemma 3.2

Consider the time interval (t_1, t_2) where there is an energy arrival at $e_j \in (t_1, t_2)$ that produces an overflow of the battery, O_j. We want to characterize which of the following solutions, $D_1(t)$ or $D_2(t)$, $t \in (t_1, t_2)$, is more efficient:

(1.) $D_1(t)$, which is such that $D_1(t) = D(t_1) + (t - t_1)r_0$, $\forall t \in (t_1, t_2)$, that is the same as saying that we are transmitting at a constant rate, r_0, in the interval (t_1, t_2).

(2.) $D_2(t)$, which is such that

$$D_2(t) = \begin{cases} D(t_1) + (t - t_1)r_1 & \text{if } t \in (t_1, e_j], \\ D_2(e_j) + (t - e_j)r_2 & \text{if } t \in (e_j, t_2), \end{cases}$$

that is the same as saying that we are transmitting at r_1, with $r_1 = r_0 + \epsilon_1$, in $(t_1, e_j]$ and at r_2, with $r_2 = r_0 - \epsilon_2$, in (e_j, t_2), where ϵ_1 and ϵ_2 are positive and such that the total transmitted data of both solutions is the same, i.e., $D_1(t_2) = D_2(t_2)$. Moreover, ϵ_1 must be small enough so that the strategy $D_2(t)$ still produces an overflow of the battery at e_j.

The problem formulated above is graphically presented in Figure 3.8, where the blue and red lines represent solutions 1 and 2, respectively. The following Lemma summarizes its solution:

Lemma 3.10. *The strategy $D_2(t)$ is more efficient than strategy $D_1(t)$ because the battery level at t_2 is higher for $D_2(t)$, i.e., $B_2(t_2) > B_1(t_2)$.*

Proof: We denote by p_0, p_1, and p_2 the powers obtained by evaluating the rates r_0, r_1, and r_2, respectively, in the power rate function $g(\cdot)$. Note that $p_2 < p_0 < p_1$ since $r_2 < r_0 < r_1$ and the power-rate function, $p(t) = g(r(t))$, is strictly increasing in $r(t)$. This implies that, at e_j, the energy expenditure of solution 2, $E_2(e_j)$, is greater than the energy expenditure of solution 1, $E_1(e_j)$. We denote the difference between the energy expenditures of the two solutions as $\Delta > 0$. Then, $E_2(e_j) = E_1(e_j) + \Delta$. Note that solution 2 reduces the overflow of the battery by Δ, therefore, the relation between the accumulated battery is

$$B_{A_2}(t, t_x) = B_{A_1}(t, t_x) + \Delta, \qquad \forall\, t > e_j, t_x > e_j. \tag{3.7}$$

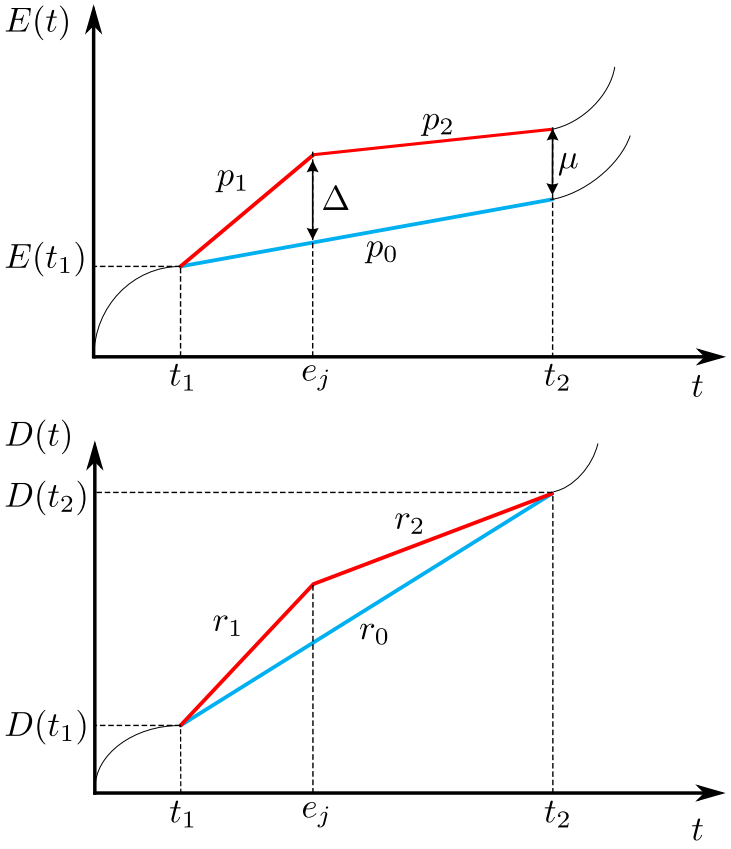

Figure 3.8: Graphical representation of the overflow problem.

However, solution 1 consumes less energy due to the fact that transmission is done at a constant rate. Let us denote the energy saving of solution 1 w.r.t. solution 2 at time instant t_2 as μ. Then, the relation between the two energy expenditure curves at t_2 is

$$E_2(t_2) = E_1(t_2) + \mu. \tag{3.8}$$

By subtracting (3.7) evaluated at t_2 and (3.8), we obtain that, $B_2(t_2) - B_1(t_2) = \Delta - \mu$, where we have used that $B(t) = B_A(t;t) - E(t)$.

We want to prove that $B_2(t_2) > B_1(t_2)$, therefore, that $\Delta - \mu > 0$. Let us first find the expressions for Δ and μ:

$$\Delta = E_2(e_j) - E_1(e_j) = (e_j - t_1)(g(r_1) - g(r_0)),$$
$$\mu = E_2(t_2) - E_1(t_2) = (e_j - t_1)g(r_1) + (t_2 - e_j)g(r_2) - (t_2 - t_1)g(r_0).$$

Finally, the expression of $\Delta - \mu$ is

$$\Delta - \mu = (t_2 - e_j)(g(r_0) - g(r_2)), \tag{3.9}$$

that is greater than zero since $g(r_0) > g(r_2)$. ■

Note that $D_2(t)$ achieves a higher battery level at t_2 due to the fact that the reduction in the battery overflow is higher than the energy saved by a constant rate transmission. Accordingly, the optimal transmission strategy increases the rate until either there is no overflow or the data buffer is emptied, which implies that the data departure curve must reach its upper bound, i.e., $D(t) = D_A(t)$. Note that the insertion of the QoS constraint, which is a lower bound on $D(t)$, does not affect the validity of the proof given above, which proves Lemma 3.2. ■

3.A.2 Rate change characterization

Proof of Lemma 3.5

The proof of the lemma is divided in two parts. We first show that if a rate change occurs at ℓ_m, then $D^{\star^{(m)}}(\ell_m) = D^{(m)}_{max}(\ell_m^-)$ and, afterwards, we show that the rate must increase.

Part 1. Let us assume that a rate change occurs at ℓ_m such that $D^{(m)}(\ell_m) < D^{(m)}_{max}(\ell_m^-)$. We will show by contradiction that this cannot be optimal. Let us consider the time interval $(\ell_m - \epsilon, \ell_m + \epsilon)$ with ϵ being positive. Note that if we select a sufficiently small ϵ, we can find a straight line with rate $r = \frac{D^{(m)}(\ell_m+\epsilon) - D^{(m)}(\ell_m-\epsilon)}{2\epsilon}$, which still satisfies the constraints, that transmits the same amount of data while having less energy expenditure. Hence, we have proved that if the rate changes at ℓ_m, then $D^{\star^{(m)}}(\ell_m) = D^{(m)}_{max}(\ell_m^-)$.

Part 2. Now we prove that when the rate changes at ℓ_m, it must increase. The procedure is the same as the one in the first part of the proof. We start by assuming that a rate decrease is optimal and then, we see that it leads to a logical contradiction. We denote r_m and r_{m+1} the rates before and after ℓ_m, respectively, where $r_m > r_{m+1}$. We consider the same time interval. In this case, we can also find a straight line whose slope is $r = \frac{D^{(m)}(\ell_m+\epsilon) - D^{(m)}(\ell_m-\epsilon)}{2\epsilon}$, which satisfies the energy and data constraints, that transmits the same amount of data while having less energy expenditure. Hence, we have proved by contradiction that if the rate changes, it must increase. ■

Proof of Lemma 3.6

This proof is similar to the proof of Lemma 3.5. The main difference is that now the discontinuity is in the lower constraint. Then, by following the same procedure we can first show that a rate change is suboptimal unless $D^{\star^{(m)}}(\ell_m) = D^{(m)}_{min}(\ell_m^+)$ and a rate decrease is produced. ■

Proof of Lemma 3.7

We know that ℓ_m comes from an energy arrival event, otherwise, the problem would not have a feasible solution as stated in Lemma 3.4. This implies that $D^{(m)}_{max}(\ell_m) = D^{(m)}_{A}(\ell_m)$ and $D^{(m)}_{min}(\ell_m) = D^{(m)}_{E_{min}}(\ell^+_m)$. Then, it is obvious that if $D^{(m)}_{A}(\ell_m) < D^{(m)}_{E_{min}}(\ell^+_m)$, an overflow is produced. Hence, the proof that $D^{\star^{(m)}}(\ell_m) = D^{(m)}_{max}(\ell_m)$ comes from Lemma 3.2, where we show that the optimal data departure curve satisfies that when an overflow is produced, all the data has been transmitted and, hence, the rate must change to zero until the following data arrival. ■

Proof of Lemma 3.8

This lemma states that two events are produced at the time instant ℓ_m. One could look at this lemma as the union of Lemmas 3.5 and 3.6 at the same time instant, hence, the proof has been already given in the aforementioned lemmas. ■

Proof of Lemma 3.9

As we pointed out, this lemma only applies when the problem has a solution. Consequently, we know that $D^{(m)}_{min}(\ell_m) = D^{(m)}_{E_{min}}(\ell^+_m)$ and that $D^{(m)}_{max}(\ell_m) = D^{(m)}_{A}(\ell^-_m)$. Hence, this is the overflow problem with the particularity that a data arrival is produced at ℓ_m. From Lemma 3.2, the optimal solution minimizes the energy lost due to overflow and, thus, $D^{\star^{(m)}}(\ell_m) = D^{(m)}_{max}(\ell^-_m)$. However, in this situation, nothing can be stated regarding the rate in the following pool. ■

3.A.3 The algorithm

In this appendix, a technical explanation of the algorithm is given. First, we introduce the maximum and minimum rates, a concept required to understand the second and third blocks of the algorithm that are presented afterwards.

3.A.3.1 Maximum and minimum rates

Let $\mathbb{R}^{(m)}_{max}$ denote the set that contains the rates obtained by joining the reference point, $(0, D^{(m)}(0) = 0)$[8], with the discontinuities from the left of $D^{(m)}_{max}(t)$, i.e., the points $(z, D^{(m)}_{max}(z^-))$, $\forall z \in \mathcal{Z}^{(m)}_{max}$, and such that the obtained curve is feasible for $t \in (0, z)$, where by feasible we mean that the curve satisfies all the constraints. Similarly, $\mathbb{R}^{(m)}_{min}$ contains the rates obtained by

[8]Note that the reference point is always $(0, 0)$ as the origin of coordinates is moved at every iteration.

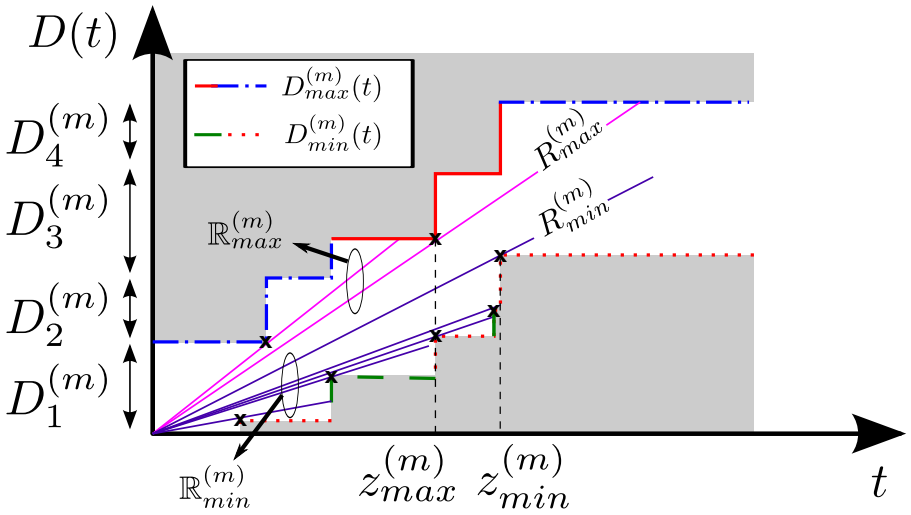

Figure 3.9: Graphical interpretation of the computation $R_{min}^{(m)}$ and $R_{max}^{(m)}$. The optimal pool rate lies within the interval $[R_{min}^{(m)}, R_{max}^{(m)}]$.

joining the reference point with the discontinuities from the right of $D_{min}^{(m)}(t)$, i.e., the points $(z, D_{min}^{(m)}(z^+))$, $\forall z \in \mathcal{Z}_{min}^{(m)}$, and such that the obtained curve is feasible in the same interval. An example of this can be seen in Figure 3.9.

Let $R_{max}^{(m)}$ denote the infimum of the set $\mathbb{R}_{max}^{(m)}$ and $R_{min}^{(m)}$ refer to the supremum of the set $\mathbb{R}_{min}^{(m)}$. Let $z_{max}^{(m)}$ and $z_{min}^{(m)}$ denote the time instants from which $R_{max}^{(m)}$ and $R_{min}^{(m)}$ have been obtained. Then, all the rates above $R_{max}^{(m)}$ and below $R_{min}^{(m)}$ are suboptimal as they would require a rate change to transmit the same amount of data.[9] Thus, the optimal rate lies within the interval $[R_{min}^{(m)}, R_{max}^{(m)}]$.

3.A.3.2 Finish transmission at a constant rate (checkFinish)

The first step, which is presented in Function 3.1, checks whether it is possible to transmit all bits by using an even power allocation in just one pool. If it is possible, which implies that transmission is ended, the algorithm returns the rate and length of the last pool, otherwise, it returns the strategy or mode that will be used in order to determine the following pool.

The function *checkFinish* first checks whether by transmitting at the maximum feasible rate, $R_{max}^{(m)}$, it is possible to transmit all the remaining data $D_{Tot}^{(m)}$ (this is done by the subroutine $getDataInCrossing$). In case it is not possible, the function returns the mode $minEnergy$. Otherwise, the function finds the time $\hat{T}_1^{(m)}$ required to transmit the remaining data $D_{Tot}^{(m)}$ with the iteration's initial battery $E_1^{(m)}$.[10] Then, it computes the equivalent rate $\hat{R}^{(m)}$ and checks whether transmitting at this rate is feasible, i.e., the following two conditions are fulfilled: (i)

[9] Note that $R_{max}^{(m)}$ is always greater than $R_{min}^{(m)}$, otherwise, either $R_{max}^{(m)}$ or $R_{min}^{(m)}$ would not be feasible.

[10] Remember that $E_1^{(m)} = B_A(i_m; i_m) - E^\star(i_m)$ as summarized in Table 3.1.

Function 3.1 checkFinish

$D_{Tot}^{(m)} = D_{max}^{(m)}(\infty)$ ▷ Remaining data
if $(getDataInCrossing(\ R_{max}^{(m)}t\ ,\ D_{max}^{(m)}(t)\ ) < D_{Tot}^{(m)})$ **then**
  return $mode = minEnergy$, $finish = 0$ ▷ It is not possible to finish yet.
else
  for $i = 1 : J^{(m)}$ **do** ▷ $J^{(m)}$ is the number of packets with $e_j > i_m$ plus one.
    $E = \sum_{j=1}^{i} E_j^{(m)}$
    $\hat{T}_i^{(m)}$ is obtained by solving $g^{-1}(E/\hat{T}_i^{(m)}) = D_{Tot}^{(m)}/\hat{T}_i^{(m)}$
    $\hat{R}^{(m)} = D_{Tot}^{(m)}/\hat{T}_i^{(m)}$ ▷ Even power allocation among all bits.
    if $R_{min}^{(m)} \leq \hat{R}^{(m)} \leq R_{max}^{(m)}$ **then** ▷ $D(t) = \hat{R}^{(m)}t$ is feasible in $[0, \hat{T}_i^{(m)}]$.
      $\mathbb{S} = \{E_j^{(m)} \mid j > i\ ,\ e_j^{(m)} \in (0, \hat{T}_i^{(m)})\}$
      if $\mathbb{S} = \emptyset$ **then** ▷ The algorithm ends and the rate and length of the last pool are returned.
        return $r = \hat{R}^{(m)}$, $\ell = \hat{T}_i^{(m)}$, $finish = 1$
      end if
    else if $(\hat{R}^{(m)} > R_{max}^{(m)})$ **then**
      return $mode = minT$, $finish = 0$
    end if
  end for
  return $mode = minEnergy$, $finish = 0$
end if

$\hat{R}^{(m)} \leq R_{max}^{(m)}$ and (ii) $\hat{R}^{(m)} \geq R_{min}^{(m)}$. In case that (i) is not fulfilled, the function returns the mode $minT$. If (ii) is not met, it is checked if, by using the following energy arrivals, a feasible curve is obtained. Finally, in case both conditions are fulfilled, it is checked whether any energy arrival has been produced in the time interval $(0, \hat{T}_1^{(m)})$. In case of no arrivals, the algorithm ends and the last pool has been found. In case there is an energy arrival in $(0, \hat{T}_1^{(m)})$, the function repeats the whole process but now using the initial battery, $E_1^{(m)}$, and the energy of the first arrival, $E_2^{(m)}$. This process is repeated until (i) becomes false or a feasible curve is found.

3.A.3.3 Get rate and length of the next pool (getPool)

This algorithm's block uses the parameter $mode$, which is obtained from the function $check-Finish$ as presented in Appendix 3.A.3.2, to compute the rate and length of the following pool.

Minimize the total completion time $(mode == minT)$*:* This strategy is used when both of the following conditions are satisfied: (i) It is possible to finish the transmission at some rate r with $r \leq R_{max}^{(m)}$. (ii) The rate obtained from an even power allocation $\hat{R}^{(m)}$ is not feasible due to $\hat{R}^{(m)} > R_{max}^{(m)}$. Hence, the objective is to find the rate that allows us to finish transmission as soon as possible, without paying attention on saving power; however, without wasting it,

either. In such case, the rate and duration of the pool are $r_m = R_{max}^{(m)}$ and $\ell_m = z_{max}^{(m)}$.

Minimize the energy expenditure $(mode == minEnergy)$*:* This strategy is used when, due to the constraints, transmission cannot be finished at a constant rate, e.g., the rate must be increased to satisfy the QoS constraints. Hence, the objective is to save as much energy as possible in order to use it when ending the transmission is feasible. Note that in this situation, the problem of obtaining the following pool is similar to the problem presented in [55] and, hence, the solution is also similar. The possible data departure curves with constant rate r, i.e., $D^{(m)}(t) = rt$, are divided in two sets. The first set $\mathbb{S}_{R_{max}}^{(m)}$ contains all the rates r such that the associated data departure curve crosses the constraint $D_{max}^{(m)}(t)$ first. Whereas the set $\mathbb{S}_{R_{min}}^{(m)}$ contains all the rates r such that the associated data departure curve crosses the constraint $D_{min}^{(m)}(t)$ first. Then, the rate of the following pool is determined as the infimum of $\mathbb{S}_{R_{max}}^{(m)}$ or, equivalently, the supremum of $\mathbb{S}_{R_{min}}^{(m)}$, i.e., $r_m = \inf\left(\mathbb{S}_{R_{max}}^{(m)}\right) = \sup\left(\mathbb{S}_{R_{min}}^{(m)}\right)$ and the duration of the pool ℓ_m can be obtained as the first time instant such that, $r_m\ell_m = D_{max}^{(m)}(\ell_m)$ or $r_m\ell_m = D_{min}^{(m)}(\ell_m)$.

3.A.3.4 Proof of the algorithm optimality

At each iteration, the algorithm checks whether Lemma 3.4 is fulfilled. In such a case, there is no solution for the problem and the algorithm ends. Otherwise, the algorithm must satisfy Lemmas 3.5-3.9 at each rate change and Lemma 3.3 by the end of the transmission. To show that these lemmas are satisfied and, hence, the algorithm computes the optimal data departure curve, we focus on the three different situations that can occur depending on the constraints of the problem: (i) The algorithm finishes transmission by using an even power allocation among all bits to be transmitted. (ii) $minT$ strategy is used to obtain the pool. (iii) The pool is computed by using $minEnergy$ mode. In the following, the optimality of these cases is proved by showing that the algorithm-chosen solution satisfies the optimality conditions and that it is unique.

Part 1 (Even power allocation). When this situation occurs, the algorithm ends transmission by transmitting at a constant rate. Hence, Lemma 3.1 is satisfied. No overflow is produced, thus, Lemma 3.2 is satisfied. Lemmas from 3.5 to 3.9 do not apply since there are no rate changes. Finally, note that Lemma 3.3 is satisfied, since the last rate r_M is obtained as the rate that allows the transmission of all the data by using all the available energy, moreover, from the properties of the function $g(\cdot)$ this rate is unique.

Part 2 ($minT$ mode). We want to demonstrate that the optimal departure curve transmits at $R_{max}^{(m)}$ during a period of time $z_{max}^{(m)}$. Let $\hat{T}^{(m)}$ be the total completion time that would be

obtained if transmitting at rate $\hat{R}^{(m)}$ was feasible, hence, $\hat{T}^{(m)} = D_{Tot}^{(m)}/\hat{R}^{(m)}$. Similarly, $T_{max}^{(m)} = D_{Tot}^{(m)}/R_{max}^{(m)}$. Note that, since $R_{max}^{(m)}$ is a feasible rate, $T_{max}^{(m)}$ is an upper bound of the remaining completion time, $T^{(m)}$, whereas $\hat{T}^{(m)}$ is a lower bound, hence:

$$\hat{T}^{(m)} < T^{(m)} < T_{max}^{(m)}. \tag{3.10}$$

Consider the data departure curve $D_1^{(m)}(t) = R_{max}^{(m)}t$. Note that any other data departure curve, $D_2^{(m)}(t)$, is suboptimal since, in order to satisfy (3.10), $D_2^{(m)}(t)$ will cross $D_1^{(m)}(t)$ for some $t_y \in (0, T_{max}^{(m)})$. Hence, at t_y, both curves have sent the same amount of data; however, $D_1^{(m)}(t)$ has consumed less energy.

Now we must show that, at $z_{max}^{(m)}$, the rate increases. Note that by transmitting at $R_{max}^{(m)}$ instead of at $\hat{R}^{(m)}$, some energy has been saved. Then, in the following pool, the available energy per bit is higher and, then, the new rate $\hat{R}^{(m+1)}$ obtained from an even power allocation among all bits in the following pool fulfills $\hat{R}^{(m+1)} > \hat{R}^{(m)} > R_{max}^{(m)}$. Hence, we have proved that at $z_{max}^{(m)}$ a rate increase is produced.

Part 3 ($minEnergy$ mode). This mode is used when it is not possible to finish transmission at any rate. Note that the algorithm can select three different kind of points, denoted as $v_x = (\ell_m, D^{\star^{(m)}}(\ell_m))$, for ending the pool depending on the constraints:

- $v_1 \mid D^{\star^{(m)}}(\ell_m) = D_{max}^{(m)}(\ell_m^-)$ where ℓ_m is either $d_i^{(m)}$ or $e_j^{(m)}$.
- $v_2 \mid D^{\star^{(m)}}(\ell_m) = D_{min}^{(m)}(\ell_m^+)$ where ℓ_m is either $q_k^{(m)}$ or $e_j^{(m)}$.
- $v_3 \mid D^{\star^{(m)}}(\ell_m) = D_{max}^{(m)}(\ell_m^-)$ where ℓ_m is $e_j^{(m)}$ and an overflow is produced.

Note that Corollary 3.1 and Lemma 3.2 are satisfied for any of the selected points and that Lemma 3.3 does not apply since transmission cannot be ended yet. Hence, we have to prove the following three conditions: (i) If a point such as v_1 is selected, there will be a rate increase (Lemma 3.5 or Lemma 3.8 for the case that at ℓ_m two events are produced). (ii) If a point such as v_2 is selected, there will be a rate decrease (Lemma 3.6 or Lemma 3.8 for the case that at ℓ_m two events are produced). And (iii), if a point such as v_3 is selected, the rate of the following pool is zero as far as ℓ_3 is not a data arrival event (Lemma 3.7 and Lemma 3.9).

Regarding (i), the rate of the pool is $r_m = \frac{D_{max}^{(m)}(\ell_m^-)}{\ell_m}$ that is the supremum of $\mathbb{S}_{R_{min}}^{(m)}$. Note that, at the following iteration, the set $\mathbb{S}_{R_{min}}^{(m+1)}$ includes all the rates in the interval $(0, r_m)$ and

the rates contained in the interval $(r_m, r_m+\epsilon)$, for some $\epsilon > 0$. Hence, the rate of the following iteration, r_{m+1}, satisfies that $r_{m+1} \geq r_m$, therefore, a rate increase is produced. A similar approach can be done for (ii), the rate of the pool is $r_m = \frac{D_{min}^{(m)}(\ell_m^+)}{\ell_m}$ that is the infimum of $\mathbb{S}_{R_{max}}^{(m)}$. At the following iteration the set $\mathbb{S}_{R_{max}}^{(m+1)}$ contains all the rates in $(r_m - \epsilon, \infty)$ for some $\epsilon > 0$. Hence, the rate of the following iteration, r_{m+1}, satisfies that $r_{m+1} \leq r_m$ and, therefore, there is a rate decrease. Finally, in case (iii), an overflow of the battery is produced. Note that the solution chosen by the algorithm satisfies that all the available data at ℓ_3^- has been transmitted. This implies that if non-data arrival is produced at ℓ_3 the rate of the following pool must be zero in order to satisfy the DCC.

Therefore, the algorithm computes the optimal solution since it satisfies all the lemmas that model the behavior of the optimal solution. ■

Chapter 4

On the precoder design of a wireless energy harvesting node in linear vector Gaussian channels with arbitrary input distribution

4.1 Introduction

In the previous chapter, we have studied a scenario where data dynamically arrives to the transmitter and thoroughly examined the effect of the DCCs and the finite buffer constraint over the optimal continuous-time transmission strategy. In opposition, the remaining chapters of this dissertation, consider the scenario in which the node has an infinite backlog of data to be transmitted through a discrete-time memoryless channel. Accordingly, in the remaining chapters, we aim to design transmission strategies that transmit the maximum amount of data by using the harvested energy.

Since the transmitter has always data to be transmitted, the DCCs are trivially satisfied and it is known that battery overflows are suboptimal. Thus, one can impose that the transmission strategy does not produce any battery overflows by imposing an additional set of constraints to the problem. Given a generic power consumption model $\mathcal{C}(\cdot)$ and the packetized model of the energy harvesting process in Figure 2.13, the ECCs and the finite battery capacity constraints can be written as

$$T_s \sum_{j=1}^{\ell} \sum_{n\in\tau_j} \mathcal{C}(p_n) \leq \sum_{j=1}^{\ell} E_j, \qquad \ell = 1,\ldots,J, \tag{4.1}$$

and

$$\sum_{j=1}^{\ell} E_j - \sum_{j=1}^{\ell-1} T_s \sum_{n\in\tau_j} \mathcal{C}(p_n) \leq C_{max}, \qquad \ell = 2,\ldots,J, \tag{4.2}$$

respectively, where $\mathcal{C}(p_n)$ is the power consumption of the n-th channel access of duration T_s and, as before, C_{max} denotes the battery capacity.

In the following chapters, we relax the finite battery capacity constraints assuming that no overflows are produced because the battery capacity is much larger than the amplitude of the harvested energy packets, i.e., $C_{max} \gg E_j$, which can be argued due to the low power levels provided by current harvesting technologies (cf. §2.1.1). We have done this relaxation because we prefer to focus on different aspects of the transmitter design; additionally, we believe that the finite battery constraint neither supposes an additional complexity to the problem (due to its similarity to the ECCs) nor provides many additional insights on the structure of the solution; finally, this assumption allows us to slightly simplify the notation of the studied problems, which is already quite involved as shown in the following chapters.

Accordingly, this chapter considers a WEHN operating in linear vector Gaussian channels with arbitrarily distributed input symbols and studies the offline (cf. §2.1.3) linear precoding strategy that maximizes the mutual information along N independent channel accesses, which must satisfy the ECCs.

It is important to recall that when the node does not have energy harvesting capabilities and the linear vector Gaussian channel can be decomposed as a set of parallel subchannels, then the optimal power allocation strategy is $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$, which has been presented in §2.2.1.2. When the channel matrix is not diagonal, the precoding strategy that maximizes the mutual information for non-harvesting nodes has been investigated in [58, 76–78] and references therein; these works are thoroughly presented in §4.3 after presenting the system model. To the best of our knowledge, none of the works in the literature has jointly considered arbitrary input distributions of the transmitted symbols and energy harvesting at the transmitter. The main contributions of this chapter are:

- Proving that, at the n-th channel use, the left singular vectors of the n-th precoder matrix are equal to the eigenvectors of the n-th channel Gram matrix.
- Deriving an expression that relates the singular values of the n-th precoder matrix with the energy harvesting profile through the MMSE matrix.
- Showing that the derivation of the optimal right singular vectors is a difficult problem and proposing a possible research direction towards the design of a numerical algorithm that computes the optimal right singular vectors. The design of this numerical algorithm is left for future research; instead, we derive the closed form power allocation that is obtained after setting the right singular vectors matrix to be the identity matrix and, in this scenario, the contributions are:

– Deriving the optimal offline power allocation, named the *MIMO Mercury Water-Flowing* solution, and providing an intuitive graphical interpretation.

– Proposing two different algorithms to compute the *MIMO Mercury Water-Flowing* solution, proving their optimality, and carrying out an exhaustive study of their computational complexity.

– Implementing an online algorithm, which does not require future knowledge of neither the channel state nor the energy arrivals, that computes a power allocation that performs close to the offline optimal *MIMO Mercury Water-Flowing* solution.

The remainder of the chapter is structured as follows. §4.2 presents the system model. In §4.3, the aforementioned offline problem is formally formulated and solved and the state of the art for non-harvesting nodes is presented. The graphical interpretation of the *MIMO Mercury Water-Flowing* solution is given in §4.4. The offline and online algorithms are introduced in §4.5 and §4.6, respectively. In §4.7, the performance of our solution is compared with different suboptimal strategies and the computational complexity of the algorithms is experimentally evaluated. Finally, the chapter is concluded in §4.8.

4.2 System model

We consider a point-to-point communication through a discrete-time linear vector Gaussian channel where the transmitter is equipped with energy harvesters. A total of N channel uses are considered where at each channel use the symbol $\mathbf{s}_n \in \mathbb{R}^{n_S}$ is transmitted.[1]

We consider that the data symbols $\{\mathbf{s}_n\}_{n=1}^N$ have independent components with unit power, i.e., $\mathbf{R}_s = \mathbb{E}\{\mathbf{s}_n\mathbf{s}_n^\mathsf{T}\} = \mathbf{I}_{n_S}$ and that they are independent and identically distributed (i.i.d.) along channel uses according to $P_S(\mathbf{s}_n)$. As shown in Figure 4.1, the symbol $\mathbf{s}_n$ is linearly processed at the transmitter by the precoder matrix $\mathbf{B}_n \in \mathbb{R}^{n_T \times n_S}$. We consider a slow-fading channel where the coherence time of the channel T_C is much larger than the symbol duration T_s, i.e., $T_s \ll T_C$. Thus, a constant channel matrix $\mathbf{G}_n \in \mathbb{R}^{n_R \times n_T}$ is considered at the n-th channel use. Let K denote the rank of the channel matrix, i.e., $K = \mathsf{rank}(\mathbf{G}_n) = \min\{n_T, n_R\}$, then

[1]The real field has been considered for the sake of simplicity. The extension to the complex case is feasible but requires the definition of the complex derivative, the generalization of the chain rule, and cumbersome mathematical derivations, which is out of the scope of this thesis. Nevertheless, the extension to the complex case can be done similarly as [78] generalized the results obtained in [76].

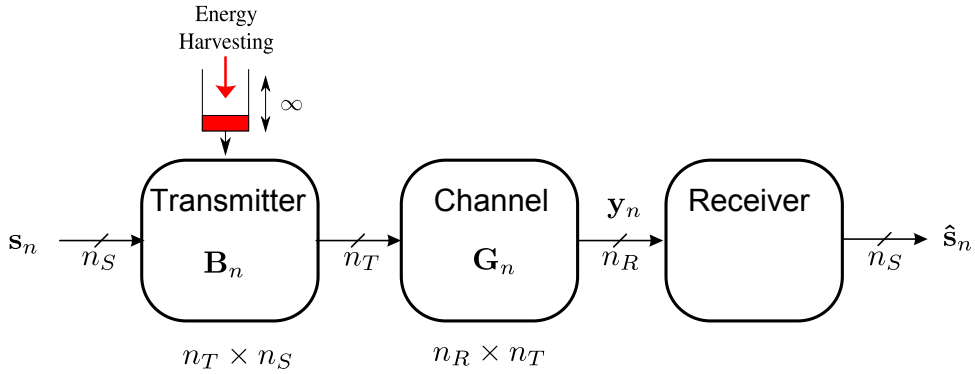

Figure 4.1: The discrete-time linear vector Gaussian channel at the n-th channel use.

we have that $n_S \leq K$.[2] Thus, the received signal at the n-th channel use is

$$\mathbf{y}_n = \mathbf{G}_n\mathbf{B}_n\mathbf{s}_n + \mathbf{w}_n, \tag{4.3}$$

where $\mathbf{w}_n$ represents the zero-mean Gaussian noise with identity covariance matrix $\mathbf{R}_{w_n} = \mathbf{I}_{n_R}$.[3] Let $\mathbf{E}_n$ denote the n-th channel use MMSE matrix, which is defined as $\mathbf{E}_n = \mathbb{E}\left\{(\mathbf{s}_n - \hat{\mathbf{s}}_n)(\mathbf{s}_n - \hat{\mathbf{s}}_n)^{\mathsf{T}}\right\}$ and $\hat{\mathbf{s}}_n = \mathbb{E}\left\{\mathbf{s}_n|\mathbf{y}_n\right\}$ is the conditional mean estimator.

Let us express the channel matrix as $\mathbf{G}_n = \mathbf{V}_{\mathbf{G}_n}\mathbf{\Delta}_n\mathbf{U}_{\mathbf{G}_n}^{\mathsf{T}}$, where $\mathbf{\Delta}_n \in \mathbb{R}^{n_S \times n_S}$ is a diagonal matrix that contains the n_S largest eigenvalues of $\mathbf{G}_n$ and $\mathbf{V}_{\mathbf{G}_n} \in \mathbb{R}^{n_R \times n_S}$ and $\mathbf{U}_{\mathbf{G}_n} \in \mathbb{R}^{n_T \times n_S}$ are semi-unitary matrices that contain the row and column associated eigenvectors, respectively. The precoder matrix $\mathbf{B}_n$ can be expressed as $\mathbf{B}_n = \mathbf{U}_{\mathbf{B}_n}\mathbf{\Sigma}_n\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}}$, where $\mathbf{U}_{\mathbf{B}_n} \in \mathbb{R}^{n_T \times n_S}$; $\mathbf{\Sigma}_n \in \mathbb{R}^{n_S \times n_S}$ is a diagonal matrix whose entries are given by the vector $\boldsymbol{\sigma}_n = [\sqrt{p_{1n}}, \ldots, \sqrt{p_{n_S n}}]^{\mathsf{T}}$; and $\mathbf{V}_{\mathbf{B}_n} \in \mathbb{R}^{n_S \times n_S}$ is a unitary matrix. Full CSI is assumed at the transmitter.

As presented in §2.1.1, the energy harvesting process at the transmitter is characterized by a packetized model, which is depicted in Figure 2.13, where J denotes the total number of packets harvested during the N channel uses. We assume that the mean time between energy arrivals, T_e, is considerably larger than the symbol duration time, i.e., $T_e \gg T_s$ and thus we can consider that packet arrival times are aligned at the beginning of a channel use.[4] First, in §4.3 - 4.5, we consider the offline approach as it provides analytical and intuitive expressions. Afterwards, in §4.6, we develop an online transmission strategy where the transmitter only has causal knowledge of the energy harvesting process, i.e., about the past and present energy

[2]We have considered that $\mathbf{G}_n$ is not rank deficient, $\forall n$, which is a realistic assumption due to the random nature of the channel.

[3]Note that if the noise is colored and its covariance matrix $\mathbf{R}_{w_n}$ is known, we can consider the whitened received signal $\mathbf{R}_{w_n}^{-1/2}\mathbf{y}_n$.

[4]In our model, the transmitter can only change its transmission strategy in a channel access basis. Accordingly, if an energy packet arrives in the middle of a channel access, we can assume that the packet becomes available for the transmitter at the beginning of the following channel access.

arrivals (cf. §2.1.3). Recall that, as presented in §2.1.1, we use the term *epoch*, $\tau_j, j = 1, \ldots, J$, to denote the set of channel accesses between two consecutive energy arrivals. As argued at the beginning of the chapter, we assume an infinite capacity battery.

4.3 Throughput maximization problem

In this section, we study the set of linear precoding matrices $\{\mathbf{B}_n\}_{n=1}^N$ that maximizes the input-output mutual information along N independent channel accesses, $\sum_{j=1}^J \sum_{n\in\tau_j} I(\mathbf{s}_n; \mathbf{y}_n)$, where $I(\mathbf{s}_n; \mathbf{y}_n)$ is the n-th channel use mutual information. The design of $\{\mathbf{B}_n\}_{n=1}^N$ is constrained to satisfy the instantaneous ECCs, which impose that energy cannot be used before it has been harvested, $T_s \sum_{j=1}^{\ell} \sum_{n\in\tau_j} ||\mathbf{B}_n \mathbf{s}_n||^2 \leq \sum_{j=1}^{\ell} E_j$, $\ell = 1, \ldots, J$. However, since in each epoch there are several channel accesses with the same channel gains (because $T_c \gg T_s$ and $T_e \gg T_s$), instead of imposing the instantaneous ECCs, we can consider the mean ECCs that become $T_s \sum_{j=1}^{\ell} \sum_{n\in\tau_j} \mathsf{Tr}(\mathbf{B}_n \mathbf{B}_n^{\mathsf{T}}) \leq \sum_{j=1}^{\ell} E_j$, $\ell = 1, \ldots, J$, which do not require prior knowledge of the transmitted symbols at each channel use as only the expectation of the symbols is needed.[5]

Therefore, the mutual information maximization problem is mathematically expressed as

$$\underset{\{\mathbf{B}_n\}_{n=1}^N}{\text{maximize}} \quad \sum_{j=1}^{J} \sum_{n\in\tau_j} I(\mathbf{s}_n; \mathbf{y}_n) \tag{4.4a}$$

$$\text{subject to} \quad T_s \sum_{j=1}^{\ell} \sum_{n\in\tau_j} \mathsf{Tr}(\mathbf{B}_n \mathbf{B}_n^{\mathsf{T}}) \leq \sum_{j=1}^{\ell} E_j, \;\; \ell = 1, \ldots, J. \tag{4.4b}$$

Before addressing the problem in (4.4), let us summarize the state of the art on the precoding strategy that maximizes the mutual information for non-harvesting nodes, which was studied in [58, 76–78] and references therein.[6] In [58], it was shown that, in general, the mutual information, $I(\mathbf{s}_n; \mathbf{y}_n)$, is not a concave function of the precoder and that depends on the precoder only through the matrix $\mathbf{Z}_n = \mathbf{B}_n^{\mathsf{T}} \mathbf{G}_n^{\mathsf{T}} \mathbf{G}_n \mathbf{B}_n$. The authors of [58] also showed that the left singular vectors of the precoder can be chosen to be equal to the eigenvectors of the channel

[5] In general, the energy harvesting and the channel state are two independent random processes. Thus, there may be situations in which only a few channel accesses separate an energy arrival from a change in the channel realization; however, note that these situations are unlikely since $T_c \gg T_s$ and $T_e \gg T_s$. In these unlikely situations, the temporal averaging is not sufficient to ensure that the fulfillment of the mean ECCs implies a fulfillment of the instantaneous ECCs; however, the averaging through the different channel dimensions brings closer the mean and instantaneous ECCs. Thus, the mean ECCs can be used instead of the instantaneous ECCs since the cases in which they differ are indeed very unlikely.

[6] When there is no energy harvesting at the transmitter, the mutual information maximization problem is the one obtained after setting $J = 1$ and $N = 1$ in (4.4). Thus, the mutual information is maximized for a single channel use under a power constraint.

Gram matrix, i.e., $\mathbf{U}_{\mathbf{B}_n} = \mathbf{U}_{\mathbf{G}_n}$. From this, $\mathbf{Z}_n = \mathbf{V}_{\mathbf{B}_n}\boldsymbol{\Sigma}_n^2\boldsymbol{\Delta}_n^2\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}}$ and the mutual information depends on the precoder only through the right eigenvectors and the associated singular values. In [76], it was shown that $I(\mathbf{s}_n;\mathbf{y}_n)$ is a concave function of the squared singular values of the precoder, $\mathsf{diag}(\boldsymbol{\Sigma}_n^2)$, when a diagonal channel matrix is considered. Finally, the authors of [58] stated that the complexity in the design of the globally optimal precoder lies in the right singular vectors of the precoder, $\mathbf{V}_{\mathbf{B}_n}$. Then, in [77], it was shown that $I(\mathbf{s}_n;\mathbf{y}_n)$ is a concave function of the matrix $\mathbf{Z}_n$ and a gradient algorithm over $\mathbf{Z}_n$ was derived to find a locally optimal precoder. References [58, 76, 77] considered a real channel model. The extension to the complex case was done in [78], where the authors pointed out that by allowing the precoder and the channel matrix to be in the complex field the mutual information can be further improved. Then, they proposed an iterative algorithm that determines the globally optimal precoder that imposes that the power constraint must be met with equality.

When energy harvesting is considered, instead of having a single power constraint, we have a set of J ECCs as in (4.4b) and it is not straightforward to determine which of the constraints must be met with equality. This fact implies that the algorithm introduced in [78] is no longer optimal when energy harvesting is considered. Altogether, (4.4) is not a convex optimization problem since the mutual information is not a concave function of the precoder and, hence, its solution is not straightforward. In the following lemma, we generalize Proposition 1 in [58] for the case of considering energy harvesting at the transmitter.

Lemma 4.1. *The left singular vectors of the n-th precoder matrix, $\mathbf{U}_{\mathbf{B}_n}$, are equal to the eigenvectors of the channel Gram matrix $\mathbf{U}_{\mathbf{G}_n}$, $\forall n$.*

Proof: See Appendix 4.A.1. ■

Thanks to Lemma 4.1, the optimal precoding matrix is $\mathbf{B}_n^\star = \mathbf{U}_{\mathbf{G}_n}\boldsymbol{\Sigma}_n^\star\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}\star}, \forall n$, and the dependence of $I(\mathbf{s}_n;\mathbf{y}_n)$ on the precoder is only through $\boldsymbol{\Sigma}_n$ and $\mathbf{V}_{\mathbf{B}_n}$.

In the following lines, we maximize the mutual information w.r.t. $\boldsymbol{\Sigma}_n$ for a given $\mathbf{V}_{\mathbf{B}_n}$. By applying Lemma 4.1 in (4.3), the next equivalent signal model is obtained

$$\mathbf{y}_n = \tilde{\mathbf{G}}_n\mathbf{s}_n + \mathbf{w}_n, \tag{4.5}$$

where $\tilde{\mathbf{G}}_n = \mathbf{V}_{\mathbf{G}_n}\boldsymbol{\Delta}_n\boldsymbol{\Sigma}_n\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}}$ and $\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}}$ is deterministic and known. To fully exploit the diversity of the channel, we assign the dimension of the input vector to be equal to the number of channel eigenmodes, i.e., $n_S = K$. It is easy to verify that the maximization of the mutual information w.r.t. $\boldsymbol{\Sigma}_n$ is not a convex optimization problem. However, if instead we maximize the mutual information w.r.t. the squared singular values of the precoder $\mathbf{p}_n = [p_{1n}, \ldots, p_{Kn}]^{\mathsf{T}}$,

the obtained problem is convex, as shown in the following lines. Thus, the problem reduces to

$$\underset{\{\mathbf{p}_n\}_{n=1}^N}{\text{maximize}} \quad \sum_{j=1}^{J}\sum_{n\in\tau_j} I(\mathbf{s}_n;\tilde{\mathbf{G}}_n\mathbf{s}_n+\mathbf{w}_n) \tag{4.6a}$$

$$\text{subject to} \quad T_s\sum_{j=1}^{\ell}\sum_{n\in\tau_j}\mathbf{1}_K^{\mathsf{T}}\mathbf{p}_n \leq \sum_{j=1}^{\ell} E_j, \quad \ell=1,\ldots,J. \tag{4.6b}$$

Observe that, at the n-th channel access, the input-output mutual information $I(\mathbf{s}_n;\tilde{\mathbf{G}}_n\mathbf{s}_n+\mathbf{w}_n)$ is concave w.r.t. $\mathbf{p}_n$, which was proved in [76]. Therefore, the objective function is concave as the sum of concave functions is concave [63]. Finally, as the constraints are affine in $\mathbf{p}_n$, (4.6) is a convex optimization problem and the KKT are sufficient and necessary optimality conditions. In particular, the optimal solution must satisfy $\mathsf{D}_{\mathbf{p}_n}\mathcal{L}=\mathbf{0}$ (the reader who is not familiar with this notation, which is presented in [116], is referred to [76, Appendix B] for a concise summary), where $\mathcal{L}$ is the Lagrangian that is

$$\mathcal{L}=\sum_{j=1}^{J}\sum_{n\in\tau_j} I(\mathbf{s}_n;\tilde{\mathbf{G}}_n\mathbf{s}_n+\mathbf{w}_n)-\sum_{\ell=1}^{J}\lambda_\ell\left(T_s\sum_{j=1}^{\ell}\sum_{n\in\tau_j}\mathbf{1}_K^{\mathsf{T}}\mathbf{p}_n-\sum_{j=1}^{\ell}E_j\right),$$

where $\{\lambda_\ell\}_{\ell=1}^J$ are the Lagrange multipliers associated with the ECCs in (4.6b). We want to remark that in all the expressions derived in the remainder of the chapter, n refers to some channel access contained in τ_j, which follows from the formulation of $\mathcal{L}$. In order to obtain $\mathsf{D}_{\mathbf{p}_n}\mathcal{L}$, we first need to determine the Jacobian matrix of the mutual information w.r.t. $\mathbf{p}_n$, which is done in the following lemma:

Lemma 4.2. *The Jacobian matrix of the mutual information w.r.t.* $\mathbf{p}_n$ *is* $\mathsf{D}_{\mathbf{p}_n}I(\mathbf{s}_n;\tilde{\mathbf{G}}_n\mathbf{s}_n+\mathbf{w}_n)=\frac{1}{2}\mathsf{diag}^{\mathsf{T}}\left(\boldsymbol{\Delta}_n^2\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}}\mathbf{E}_n\mathbf{V}_{\mathbf{B}_n}\right)$.

Proof: See Appendix 4.A.2. ■

With this result, we can proceed to solve the KKT condition $\mathsf{D}_{\mathbf{p}_n}\mathcal{L}=\mathbf{0}$:

$$\mathsf{D}_{\mathbf{p}_n}\mathcal{L}=\frac{1}{2}\mathsf{diag}^{\mathsf{T}}\left(\boldsymbol{\Delta}_n^2\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}}\mathbf{E}_n\mathbf{V}_{\mathbf{B}_n}\right)-T_s\sum_{\ell=j}^{J}\lambda_\ell\mathbf{1}_K^{\mathsf{T}}=\mathbf{0} \quad \Rightarrow$$

$$\left[\boldsymbol{\Delta}_n^2\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}}\mathbf{E}_n\mathbf{V}_{\mathbf{B}_n}\right]_{kk}=\frac{1}{W_j}, \quad k=1,\ldots,K, \quad n\in\tau_j, \tag{4.7}$$

where W_j is the j-th epoch water level, i.e.,

$$W_j=\frac{1}{2T_s\sum_{\ell=j}^{J}\lambda_\ell}. \tag{4.8}$$

From (4.7), at each channel use, we obtain a set of K conditions that relate the power allocation in each stream (through the MMSE matrix) with the energy harvesting profile (through the epoch's water level). Some properties of the water level W_j can be derived from the KKT optimality conditions:

$$\lambda_\ell \geq 0, \qquad \forall \ell, \tag{4.9}$$

$$\lambda_\ell \left(T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} \mathbf{1}_K^{\mathsf{T}} \mathbf{p}_n - \sum_{j=1}^{\ell} E_j \right) = 0, \qquad \forall \ell. \tag{4.10}$$

Plugging (4.9) in (4.8), it is straightforward to obtain the following property:

Property 4.1. The water level is non-decreasing in time.[7]

From (4.10), we can get more insights in the solution. There are two possibilities to fulfill (4.10):

- Empty battery: This situation occurs when, at the end of the ℓ-th epoch, the node has consumed all the energy, i.e., $T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} \mathbf{1}_K^{\mathsf{T}} \mathbf{p}_n - \sum_{j=1}^{\ell} E_j = 0$.
- Energy flow: This situation occurs when, at the end of the ℓ-th epoch, the node has some remaining energy in the battery, which will be used in the following epochs. When this happens $\lambda_\ell = 0$ and, hence, $W_{\ell+1} = W_\ell$.

Property 4.2. Changes on the water level are only produced when at the end of the previous epoch the node has consumed all the available energy.

Note that the ECCs take into account the energy spent by the node over all the dimensions. Thus, these two properties also hold in a scalar channel model as proved in [85, Theorem 3].

Since the problem in (4.6) is convex, by using (4.7) and Properties 4.1 and 4.2, we can construct efficient numerical algorithms to compute the optimal power allocation, $\{\mathbf{p}_n^\star\}_{n=1}^N$, for a given $\mathbf{V}_{\mathbf{B}_n}$. The maximization of the mutual information w.r.t. $\mathbf{V}_{\mathbf{B}_n}$ is indeed much more complicated as pointed out in [58] for the non-harvesting scenario. In this context, in this work, we focus on the particular case in which $\mathbf{V}_{\mathbf{B}_n} = \mathbf{I}_K$ because, in spite of not being necessarily the globally optimal precoder, it leads to an analytical closed form power allocation that allows an intuitive graphical representation of the solution, as it is explained in the next section. Observe that for any other choice $\mathbf{V}_{\mathbf{B}_n} \neq \mathbf{I}_K$, we must resort to numerical methods to compute the optimal power allocation.

[7] This property is only valid under an infinity battery capacity assumption. When a finite battery is considered the water level may increase or decrease [84].

The design and development of a numerical algorithm that computed the globally optimal precoder at each channel access would be an interesting research problem in its own and is left for future research. We believe that a possible starting point would be to analyze how to expand the algorithms presented in [77] and [78], which exploit the concavity of the mutual information w.r.t. the matrix $\mathbf{Z}_n$ and the fact that the power constraint must be met with equality, to the energy harvesting scenario. Note that if we knew the optimal total power allocation in each channel access, we could run N times the algorithm proposed in [78] to obtain the globally optimal precoder in each channel access; however, this approach has two major drawbacks. First, the optimal total power allocation in each channel access is not known a priori and its computation is not straightforward since the total power consumptions of the different channel accesses belonging to the same epoch must simultaneously satisfy (4.7). The second drawback is the required computational burden since any iterative approach requires a new estimation of the MMSE matrix, $\mathbf{E}_n$, at every iteration since it depends on $\mathbf{V}_{\mathbf{B}_n}$ and $\mathbf{\Sigma}_n$. These two reasons make challenging the applicability of the proposed approach and, hence, different alternatives to find the globally optimal precoder may be required. Altogether, we believe that the development of a numerical algorithm that computes the globally optimal precoder for a WEHN is the object of a new work in its own and is left for future research.

4.4 The MIMO Mercury Water-Flowing solution

In the remainder of the chapter, we consider a communication system in which the precoder is constrained to satisfy $\mathbf{V}_{\mathbf{B}_n} = \mathbf{I}_K$ or, equivalently, a communication system such that both the precoder and channel matrices are diagonal. In spite of the fact that the total achievable mutual information is reduced by forcing $\mathbf{V}_{\mathbf{B}_n} = \mathbf{I}_K$, we consider that it is interesting to study this scenario for the following three reasons:

- The system $\mathbf{y}'_n = \mathbf{\Delta}_n\mathbf{\Sigma}_n\mathbf{s}_n + \mathbf{w}'_n$, with $\mathbf{w}'_n$ being the observed noise at the receiver, is commonly encountered in practical systems where, for simplicity at the decoder, independent symbols are transmitted in each dimension (e.g., in multi-tone transmissions like OFDM), and it has been broadly considered in the literature, indeed, the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ solution was derived for such an input-output system model in [57].

- The optimal power allocation, which is named *Mercury Water-Flowing*, accepts a closed form expression and an intuitive graphical representation.

- We believe that the intuition gained thanks to the *Mercury Water-Flowing* graphical interpretation may help for the design of the algorithm that computes the globally optimal precoder of the problem in (4.4).

In this context, the input-output model $\mathbf{y}'_n = \boldsymbol{\Delta}_n\boldsymbol{\Sigma}_n\mathbf{s}_n + \mathbf{w}'_n$ can be obtained from the general model in (4.5) by setting $\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}} = \mathbf{I}_K$ and $\mathbf{y}'_n = \mathbf{V}_{\mathbf{G}_n}^{\mathsf{T}}\mathbf{y}_n$. From this, we obtain that the equivalent noise is $\mathbf{w}'_n = \mathbf{V}_{\mathbf{G}_n}^{\mathsf{T}}\mathbf{w}_n$. Thus, a set of K independent parallel streams are observed at each channel use. The received signal in the k-th stream is $y'_{kn} = \sqrt{h_{kn}p_{kn}}s_{kn} + w'_{kn}$, where the transmitted symbol is the k-th component of $\mathbf{s}_n$, i.e., $s_{kn} = [\mathbf{s}_n]_k$; $w'_{kn} = [\mathbf{w}'_n]_k$ is the observed noise; $h_{kn} = [\boldsymbol{\Delta}_n^2]_{kk}$ is the channel gain; and $p_{kn} = [\boldsymbol{\Sigma}_n^2]_{kk}$ is the transmission radiated power. Therefore, in this section we solve the following optimization problem:

$$\underset{\{\{p_{kn}\}_{k=1}^K\}_{n=1}^N}{\text{maximize}} \quad \sum_{j=1}^{J}\sum_{n\in\tau_j} I(\mathbf{s}_n;\mathbf{y}'_n), \quad \text{subject to} \quad \text{(4.6b)}. \tag{4.11}$$

Note that $I(\mathbf{s}_n;\mathbf{y}'_n) = I(\mathbf{s}_n;\mathbf{y}_n)$ since a linear unitary rotation in the received signal does not affect the input-output mutual information [78]. Thus, the power allocation that maximizes (4.11) is equal to the one that maximizes (4.6) and it can be obtained by particularizing (4.7) with $\mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}} = \mathbf{I}_K$, i.e., $[\mathbf{E}_n]_{kk} = \frac{1}{h_{kn}W_j}$. From where, it follows that

$$p_{kn} = \frac{1}{h_{kn}} mmse_k^{-1}\left(\min\left\{1, \frac{1}{W_j h_{kn}}\right\}\right), \quad \forall k, \forall j, \forall n \in \tau_j, \tag{4.12}$$

where $mmse_k^{-1}(\cdot)$ is the inverse MMSE function, defined as in the $\mathcal{H}_{\text{g}}$WF power allocation [57] (cf. §2.2.1.2), that returns the SNR of the k-th stream for a given MMSE, which depends on the probability distribution of s_{kn}.

To present the graphical interpretation of the solution, we need to reformulate (4.12) as

$$p_{kn} = \left[W_j - \frac{1}{h_{kn}} G_k\left(\frac{1}{W_j h_{kn}}\right)\right]^+, \quad \forall k, \forall j, \forall n \in \tau_j, \tag{4.13}$$

where $G_k(\psi)$ is defined in (2.14) [57], depends on the modulation used, and satisfies the next lemma:

Lemma 4.3. *The function $G_k(\psi)$ is monotonically decreasing in ψ.*

Proof: See Appendix 4.A.3. ■

Remark 4.1. To demonstrate the validity of the graphical representation presented in this section, we need to analytically demonstrate that $G_k(\psi)$ is monotonically decreasing in ψ. In [57], it was already stated that $G_k(\cdot)$ is decreasing; however, the authors did not provide an analytical proof for their statement. Therefore, we consider that Lemma 4.3 and its explicit proof are crucial to validate the graphical representation introduced in this section.

Observe the similarity of the power allocation found in (4.13) with the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ in (2.13) [57]. The main difference of our solution is that, due to the nature of the energy harvesting process, the water level depends on the channel access. Indeed, from Properties 4.1 and 4.2 we have seen that the node is able to increase the water level as energy is being harvested.

Moreover, observe that if we particularize (4.13) for Gaussian distributed inputs, which have $G_k(\psi) = 1$, $\forall\psi$, (see [57]), the DWF solution in (2.21) (cf. §2.3.1) is recovered. Therefore, the mercury factor gives a measure of how the power allocation is modified when using non-Gaussian input distributions.

Let $\mathcal{H}_g^{\{kn\}}(W_j)$ be the mercury level of the k-th stream at the n-th channel use, i.e.,

$$\mathcal{H}_g^{\{kn\}}(W_j) = \frac{1}{h_{kn}} G_k\left(\frac{1}{W_j h_{kn}}\right), \quad \forall k, \forall j, \forall n \in \tau_j, \tag{4.14}$$

which depends on the gain and water level of the channel use. Then, the power allocated in a certain stream is the difference between the water and mercury levels, i.e.,

$$p_{kn} = \left[W_j - \mathcal{H}_g^{\{kn\}}(W_j)\right]^+.$$

The solution interpretation presented in this section is based on the fact that the mercury level is monotonically increasing in W_j, which follows directly from Lemma 4.3, and generalizes both the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ and the DWF solutions derived in [57] and [84], respectively, which have been presented in §2.2.1.2 and §2.3.1. The *MIMO Mercury Water-Flowing* interpretation, depicted in Figure 4.2, is the following:

(1.) Each parallel channel is represented with a water-porous mercury-nonporous vessel with a base of area T_s[8].

(2.) Then, each vessel is filled with a solid substance up to a height equal to h_{kn}^{-1}.

(3.) A water *right-permeable* material is used to separate the different epochs.

(4.) Each vessel has a faucet that controls the rhythm at which mercury is poured. The faucet modifies the mercury flow so that the relation between mercury and water levels in (4.14) is always satisfied.

(5.) Simultaneously,

- The water level is progressively increased to all epochs at the same time, adding the necessary amount of water to each epoch. The maximum amount of water that can

[8]The vessel boundaries are not depicted in Figure 4.2 for the sake of simplicity.

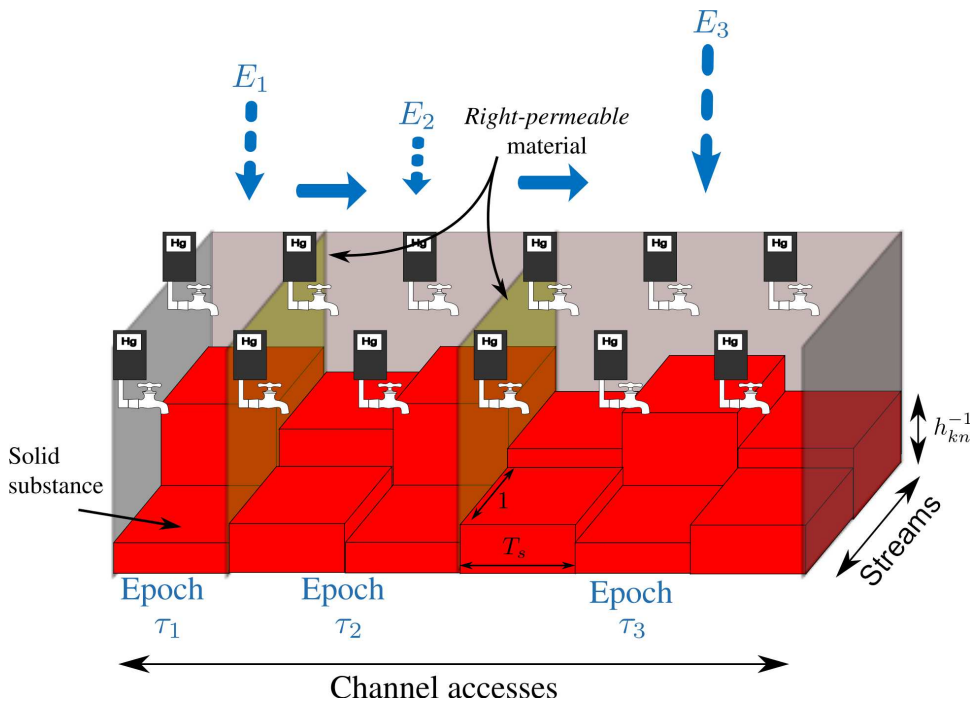

Figure 4.2: Graphical interpretation of the *MIMO Mercury Water-Flowing* solution, where $N = 6$, $J = 3$, and $K = 2$.

be externally added at some epoch is given by the epoch's harvested energy, E_j. Let the water freely flow right through the different epochs.

- Mercury is added to each of the vessels at a different rhythm which is controlled by the vessel's faucet.

(6.) The optimal power allocation in each parallel channel is found when all the epochs have used all the harvested energy and is obtained as the difference between the water and mercury levels.

4.5 MIMO Mercury Water-Flowing offline algorithms

We have designed two different algorithms to compute the optimal *MIMO Mercury Water-Flowing* solution, namely, the Non Decreasing water level Algorithm (NDA) and the Forward Search Algorithm (FSA), which are presented in §4.5.1 and §4.5.2, respectively. Afterwards, in §4.5.3, we prove the algorithms' optimality and analyze their computational complexity.

As shown by the KKT optimality conditions, the water in a certain epoch may flow to the epochs at its right (i.e., from prior to later time instants). This way, the water level over a consecutive set of epochs may be equalized. This set of constant water level epochs is referred to as a *pool*, $\mathbb{P}_m, m = 1, \ldots, M$, where M is the total number of pools and it is unknown

a priori.[9] Note that, since the epochs are a partition of the pools, a certain epoch τ_j is only contained in one pool. However, a pool may contain several epochs, therefore, $M \leq J$.

To compute the power allocation in (4.13), we just need to determine which epochs are contained in each pool $\mathbb{P}_m$ as, once the pools are known, the optimal power allocation of the pool can be found by performing the Mercury/Water-Filling Algorithm ($\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$) introduced in [57], where the m-th pool water level, $\bar{W}_m$, is found by forcing that the energy expended in the pool has to be equal to the energy harvested, which follows from Property 4.2.

The following two algorithms use a different approach two determine the pools:

4.5.1 NDA

The NDA uses the fact that a water level decrease is suboptimal, which follows from the KKT conditions (see Property 4.1), to compute the optimal power allocation as follows:

(1.) Initially, set $M := J$, i.e., every pool contains one epoch $\mathbb{P}_m := \{\tau_m\}$, $m = 1, \ldots, M$.

(2.) Perform the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ in [57] to every pool to obtain the water level, $\bar{W}_m$, in each pool.

(3.) Look for some pool, m', at which the water level decreases, i.e., $\bar{W}_{m'} > \bar{W}_{m'+1}$:

- If some pool is found, merge this pool with the following pool, i.e., $\mathbb{P}_{m'} := \mathbb{P}_{m'} \cup \mathbb{P}_{m'+1}$. The harvested energy of the resulting pool is the sum of the two original pools. Then, the total number of pools has been reduced by one, i.e., $M := M - 1$. Perform the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ to obtain the new water level of the m'-th pool, i.e., $\bar{W}_{m'}$, and go back to 3.
- If no pool is found, the optimal M has been found along with the optimal power allocation.

4.5.2 FSA

The FSA determines the different pools by finding the optimal *transition epochs*, $\{\mathcal{T}_m^\star\}_{m=1}^M$, that are defined as the first epoch of each pool. As stated before, once the pools are known the optimal power allocation is determined by applying the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ to each pool.[10] To determine $\{\mathcal{T}_m^\star\}_{m=1}^M$, we have designed a forward-search algorithm that extends the algorithm introduced

[9] In the previous chapter, where a static channel was considered, a pool denoted the time interval with constant power transmission. Note that, when the channel is static in all the dimensions, constant water level implies constant power and the term pool is consistent with the definition given in the previous chapter.

[10] Observe that, by definition, $\mathcal{T}_1^\star$ is the first epoch.

in [85] to take into account arbitrary input distributions. We explain how to obtain $\mathcal{T}_2^\star$ and the others are found in the same manner:

(1.) Assume that the first pool contains all the epochs, $\mathbb{P}_1 := \{\tau_1, \tau_2, \ldots, \tau_J\}$.

(2.) Perform the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ in [57] to the pool.

(3.) Check whether all the ECCs within the pool are fulfilled:

- If they are not fulfilled, remove the last epoch from $\mathbb{P}_1$ and go back to step 2.
- If they are fulfilled, the optimal transition epoch, $\mathcal{T}_2^\star$, is the first epoch not included in the pool.

The same procedure is repeated to determine the following transition epochs until the J-th epoch is included in some pool. When this happens, the optimal power allocation has been found for all the channel accesses and streams.

4.5.3 Optimality and performance characterization of the offline algorithms

In this section, first, we demonstrate the optimality of the NDA and the FSA, which is presented in Theorem 4.1 and, afterwards, we characterize their associated computational complexity.

Theorem 4.1. *Both the NDA and the FSA compute the optimal power allocation given in* (4.13).

Proof: See Appendix 4.A.4. ■

With the previous theorem, we have demonstrated that both algorithms compute the optimal power allocation; however, the computational cost of such a computation may be very different. To evaluate this, in Appendix 4.A.5, we have conducted an exhaustive study on the computational complexity of each of these two algorithms.

Our performance analysis is three-fold, namely, the best, worst, and average computational complexities are computed. Note that both algorithms internally call the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ a certain number of times to find the optimal solution. Let $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$ denote the number of calls to the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ required to compute the *MIMO Mercury Water-Flowing* solution, which depends on the algorithm itself and on the dynamics of the energy harvesting process. In this context, the best or worst computational complexity is the performance when the minimum or maximum number of calls to $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ are required, respectively. The average computational complexity

	Best		Worst		Average
	(Op.)	($C_{\mathcal{H}_\mathrm{g}\mathrm{WFA}}$)	(Op.)	($C_{\mathcal{H}_\mathrm{g}\mathrm{WFA}}$)	($C_{\mathcal{H}_\mathrm{g}\mathrm{WFA}}$)
NDA	$\hat{a}NK$	J	$\mathcal{O}(\frac{\hat{a}}{2}KNJ)$	$2J-1$	$\mathbb{E}\{C^{NDA}_{\mathcal{H}_\mathrm{g}\mathrm{WFA}}\} = J(\hat{q}+1)-\hat{q}$ $\mathsf{var}\{C^{NDA}_{\mathcal{H}_\mathrm{g}\mathrm{WFA}}\} = (J-1)\hat{q}(1-\hat{q})$
FSA	$\hat{a}NK$	1	$\mathcal{O}(\frac{\hat{a}}{6}KNJ^2)$	$\frac{J^2+J}{2}$	$\mathbb{E}\{C^{FSA}_{\mathcal{H}_\mathrm{g}\mathrm{WFA}}\} = \left(\frac{J^2}{2}+\frac{J}{2}-1\right)\hat{p}+1$ $\mathsf{var}\{C^{FSA}_{\mathcal{H}_\mathrm{g}\mathrm{WFA}}\} = \left(\frac{J}{2}+1\right)^2(J-1)\hat{p}(1-\hat{p})$

Table 4.1: Computational complexity of the NDA and the FSA in the best, worst and average case scenarios.

uses a probabilistic model to compute the average number of calls to $\mathcal{H}_\mathrm{g}\mathrm{WFA}$. Basically, for the NDA we assume that there is a fixed probability $\hat{q}$ that the water level decreases from pool to pool, whereas, for the FSA we assume that there is a fixed probability $\hat{p}$ that a certain ECC is not satisfied. Both $\hat{p}$ and $\hat{q}$ can be experimentally adjusted depending on the energy harvesting profile. The computational complexity in terms of operations (Op.), as well as, in terms of $C_{\mathcal{H}_\mathrm{g}\mathrm{WFA}}$ is summarized in Table 4.1, where $\hat{a}$ is a constant parameter that depends, among others, on the size of the MMSE table required to compute the inverse MMSE function and on the tolerance used in the stopping criteria of the $\mathcal{H}_\mathrm{g}\mathrm{WFA}$. The details of the derivations of the different computational complexities can be found in Appendix 4.A.5. In §4.7.2, the theoretical results on the algorithms' computational complexities are compared with the ones obtained through simulation.

4.6 Online algorithm

Up to now, we have assumed that the transmitter has non-causal knowledge of both the CSI and the EHI, which might be realistic in scenarios in which the energy source is predictable, cf., §2.1.3. Therefore, the *MIMO Mercury Water-Flowing* solution provides an upper bound on the achievable mutual information of practical schemes in which $\mathbf{V}_{\mathbf{B}_n} = \mathbf{I}_K$. In this section, we develop an online algorithm, which is strongly based on the optimal offline solution, the *MIMO Mercury Water-Flowing* power allocation, but that does not require future knowledge of neither the energy arrivals nor the channel state, that computes a suboptimal power allocation of the problem in (4.11).

Let F_w be the *flowing window* that is an input parameter of the online algorithm that refers to the number of channel accesses in which the water is allowed to flow, which can be obtained by a previous training under the considered energy harvesting profile, and let an *event* denote a channel access in which a change in the channel state is produced or an energy packet is

harvested, i.e.,

$$\mathcal{E}_t = \{n|\mathbf{\Delta}_{n-1} \neq \mathbf{\Delta}_n\} \cup \{n|n = e_j, j = 1, \ldots, J\}, \quad t = 1, \ldots, S,$$

where $S \in [J, N]$. In this context, the proposed online algorithm proceeds as follows:

(1.) The initial energy in the battery, E_1, is allocated to the different streams of the first F_w channel accesses according to the $\mathcal{H}_{\text{g}}$WFA where the channel is expected to be static and equal to the gain of the first channel use $\mathbf{\Delta}_n = \mathbf{\Delta}_1, \forall n \in [1, F_w]$.

(2.) When the transmitter detects an event, it updates the allocated power of the channel accesses $n \in [\mathcal{E}_t, \min\{\mathcal{E}_t + F_w - 1, N\}]$ by using the $\mathcal{H}_{\text{g}}$WFA with the remaining energy in the battery and with the energy of the harvested packet (if the event is an energy arrival), i.e., $\sum_{j|e_j \leq \mathcal{E}_t} E_j - T_s \sum_{n=1}^{\mathcal{E}_t - 1} \sum_k p_{kn}$, and by assuming that the channel remains constant during the flowing window, i.e., $\mathbf{\Delta}_n = \mathbf{\Delta}_{\mathcal{E}_t}, \forall n \in [\mathcal{E}_t, \min\{\mathcal{E}_t + F_w - 1, N\}]$. Note that the transmitter may stay silent in some channel accesses if the difference between two consecutive incoming energy packets is greater than the flowing window, $e_j - e_{j-1} > F_w$.[11]

(3.) Step (2.) is repeated until the N-th channel access is reached. The proposed online algorithm satisfies the ECCs and, as pointed out, does not require future information of neither the channel state nor the energy arrivals.

The performance in terms of achieved mutual information depends on the correctness of the estimation of the flowing window, F_w, as discussed with the numerical analysis in §4.7. In summary, this online algorithm provides us a lower bound on the mutual information that can be achieved with sophisticated online algorithms that make use of precise statistical models of the energy harvesting process and channel state. A myriad of works have dealt with channel modeling; however, as it has been argued in §2.1.1, having a precise statistical model of the energy harvesting process is indeed not trivial as it depends on many factors such as the harvester used by the node (e.g., a solar panel, piezoelectric generator, etc.), the node's placement, mobility, etc. According to this, to develop better online algorithms it is key that during the following years statistical models of the energy harvesting process are developed, which must capture the energy variability due to environmental and climatic conditions, and the node's placement and mobility.

[11] This situation rarely takes place in practice since, in most common situations, F_w is several times the mean number of channel accesses per epoch. For example, in the simulated framework presented in §4.7, we have obtained that F_w is 4.4 times the mean number of channel accesses per epoch.

4.7 Results

This section first evaluates the gain of the proposed *MIMO Mercury Water-Flowing* solution w.r.t. other suboptimal solutions and, secondly, it presents an analysis through simulation of the computational complexity of the NDA and FSA.

4.7.1 Results on the MIMO Mercury Water-Flowing solution

In this section, we evaluate the mutual informations obtained with the optimal offline solution, the *MIMO Mercury Water-Flowing* ($\mathcal{H}_{\mathrm{g}}$-WFlow), and with the online policy presented in §4.6.

To the best of our knowledge, there are no offline algorithms in the literature that maximize the mutual information by jointly considering energy harvesting at the transmitter and arbitrary distributed input symbols. In this context, we use the following three algorithms, which are optimal in different setups and have been adapted to the energy harvesting scenario, as a reference to evaluate the mutual information achieved by the proposed offline and online solutions:

- The DWF solution in (2.21) that is the optimal offline power allocation for a WEHN when the distribution of the input symbols is Gaussian.
- Epoch-by-Epoch Water-Filling ($\mathrm{EbE\text{-}WF}$) that uses the CWF power allocation in (2.10) by forcing that the harvested energy in a certain epoch is expended in the channel accesses of that same epoch.
- Epoch-by-Epoch Mercury/Water-Filling ($\mathrm{EbE\text{-}}\mathcal{H}_{\mathrm{g}}\mathrm{WF}$) where the power allocation is obtained by using the $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ solution in (2.13) and forcing that the harvested energy in a certain epoch is expended in the channel accesses of that same epoch.

We have considered a channel matrix of rank $K = 4$, where the channel gains are uniformly distributed in $(0, 1)$. The modulations used in each stream are BPSK, 4-PAM, 16-PAM, and 32-PAM, respectively. The symbol duration is $T_s = 10$ ms and $N = 100$ channel accesses have been considered during which a total of $J = 40$ energy packets are harvested. Energy arrivals are uniformly distributed along the channel accesses and with random amounts of energy, which are normalized according to the total harvested energy that varies along the x-axis of Figure 4.3. The y-axis shows the mutual information obtained with the different strategies. After some training in this scenario, we have obtained that the optimal flowing window is $F_w = 11$ channel accesses.

As shown in Figure 4.3, our proposed solution, the $\mathcal{H}_{\mathrm{g}}$-WFlow, outperforms all the suboptimal strategies. The improvement of the $\mathcal{H}_{\mathrm{g}}$-WFlow w.r.t. the $\mathrm{EbE\text{-}}\mathcal{H}_{\mathrm{g}}\mathrm{WF}$ comes from letting

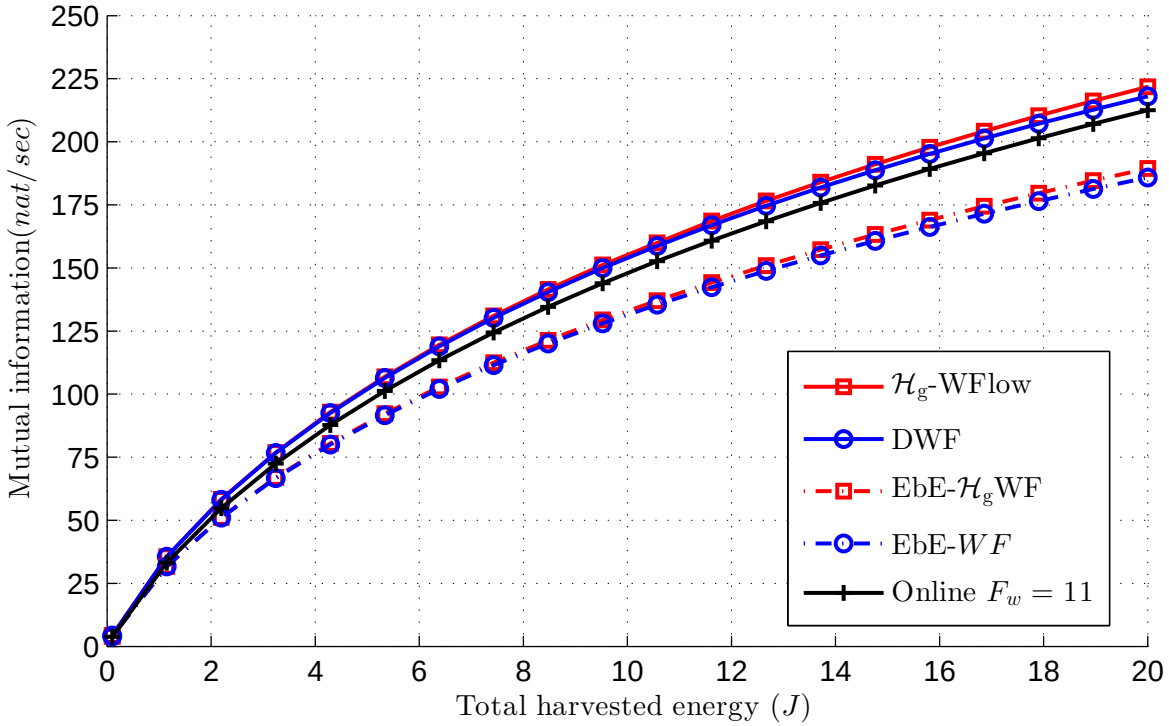

Figure 4.3: Mutual information for the different transmission strategies versus total harvested energy.

the water to flow across epochs and, hence, it directly depends on the parameter J since the higher is the number of epochs, the higher is the mutual information gain that can be achieved by letting the water flow.[12] The same happens with the improvement of the DWF w.r.t. EbE-WF. On the other hand, the mutual information gain of the EbE-$\mathcal{H}_g$WF and $\mathcal{H}_g$-WFlow w.r.t. their respective CWF strategies, EbE-WF and DWF, comes from the use of mercury in the resource allocation. Thus, when the energy availability is low, both perform similarly because the node is working in the low SNR regime in which the mutual information of finite alphabets is well approximated by the mutual information of the Gaussian distribution [117]. However, when the energy availability is high, the EbE-$\mathcal{H}_g$WF and $\mathcal{H}_g$-WFlow achieve a higher mutual information than their respective CWF strategies since the mutual information of finite constellations asymptotically saturates (not more than $\log_2 \mathcal{M}$ bits of information can be sent per channel use), cf., §2.2.1.2. Finally, note that, in spite of not having knowledge of the energy arrivals nor channel state, the online power allocation performs close to the offline optimal $\mathcal{H}_g$-WFlow in the low SNR regime. When the available energy increases, the gap between the *MIMO Mercury Water-Flowing* and the proposed online algorithm also increases, nevertheless the online algorithm still presents a reasonably good mutual information outperforming any Epoch-by-Epoch strategy.

[12] When $J = 1$, the solid and dashed curves overlap since there is only one epoch.

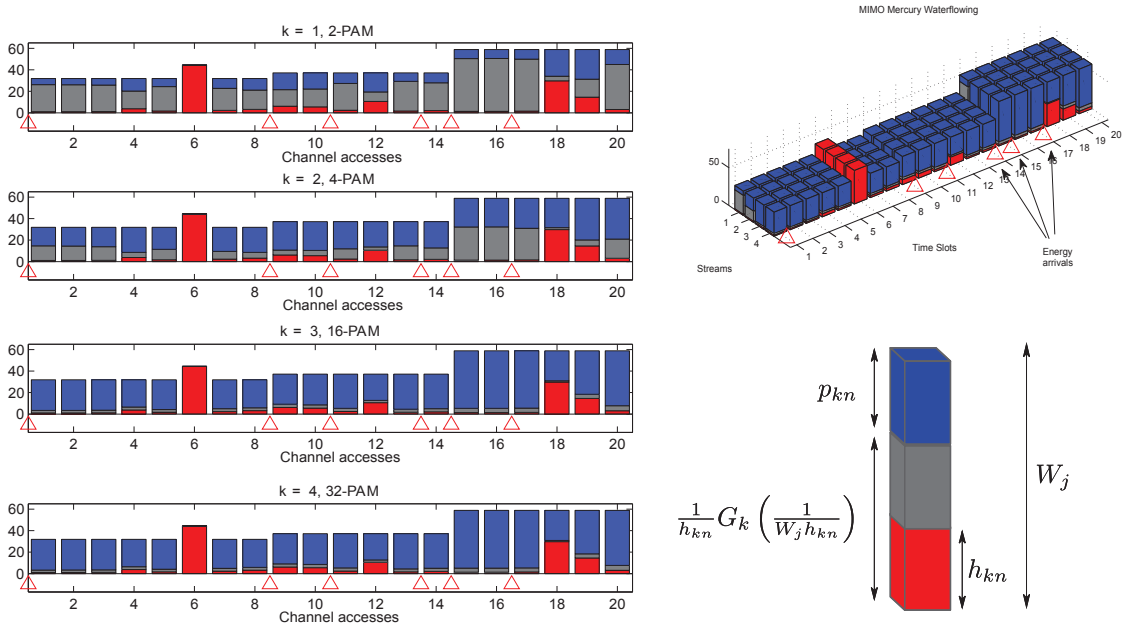

Figure 4.4: Graphical representation of the *MIMO Mercury Water-Flowing* solution. The red, gray, and blue solid bars represent the inverse of the channel gain, the mercury and the water levels, respectively. The allocated power is obtained as the difference between the water level and the mercury level.

The study of the performance in the static scenario is of special interest because the assumption of having future knowledge of the channel state, which has been used for the design of the optimal offline solution, becomes realistic when the channel is static. We have evaluated the achieved mutual information in the above setup for the static channel case and we have obtained similar results to the ones in Figure 4.3, where the only difference is that the achieved mutual information of the different algorithms in the static case is slightly lower since there is less channel gains diversity to assign the available energy.[13]

In Figure 4.4, the power allocation obtained by the *MIMO Mercury Water-Flowing* solution in a single simulation is shown for $N = 20$ and $K = 4$, where the modulations used in the streams 1-4 are BPSK, 4-PAM, 16-PAM and 32-PAM, respectively. Six energy arrivals are produced at the beginning of the channel accesses marked with a triangle. The gains have been generated randomly along channel uses, but fixed constant along streams to ease the observation of the mercury level obtained for the different modulations. As expected from Property 4.1, the obtained water level is an increasing stepwise function. Observe that the solution contains three pools, i.e., three different water levels, where the epochs contained in each pool are $\mathbb{P}_1 = \{\tau_1\}$, $\mathbb{P}_2 = \{\tau_2, \tau_3, \tau_4\}$, and $\mathbb{P}_3 = \{\tau_5, \tau_6\}$. Moreover, observe that under the same channel gain and water level, the mercury level decreases as the modulation dimension increases.

[13]The figure of the static scenario has been omitted for the sake of brevity.

4.7.2 Results on the algorithms' performance

In §4.5.3, we have given a summary of the computational complexity of the NDA and FSA (Table 4.1 summarizes the obtained results). In this section, we compare the theoretical and experimental performance of both algorithms.

From the simulations, we confirm that, in the best and worst case scenarios, the experimental computational complexity shown in Figure 4.5 fits the theoretical results presented in Table 4.1. Regarding the average case scenario, the mean number of calls to $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ of the NDA fits the analytical expression $\mathbb{E}\{C^{NDA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}\} = J(\hat{q}+1)-\hat{q}$ for a value of $\hat{q} = 0.98$. Regarding the FSA, the mean obtained through simulation and the analytically computed expression $\mathbb{E}\left\{C^{FSA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}\right\} = \left(\frac{J^2}{2} + \frac{J}{2} - 1\right)\hat{p} + 1$ differ from one another. Observe that the quadratic and linear terms of J have the same weight independently of the value of $\hat{p}$. However, it is easy to observe in Figure 4.5 that the linear component dominates over the quadratic. Therefore, there is a mismatch between the analytical and experimentally obtained expressions. We believe that this mismatch is due to the fact that in order to obtain some tractable model (see Appendix 4.A.5), we have assumed that all the ECCs have the same probability $\hat{p}$ of not being satisfied; however, in reality this probability is not necessarily equal but depends on the dynamics of the energy harvesting process.

Regardless of the aforementioned mismatch, we observe that, in our simulated energy harvesting set up (the amount of energy in the packets is uniformly distributed), both algorithms have a similar performance in the average case scenario. Note that the difference between the best and worst case scenario is much smaller for the NDA than for the FSA. This comes from the fact that, in the worst case scenario, the FSA has a quadratic dependence on J, whereas, for the NDA the dependence is linear. This makes the NDA more robust in front of changes in the energy harvesting profile. In other words, if the energy harvesting profile changes, the FSA has more margin to either improve or degrade its performance. For instance, consider a node that harvest energy through a solar panel. Then, if the node's initial battery is very high and it is operating in the sunset (the amount of harvested energy at the beginning of the transmission duration is higher than the amount harvested at the end), it is likely that the performance of the FSA is close to the best case scenario, i.e., a single call to the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$. On the other hand, if the battery is almost empty at the beginning and the node operates in the sunrise, the performance of the FSA will be very poor.

To conclude the discussion between the NDA and the FSA, we want to highlight again that the NDA is more robust to changes in the energy harvesting profile. However, the FSA may be preferable in certain profiles as in its best case performance just requires a call to the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$. Therefore, we believe that the algorithm selection must be done by taking into account the energy harvesting profile and the environmental conditions in which the node is operating.

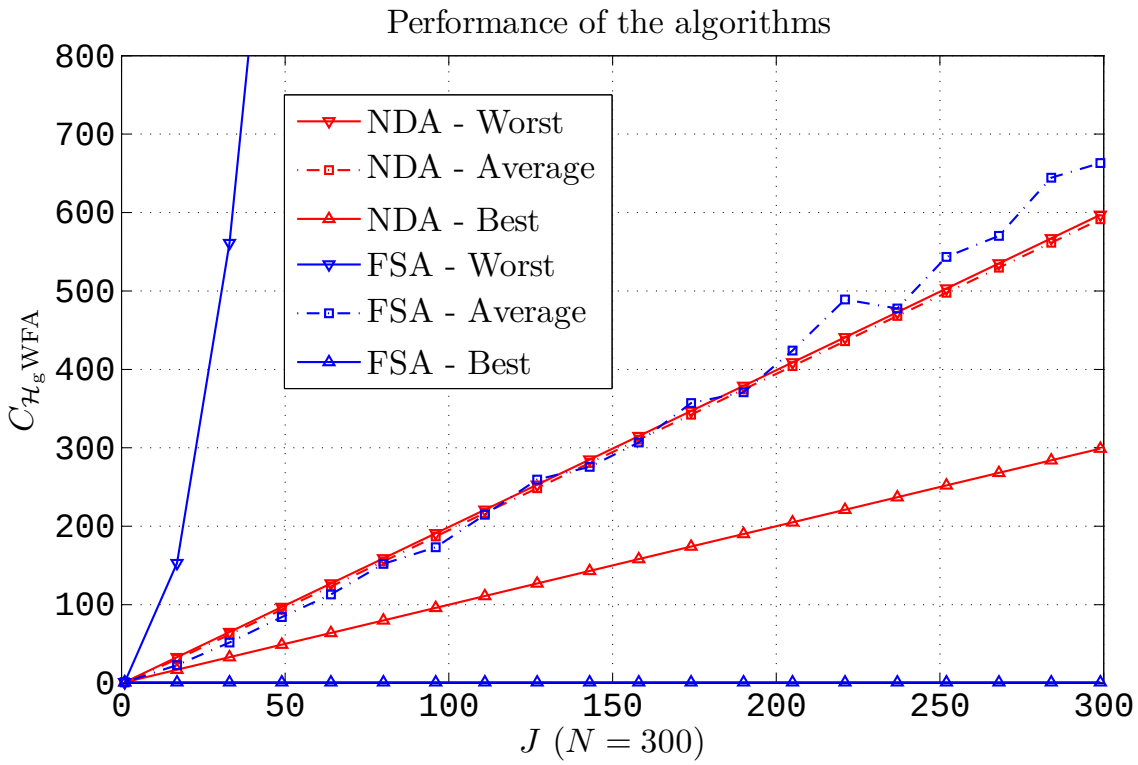

Figure 4.5: Analysis of the performance of the NDA and FSA in terms of $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$.

4.8 Conclusions

In this chapter, we have considered a WEHN transmitting arbitrarily distributed symbols in a discrete-time linear vector Gaussian channel. We have studied the precoding strategy that maximizes the mutual information by taking into account causality constraints on the use of energy. We have proved that the optimal left singular vectors of the precoder matrix diagonalize the channel, similarly as in the optimal precoder for the case of non-harvesting nodes. We have derived the expression $\left[\boldsymbol{\Delta}_n^2 \mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}} \mathbf{E}_n \mathbf{V}_{\mathbf{B}_n}\right]_{kk} = \frac{1}{W_j}$ that relates the singular values of the precoder (through the MMSE matrix) with the energy harvesting profile (through the different water levels). The derivation of the optimal right singular vectors, $\mathbf{V}_{\mathbf{B}_n}^{\star}$, is left as an open problem. Then, we have derived the *MIMO Mercury Water-Flowing* solution, the optimal power allocation when $\mathbf{V}_{\mathbf{B}_n} = \mathbf{I}_K$, which can be expressed in closed form and accepts an intuitive graphical interpretation based on the fact that the power allocation in a certain stream is the difference between the water level and the mercury level, which, as shown in this chapter, is a monotonically increasing function of the water level. Additionally, we have developed two different algorithms that compute the *MIMO Mercury Water-Flowing* solution and we have analytically and experimentally evaluated their computational complexity. We have also proposed an online algorithm that only requires causal knowledge of the energy harvesting process and channel state. Finally, through numerical simulations, we have shown a substantial increase in

the mutual information w.r.t. other suboptimal offline strategies, which do not account for the shape, size and distribution of the input symbol or do not exploit the water level equalization across epochs, and we have seen that the mutual information achieved with the online algorithm is close to the one of the *MIMO Mercury Water-Flowing* solution.

4.A Appendix

4.A.1 Proof of Lemma 4.1

Let us assume that the optimal precoding matrices of the channel accesses $n = 2, \ldots, N$ are known, i.e., $\{\mathbf{B}_n^\star\}_{n=2}^N$. Then, we focus on finding the optimal precoding matrix of the first channel use $\mathbf{B}_1^\star$. The problem in (4.4) is equivalent to

$$\begin{aligned} \underset{\mathbf{B}_1}{\text{maximize}} \quad & I(\mathbf{s}_1; \mathbf{y}_1) + a && \text{(4.15a)} \\ \text{subject to} \quad & T_s \mathsf{Tr}(\mathbf{B}_1 \mathbf{B}_1^\mathsf{T}) + b + c(\ell) \leq \sum_{j=1}^{\ell} E_j, \quad \ell = 1, \ldots, J, \end{aligned}$$

where a, b and $c(\ell)$ do not depend on $\mathbf{B}_1$. By only keeping the most restrictive constraint, which is denoted by $\bar{P}$, the previous optimization problem reduces to

$$\begin{aligned} \underset{\mathbf{B}_1}{\text{maximize}} \quad & I(\mathbf{s}_1; \mathbf{y}_1) && \text{(4.16a)} \\ \text{subject to} \quad & \mathsf{Tr}(\mathbf{B}_1 \mathbf{B}_1^\mathsf{T}) \leq \bar{P}. && \text{(4.16b)} \end{aligned}$$

Finally, once the problem is expressed as (4.16), it is known from [58, Prp. 1] that the left singular vectors of $\mathbf{B}_1^\star$ can be chosen to coincide with the eigenvectors of the channel Gram matrix, i.e., $\mathbf{U}_{\mathbf{B}_1} = \mathbf{U}_{\mathbf{G}_1}$. A similar approach can be applied to show that $\{\mathbf{B}_n^\star\}_{n=2}^N$ diagonalize their respective channels. ■

4.A.2 Proof of Lemma 4.2

By applying the chain rule, we have that

$$\mathsf{D}_{\mathbf{p}_n} I(\mathbf{s}_n; \tilde{\mathbf{G}}_n \mathbf{s}_n + \mathbf{w}_n) = \mathsf{D}_{\tilde{\mathbf{G}}_n} I(\mathbf{s}_n; \tilde{\mathbf{G}}_n \mathbf{s}_n + \mathbf{w}_n) \; \mathsf{D}_{\mathbf{p}_n} \tilde{\mathbf{G}}_n.$$

The first term in the previous equation can be easily derived from (2.12) as

$$\mathsf{D}_{\tilde{\mathbf{G}}_n} I(\mathbf{s}_n; \tilde{\mathbf{G}}_n \mathbf{s}_n + \mathbf{w}_n) = \mathsf{vec}^{\mathsf{T}}(\tilde{\mathbf{G}}_n \mathbf{E}_n).$$

The second term, $\mathsf{D}_{\mathbf{p}_n} \tilde{\mathbf{G}}_n$, is

$$\begin{aligned}
\mathsf{D}_{\mathbf{p}_n} \tilde{\mathbf{G}}_n &= \frac{1}{2}(\mathbf{V}_{\mathbf{B}_n} \otimes \mathbf{V}_{\mathbf{G}_n} \boldsymbol{\Delta}_n) \bar{\mathbf{S}}_K \mathsf{Diag}(\boldsymbol{\sigma}_n^{-1}) && (4.17)\\
&= \frac{1}{2}(\mathbf{V}_{\mathbf{B}_n} \otimes \mathbf{V}_{\mathbf{G}_n} \boldsymbol{\Delta}_n) \bar{\mathbf{S}}_K \mathsf{Diag}(\mathsf{diag}(\boldsymbol{\Sigma}_n^{-1})) && (4.18)\\
&= \frac{1}{2}(\mathbf{V}_{\mathbf{B}_n} \otimes \mathbf{V}_{\mathbf{G}_n} \boldsymbol{\Delta}_n) \bar{\mathbf{S}}_K \bar{\mathbf{S}}_K^T (\mathbf{I}_K \otimes \boldsymbol{\Sigma}_n^{-1}) \bar{\mathbf{S}}_K && (4.19)\\
&= \frac{1}{2}(\mathbf{V}_{\mathbf{B}_n} \otimes \mathbf{V}_{\mathbf{G}_n} \boldsymbol{\Delta}_n)(\mathbf{I}_K \otimes \boldsymbol{\Sigma}_n^{-1}) \bar{\mathbf{S}}_K, && (4.20)
\end{aligned}$$

where (4.17) can be proved in a similar manner to $\mathsf{D}_{\boldsymbol{\lambda}} \mathbf{P}$ in [76, Proof of Theorem 5] with $\bar{\mathbf{S}}_K \in \mathbb{R}^{K^2 \times K}$ being the reduction matrix introduced in [76] (see Appendix 4.A.6 for a concise summary on the properties of $\bar{\mathbf{S}}_K$). In (4.19) and (4.20), we have applied Properties 6 and 8 in Appendix 4.A.6, respectively.

Therefore, we have that $\mathsf{D}_{\mathbf{p}_n} I(\mathbf{s}_n; \tilde{\mathbf{G}}_n \mathbf{s}_n + \mathbf{w}_n)$ is

$$\begin{aligned}
\mathsf{D}_{\mathbf{p}_n} I(\mathbf{s}_n; \tilde{\mathbf{G}}_n \mathbf{s}_n + \mathbf{w}_n) &= \frac{1}{2} \mathsf{vec}^{\mathsf{T}}(\tilde{\mathbf{G}}_n \mathbf{E}_n) \left(\mathbf{V}_{\mathbf{B}_n} \otimes \mathbf{V}_{\mathbf{G}_n} \boldsymbol{\Delta}_n\right) \left(\mathbf{I}_K \otimes \boldsymbol{\Sigma}_n^{-1}\right) \bar{\mathbf{S}}_K && (4.21)\\
&= \frac{1}{2} \mathsf{vec}^{\mathsf{T}}(\tilde{\mathbf{G}}_n \mathbf{E}_n) \left(\mathbf{V}_{\mathbf{B}_n} \otimes \mathbf{V}_{\mathbf{G}_n} \boldsymbol{\Delta}_n \boldsymbol{\Sigma}_n^{-1}\right) \bar{\mathbf{S}}_K && (4.22)\\
&= \frac{1}{2} \mathsf{vec}^{\mathsf{T}} \left((\mathbf{V}_{\mathbf{G}_n} \boldsymbol{\Delta}_n \boldsymbol{\Sigma}_n^{-1})^{\mathsf{T}} \tilde{\mathbf{G}}_n \mathbf{E}_n \mathbf{V}_{\mathbf{B}_n}\right) \bar{\mathbf{S}}_K && (4.23)\\
&= \frac{1}{2} \mathsf{vec}^{\mathsf{T}} \left(\boldsymbol{\Delta}_n^2 \mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}} \mathbf{E}_n \mathbf{V}_{\mathbf{B}_n}\right) \bar{\mathbf{S}}_K = \frac{1}{2} \mathsf{diag}^{\mathsf{T}} \left(\boldsymbol{\Delta}_n^2 \mathbf{V}_{\mathbf{B}_n}^{\mathsf{T}} \mathbf{E}_n \mathbf{V}_{\mathbf{B}_n}\right), && (4.24)
\end{aligned}$$

where, in (4.22) and (4.23), we have used that $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ and $\mathsf{vec}(\mathbf{ABC}) = (\mathbf{C}^{\mathsf{T}} \otimes \mathbf{A})\mathsf{vec}\mathbf{B}$ for any matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ such that the matrix products $\mathbf{AC}$, $\mathbf{BD}$, and $\mathbf{ABC}$ are well defined [116]. Finally, (4.24) follows from the definition of the reduction matrix (see Appendix 4.A.6). This concludes the proof. ■

4.A.3 Proof of Lemma 4.3

Let ψ be some fixed MMSE that can be obtained as

$$\psi = mmse_G(snr_G) = mmse_A(snr_A), \tag{4.25}$$

where $mmse_G(snr_G)$ and $mmse_A(snr_A)$ give the MMSE as a function of the SNR for a Gaussian and for an arbitrary input distribution, respectively. Thus, snr_G and snr_A are the associated required SNRs to achieve the error ψ for these distributions.

Similarly, the required SNR to obtain a certain error can be computed by the inverse MMSE function as $snr_G = mmse_G^{-1}(\psi)$ and $snr_A = mmse_A^{-1}(\psi)$.

For the Gaussian case, it is broadly known that $\psi = mmse_G(snr_G) = \frac{1}{1+snr_G}$ with derivative $\frac{\mathrm{d}mmse_G(snr_G)}{\mathrm{d}snr_G} = \frac{-1}{(1+snr_G)^2}$. Similarly, $snr_G = mmse_G^{-1}(\psi) = \frac{1}{\psi} - 1$ and $\frac{\mathrm{d}mmse_G^{-1}(\psi)}{\mathrm{d}\psi} = \frac{-1}{\psi^2}$.

Note that for any generic function $f(x)$, such that $f(x)$ and $f^{-1}(x)$ are differentiable, it is verified that $\frac{\mathrm{d}f^{-1}(f(x))}{\mathrm{d}x} = \frac{\mathrm{d}f^{-1}(f(x))}{\mathrm{d}f(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x} = 1$. By applying the previous property, the following relation is obtained:

$$\frac{\mathrm{d}mmse_G^{-1}(\psi)}{\mathrm{d}\psi}\frac{\mathrm{d}mmse_G(snr_G)}{\mathrm{d}snr_G} = \frac{\mathrm{d}mmse_A^{-1}(\psi)}{\mathrm{d}\psi}\frac{\mathrm{d}mmse_A(snr_A)}{\mathrm{d}snr_A}.$$

Recall that $G(\psi) = \frac{1}{\psi} - mmse_A^{-1}(\psi)$ as $\psi \in [0,1]$. Then, its derivative is

$$\begin{aligned}\frac{\mathrm{d}G(\psi)}{\mathrm{d}\psi} &= \frac{-1}{\psi^2} - \frac{\mathrm{d}mmse_A^{-1}(\psi)}{\mathrm{d}\psi} = \frac{\mathrm{d}mmse_G^{-1}(\psi)}{\mathrm{d}\psi} - \frac{\mathrm{d}mmse_A^{-1}(\psi)}{\mathrm{d}\psi} \\ &= \frac{\frac{\mathrm{d}mmse_G^{-1}(\psi)}{\mathrm{d}\psi}}{\frac{\mathrm{d}mmse_A(snr_A)}{\mathrm{d}snr_A}}\left(\frac{\mathrm{d}mmse_A(snr_A)}{\mathrm{d}snr_A} - \frac{\mathrm{d}mmse_G(snr_G)}{\mathrm{d}snr_G}\right).\end{aligned}$$

In [118], it was recently shown that $mmse_A(snr_A) = E\{M_2\}$ and $\frac{\mathrm{d}mmse_A(snr_A)}{\mathrm{d}snr_A} = -E\{M_2^2\}$, where $M_2 = \mathsf{var}\{x|\sqrt{snr_A}x + n\}$. Therefore, the first term of the previous equation is always nonnegative since both the MMSE and the inverse MMSE functions are decreasing for any distribution. Accordingly, we have that

$$\begin{aligned}\mathsf{sign}\left(\frac{\mathrm{d}G(\psi)}{\mathrm{d}\psi}\right) &= \mathsf{sign}\left(\frac{\mathrm{d}mmse_A(snr_A)}{\mathrm{d}snr_A} - \frac{\mathrm{d}mmse_G(snr_G)}{\mathrm{d}snr_G}\right) && (4.26)\\ &= \mathsf{sign}\left(\frac{\mathrm{d}mmse_A(snr_A)}{\mathrm{d}snr_A} + \frac{1}{(1+snr_G)^2}\right) && (4.27)\\ &= \mathsf{sign}\left(\frac{\mathrm{d}mmse_A(snr_A)}{\mathrm{d}snr_A} + mmse_A(snr_A)^2\right) && (4.28)\\ &= \mathsf{sign}\left(-E\{M_2^2\} + (E\{M_2\})^2\right) = \mathsf{sign}\left(-\mathsf{var}\{M_2\}\right), && (4.29)\end{aligned}$$

where in (4.28), we have used that $snr_G = \frac{1}{mmse_A(snr_A)} - 1$, which follows from (4.25). In (4.29), we have used the recently found expressions of the MMSE and its derivative [118]. Finally, as the variance is nonnegative, $\frac{\mathrm{d}G(\psi)}{\mathrm{d}\psi} \leq 0$ and $G(\psi)$ is a monotonically decreasing function.

■

4.A.4 Proof of Theorem 4.1

The optimality of the algorithms is proved by demonstrating that the power allocation obtained by means of each of the algorithms satisfies the KKT sufficient optimality conditions:

(1.) $\frac{\partial \mathcal{L}}{\partial p_{kn}} = 0, \forall k, n.$

(2.) $T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} \sum_{k=1}^{K} p_{kn} \leq \sum_{j=1}^{\ell} E_j, \ell = 1, \ldots, J.$

(3.) $\lambda_\ell \geq 0, \ell = 1, \ldots, J.$

(4.) $\lambda_\ell \big(T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} \sum_{k=1}^{K} p_{kn} - \sum_{j=1}^{\ell} E_j \big) = 0, \ell = 1, \ldots, J.$

Moreover, we know that by the end of the transmission the battery must be empty since, otherwise, the remaining energy in the battery can be used to increase the total mutual information. Thus, (2.) must be met with equality for $\ell = J$. Note that both algorithms compute a power allocation strategy that satisfies the ECCs and that by the end of the last channel access all the energy has been used. Therefore, (2.) is satisfied $\forall \ell$ and it is satisfied with equality for $\ell = J$. From Property 4.1, if the water level is non-decreasing in time, then (3.) can be verified. In the NDA, the water level is clearly non-decreasing in time. Regarding the FSA, if some ECC is not satisfied, it is because the water level must be reduced before the point where the ECC is not satisfied and increased afterwards. Indeed, this is what the algorithm does in the procedure of finding the optimal pools. Therefore, (3.) is also satisfied in the FSA. Finally, since both algorithms compute the optimal power allocation within a pool by using the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$, where the water level is found by forcing that all the available energy must be used by the end of the pool, conditions (1.) and (4.) are satisfied. With this, we have demonstrated that the power allocation computed by the NDA and the FSA is the optimal power allocation. ■

4.A.5 Computational complexity of the algorithms

In this appendix, we study the performance of the two algorithms that compute the *MIMO Mercury Water-Flowing* solution, the NDA and the FSA.

We have carried out a three-fold analysis, namely, the best, worst and average computational complexity. As mentioned before, both algorithms internally call the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ a certain number of times to find the optimal solution. The performance is evaluated in terms of operations and number of calls to the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ required to compute the *MIMO Mercury Water-Flowing* solution, $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$.

Before getting into the complexity of each of the aforementioned scenarios, let us first compute the complexity of the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ when the algorithm computes the power allocation of NK parallel channels, i.e.,

$$CC_{\mathcal{H}_{\mathrm{g}}\mathrm{WF}}(N,K) = \hat{a}NK, \tag{4.30}$$

where $\hat{a}$ is a constant parameter that depends, among others, on the size of the MMSE table required to compute the inverse MMSE function $mmse_k^{-1}(\cdot)$ and on the tolerance used in the stopping criteria of the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$. Now, let us proceed to compute the computational complexity of the NDA and FSA.

4.A.5.1 Computational complexity in the best case scenario

NDA: The best case scenario for the NDA occurs when the resulting water levels of applying the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ at each epoch are non-decreasing throughout all the transmission. Thus, the best case computational complexity for the NDA is

$$CC^{B}_{NDA}(N,K,J) = \sum_{j=1}^{J} CC_{\mathcal{H}_{\mathrm{g}}\mathrm{WF}}(L_j,K) = \sum_{j=1}^{J} \hat{a}L_jK = \hat{a}NK, \tag{4.31}$$

where L_j is the number of channel accesses contained in τ_j and, accordingly, $\sum_{j=1}^{J} L_j = N$. Note that the number of calls to the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ is $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}} = J$.

FSA: Regarding the FSA the best performance is obtained when the algorithm can stop at the first iteration, i.e., after applying the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ to the N channel accesses it is observed that the resulting power allocation satisfies all the ECCs. Thus, we have that

$$CC^{B}_{FSA}(N,K,J) = CC_{\mathcal{H}_{\mathrm{g}}\mathrm{WF}}(N,K) = \hat{a}NK. \tag{4.32}$$

Note that the number of calls to the $\mathcal{H}_{\mathrm{g}}\mathrm{WFA}$ for the FSA in the best case scenario is $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}} = 1$.

Observe that, even though $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$ differs from one algorithm to another, they achieve the same computational complexity in terms of operations in the best case scenario. However, note that the best case scenario for the FSA occurs when the water level of the optimal power allocation remains constant throughout all the transmission time, in other words, there is a single pool. However, the best case scenario for the NDA is completely the opposite, the water level is different at every epoch and, thus, the total number of pools is J.

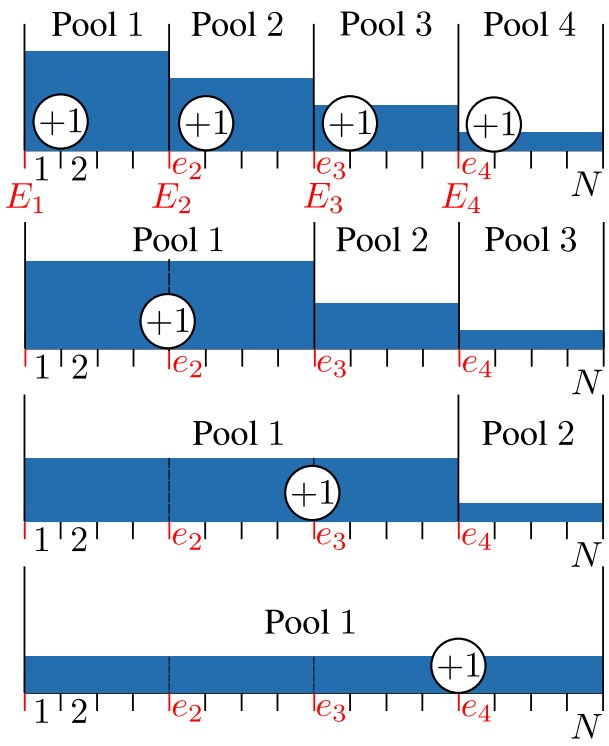

Figure 4.6: Representation of the NDA algorithm.

4.A.5.2 Computational complexity in the worst case scenario

NDA: The worst case computational complexity for the NDA is produced when at every iteration of the algorithm it is observed that the water level is decreasing in some epoch transition. Figure 4.6 shows an example of how the algorithm proceeds for $J = 4$. In the first iteration a total of J calls to the $\mathcal{H}_{\text{g}}$WFA are required. Then, in the second iteration, an additional call is performed to merge the first two pools where it is observed that the water level is decreasing. As we are considering the worst case scenario, the resulting water levels will be decreasing at some pool transition and an additional call is required until all epochs have been merged in a single pool. Therefore, the worst case computational complexity for the NDA is

$$CC^{W}_{NDA}(N,K,J) = \sum_{j=1}^{J} CC_{\mathcal{H}_{\text{g}}\text{WF}}(L_j,K) + \sum_{j=2}^{J} CC_{\mathcal{H}_{\text{g}}\text{WF}}(jL_j,K) \tag{4.33}$$

$$= \hat{a}KN + \sum_{j=2}^{J} \hat{a}KjN/J \tag{4.34}$$

$$= \mathcal{O}\left(\frac{\hat{a}}{2}KNJ\right), \tag{4.35}$$

where the first summation comes from the first iteration of the algorithm and the second one comes from merging the epochs with decreasing water level, i.e., iterations from 2 to J. In

Iteration	Pool	Complexity
1	$\mathbb{P}_1 = \{\tau_1, \ldots, \tau_J\}$	$\hat{a}K \sum_{j=1}^{J} L_j$
2	$\mathbb{P}_1 = \{\tau_1, \ldots, \tau_{J-1}\}$	$\hat{a}K \sum_{j=1}^{J-1} L_j$
$\vdots$	$\vdots$	$\vdots$
J	$\mathbb{P}_1 = \{\tau_1\}$	$\hat{a}KL_1$
		Total $\mathbb{P}_1^\star = \hat{a}K[JL_1 + (J-1)L_2 + \cdots + L_J]$
$J+1$	$\mathbb{P}_2 = \{\tau_2, \ldots, \tau_J\}$	$\hat{a}K \sum_{j=2}^{J} L_j$
$\vdots$	$\vdots$	$\vdots$
$2J-1$	$\mathbb{P}_2 = \{\tau_2\}$	$\hat{a}KL_2$
		Total $\mathbb{P}_2^\star = \hat{a}K[(J-1)L_2 + (J-2)L_3 + \cdots + L_J]$
		$\vdots$
$\frac{J(J+1)}{2}$	$\mathbb{P}_J = \{\tau_J\}$	$\hat{a}KL_J$
		Total $\mathbb{P}_J^\star = \hat{a}KL_J$

Table 4.2: Computational complexity of the FSA in the worst case scenario.

(4.34), we have made the simplification of having equal length epochs, i.e., $L_j = N/J$, $\forall j$. The number of calls to the $\mathcal{H}_{\mathrm{g}}$WFA is $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}} = 2J - 1$.

FSA: The FSA starts by assuming that the first pool contains all the epochs, then, it performs $\mathcal{H}_{\mathrm{g}}$WF and checks whether the ECCs are satisfied, which are not as we are considering the worst case scenario. Then, it removes the last epoch from $\mathbb{P}_1$ and tries again and so forth until $\mathbb{P}_1$ just contains one epoch and then the constraints must be satisfied. Therefore, a total of J iterations are required to determine $\mathbb{P}_1^\star$. Similarly, $J-1$ iterations are required to determine $\mathbb{P}_2^\star$. The computational complexity at each iteration is summarized in Table 4.2 from where we can conclude that the worst case computational complexity of the FSA is

$$CC_{FSA}^{W}(N,K,J) = \sum_{j=1}^{J} \hat{a}KL_j(J-j+1)j \tag{4.36}$$

$$= \hat{a}K\frac{N}{J}\sum_{j=1}^{J}(jJ - j^2 + j) = \mathcal{O}\left(\frac{\hat{a}}{6}KNJ^2\right), \tag{4.37}$$

where in (4.37) we have made the simplification of having equal length epochs, i.e., $L_j = N/J$, $\forall j$. As every iteration performs a call to $\mathcal{H}_{\mathrm{g}}$WF, the total number of calls is $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}} = \frac{J(J+1)}{2}$.

4.A.5.3 Computational complexity in the average case scenario

For the average case scenario, due to the inherent difficulty of determining the computational complexity measured in operations, we have just derived the complexity in terms of calls to the $\mathcal{H}_{\mathrm{g}}$WFA, i.e., $C_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$. By doing this, we can see how the computational complexity is affected by the number of energy arrivals J.

NDA: We start by analyzing the average performance of the NDA. Let $\hat{q}_j$, $j = 1 \dots J-1$, be the probability that the water level decreases at some epoch transition. Let us assume equal probability at all the transition $\hat{q}_j = \hat{q}$, $\forall j$. Let $C^{NDA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$ be a random variable that, for a certain call to the NDA algorithm, denotes the number of calls to the $\mathcal{H}_{\mathrm{g}}$WFA. Note that the minimum number of calls to the $\mathcal{H}_{\mathrm{g}}$WFA is J and, from here, an additional call is produced every time that a water level decrease is produced. Observe that this additional number of calls is a binomial distribution of parameters $J-1$ and $\hat{q}$, i.e., $\mathbb{B}\left(J-1,\hat{q}\right)$. Therefore, $C^{NDA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}} = J + \mathbb{B}\left(J-1,\hat{q}\right)$ and the mean and variance are

$$\mathbb{E}\left\{C^{NDA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}\right\} = J + \mathbb{E}\{\mathbb{B}\left(J-1,\hat{q}\right)\} = J + (J-1)\hat{q} = J(\hat{q}+1) - \hat{q}, \tag{4.38}$$

$$\mathsf{var}\{C^{NDA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}\} = \mathsf{var}\{\mathbb{B}\left(J-1,\hat{q}\right)\} = (J-1)\hat{q}(1-\hat{q}). \tag{4.39}$$

FSA: Similarly for the FSA, let $\hat{p}_j$, $j = 1 \dots J-1$, denote the probability that the j-th ECC of the FSA is not satisfied. We assume that this probability is equal for all the constraints $\hat{p}_j = \hat{p}$, $\forall j$. Let $C^{FSA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$ be a random variable that, for a certain call to the FSA algorithm, denotes the number of calls to the $\mathcal{H}_{\mathrm{g}}$WFA. To determine $\mathbb{E}\left\{C^{FSA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}\right\}$ for a general J, we first obtain in Table 4.3, the number of calls to the $\mathcal{H}_{\mathrm{g}}$WFA, $C^{FSA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$, for some specific values of J as a function of the unfulfilled constraints; note that up to $J-1$ constraints can be non satisfied. In Table 4.3, ✔ and ✘ denote that a certain constraint is satisfied or not, respectively. For example, when $J = 3$, the ECCs that can be unfulfilled are in the transitions of $\tau_1 \rightarrow \tau_2$, which is depicted in the first column, and $\tau_2 \rightarrow \tau_3$, in the second column. After carefully examining the previous table, one may realize that there exists a fixed cost that depends on the number of unfulfilled constraints $\hat{b}$ that is $\hat{b}+1$ (at least, one call to the $\mathcal{H}_{\mathrm{g}}$WFA is required before and after the unfulfilled constraint) and a variable cost that depends on the placement of each unfulfilled constraint. If the position of a given unfulfilled constraint is the last one, the associated variable cost is 1. If it is the one before the last one, the variable cost is 2 and so forth up to the case in which the unfulfilled constraint is the first ECC where the variable cost is $J-1$.

J=3

Constraint		$C^{FSA}_{\mathcal{H}_g\text{WFA}}$	Probability
✔	✔	1	$(1-\hat{p})^2$
✔	✘	3	$(1-\hat{p})\hat{p}$
✘	✔	4	$(1-\hat{p})\hat{p}$
✘	✘	6	$\hat{p}^2$

$\mathbb{E}\{C^{FSA}_{\mathcal{H}_g\text{WFA}}\} = (1-\hat{p})^2 + 7(1-\hat{p})\hat{p} + 6\hat{p}^2$

J=4

Constraint			$C^{FSA}_{\mathcal{H}_g\text{WFA}}$	Probability
✔	✔	✔	1	$(1-\hat{p})^3$
✔	✔	✘	3	$(1-\hat{p})^2\hat{p}$
✔	✘	✔	4	$(1-\hat{p})^2\hat{p}$
✘	✔	✔	5	$(1-\hat{p})^2\hat{p}$
✔	✘	✘	6	$(1-\hat{p})\hat{p}^2$
✘	✔	✘	7	$(1-\hat{p})\hat{p}^2$
✘	✘	✔	8	$(1-\hat{p})\hat{p}^2$
✘	✘	✘	10	$\hat{p}^3$

$\mathbb{E}\{C^{FSA}_{\mathcal{H}_g\text{WFA}}\} = (1-\hat{p})^3 + 12(1-\hat{p})^2\hat{p} + 21(1-\hat{p})\hat{p}^2 + 10\hat{p}^3$

J=5

Constraint				$C^{FSA}_{\mathcal{H}_g\text{WFA}}$	Probability
✔	✔	✔	✔	1	$(1-\hat{p})^4$
✔	✔	✔	✘	3	$(1-\hat{p})^3\hat{p}$
✔	✔	✘	✔	4	$(1-\hat{p})^3\hat{p}$
✔	✘	✔	✔	5	$(1-\hat{p})^3\hat{p}$
✘	✔	✔	✔	6	$(1-\hat{p})^3\hat{p}$
✔	✔	✘	✘	6	$(1-\hat{p})^2\hat{p}^2$
✔	✘	✔	✘	7	$(1-\hat{p})\hat{p}^2$
✘	✔	✔	✘	8	$(1-\hat{p})^2\hat{p}^2$
✔	✘	✘	✔	8	$(1-\hat{p})^2\hat{p}^2$
✘	✔	✘	✔	9	$(1-\hat{p})^2\hat{p}^2$
✘	✘	✔	✔	10	$(1-\hat{p})^2\hat{p}^2$
✔	✘	✘	✘	10	$(1-\hat{p})\hat{p}^3$
✘	✔	✘	✘	11	$(1-\hat{p})\hat{p}^3$
✘	✘	✔	✘	12	$(1-\hat{p})\hat{p}^3$
✘	✘	✘	✔	13	$(1-\hat{p})\hat{p}^3$
✘	✘	✘	✘	15	$\hat{p}^4$

$\mathbb{E}\{C^{FSA}_{\mathcal{H}_g\text{WFA}}\} = (1-\hat{p})^4 + 18(1-\hat{p})^3\hat{p} + 48(1-\hat{p})^2\hat{p}^2 + 46(1-\hat{p})\hat{p}^3 + 15\hat{p}^4$

Table 4.3: Computational complexity of the FSA in the average case scenario (in terms of calls to the $\mathcal{H}_g$WFA) for different values of J. ✔ and ✘ denote that a certain constraint is satisfied or not, respectively.

From this observation we can find $\mathbb{E}\left\{C^{FSA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}\right\}$ for a general J as

$$\begin{aligned}\mathbb{E}\left\{C^{FSA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}\right\} &= \sum_{b=0}^{J-1}\left[\binom{J-1}{\hat{b}}(\hat{b}+1)+\binom{J-2}{\hat{b}-1}\frac{(J-1)J}{2}\right]\hat{p}^{\hat{b}}(1-\hat{p})^{J-1-\hat{b}} \quad (4.40)\\ &= \sum_{\hat{b}=0}^{J-1}\binom{J-1}{\hat{b}}\left(\hat{b}\left(\frac{J}{2}+1\right)+1\right)\hat{p}^{\hat{b}}(1-\hat{p})^{J-1-\hat{b}}, \quad (4.41)\\ &= \left(\frac{J}{2}+1\right)(J-1)\hat{p}+1=\left(\frac{J^2}{2}+\frac{J}{2}-1\right)\hat{p}+1, \quad (4.42)\end{aligned}$$

where in (4.42), we have used that the mean of a binomial distribution with parameters n and $\hat{p}$ is $n\hat{p}$. Similarly, the variance of $C^{FSA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}$ can be obtained through the variance of a binomial distribution as

$$\mathsf{var}\{C^{FSA}_{\mathcal{H}_{\mathrm{g}}\mathrm{WFA}}\}=\left(\frac{J}{2}+1\right)^2(J-1)\hat{p}(1-\hat{p}). \quad (4.43)$$

This concludes the analysis of the computational complexity of the algorithms.

4.A.6 Properties of the reduction matrix

The reduction matrix, $\bar{\mathbf{S}}_K \in \mathbb{R}^{K^2\times K}$, was introduced in [76] and is defined as:

$$[\bar{\mathbf{S}}_K]_{i+(j-1)k,z}=\bar{\delta}_{ijz}, \quad \{i,j,z\}\in[1,K], \quad (4.44)$$

with $\bar{\delta}_{ijz}=1$ if $i=j=z$ and $\bar{\delta}_{ijz}=0$, otherwise.

Note that from the structure of $\bar{\mathbf{S}}_K$, in each column there is only one entry different than zero and it is equal to one. For instance, the matrices for $K=2$ and $K=3$ are:

$$\bar{\mathbf{S}}_2=\begin{pmatrix}1&0\\0&0\\0&0\\0&1\end{pmatrix}, \quad \text{and} \quad \bar{\mathbf{S}}_3=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\\0&0&0\\0&1&0\\0&0&0\\0&0&0\\0&0&0\\0&0&1\end{pmatrix}.$$

The reduction matrix is designed so that

$$\bar{\mathbf{S}}_K^{\mathsf{T}}\mathsf{vec}(\mathbf{A}) = \mathsf{diag}(\mathbf{A}) \tag{4.45}$$

for $\mathbf{A} \in \mathbb{R}^{K\times K}$. In this appendix, we summarize some additional properties of the reduction matrix:

Property 4.3. Multiplication properties:

- Let $\mathbf{A} \in \mathbb{R}^{K^2\times R}$, then the multiplication $\bar{\mathbf{S}}_K^{\mathsf{T}}\mathbf{A}$ removes $K^2 - K$ rows of $\mathbf{A}$.
- Let $\mathbf{A} \in \mathbb{R}^{K\times R}$, then the multiplication $\bar{\mathbf{S}}_K\mathbf{A}$ adds $K^2 - K$ rows of zeros to $\mathbf{A}$.
- Let $\mathbf{A} \in \mathbb{R}^{R\times K}$, then the multiplication $\mathbf{A}\bar{\mathbf{S}}_K^{\mathsf{T}}$ adds $K^2 - K$ columns of zeros to $\mathbf{A}$.
- Let $\mathbf{A} \in \mathbb{R}^{R\times K^2}$, then the multiplication $\mathbf{A}\bar{\mathbf{S}}_K$ removes $K^2 - K$ columns of $\mathbf{A}$.

Proof: The proof follows directly from the structure of the reduction matrix. ■

Property 4.4. Let $\mathbf{A} \in \mathbb{R}^{K\times R}$, $\mathbf{B} \in \mathbb{R}^{K\times R}$, then $\bar{\mathbf{S}}_K^{\mathsf{T}}\left(\mathbf{A}\otimes\mathbf{B}\right)\bar{\mathbf{S}}_K = \mathbf{A}\circ\mathbf{B}$.

Proof: See [76, Lemma A.2]. ■

Property 4.5. $\bar{\mathbf{S}}_K^{\mathsf{T}}\bar{\mathbf{S}}_K = \mathbf{I}_K$.

Proof: The proof directly follows from setting $\mathbf{A} = \mathbf{I}_K$ and $\mathbf{B} = \mathbf{I}_K$ in Property 4.4. ■

Property 4.6. Let $\mathbf{A} \in \mathbb{R}^{K\times K}$, then $\bar{\mathbf{S}}_K^{\mathsf{T}}(\mathbf{A}\otimes\mathbf{I}_K)\bar{\mathbf{S}}_K = \mathsf{Diag}(\mathsf{diag}(A))$.

Proof: The proof directly follows from setting $\mathbf{B} = \mathbf{I}_K$ in Property 4.4. ■

Property 4.7. Let $\mathbf{v} \in \mathbb{R}^K$, then $\bar{\mathbf{S}}_K^{\mathsf{T}}(\mathbf{v}\otimes\mathbf{I}_K) = \mathsf{Diag}(\mathbf{v})$.

Proof: The Kronecker product expands the vector $\mathbf{v}$ in a $K^2 \times K$ matrix that stacks K diagonal matrices. Then, the multiplication by $\bar{\mathbf{S}}_K^{\mathsf{T}}$ eliminates rows (see Property 4.3) so that the resulting matrix is $\mathsf{Diag}(\mathbf{v})$. ■

Property 4.8. Let $\mathbf{A} \in \mathbb{R}^{K^2\times K^2}$ be a diagonal matrix, then $\bar{\mathbf{S}}_K\bar{\mathbf{S}}_K^{\mathsf{T}}\mathbf{A}\bar{\mathbf{S}}_K = \mathbf{A}\bar{\mathbf{S}}_K$

Proof: From Property 4.3, $\bar{\mathbf{S}}_K^{\mathsf{T}}\mathbf{A}$ removes rows from $\mathbf{A}$. Then, the product by the left by $\bar{\mathbf{S}}_K$ adds rows of zeros. As a result, $\bar{\mathbf{S}}_K\bar{\mathbf{S}}_K^{\mathsf{T}}\mathbf{A} \in \mathbb{R}^{K^2\times K^2}$ zeroes $K^2 - K$ rows of $\mathbf{A}$. Finally, the product with $\bar{\mathbf{S}}_K$ from the right removes $K^2 - K$ columns. As $\mathbf{A}$ is diagonal, the entries that are modified by multiplying from the left by $\bar{\mathbf{S}}_K\bar{\mathbf{S}}_K^{\mathsf{T}}$ are later removed by multiplying from the right by $\bar{\mathbf{S}}_K$. Therefore, $\bar{\mathbf{S}}_K\bar{\mathbf{S}}_K^{\mathsf{T}}\mathbf{A}\bar{\mathbf{S}}_K$ is equal to $\mathbf{A}\bar{\mathbf{S}}_K$, which directly removes the columns. ■

Chapter 5

On the optimal resource allocation for a wireless energy harvesting node considering the circuitry power consumption

5.1 Introduction

Most transmission strategies like the *MIMO Mercury Water-Flowing* presented in the previous chapter or the well-known CWF, $\mathcal{H}_{\mathrm{g}}\mathrm{WF}$, and DWF (cf. §2.2.1 and §2.3 [57, 84]) consider that the radiated power is the unique source of energy consumption at the transmitter. Nevertheless, as it has been argued in §2.1.4, this is a reasonable assumption only when the link distance is large; in short-range communications, the remaining energy sinks at the transmitter, such as the energetic cost associated with having the transceiver "on" in a certain channel access, must be also accounted for in the design, as it was done in the Glue pouring strategy for non-harvesting nodes (cf. §2.2.1.3 [50, 51]).

According to this, the aim of this chapter is to study and analyze the impact of other sources of energy consumption at the transmitter (apart from the radiated power) over the transmission strategy that maximizes the mutual information in a point-to-point link through a discrete-time fading channel composed of several parallel independent streams.

In particular, since we consider several parallel subchannels at each channel access, we generalize the power consumption model $\mathcal{C}^1$ presented in (2.4), to account not only for the channel access activation cost but also for the activation cost of each parallel subchannel.

As it has been argued in the state of the art in Chapter 2, recently, a few works have studied the impact of the channel access activation cost in point-to-point links with energy harvesting at

	Continuous channel model		Discrete channel model
	Static	Temporal fading	Temporal fading
SISO	[49]	[93]	**This chapter**
Frequency/space parallel channels	[49]: Cost per channel access only	**This chapter**: Circuitry cost per temporal and frequency/spatial accesses	

Table 5.1: State of the art: circuitry energy consumption in WEHNs

the transmitter, which are summarized in Table 5.1. In [49] and [93], a continuous-time channel was considered: the authors of [93] studied a Single-Input Single-Output (SISO) channel and showed that the mutual information maximization problem is convex when the channel is continuous in time; whereas in [49], a system composed by multiple parallel AWGN channels was studied, but the channel was considered static along time, which substantially simplifies the analysis since, when the channel is static, there is no tradeoff between channel gain and energy availability (see §5.2). Additionally, Xu et al. [49] considered a power consumption model that has a fixed cost for activating the transmitter in a given time instant independently of the number of active parallel channels. Due to this, their model is only applicable to a limited set of transmitter architectures as it is argued in §5.2.

In opposition to [49] and [93], we consider a WEHN operating in a *discrete-time* channel, composed of multiple parallel streams at each channel use, and that is affected by temporal and spatial/frequency fading. The fact of considering a discrete-time channel model is key because it is the actual channel model that is being used in current digital communication systems, e.g., in OFDM. As it is later shown, the discreteness of the channel and the temporal variations of the channel coefficients substantially complicate the problem since it is no longer convex.

According to this, the major contributions of this chapter are:

- Generalizing the power consumption model in (2.4) to consider multiple parallel AWGN channels and showing its applicability in practical transmitter architectures.
- Studying the resource allocation that maximizes the mutual information over N channel accesses when there are multiple parallel data streams by jointly considering energy harvesting and the different sources of energy consumption at the transmitter.
- Deriving an upper bound of the achievable mutual information and two asymptotically optimal solutions of the offline maximization problem, i.e., solutions that tend to the optimal when the number of streams or channel accesses grows without bound.
- Proposing an intuitive graphical representation of the asymptotically optimal offline solution, named *Boxed Water-Flowing*.

- Implementing an online algorithm that achieves a mutual information that is close to the one achieved by the optimal offline solution.
- Evaluating and comparing the computational complexities of the proposed strategies.

This chapter is structured as follows. In §5.2, the system model and problem formulation are presented. The offline resource allocation problem is studied in §5.3 from two perspectives: integer relaxation (§5.3.1) and through the dual problem (§5.3.2). A graphical interpretation of the offline solution is presented in §5.3.3. In §5.4, the online solution is presented. The mutual information and computational requirements of the different algorithms are evaluated in §5.5. Finally, the chapter is concluded in §5.6.

5.2 System model and problem formulation

We consider a WEHN transmitting in a point-to-point link in which, at each channel access, the communication channel can be decomposed into a set of K parallel non-interfering streams by performing some joint signal processing at the transmitter and receiver, e.g., by using OFDM or by diagonalizing a MIMO channel.

Let y_{kn} be the channel output of the k-th stream at the n-th channel access, i.e.,

$$y_{kn} = \sqrt{p_{kn}} g_{kn} x_{kn} + w_{kn}, \qquad k = 1, \dots, K, \quad n = 1, \dots, N,$$

where x_{kn} is the input symbol with $E\{||x_{kn}||^2\} = 1$; p_{kn} is the radiated power; g_{kn} is the complex channel response with $h_{kn} = ||g_{kn}||^2$ being the channel power gain; and $w_{kn} \sim \mathcal{CN}(0,1)$ is the noise. First, in §5.3, we assume that the transmitter has non-causal knowledge of all the channel gains. This assumption is removed in §5.4 for the design of the online algorithm.

As in the previous chapters, we characterize the energy harvesting process at the transmitter with the packetized model introduced in §2.1.1, which is depicted in Figure 2.13, and assume an ideal energy storage that has an infinite capacity and no imperfections.

The power consumption at the transmitter depends on its hardware and software architecture. As we focus on architectures with multiple data streams, it naturally follows that the transmitter may experience either a power consumption associated with the channel access activation, α, or a power consumption associated with the activation of each of the streams, β, or both simultaneously.

In this context, we propose a power consumption model that, as shown later, can be applied to several transmitter architectures, which is a generalization of $\mathcal{C}^1$ in (2.5) to scenarios in which

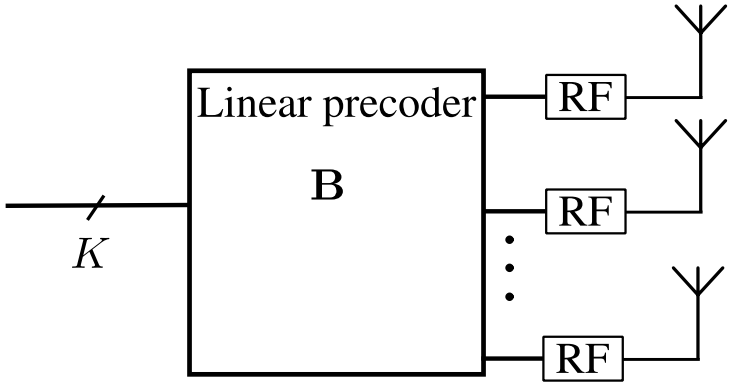

Figure 5.1: Example of MIMO and V-BLAST ($\mathbf{B} = \mathbf{I}_K$) transmitter architectures.

there exist several parallel streams. The total consumed power at the n-th channel access is modeled as:

$$\mathcal{C}^2(\mathbf{p}_n) = \underbrace{\delta}_{\text{Idle consumption}} + \underbrace{\frac{\xi}{\eta}\sum_{k=1}^{K} p_{kn}}_{\text{Transmission power}} + \underbrace{(\alpha-\delta)\mathcal{H}^{\ell}\left(\sum_{k=1}^{K} p_{kn}\right)}_{\text{Consumption per active slot}} + \sum_{k=1}^{K} \underbrace{\beta\mathcal{H}^{\ell}(p_{kn})}_{\text{Consumption per active stream}}, \tag{5.1}$$

where $\mathbf{p}_n = [p_{1n}, \ldots, p_{Kn}]^{\mathsf{T}}$ is the vector of transmission radiated powers at the n-th channel access; as argued in §2.1.4, ξ and η are the power amplifier output back-off and drain efficiency, respectively [48]; and the constants δ, α and β model the idle state power consumption and the power consumptions associated with the channel access and stream activation, respectively. These constants have to accurately capture the different sources of energy consumption and are dependent on the transmitter hardware/software architecture. For example, consider the following architectures:

MIMO linear precoding: The information of the different streams is linearly processed by a precoding matrix, $\mathbf{B}$, and transmitted over the different antennas (see Figure 5.1). If the channel access is active, the precoding operation usually[1] activates all the RF chains at the transmitter independently of the number of active streams. Thus, α would account for the circuitry power consumption of all RF chains, whereas, the number of active streams affects on the number of products and summations required for the linear precoding. In this context, β models the power consumption at the base band processing boards of one of these products and additions.

V-BLAST: The information of the different streams is directly sent over the channel without performing any linear precoding, i.e., in Figure 5.1, set $\mathbf{B} = \mathbf{I}_K$. In this architecture, the

[1]Excluding some specific precoder designs such as $\mathbf{B} = \mathbf{I}_K$, the precoder is generally designed to transmit through the channel eigenmodes and as a result, if one stream is active, all the elements at the output of the precoding matrix are "active" (different than zero).

number of active RF chains is equal to the number of active streams and thus β accounts for the power consumption of each RF chain and $\alpha = 0$ since there is no additional cost per channel access.

OFDM: The number of operations required for the Inverse Fast Fourier Transform (IFFT) depends on the number of active streams. Thus, the cost per operation of the IFFT can be modeled through β and the cost of performing the serial to parallel conversion is mapped in α.

As mentioned in the introduction, the previous results for multi stream communications with WEHNs [49], only considered architectures in which $\alpha \neq 0$ and $\beta = 0$. Moreover, the results in [49] only apply for time-static channels. Thus, we extend the results obtained in [49] to a broader class of transmitter architectures and to consider time-varying fading channels.

The goal of this chapter is to derive the power allocation strategy $\mathbf{P} \in \mathbb{R}_{+}^{K \times N}$ with $\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_N]$ that maximizes the mutual information $I(\mathbf{P})$ while satisfying the ECCs obtained with the power consumption model $\mathcal{C}^2$ in (5.1) that impose that the battery level by the end of the ℓ-th epoch is nonnegative, i.e.,

$$[\boldsymbol{\mathcal{B}}(\mathbf{P})]_\ell \triangleq \sum_{j=1}^{\ell} E_j - T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} \mathcal{C}^2(\mathbf{p}_n) \geq 0, \qquad \ell = 1, \ldots, J.$$

Accordingly, the mutual information maximization problem reads as

$$\mathcal{I}^\star = \underset{\mathbf{P}}{\text{maximize}} \qquad I(\mathbf{P}) \tag{5.2a}$$

$$\text{subject to} \qquad \boldsymbol{\mathcal{B}}(\mathbf{P}) \succeq \mathbf{0}, \tag{5.2b}$$

where, assuming Gaussian distributed input symbols, the mutual information at the k-th stream of the n-th channel use is $\log(1 + h_{kn}p_{kn})$ and the accumulated mutual information is $I(\mathbf{P}) = \sum_{n=1}^{N} \sum_{k=1}^{K} \log(1 + h_{kn}p_{kn})$.

The problem in (5.2) is nonsmooth and nonconvex due to the presence of the unit step functions $\mathcal{H}^\ell(x)$ within the ECCs; accordingly, its solution cannot be directly derived through classical convex optimization techniques.

In this chapter, in order to tackle the nonconvex and nonsmooth problem in (5.2), we substitute the step functions by an additional set of optimization variables that are named *Indicator Variables (IVs)*; whereas, in Chapter 6, we take a different approach based on a Successive Smooth Approximation (SSA) of the step functions. Accordingly, the stream IV,

ψ_{kn}, denotes whether a certain stream is active, i.e.,

$$\psi_{kn} = \begin{cases} 1 & \text{if } p_{kn} > 0, \\ 0 & \text{if } p_{kn} = 0. \end{cases}$$

Similarly, the channel access IV, ρ_n, denotes when a certain channel access is active, i.e.,

$$\rho_n = \begin{cases} 1 & \text{if } \sum_k p_{kn} > 0, \\ 0 & \text{if } \sum_k p_{kn} = 0. \end{cases}$$

Let $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N]$ be the $K \times N$ matrix that contains the stream IVs, where the vector $\boldsymbol{\psi}_n = [\psi_{1n}, \ldots, \psi_{Kn}]^{\mathsf{T}}$ stacks the IVs of the different streams at the n-th channel access. Similarly, $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_N]$ is a $1 \times N$ vector that contains the channel access IVs. Given these new optimization variables, the ECCs in (5.2b) can be equivalently written as $\boldsymbol{\mathcal{B}}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) \succeq \mathbf{0}$ with

$$[\boldsymbol{\mathcal{B}}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})]_\ell \triangleq \sum_{j=1}^{\ell} E_j - T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} \left(\alpha \rho_n + \sum_{k=1}^{K} (p_{kn} + \beta \psi_{kn}) \right), \quad \ell = 1, \ldots, J, \tag{5.3}$$

where to simplify the problem notation and without loss of generality we have assigned $\frac{\xi}{\eta} = 1$ and $\delta = 0$ (note that we can scale the constants α, β, and E_j to have $\frac{\xi}{\eta} = 1$ and $\delta = 0$).[2]

Similarly, the mutual information can be equivalently written in terms of the new optimization variables as

$$I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) = \sum_{j=1}^{J} \sum_{n \in \tau_j} \rho_n \sum_{k=1}^{K} \psi_{kn} \log(1 + h_{kn} p_{kn}).$$

Accordingly, the remainder of this chapter deals with deriving the resource allocation $\mathbf{P}$, $\boldsymbol{\rho}$, and $\boldsymbol{\Psi}$ that maximizes the mutual information $I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})$ while satisfying the ECCs,

$$\mathcal{I}^\star = \underset{\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi} \in \mathcal{X}}{\text{maximize}} \quad I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) \tag{5.4a}$$

$$\text{subject to} \quad \boldsymbol{\mathcal{B}}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) \succeq \mathbf{0}, \tag{5.4b}$$

[2] We assume that the node has enough energy to at least be kept in the idle state during the whole transmission duration, i.e., $T_s \sum_{j=1}^{\ell} \sum_{n \in \tau_j} \delta \leq \sum_{j=1}^{\ell} E_j, \forall \ell = 1, \ldots, J$, as otherwise the problem would not have a feasible solution.

where the feasible set of the optimization variables is

$$\mathcal{X} = \{p_{kn} \geq 0, \rho_n \in \{0,1\}, \psi_{kn} \in \{0,1\}, \forall k, n\}\,.$$

To simplify the notation and without loss of generality, the remainder of the chapter assumes that, within a given time index n, the streams are sorted with non-increasing channel gains, i.e., $h_{1n} \geq h_{2n} \cdots \geq h_{Kn}, \forall n$.

Note that (5.4) is not a convex optimization problem either as the feasible set $\mathcal{X}$ is not a convex set and the objective function is not jointly convex in the optimization variables. The complexity of (5.4) lies in the selection of the active channel accesses and streams ($\boldsymbol{\rho}$ and $\boldsymbol{\Psi}$), because once $\boldsymbol{\rho}$ and $\boldsymbol{\Psi}$ are fixed, the optimal power allocation in the active streams is given by the DWF in (2.21) (see [10, 11] for more details).

The optimal stream and channel access selection depends on the tradeoff between the magnitude of the *channel gain* and the *energy availability* and is a hard combinatorial problem [10, 11]. For instance, assuming that we knew in advance that a single channel access and stream is active, then we could wonder which would be the pair of indices (k, n) among the KN possibilities that provides the highest mutual information: the one with the best gain or some other pair that has the highest energy availability but worse gain? The answer depends on the specific values of the channel gains and the energy arrival distribution and, hence, the derivation of the optimal solution to (5.4) is not straight forward.

Remark 5.1. If the transmitter does not have energy harvesting capabilities (which means that it is only powered by the initial energy in the battery), then the presented system model still applies by particularizing $J = 1$. To the best of our knowledge, even for the particular case $J = 1$, this chapter is the first work to derive an asymptotically optimal power allocation for battery operated nodes in a fading channel by considering both the channel access and stream activation costs (α and β).

Remark 5.2. The system model and problem formulation could also include: (i) instantaneous mask constraints on the transmission power (cf. §2.1.5); and (ii) concave non-linearities of the RF amplifier.[3] Although the structure of the solution and its graphical interpretation depend on the considered scenario, the numerical algorithms proposed in the remaining of the chapter can be trivially extended to include (i) and (ii).

In this context, in the following section we study two different offline feasible solutions

[3]The objective function should be modified to $\sum_{j=1}^{J} \sum_{n \in \tau_j} \rho_n \sum_{k=1}^{K} \psi_{kn} \log(1 + h_{kn} g(p_{kn}))$, where $g(\cdot)$ is the non-linear concave function that returns the output power at the RF amplifier as a function of the input power. Thus, the design variable would be the input power at the RF amplifier.

that perform close to $\mathcal{I}^\star$, whereas, in §5.4 we propose an online resource allocation.

5.3 Offline resource allocation

In this section, we analyze the offline resource allocation from two different perspectives, namely, integer relaxation (§5.3.1) and duality (§5.3.2). In §5.3.3, we present a graphical interpretation of the asymptotically optimal offline resource allocation.

5.3.1 Integer relaxation

In this section, we relax the original problem in (5.4) and formulate a similar convex optimization problem whose solution upper bounds the solution to (5.4). Moreover, from the optimal resource allocation of the relaxed problem, we derive a feasible solution to the original problem in (5.4) whose mutual information is close to $\mathcal{I}^\star$.

In this context, we have modified the objective function in (5.4) so that the new objective function, i.e.,

$$\tilde{I}(\mathbf{P}, \mathbf{\Psi}) = \sum_{j=1}^{J} \sum_{n \in \tau_j} \sum_{k=1}^{K} \psi_{kn} \log\left(1 + \frac{h_{kn} p_{kn}}{\psi_{kn}}\right),$$

is jointly concave in the optimization variables. Additionally, we have relaxed the binary constraint in the IVs, i.e., letting ρ_n and ψ_{kn} to be in the interval $[0, 1]$.[4] The relaxed problem to (5.4) is mathematically expressed as

$$\begin{aligned}
\tilde{\mathcal{I}}^\star = \underset{\mathbf{P}, \boldsymbol{\rho}, \mathbf{\Psi}}{\text{maximize}} \quad & \tilde{I}(\mathbf{P}, \mathbf{\Psi}) && && \text{(5.5a)}\\
\text{subject to} \quad & \boldsymbol{\mathcal{B}}(\mathbf{P}, \boldsymbol{\rho}, \mathbf{\Psi}) \succeq \mathbf{0}, && && \text{(5.5b)}\\
& \psi_{kn} \leq \rho_n, && \forall k, n, && \text{(5.5c)}\\
& \psi_{kn} \leq 1, \;\; -\psi_{kn} \leq 0, && \forall k, n, && \text{(5.5d)}\\
& -p_{kn} \leq 0, && \forall k, n. && \text{(5.5e)}
\end{aligned}$$

Note that in order to have a jointly concave objective function, we have removed the dependency on ρ_n from the objective function; however, we have included the channel access activation constraint in (5.5c) to ensure that the channel access IV is at least as large as the pointwise maximum of the stream IVs. Therefore, with (5.5c), we force that if any stream is active,

[4] In this section, ρ_n can be viewed as if it represented the n-th channel access usage fraction rather than just indicating if the channel access is "on" or "off" and, similarly, ψ_{kn} can be interpreted as the usage fraction of the k-th stream.

$\psi_{kn} > 0$, the associated channel access circuitry consumption $\rho_n \alpha$ is correctly accounted for in the ECCs in (5.5b), where $[\boldsymbol{\mathcal{B}}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})]_\ell$ is defined in (5.3). In (5.5d) and (5.5e), we ensure that ψ_{kn} and p_{kn} lie in their respective feasible sets. We do not have any constraint in the feasible set of ρ_n; however, note that the value of the optimal ρ_n is always in $[0, 1]$ since, in order to reduce the term $\rho_n \alpha$ in the ECCs in (5.5b) (see (5.3)), the optimal ρ_n takes the minimum allowed value in (5.5c) that is the pointwise maximum of the optimal ψ_{kn}, i.e., $\tilde{\rho}_n^\star = \max\{\tilde{\psi}_{1,n}^\star, \ldots, \tilde{\psi}_{K,n}^\star\}$, therefore, $\tilde{\rho}_n^\star$ is also in the interval $[0, 1]$. Moreover, observe that if $\psi_{kn} = \{0, 1\}$ and $\rho_n = \{0, 1\}$, $\forall k, n$, then the value of the new objective function $\tilde{I}(\mathbf{P}, \boldsymbol{\Psi})$ is equal to the value of the original function $I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})$. This implies that the optimal solution to (5.4) is a feasible solution to (5.5) and, hence, the solution to (5.5) upper bounds the solution to (5.4).

The Lagrangian of (5.5) is

$$\begin{aligned}\tilde{\mathcal{L}} = \quad & \tilde{I}(\mathbf{P}, \boldsymbol{\Psi}) + \sum_{\ell=1}^{J} \lambda_\ell \left[\boldsymbol{\mathcal{B}}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})\right]_\ell \\ & + \sum_{j=1}^{J} \sum_{n \in \tau_j} \sum_{k=1}^{K} \left(-\mu_{kn}(\psi_{kn} - \rho_n) - \check{\eta}_{kn}(\psi_{kn} - 1) + \hat{\eta}_{kn}\psi_{kn} + \xi_{kn} p_{kn}\right),\end{aligned}$$

where λ_ℓ and μ_{kn} are the Lagrange multipliers associated with the ECCs and the channel access activation constraints, respectively; $\hat{\eta}_{kn}$, $\check{\eta}_{kn}$ are the multipliers associated with the feasible set of ψ_{kn}; and ξ_{kn} is the multiplier associated with the feasible set of p_{kn}.

Since $\tilde{I}(\mathbf{P}, \boldsymbol{\Psi})$ is jointly concave (it is easy to check that its Hessian matrix is negative semidefinite) and the constraints are affine, (5.5) is a convex optimization problem and can be solved by, e.g., interior point methods [63]. In the following lines, we study the KKT sufficient optimality conditions, which are given in Table 5.2, to gain some knowledge on the structure of the optimal solution to (5.5). This structural knowledge of the solution is later used in §5.3.3 to devise the graphical interpretation of the asymptotically optimal solution and in §5.4 to design the online resource allocation algorithm. From (5.6a), we obtain that

$$p_{kn} = \psi_{kn} \left(\frac{1}{-\xi_{kn} + T_s \sum_{\ell=j}^{J} \lambda_\ell} - h_{kn}^{-1} \right) = \psi_{kn} \left[W_j - h_{kn}^{-1} \right]^+, \tag{5.7}$$

where $W_j = \frac{1}{T_s \sum_{\ell=j}^{J} \lambda_\ell}$ is the j-th epoch water level that is equal for all the active streams contained in some channel access $n \in \tau_j$ and where we have used the slackness condition (5.6e).

The j-th epoch water level, W_j, is related to the available energy at the transmitter through the dependence with the Lagrange multipliers λ_ℓ, $\ell = j, \ldots, J$. As it is later shown in Lemma 5.1, when the available energy is very low, then $W_j \to 0$ to satisfy the ECCs, the n-th channel

$$\forall\, k,n: \qquad \frac{\partial \tilde{\mathcal{L}}}{\partial p_{kn}} = \frac{\psi_{kn} h_{kn}}{\psi_{kn} + h_{kn} p_{kn}} + \xi_{kn} - T_s \sum_{\ell=j}^{J} \lambda_\ell = 0, \tag{5.6a}$$

$$\forall\, k,n: \qquad \frac{\partial \tilde{\mathcal{L}}}{\partial \psi_{kn}} = \log\left(1 + \frac{h_{kn} p_{kn}}{\psi_{kn}}\right) - \frac{h_{kn} p_{kn}}{\psi_{kn} + h_{kn} p_{kn}} - T_s \sum_{\ell=j}^{J} \lambda_\ell \beta - \mu_{kn} - \check{\eta}_{kn} + \hat{\eta}_{kn} = 0. \tag{5.6b}$$

$$\forall\, n: \qquad \frac{\partial \tilde{\mathcal{L}}}{\partial \rho_n} = -T_s \sum_{\ell=j}^{J} \lambda_\ell \alpha + \sum_{k=1}^{K} \mu_{kn} = 0. \tag{5.6c}$$

$$\forall\, \ell = 1,\ldots,J: \qquad \lambda_\ell \left[\boldsymbol{\mathcal{B}}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi})\right]_\ell = 0. \tag{5.6d}$$

$$\forall\, k,n: \quad \xi_{kn} p_{kn} = 0, \tag{5.6e}$$

$$\hat{\eta}_{kn} \psi_{kn} = 0, \tag{5.6f}$$

$$\check{\eta}_{kn}(\psi_{kn} - 1) = 0, \tag{5.6g}$$

$$\mu_{kn}(\psi_{kn} - \rho_n) = 0. \tag{5.6h}$$

$$\forall\, \ell = 1,\ldots,J, \quad \forall\, k,n: \qquad \lambda_\ell \geq 0, \quad \xi_{kn} \geq 0, \quad \hat{\eta}_{kn} \geq 0, \quad \check{\eta}_{kn} \geq 0, \quad \mu_{kn} \geq 0 \tag{5.6i}$$

Table 5.2: KKT optimality conditions of problem (5.5).

access is "off" ($\rho_n = 0$), and the mutual information of the n-th channel access is zero. Then, if the available energy grows, W_j increases and there is a point that we refer to as the n-th *channel access cutoff water level*, $\hat{W}_n(M_n^\star)$, in which the obtained reward in terms of mutual information becomes higher than the activation cost. As it is shown next, $\hat{W}_n(M_n^\star)$ depends on α, β, the number of streams that contribute to the channel access activation, $M_n^\star \in [1, K]$, which is a priori unknown, and the channel gains of these streams. Thus, when the available energy and the other system parameters are such that $W_j = \hat{W}_n(M_n^\star)$, the n-th channel access becomes "partially active", i.e., $\rho_n \in (0, 1)$. Finally, if the available energy is very high, the channel access is completely active, i.e., $\rho_n = 1$. In the following lemma, we derive the expression of the channel access cutoff water level as a function of $M_n^\star$ and later, in Proposition 5.1, we propose a low complexity method to obtain $M_n^\star$.

Lemma 5.1. *The optimal channel access IV satisfies that*

$$\tilde{\rho}_n^\star = \begin{cases} 1 & \text{if } W_j > \hat{W}_n(M_n^\star), \\ (0,1) & \text{if } W_j = \hat{W}_n(M_n^\star), \\ 0 & \text{if } W_j < \hat{W}_n(M_n^\star), \end{cases} \qquad n \in \tau_j,$$

where the n-th channel access cutoff water level reads as

$$\hat{W}_n(M_n) = \frac{\frac{1}{M_n}(\alpha + M_n\beta - \sum_{k=1}^{M_n} h_{kn}^{-1})}{\mathcal{W}_0\left(\frac{\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}}}{eM_n}(\alpha + M_n\beta - \sum_{k=1}^{M_n} h_{kn}^{-1})\right)}, \tag{5.8}$$

and depends on α, β, the number of streams that contribute to the channel access activation M_n and on the channel gains of these streams. $\mathcal{W}_0(\cdot)$ is the positive branch of the Lambert function [80]. Thus, the optimal resource allocation of the streams $k \in [1, M_n^\star]$ satisfies that

$$\tilde{p}_{kn}^\star = \begin{cases} \left(W_j - h_{kn}^{-1}\right) & \text{if } W_j > \hat{W}_n(M_n^\star), \\ \tilde{\psi}_{kn}^\star \left(W_j - h_{kn}^{-1}\right) & \text{if } W_j = \hat{W}_n(M_n^\star), \\ 0 & \text{if } W_j < \hat{W}_n(M_n^\star), \end{cases}$$

$$\tilde{\psi}_{kn}^\star = \begin{cases} 1 & \text{if } W_j > \hat{W}_n(M_n^\star), \\ (0,1) & \text{if } W_j = \hat{W}_n(M_n^\star), \\ 0 & \text{if } W_j < \hat{W}_n(M_n^\star), \end{cases} \quad n \in \tau_j.$$

Proof: See Appendix 5.A.1. ■

Remark 5.3. $\hat{W}_n(M_n)$ increases with both α and β, and decreases with h_{kn}, $\forall k \in [1, M_n]$ (the proof follows from Lemma 5.4.b in Appendix 5.A.3).

Note that, in Lemma 5.1, we have used that the $M_n^\star$ streams that contribute to the channel access activation are the ones with the best channel gains, i.e., $h_{1n}, \ldots, h_{M_n^\star n}$, because these streams are the ones that contribute the most to the objective function. Intuitively, the $M_n^\star$ streams that become active first share the cost of using the channel access α. Once the channel access is being used, the remaining streams, $k > M_n^\star$, may become active by just paying their own stream circuitry cost, β. As a result of this, the streams $k \in (M_n^\star, K]$ experience different activation water levels as shown in the following lemma:

Lemma 5.2. *The optimal resource allocation of the streams $k \in (M_n^\star, K]$ satisfies that*

$$\tilde{p}_{kn}^\star = \begin{cases} \left(W_j - h_{kn}^{-1}\right) & \text{if } W_j > \bar{W}_{kn}, \\ \tilde{\psi}_{kn}^\star \left(W_j - h_{kn}^{-1}\right) & \text{if } W_j = \bar{W}_{kn}, \\ 0 & \text{if } W_j < \bar{W}_{kn}, \end{cases} \quad \tilde{\psi}_{kn}^\star = \begin{cases} 1 & \text{if } W_j > \bar{W}_{kn}, \\ (0,1) & \text{if } W_j = \bar{W}_{kn}, \\ 0 & \text{if } W_j < \bar{W}_{kn}, \end{cases} \quad n \in \tau_j,$$

where the k-th stream cutoff water level at the n-th channel use reads as

$$\bar{W}_{kn} = \frac{\beta - h_{kn}^{-1}}{\mathcal{W}_0\left(\frac{\beta h_{kn}-1}{e}\right)}, \quad k \in (M_n^\star, K], \tag{5.9}$$

and depends on the stream circuitry consumption, β, and the stream gain h_{kn}.

Proof: See Appendix 5.A.2. ■

Remark 5.4. When $\beta \to 0$, the cutoff water level in DWF is recovered, i.e., $\bar{W}_{kn} = h_{kn}^{-1}$ (see (2.21)). Moreover, $\bar{W}_{kn}$ increases with β and decreases with h_{kn} (this can be proved similarly to the proof of Lemma 5.4.b in Appendix 5.A.3).

Note that for coherence, $\bar{W}_{kn} > \hat{W}_n(M_n^\star)$, $\forall k > M_n^\star$, which implies that the streams with higher gains are activated first. However, from the expressions (5.8) and (5.9), this is not obvious. Indeed, if $\bar{W}_{k',n} < \hat{W}_n(M_n^\star)$ for some $k' > M_n^\star$, then the stream k' would become active before the channel access was active, which is a logical contradiction. If such a situation happens, the stream k' should also contribute to activate the channel access, which means that actually $M_n^\star$ is not the optimal number of streams to activate the channel access. Since the mutual information of the n-th channel access is zero until the channel access becomes active, the optimal number of active streams at the channel access cutoff water level is the one that allows to activate the channel access with the lowest water level, i.e., $M_n^\star = \arg\operatorname{minimize}_{M_n} \hat{W}_n(M_n)$. To find $M_n^\star$, an exhaustive search over M_n could be performed. However, this may require a high computational complexity (especially when $K \gg 1$) that can be reduced by means of the following procedure:

Proposition 5.1. *The n-th channel access cutoff water level, $\hat{W}_n(M_n^\star)$, can be found by performing a forward search over M_n, i.e.,*

(1.) Initially, set $M_n := 1$.

(2.) Compute $\hat{W}_n(M_n)$ and $\bar{W}_{(M_n+1)n}$.

(3.) Check if $\hat{W}_n(M_n) < \bar{W}_{(M_n+1)n}$: if the condition is true, then $M_n^\star = M_n$ and the algorithm ends; otherwise, increase M_n, i.e., $M_n := M_n + 1$ and go back to step 2.

Proof: See Appendix 5.A.3. ■

Until now, we have derived Lemmas 5.1, 5.2 and Proposition 5.1 to gain some knowledge on the structure of the optimal solution to (5.5). As mentioned before, since (5.5) is a convex optimization problem, the resource allocation that maximizes (5.5), $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$, can be found

Algorithm 5.1 Close to optimal solution to (5.4) from integer relaxation

Require: $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ and $\Gamma \in (0,1)$

1: Let $\mathcal{S}$ be the set that contain the partially used streams, i.e., $\mathcal{S} = \{\{k,n\} | [\tilde{\psi}^\star]_{kn} \in (0,1)\}$ and let $\mathcal{S}_1 = \{\{k,n\} | [\tilde{\boldsymbol{\Psi}}^\star]_{kn} \in [\Gamma, 1)\}$ and $\mathcal{S}_0 = \{\{k,n\} | [\tilde{\boldsymbol{\Psi}}^\star]_{kn} \in (0, \Gamma)\}$ be a partition of $\mathcal{S}$, where Γ is a constant in $(0,1)$.
2: **if** $|\mathcal{S}| = 0$ **then**
3: $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\} = \{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ ▷ $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$ is the optimal solution to (5.4).
4: **else**
5: $[\hat{\mathbf{P}}]_{kn} = [\tilde{\mathbf{P}}^\star]_{kn}$, $[\hat{\boldsymbol{\Psi}}]_{kn} = [\tilde{\boldsymbol{\Psi}}^\star]_{kn}, \forall \{k,n\} \notin \mathcal{S}$. ▷ Assign the resource allocation in $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ to all the streams not contained in $\mathcal{S}$.
6: $[\hat{\mathbf{P}}]_{kn} = [\tilde{\mathbf{P}}^\star]_{kn}$ and $[\hat{\boldsymbol{\Psi}}]_{kn} = 1, \forall \{k,n\} \in \mathcal{S}_1$; ▷ Round up or down the stream IVs.
$[\hat{\mathbf{P}}]_{kn} = 0$ and $[\hat{\boldsymbol{\Psi}}]_{kn} = 0, \forall \{k,n\} \in \mathcal{S}_0$.
7: $[\hat{\boldsymbol{\rho}}]_n = \max_k [\hat{\boldsymbol{\Psi}}]_{kn}, \forall n = 1, \ldots, N$. ▷ Compute the channel access IVs.
8: Ensure the feasibility of $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$ by scaling down the transmission radiated power of the channel accesses that produce some ECC violation.
9: **end if**
10: **return** $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$

by, e.g., interior point methods [63]. In Algorithm 5.1, we propose a procedure to derive a feasible resource allocation of (5.4) , $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$, from the solution to (5.5), $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$, whose mutual information, $\hat{\mathcal{I}} = I(\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}})$, performs close to $\mathcal{I}^\star$, as argued in the following lines.

Note that in general $\hat{\mathcal{I}} \leq \mathcal{I}^\star \leq \tilde{\mathcal{I}}^\star$. However, these inequalities are tight ($\hat{\mathcal{I}} = \mathcal{I}^\star = \tilde{\mathcal{I}}^\star$), when, in Algorithm 5.1, we have that $\mathcal{S} = \{\emptyset\}$, or, equivalently, if $W_j \neq \hat{W}_n(M_n^\star)$ and $W_j \neq \bar{W}_{kn}, \forall n \in \tau_j, \forall j, \forall k > M_n^\star$. This means that $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ is the optimal resource allocation to (5.4). Alternatively, when $\mathcal{S} \neq \{\emptyset\}$, we know that the optimality gap, i.e., $\mathcal{I}^\star - \hat{\mathcal{I}}$, is at most $\tilde{\mathcal{I}}^\star - \hat{\mathcal{I}}$ and is closely related to the cardinality of $\mathcal{S}$. Since for most of the streams and channel accesses the water level is different to the cutoff water level, we know that $|\mathcal{S}| \ll KN$, which implies that the optimal resource allocation to (5.5) is used in the majority ($KN - |\mathcal{S}|$) of the streams. This discussion is later continued in Remark 5.7 once the graphical representation of the asymptotically optimal solution is presented.

Remark 5.5. Observe that $\{\tilde{\mathbf{P}}^\star, \tilde{\boldsymbol{\rho}}^\star, \tilde{\boldsymbol{\Psi}}^\star\}$ is the optimal solution to the time continuous channel problem. Hence, if we particularize $K = 1$, then our solution reduces to the directional glue pouring algorithm introduced in [93].

In the following section, we solve the dual problem to (5.4). Interestingly, the concept of the cutoff water level also appears when solving the dual problem, which is indeed surprising due to the great difference between the relaxed and dual problem approaches.

5.3.2 Duality

In this section, we study the Lagrange dual problem to (5.4) and show that, even though (5.4) is not a convex optimization problem, the duality gap tends asymptotically to zero as the number of streams or channel accesses per epoch grows without bound.

The Lagrangian of (5.4) is

$$\mathcal{L}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}, \boldsymbol{\lambda}) = I(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}) + \boldsymbol{\lambda}^{\mathsf{T}} \boldsymbol{\mathcal{B}}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}),$$

where $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_J]^{\mathsf{T}}$ is the dual variable that contains the Lagrange multipliers associated with the ECCs. The dual function is defined for $\boldsymbol{\lambda} \succeq \mathbf{0}$ as

$$g(\boldsymbol{\lambda}) = \underset{\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi} \in \mathcal{X}}{\text{maximize}} \, \mathcal{L}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}, \boldsymbol{\lambda})$$

(see [63]) and yields to upper bounds to the maximum achievable mutual information $\mathcal{I}^\star$ obtained by maximizing the primal problem (5.4), i.e., $\mathcal{I}^\star \leq g(\boldsymbol{\lambda})$. The Lagrange dual problem,

$$\mathcal{D}^\star = \underset{\boldsymbol{\lambda} \succeq \mathbf{0}}{\text{minimize}} \, g(\boldsymbol{\lambda}),$$

is a convex program that determines the best upper bound on $\mathcal{I}^\star$ as $\mathcal{I}^\star \leq \mathcal{D}^\star \leq g(\boldsymbol{\lambda})$. The duality gap is defined as $\mathcal{D}^\star - \mathcal{I}^\star$ and it is zero if Slater qualification constraints are satisfied. However, in our problem the Slater qualification constraints are not satisfied since the feasible set $\mathcal{X}$ is not convex and, therefore, the duality gap might not be zero.

The *time-sharing condition* introduced in [119] provides a condition under which the duality gap is zero even though the primal optimization problem is not convex. In the following proposition, we demonstrate that the *time-sharing condition* is asymptotically satisfied as the number of streams or channel accesses per epoch grows without bound.

Proposition 5.2. *The time-sharing condition is asymptotically satisfied when, within each epoch, every channel realization is observed a sufficiently large number of times.*[5]

Proof: See Appendix 5.A.6. ■

[5]The time sharing condition has been broadly used in different non-harvesting scenarios where the nodes have to satisfy a single sum-power constraint. In such non-harvesting scenarios, the requirement for the asymptotic fulfillment of the time sharing condition is that every channel realization must be observed a large number of times [119]. When energy harvesting is considered, the problem is constrained by a set of ECCs and the time-sharing condition is asymptotically satisfied if, within each epoch, every channel realization is observed a sufficiently large number of times. When $K = 1$, it is necessary that every channel realization is observed in a sufficiently large number of channel accesses. This situation happens, for instance, when the number of channel accesses per epoch is large, i.e., $e_{j+1} - e_j \gg 1$, $\forall j$, and $T_c \gg T_s$. Whereas when $K > 1$, this condition is more likely to be fulfilled due to the additional (space or frequency) dimension.

Algorithm 5.2 Projected subgradient

Initialization:

Set $q := 0$ and initialize $\boldsymbol{\lambda}^{(0)}$ to any value such that $\boldsymbol{\lambda}^{(0)} \succeq \mathbf{0}$.

For all $n = 1, \ldots, N$, compute $\hat{W}_n(M_n^\star)$ according to (5.8) with $M_n^\star$ obtained from the forward search in Proposition 5.1.

Step 1: If a termination condition is met, the algorithm stops.

Step 2: Compute the optimal primal variables at the q-th iteration that are

$[\mathbf{P}^{(q)}, \boldsymbol{\rho}^{(q)}, \boldsymbol{\Psi}^{(q)}] = \arg\text{maximize}_{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{X}}\, \mathcal{L}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}, \boldsymbol{\lambda}^{(q)})$

by means of Algorithm 5.3 that requires $\boldsymbol{\lambda}^{(q)}$ and $\hat{W}_n(M_n^\star)$.

Step 3: Update the dual variable following the subgradient, i.e., $\forall j = 1, \ldots, J$,

$$[\boldsymbol{\lambda}^{(q+1)}]_j = \lambda_j^{(q+1)}, \quad \text{with} \quad \lambda_j^{(q+1)} = \left[\lambda_j^{(q)} - s^{(q)}[\boldsymbol{\mathcal{B}}(\mathbf{P}^{(q)}, \boldsymbol{\rho}^{(q)}, \boldsymbol{\Psi}^{(q)})]_j\right]^+.$$

Step 4: Set $q := q + 1$ and go to Step 1.

Thanks to the previous proposition, when the number of streams or channel accesses per epoch is high and the channel variations in one of the dimensions (time, space or frequency) are slow, the duality gap tends to zero and, consequently, the solution to $\mathcal{D}^\star$ asymptotically tends to $\mathcal{I}^\star$. At this stage, it is important to highlight that, in practice, it is not necessary that the number of streams or channel accesses per epoch grows without bound; a small duality gap is already observed for small values of these magnitudes as verified in the simulations results (see §5.5) where $K = 8$ and the mean number of channel accesses per epoch is 5. This behaviour was previously observed in scenarios without energy harvesting in, e.g., [119].

To solve the dual problem we have implemented the projected subgradient method [120], which is presented in Algorithm 5.2, that guarantees convergence if the updating step size $s^{(q)}$ is correctly chosen. In this context, we have used $s^{(q)} = \frac{\mathcal{A}}{\sqrt{q}||\boldsymbol{\mathcal{B}}(\mathbf{P}^{(q)},\boldsymbol{\rho}^{(q)},\boldsymbol{\Psi}^{(q)})||}$ that satisfies the diminishing conditions $s^{(q)} \geq 0$, $\lim_{q\to\infty} s^{(q)} = 0$ and $\sum_{q=1}^{\infty} s^{(q)} = \infty$ [120], where $\mathcal{A} > 0$ is an arbitrary constant. When the algorithm converges to the optimal dual variable, $\boldsymbol{\lambda}^\star$, all the ECCs are satisfied, which is ensured by the termination condition in Step 1. In the next subsection, we explain Step 2 of Algorithm 5.2, i.e., how to obtain the primal variables $\mathbf{P}^{(q)}$, $\boldsymbol{\rho}^{(q)}$ and $\boldsymbol{\Psi}^{(q)}$ at the q-th iteration of the subgradient method.

5.3.2.1 Maximizing the Lagrangian for a given $\boldsymbol{\lambda}^{(q)}$

At every iteration of the subgradient algorithm, it is necessary to compute the optimal primal variables given the dual variables of the iteration, i.e., $\boldsymbol{\lambda}^{(q)}$. From the expression of $g(\boldsymbol{\lambda})$, the optimal primal variables at the q-th iteration are

$$[\mathbf{P}^{(q)}, \boldsymbol{\rho}^{(q)}, \boldsymbol{\Psi}^{(q)}] = \arg\underset{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{X}}{\text{maximize}}\, \mathcal{L}(\mathbf{P}, \boldsymbol{\rho}, \boldsymbol{\Psi}, \boldsymbol{\lambda}^{(q)}).$$

Note that the maximization of the Lagrangian is not a convex problem as $\mathcal{X}$ is not a convex set. To solve the maximization of the Lagrangian we apply decomposition as follows:

$$\begin{aligned}
&\underset{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{X}}{\text{maximize}}\, \mathcal{L}(\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi},\boldsymbol{\lambda}^{(q)}) \\
= \;& \underset{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{X}}{\text{maximize}}\, \mathcal{I}(\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}) + \sum_{\ell=1}^{J}\lambda_\ell^{(q)}\left(\sum_{j=1}^{\ell}(E_j) - T_s\sum_{j=1}^{\ell}\sum_{n\in\tau_j}\left(\rho_n\alpha + \sum_{k=1}^{K}p_{kn} + \psi_{kn}\beta\right)\right) \\
= \;& \sum_{\ell=1}^{J}\left(\lambda_\ell^{(q)}\sum_{j=1}^{\ell}E_j\right) + \\
& \underset{\mathbf{P},\boldsymbol{\rho},\boldsymbol{\Psi}\in\mathcal{X}}{\text{maximize}} \sum_{j=1}^{J}\sum_{n\in\tau_j}\left[\rho_n\Big(\sum_{k=1}^{K}\psi_{kn}\log(1+h_{kn}p_{kn})\Big) - T_s\Big(\rho_n\alpha + \sum_{k=1}^{K}p_{kn} + \psi_{kn}\beta\Big)\sum_{\ell=j}^{J}\lambda_\ell^{(q)}\right] \\
= \;& \sum_{\ell=1}^{J}\left(\lambda_\ell^{(q)}\sum_{j=1}^{\ell}E_j\right) + \sum_{j=1}^{J}\sum_{n\in\tau_j}\underset{\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n\in\mathcal{X}}{\text{maximize}}\, g_n(\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n).
\end{aligned}$$

We have reordered the sums over j and ℓ to decompose the Lagrangian maximization in N independent maximization problems, one for each channel use, where the objective function is

$$g_n(\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n) = \rho_n\left(\sum_{k=1}^{K}\psi_{kn}\log(1+h_{kn}p_{kn})\right) - \frac{1}{W_j^{(q)}}\left(\rho_n\alpha + \sum_{k=1}^{K}p_{kn} + \psi_{kn}\beta\right)$$

with $W_j^{(q)} = \frac{1}{T_s\sum_{\ell=j}^{J}\lambda_\ell^{(q)}}$ being the water level of the j-th epoch at the q-th iteration.

Note that the problem

$$\underset{\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n\in\mathcal{X}}{\text{maximize}} \quad g_n(\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n)$$

is still a nonconvex problem due to the binary variables. However, after applying decomposition, it is feasible to perform an exhaustive search over ρ_n as there are only two possibilities either $\rho_n = 0$ or $\rho_n = 1$. Thus, we can solve two separated maximization problems and select the pointwise maximum of the two, i.e.,

$$\underset{\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n\in\mathcal{X}}{\text{maximize}}\, g_n(\mathbf{p}_n,\rho_n,\boldsymbol{\psi}_n) = \max\Big\{\underbrace{\underset{\mathbf{p}_n,\boldsymbol{\psi}_n\in\mathcal{X}}{\text{maximize}}\, g_n(\mathbf{p}_n,0,\boldsymbol{\psi}_n)}_{\text{(SP 1)}}, \underbrace{\underset{\mathbf{p}_n,\boldsymbol{\psi}_n\in\mathcal{X}}{\text{maximize}}\, g_n(\mathbf{p}_n,1,\boldsymbol{\psi}_n)}_{\text{(SP 2)}}\Big\}. \tag{5.10}$$

These two problems are solved in the following lines and Table 5.3 summarizes the obtained results.

Subproblem	Maximum value	Optimal $\mathbf{p}_n$	Optimal $\boldsymbol{\psi}_n$
(SP 1)	0	$\mathbf{p}_n^{(\text{SP 1})} = \mathbf{0}$	$\boldsymbol{\psi}_n^{(\text{SP 1})} = \mathbf{0}$
(SP 2)	$-\frac{\alpha}{W_j^{(q)}} + \sum_{k=1}^{A_n^{(q)}} \Big(\log(W_j^{(q)} h_{kn})$ $-1 + \frac{1}{W_j^{(q)} h_{kn}} - \frac{\beta}{W_j^{(q)}} \Big)$	$[\mathbf{p}_n^{(\text{SP 2})}]_{k\in[1,A_n^{(q)}]} = W_j^{(q)} - h_{kn}^{-1}$ $[\mathbf{p}_n^{(\text{SP 2})}]_{k\in(A_n^{(q)},K]} = 0$	$[\boldsymbol{\psi}_n^{(\text{SP 2})}]_{k\in[1,A_n^{(q)}]} = 1$ $[\boldsymbol{\psi}_n^{(\text{SP 2})}]_{k\in(A_n^{(q)},K]} = 0$

Table 5.3: Optimal solution to the subproblems in (5.10).

Subproblem	Maximum value	Optimal p_{kn}
(SP 2.1)	0	$p_{kn}^{(\text{SP 2.2})} = 0$
(SP 2.2)	$\log(W_j^{(q)} h_{kn}) - 1 + \frac{1}{W_j^{(q)} h_{kn}} - \frac{\beta}{W_j^{(q)}}$	$p_{kn}^{(\text{SP 2.2})} = W_j^{(q)} - h_{kn}^{-1}$

Table 5.4: Optimal solution to the subproblems in (5.11).

Solution to (SP 1): By observing the objective function of (SP 1), i.e.,

$$g_n(\mathbf{p}_n, 0, \boldsymbol{\psi}_n) = -\frac{1}{W_j^{(q)}} \left(\sum_{k=1}^{K} p_{kn} + \psi_{kn}\beta \right)$$

and by noting that $W_j^{(q)}$ is positive, it is straight-forward to show that the optimal transmitted powers and stream IVs of the n-th channel access are $\mathbf{p}_n^{(\text{SP 1})} = \boldsymbol{\psi}_n^{(\text{SP 1})} = \mathbf{0}$ and the maximum value of the objective function is 0, as expected since $\rho_n = 0$.

Solution to (SP 2): To solve the second subproblem, which is nonconvex due to the stream IVs, we can again apply decomposition as follows:

$$\underset{\mathbf{p}_n, \boldsymbol{\psi}_n \in \mathcal{X}}{\text{maximize}}\, g_n(\mathbf{p}_n, 1, \boldsymbol{\psi}_n) = -\frac{\alpha}{W_j^{(q)}} + \sum_{k=1}^{K} \underset{p_{kn}, \psi_{kn} \in \mathcal{X}}{\text{maximize}}\, g_{kn}(p_{kn}, \psi_{kn}),$$

where $g_{kn}(p_{kn}, \psi_{kn}) = \psi_{kn} \log(1 + h_{kn} p_{kn}) - \frac{1}{W_j^{(q)}}(p_{kn} + \psi_{kn}\beta)$. As before, after applying decomposition, we can perform an exhaustive search over ψ_{kn} since there are only two possibilities, i.e., either $\psi_{kn} = 0$ or $\psi_{kn} = 1$. Thus,

$$\underset{p_{kn}, \psi_{kn} \in \mathcal{X}}{\text{maximize}}\, g_{kn}(p_{kn}, \psi_{kn}) = \max \Big\{ \underbrace{\underset{p_{kn} \in \mathcal{X}}{\text{maximize}}\, g_{kn}(p_{kn}, 0)}_{\text{(SP 2.1)}}, \underbrace{\underset{p_{kn} \in \mathcal{X}}{\text{maximize}}\, g_{kn}(p_{kn}, 1)}_{\text{(SP 2.2)}} \Big\}. \tag{5.11}$$

Now, both subproblems are convex and can be easily solved. Table 5.4 summarizes the maximum achieved value and the optimal transmission power of each subproblem. Thus, the k-th stream is active if

$$\log(W_j^{(q)} h_{kn}) - 1 + \frac{1}{W_j^{(q)} h_{kn}} - \frac{\beta}{W_j^{(q)}} > 0.$$

Solving the previous equation (set $M := 1$, $\hat{W} := \bar{W}_{kn}$, $H_1 := h_{kn}$ and $P_c := \beta$ in Appendix 5.A.7), we obtain an equivalent condition for the k-th stream activation, i.e., $W_j^{(q)} > \bar{W}_{kn}$, where $\bar{W}_{kn}$ is the stream cutoff water level given in (5.9). Thus, after evaluating the condition $W_j^{(q)} > \bar{W}_{kn}$, $\forall k$, we obtain the number of streams that are activated if the channel access is active, $A_n^{(q)}$.

Now that both (SP 1) and (SP 2) are solved (Table 5.3 summarizes the obtained results), we can conclude that the n-th channel access is active if

$$-\frac{\alpha}{W_j^{(q)}} + \sum_{k=1}^{A_n^{(q)}} \left(\log(W_j^{(q)} h_{kn}) - 1 + \frac{1}{W_j^{(q)} h_{kn}} - \frac{\beta}{W_j^{(q)}} \right) > 0$$

or, equivalently, if $W_j^{(q)} > \hat{W}_n(A_n^{(q)})$ (set $M := A_n^{(q)}$, $\hat{W} := \hat{W}_n(A_n^{(q)})$, $H_k := h_{kn}$, and $P_c := A_n^{(q)}\beta + \alpha$ in Appendix 5.A.7 to show this equivalence).[6]

In summary, the optimal primal variables at the q-th iteration of the subgradient can be obtained by checking the condition $W_j^{(q)} > \bar{W}_{kn}$, $\forall k$, to obtain the number of streams that would be active if the channel access was active, $A_n^{(q)}$, and then checking the condition $W_j^{(q)} > \hat{W}_n(A_n^{(q)})$ to find out whether the channel access is active or not. If the channel access is active, the optimal primal variables at the q-th iteration of the subgradient are $\mathbf{p}_n^{(q)} = \mathbf{p}_n^{(SP2)}$, $\boldsymbol{\psi}_n^{(q)} = \boldsymbol{\psi}_n^{(SP2)}$, and $\rho_n^{(q)} = 1$. Otherwise, we have that $\mathbf{p}_n^{(q)} = \mathbf{0}$, $\boldsymbol{\psi}_n^{(q)} = \mathbf{0}$, and $\rho_n^{(q)} = 0$. However, this procedure might be quite inefficient when the number of streams is large ($K \gg 1$) and it can be avoided by first checking whether the channel access is active. Note that for any value of $A_n^{(q)}$ the channel access is active if and only if $W_j^{(q)} > \hat{W}_n(M_n^\star)$.[7] In this context, the procedure in Algorithm 5.3 is equivalent to the proposed above, but more computationally efficient.

Remark 5.6. The resource allocation obtained by solving the dual problem is almost equal to the one obtained by means of the relaxed problem, which is given in Lemmas 5.1 and 5.2. The channel access or stream activation conditions obtained in this section only differ from the ones in Lemmas 5.1 and 5.2 when the water level is equal to the stream or channel access cutoff water levels. In this section we have seen that if $W_j = \hat{W}_n(M_n^\star)$ (or $W_j = \bar{W}_{kn}$ for $k \in (M_n^\star, K]$), it is indifferent to have the channel access (or stream) active or inactive since both situations achieve the same value of the dual function, whereas, in the relaxed problem we

[6]Note that when $W_j^{(q)} = \bar{W}_{kn}$ (or $W_j^{(q)} = \hat{W}_n(A_n^{(q)})$) it is equivalent to activate or not the stream (or channel access) since both achieve the same value of the objective function.

[7]By using the definition of $M_n^\star$ and Lemmas 5.3 and 5.4 in Appendix 5.A.3, it is easy to show that: if $A_n^{(q)} > M_n^\star$, then $\hat{W}_n(M_n^\star) < \hat{W}_n(A_n^{(q)}) < \bar{W}_{A_n^{(q)},n} < W_j$; and if $A_n^{(q)} < M_n^\star$, then $\hat{W}_n(A_n^{(q)}) \geq \hat{W}_n(M_n^\star) \geq \bar{W}_{M_n^\star,n} \geq W_j > \bar{W}_{A_n^{(q)},n}$. Thus, from these inequalities, we can compare W_j directly with $\hat{W}_n(M_n^\star)$ to determine whether the channel access is active or not.

Algorithm 5.3 Maximization of the Lagrangian

Data: $\boldsymbol{\lambda}^{(q)}$, $\hat{W}_n(M_n^\star)$.

Compute $W_j^{(q)} = \frac{1}{T_s \sum_{\ell=j}^{J} \lambda_\ell^{(q)}}, \forall j = 1, \ldots, J.$ ▷ Compute the water level in all the epochs given $\boldsymbol{\lambda}^{(q)}$.
for $n \in \tau_j$, $j := 1, \ldots, J$ **do** ▷ For all the channel accesses.
  if $W_j^{(q)} > \hat{W}_n(M_n^\star)$ **then** ▷ Check if the channel access is active.
    $A_n^{(q)} := M_n^\star$; ▷ The channel access is active; Then, count the number of active streams.
    for $k := M_n^\star + 1, \ldots, K$ **do**
      if $W_j^{(q)} > \bar{W}_{kn}$ **then**
        $A_n^{(q)} := A_n^{(q)} + 1$; ▷ The k-th stream is active at the q-th iteration water level.
      end if
    end for ▷ End of counting.
    $\mathbf{p}_n^{(q)} = \mathbf{p}_n^{(SP2)}$, $\boldsymbol{\psi}_n^{(q)} = \boldsymbol{\psi}_n^{(SP2)}$, and $\rho_n^{(q)} = 1$;
  else
    $\mathbf{p}_n^{(q)} = \mathbf{0}$, $\boldsymbol{\psi}_n^{(q)} = \mathbf{0}$, and $\rho_n^{(q)} = 0$; ▷ The n-th channel access is turned off.
  end if
end for

obtained that a partial use of the channel access (or the stream) is optimal, which is not allowed in the problem considered in this section due to the binary feasible set of the IVs.

5.3.3 The Boxed Water-Flowing interpretation

In this section, we provide a graphical representation of the asymptotically optimal offline solution named the *Boxed Water-Flowing* interpretation, which is depicted in Figure 5.2. This interpretation follows directly from the concept of the cutoff water levels and it generalizes the DWF interpretation in §2.3.1 [84] by considering the different sources of energy consumption at the transmitter. The interpretation is the following:

(1.) Each stream is represented with a water-porous vessel with base equal to T_s.[8] There are two types of boxes, namely, the *channel access box* and the *stream box*. At the n-th channel access, the *channel access box* with height $\hat{W}_n(M_n^\star)$ is shared among the streams $k = 1, \ldots, M_n^\star$. The remaining streams, i.e., $k > M_n^\star$, have their own stream box with height equal to $\bar{W}_{kn}$.[9] A water *right-permeable* material is used to separate the different epochs.

(2.) Each box is filled by a solid substance up to a height equal to h_{kn}^{-1} and the boxes are closed by a lid. The cost (in terms of water) of opening the channel access box is $(\alpha + M_n^\star \beta)T_s$, whereas, the cost of opening each stream box is βT_s.

[8] The vessel boundaries are not depicted in Figure 5.2 for the sake of simplicity.

[9] Thus, the n-th channel access has $K - M_n^\star$ stream boxes and one channel access box.

(3.) The water level is progressively increased to all epochs at the same time by adding the necessary amount of water to each epoch. The maximum amount of water that can be externally added at some epoch is given by the epoch's harvested energy (depicted with the top-down arrows in Figure 5.2).[10] When some epoch runs out of water, it uses water that flows from previous epochs (if any is available) in order to continue increasing the water level simultaneously. When the water level reaches the lid of some box, check if there is enough available water (in the current and previous epochs) to pay the cost of opening the lid and to fill in the whole box with water. If there is enough water, remove the lid (which means that the amount of water associated with the lid opening cost is lost), let the water fill the box and go back to Step 3; otherwise, keep the lid in the box and go back to Step 3.

(4.) When all the available water has been poured, the optimal power allocation is found as the amount of water in each of the vessels divided by T_s or, equivalently, as the height of the water in each vessel, i.e., $p_{kn} = [W_j - h_{kn}^{-1}]^+$.

Interestingly, by particularizing the *Boxed Water-Flowing* interpretation to the case in which there is no circuitry consumption ($\alpha = 0$ and $\beta = 0$), the heights of the boxes reduce to its minimum possible value, i.e., h_{kn}^{-1}, (set $\alpha = 0$ and $\beta = 0$ in Lemmas 5.1 and 5.2) and the DWF graphical interpretation in [84] is recovered, which has been presented in §2.3.1.

Remark 5.7. Having the graphical representation of the asymptotically optimal solution in mind, it is easy to understand why its performance is close to $\mathcal{I}^\star$. In the representation shown in Figure 5.2, all the streams are using the optimal resource allocation to (5.5), except the second stream of τ_1. After solving the integer relaxation problem, we would have obtained that a fractional use of this stream would be optimal. However, as this fractional use is not allowed in a discrete channel model, we do not know where to optimally allocate the small remaining energy in τ_1. In summary, the following arguments justify why the optimality gap of the *Boxed Water-Flowing* solution is small:

- As mentioned before, if the water level is different than the cutoff water level of all the boxes ($\tilde{\rho}_n^\star \in \{0, 1\}$ and $\tilde{\psi}_{kn}^\star \in \{0, 1\}$, $\forall k, n$), then the *Boxed Water-Flowing* solution is optimal and the optimality gap is zero.

- Otherwise, when some stream or slot is partially used ($\tilde{\rho}_n^\star \in (0, 1)$ or $\tilde{\psi}_{kn}^\star \in (0, 1)$), the remaining energy in the epoch is very small and can be allocated in any of the active

[10] The amount of water corresponds to energy, whereas, the water level, i.e., the height of the water, corresponds to power.

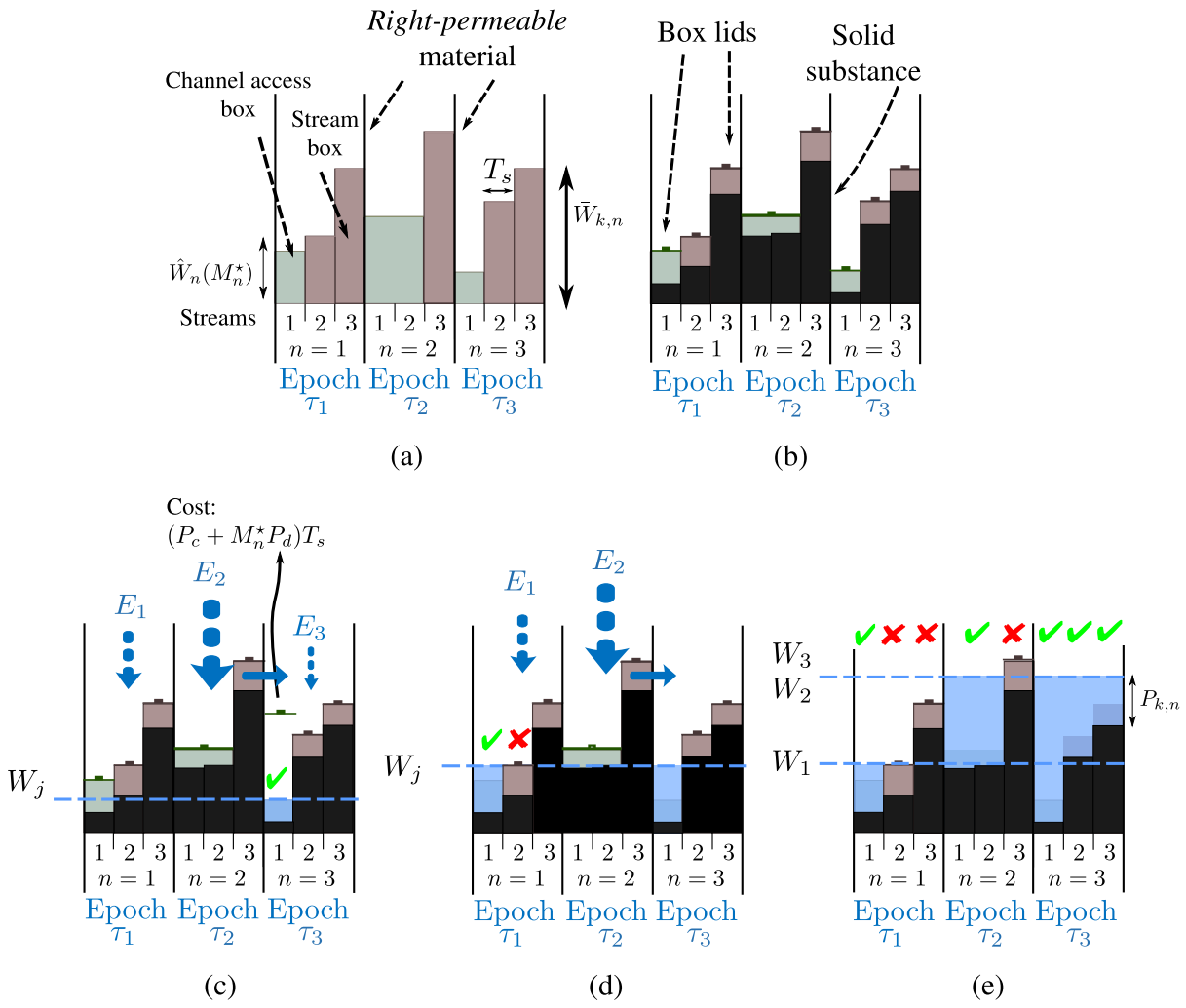

Figure 5.2: The *Boxed Water-Flowing* interpretation. For graphical simplicity, only one channel access is contained in each epoch. Figures (a) and (b) depict steps 1 and 2 of the explanation, respectively. Figures (c) to (e) depict step 3 where the water level is progressively increased and different situations occur: in (c), the box is opened because, by using water from τ_2 and τ_3, there is enough water to fill the box and pay the opening cost of the channel access box; in (d), when the water level reaches the stream box ($n = 1, k = 2$) the remaining water in the first epoch is not enough to pay the opening cost and fill in the whole box with water (see Remark 5.7 for a discussion on what happens with this remaining water); finally, (e) depicts the obtained resource allocation once all the available water has been poured.

channels without having a relevant impact on the total mutual information.

- Within each epoch, the water level can only be equal to one specific box height, which might be present in different channels as far as they have the same channel gain. This means that all the channels with different box heights (or different channel gains) are using the optimal resource allocation to (5.5).
- If this specific box height is present in several channel accesses or streams of that same epoch, then the time sharing argument can be used to allocate the little remaining energy in the epoch, i.e., a fraction $\tilde{\rho}_n^\star$ ($\tilde{\psi}_{kn}^\star$) of channel access (stream) boxes are opened and the remaining ones are kept closed.

Due to this, we can expect a small optimality gap as it is confirmed by the conducted experimental results presented in §5.5.

5.4 Online resource allocation

Up to now, we have assumed that the transmitter has non-causal knowledge of both the channel state and the energy harvesting process, which is only a realistic assumption under very specific scenarios, e.g., when the channel is static and the energy source is controllable, cf., §2.1.3. In this section, we develop an online algorithm, which does not require future knowledge of neither the energy arrivals nor the channel state, that is based on the structure of the *Boxed Water-Flowing*, the asymptotically optimal offline resource allocation that we derived in the previous section.

Similarly to Chapter 4, the proposed online algorithm has an input parameter, F_w, which is named the *flowing window*, that controls the number of channel accesses in which the water is allowed to flow, which can be obtained by a previous training under the considered (or measured) energy harvesting profile. Let an *event*, $\mathcal{E}_t$, denote the time index of a channel access in which either a change in the channel state is produced or an energy packet is harvested (or both events take place at the same time), i.e.,

$$\mathcal{E}_t = \cup_{k=1}^{K}\{n|h_{k(n-1)} \neq h_{kn}\} \cup \{n|n = e_j, j = 1, \ldots, J\}, \quad t = 1, \ldots, S,$$

where $S \in [J, N]$. In this context, the proposed online algorithm proceeds as follows:

(1.) The initial energy in the battery, E_1, is allocated to the different streams of the first F_w channel accesses according to the *Boxed Water-Flowing* where the channel is expected

to be static and equal to the observed channel at the first channel use, i.e., we assume as if $h_{kn} = h_{k1}, \forall n \in [1, F_w], \forall k$.

(2.) When the transmitter detects an event, it updates the allocated power of the channel accesses $n \in [\mathcal{E}_t, \min\{\mathcal{E}_t + F_w - 1, N\}]$ by using the *Boxed Water-Flowing* with the remaining energy in the battery and with the energy of the harvested packet (if the event is an energy arrival), i.e., $\sum_{j|e_j \leq \mathcal{E}_t} E_j - T_s \sum_{n=1}^{\mathcal{E}_t - 1} \mathcal{C}^2(\mathbf{p}_n)$, and by assuming that the channel remains constant during the flowing window, i.e., $h_{kn} = h_{k\mathcal{E}_t}, \forall n \in [\mathcal{E}_t, \min\{\mathcal{E}_t + F_w - 1, N\}], \forall k$.[11]

(3.) Step (2.) is repeated until the N-th channel access is reached. A natural requirement of WEHNs is that they operate perpetually. Note that the proposed online algorithm can operate in an infinite time window, i.e., $N \to \infty$, where the algorithm continuously remains in Step (2.).

The proposed online algorithm satisfies the ECCs and, as pointed out, does not require future information of neither the channel state nor the energy arrivals.

The mutual information that can be achieved by any online algorithm is inherently limited by the partial knowledge of the harvested energy and channel state. By using sophisticated statistical models of the Energy Harvesting (EH) process, one can design online algorithms that perform close to the optimal offline algorithm. Currently, the design of sophisticated online algorithms is limited by the following factors:

- There is a lack of models of the energy harvesting process (cf. §2.1.1).
- The computational complexity required by the resulting online algorithms has to be as low as possible since the energy spent in the computation of the online strategy cannot be used for transmission, which directly affects the achievable mutual information.

We believe that our proposed online algorithm correctly balances these two points since it is a low-complexity online algorithm (the estimation of F_w can be done during the node deployment when the node is not limited by the harvested energy) that achieves a remarkably high fraction of the mutual information as shown by numerical simulation in the next section.

[11] Note that the transmitter may stay silent in some channel accesses if the difference between two consecutive incoming energy packets is greater than the flowing window, $e_j - e_{j-1} > F_w$. This situation rarely takes place in practice since, in most common situations, F_w is several times the mean number of channel accesses per epoch. For example, in the simulated framework presented in §5.5, we have obtained that F_w is 5 times the mean number of channel accesses per epoch.

5.5 Simulation results

By numerical simulation, in this section we evaluate the performance of the different solutions presented in the previous sections. We have considered a total of $N = 100$ channel accesses in which symbols are transmitted through $K = 8$ parallel streams. The channel access duration is $T_s = 10\,\mathrm{ms}$. The power consumptions associated with the channel access and stream activation are $\alpha = 100\ \mathrm{mW}$ and $\beta = 10\ \mathrm{mW}$, respectively [49]. A Rayleigh fading channel has been considered where the channel power gain satisfies $\mathbb{E}\{h_{kn}\} = 1$. The energy harvesting process is modeled as a compound Poisson process as done in [84], where the packet arrival instants follow a Poisson distribution with rate $\frac{1}{5}$ and the energy in the packets is drawn from a uniform distribution and normalized by the total harvested energy that varies along the x-axis of Figures 5.3-5.7.

In the setup above, Figure 5.3 shows the achieved mutual information with the different presented resource allocation strategies: $\tilde{\mathcal{I}}^{\star}$ is the upper bound obtained in §5.3.1 by relaxing the feasible set of the stream and channel access IVs to the integer interval $[0, 1]$; $\hat{\mathcal{I}}$ is the mutual information achieved by the feasible the resource allocation $\{\hat{\mathbf{P}}, \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\Psi}}\}$ that is obtained by projecting ρ_n and ψ_{kn} into the set $\{0, 1\}$ as explained in §5.3.1, where we have used $\Gamma = 0.5$ as it provides a good performance; *Duality* shows the mutual information achieved by solving the dual problem as explained in §5.3.2; and *Online* depicts the mutual information achieved by the online algorithm presented in §5.4. Additionally, to assess the impact of energy harvesting versus traditional non-harvesting nodes, we have evaluated the performance of a virtual non-harvesting node in which the battery of the node is replaced by a new battery containing E_j Joules at the channel access e_j. Although this battery replacement is not feasible in practice, it allows us to fairly compare the performance of the energy harvesting node and the virtual battery operated node since both nodes have the same energy levels. For the non-harvesting node, we have designed a resource allocation strategy, named *Epoch by Epoch* (*EbE*), that uses the *Boxed Water-Flowing* in each epoch independently, i.e., water is not allowed to flow between epochs (due to the virtual battery replacement). Finally, we also compare our strategies with *DWF & PP* that uses the DWF in (2.21) with an additional post processing stage that scales the transmission powers to guarantee that the ECCs containing the circuitry power consumption are satisfied.

In the magnified plot in Figure 5.3, one can observe that the optimality gap is almost zero since the difference between the upper bound, $\tilde{\mathcal{I}}^{\star}$, and the strategies $\hat{\mathcal{I}}$ and *Duality* is almost zero (remember that $\mathcal{I}^{\star} - \hat{\mathcal{I}} \leq \tilde{\mathcal{I}}^{\star} - \hat{\mathcal{I}}$ and the same applies for *Duality*). As expected, the proposed online algorithm has performance loss in comparison to the optimal offline solution as it has no knowledge of the future channel state and energy arrivals. This performance loss

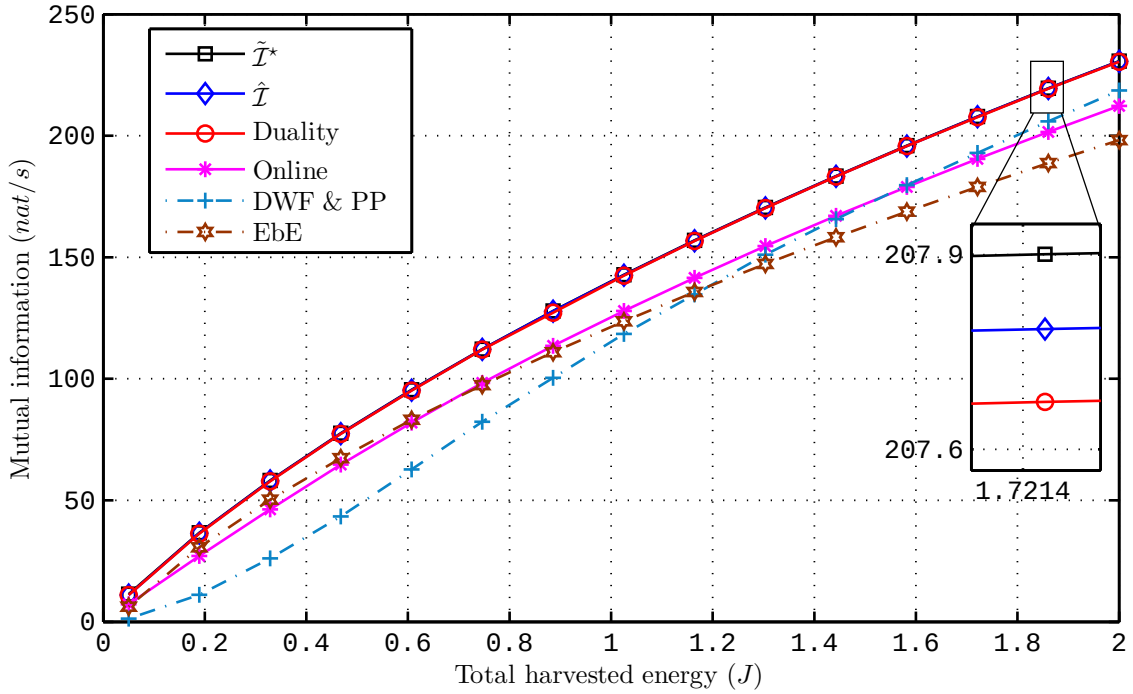

Figure 5.3: Achieved mutual information versus total harvested energy for the different algorithms.

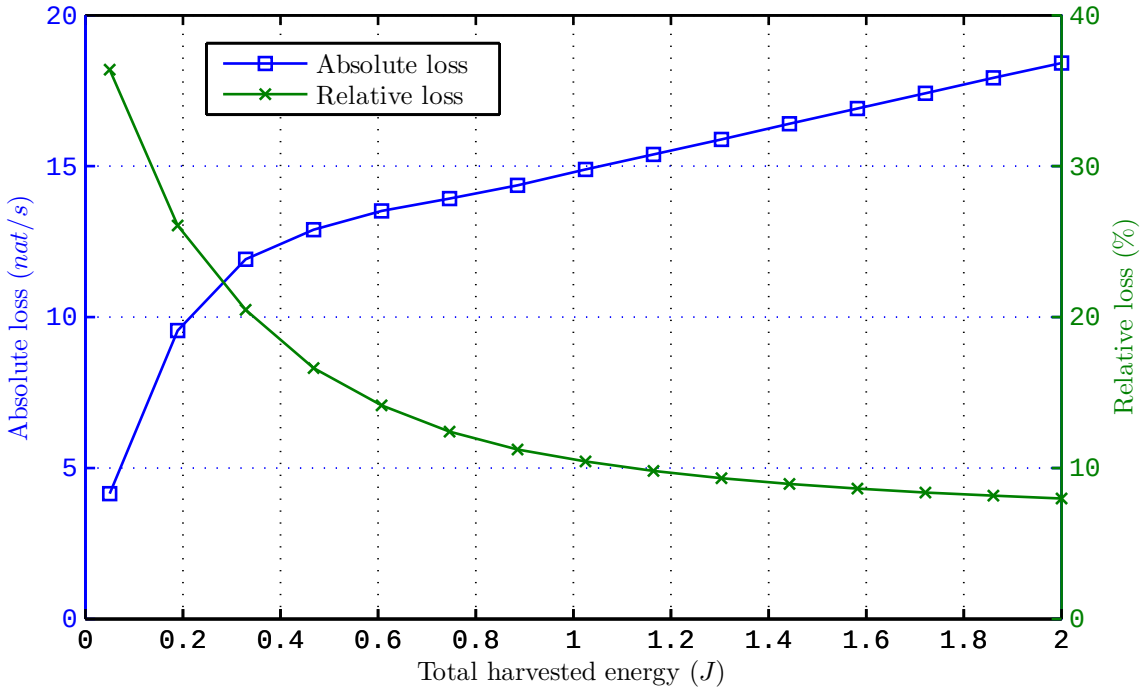

Figure 5.4: Performance loss of the proposed online algorithm versus the optimal offline solution.

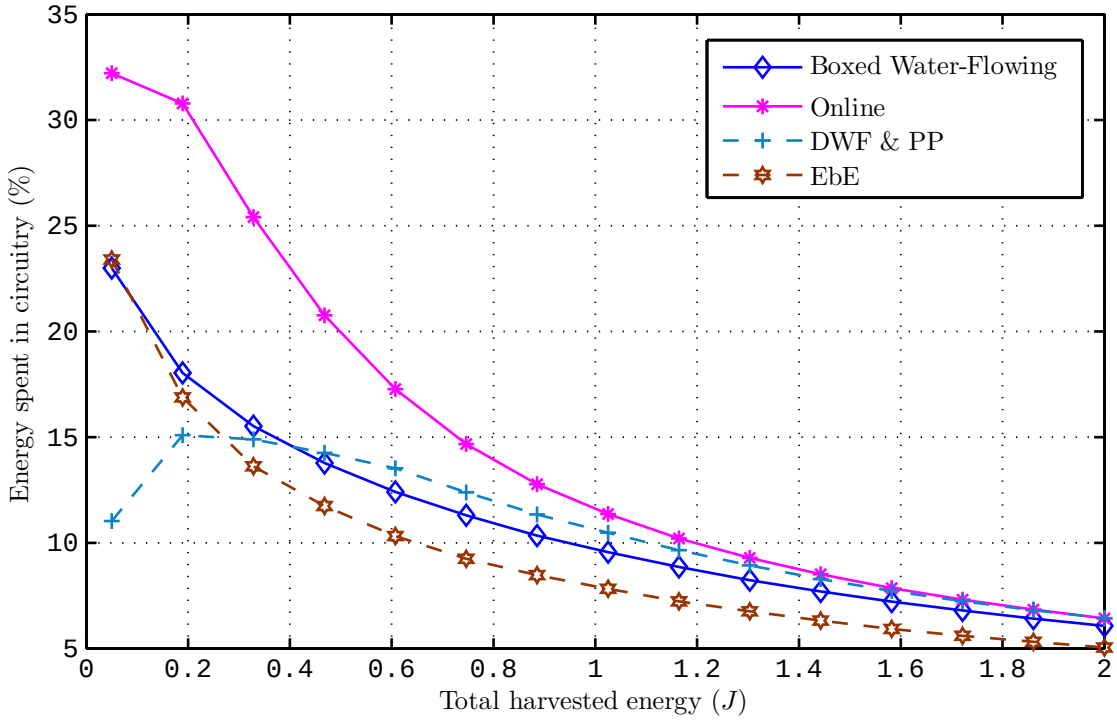

Figure 5.5: Percentage of the total harvested energy expended in the circuitry.

is evaluated in Figure 5.4 both in absolute (left y-axis) and relative (right y-axis) terms. It is observed that when the harvested energy is above 1 J, the relative performance loss is around 10%. The *Boxed Water-Flowing* solutions ($\mathcal{I}^\star$ and *Duality*) also outperform the *EbE* scheme (for the battery operated node) and the *DWF & PP*. It is observed that the performance of the *EbE* scheme drops down for high energy levels because water is not allowed to flow across epochs. In opposition, the *DWF & PP* behaves better for high energy levels since water is allowed to flow across epochs but has a very poor performance for low energy levels since the circuitry power consumption has not been accounted for in the optimization and plays an important role.

Figure 5.5 shows the percentage of the total harvested energy that is expended in the circuitry, i.e., $T_s \sum_{n=1}^{N} \left(\alpha \rho_n + \sum_{k=1}^{K} \beta \psi_{kn} \right) \frac{100}{\sum_{j=1}^{J} E_j}$. It is observed that when the harvested energy is low, the amount of energy spent in the circuitry components is a relatively high fraction of the total harvested energy. Additionally, when the harvested energy is high the different strategies show a similar percentage of circuitry energy consumption. However, the *Boxed Water-Flowing* strategies ($\mathcal{I}^\star$ and *Duality*) achieve a higher mutual information as seen in Figure 5.3. This is because the *Boxed Water-Flowing* solutions are able to activate the channel accesses and streams that contribute the most to the mutual information.

The computational complexities of the different strategies are compared in Figure 5.6 in

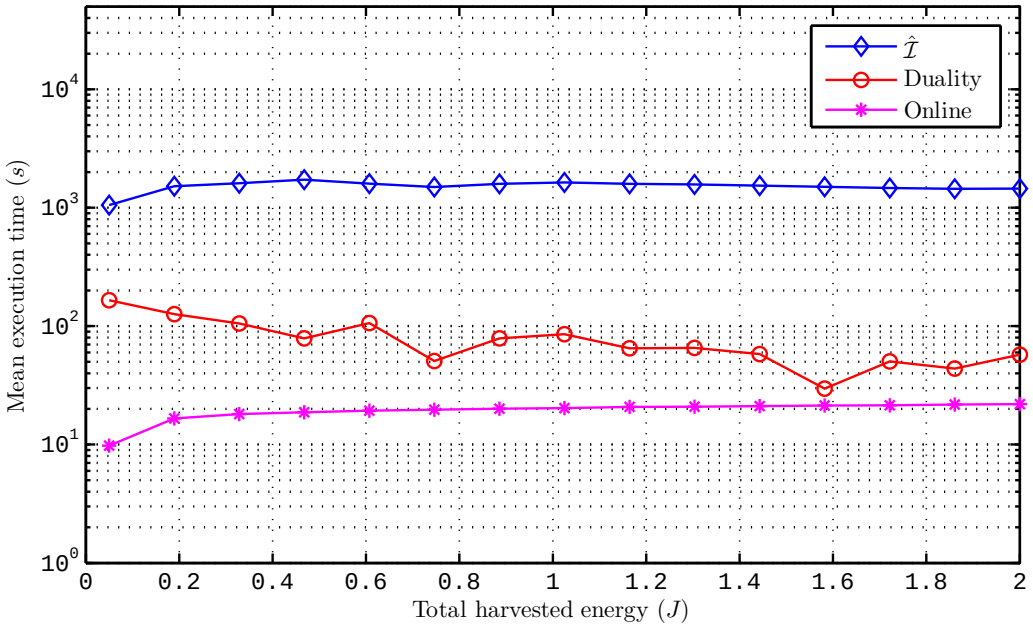

Figure 5.6: Mean execution time versus total harvested energy.

terms of measured execution time versus harvested energy.[12] Observe that *Duality* requires a much lower execution time than $\hat{\mathcal{I}}$. However, to obtain such a good performance, the step size to update the dual variable must be carefully selected depending on the energy harvesting profile. In some manner, $\hat{\mathcal{I}}$ is more robust to variations of the energy harvesting profile; however, at a cost of having a higher computational complexity. Moreover, the complexity of the proposed *Online* solution is remarkably low, which makes it a good candidate solution to be implemented in wireless devices.

As pointed out in Remark 5.2, the *Boxed Water-Flowing* algorithms in Algorithms 5.1 and 5.2, can be trivially extended to include mask constraints. Figure 5.7 evaluates the impact of two different mask constraints on the achieved mutual information for $K = 80$ parallel streams where $\beta = 1$ mW (the remaining system parameters are the ones mentioned above). We have considered two different mask constraints: the first mask, *Mask 1*, limits the transmission power in each stream as $p_{kn} \leq 25$ mW, $\forall k, n$; in the second mask, *Mask 2*, the transmission power of the external streams is further limited to avoid interferences to other possible transmissions, i.e., $p_{kn} \leq 5$ mW, $k \in [1, 20] \cup [61, 80], \forall n$, and $p_{kn} \leq 25$ mW, $\forall k \in [21, 60], \forall n$. As expected, at low energy levels, the mask constraint does not have a significant impact on the achieved mutual information because the transmission power in the different subchannels is low; however, when the harvested energy increases, the mask constraint limits the transmission power in the different subchannels and, as a result, the mutual information is reduced.

[12] Note that the execution time is approximately proportional to the algorithmic computational complexity.

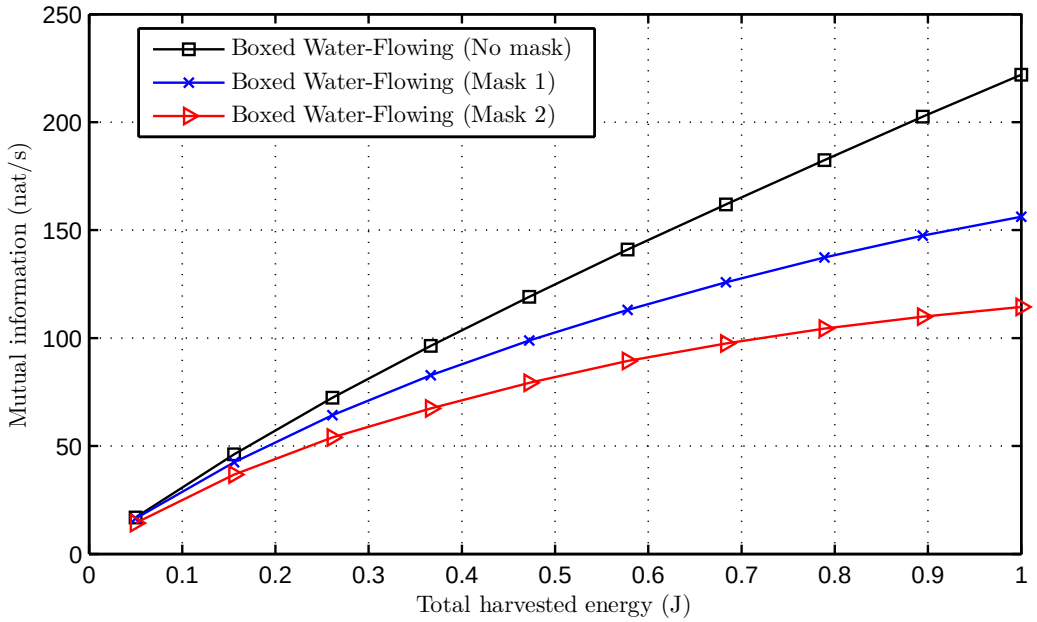

Figure 5.7: Mutual information of the *Boxed Water-Flowing* solution given different mask constraints.

5.6 Conclusions

In this chapter, we have studied the resource allocation for a WEHN that maximizes the mutual information along N independent channel accesses in which symbols are sent through K parallel streams. The main contribution w.r.t. previous works is that we not only account for the transmission radiated power but we also consider the channel access and stream activation costs. First, we have studied the offline maximization problem (where the transmitter has full knowledge of the harvested energy and channel state) and we have shown that it is not a convex optimization problem. Due to this lack of convexity, we have proposed and studied two different problems (the integer relaxation and the dual problem) from which we have obtained two suboptimal solutions of the offline maximization problem that asymptotically tend to the optimal solution when the number of channel accesses or streams per epoch is large. From these two problems, we have obtained a common condition for the activation of the channel access and streams, i.e., if the epoch water level is greater than the corresponding cutoff water level. Based on the cutoff water level concept, we have devised the *Boxed Water-Flowing*, a novel graphical representation of the asymptotically optimal offline resource allocation. Additionally, we have proposed a practical online algorithm that does not require knowledge of the future energy arrivals nor the channel state. From the simulation results, we have confirmed that the *Boxed Water-Flowing* resource allocation is the asymptotically optimal offline resource allocation and that the performance loss of the proposed *online* solution is very small. Moreover, we

have evaluated the computational complexity of the different resource allocation strategies and obtained that the online solution is the one that requires the lowest execution time.

5.A Appendix

5.A.1 Proof of Lemma 5.1

We know that the $M_n^\star$ streams that contribute to the channel access activation are the ones with the best channel gains, i.e., $h_{1n}, \ldots, h_{M_n^\star n}$, because these streams are the ones that contribute the most to the objective function. Given $M_n^\star$, the following two conditions must be satisfied at the cutoff water level, i.e., when $W_j = \hat{W}_n(M_n^\star)$:

(C1) $\forall k \leq M_n^\star$, the following relations must be fulfilled: (i) $p_{kn} > 0$; (ii) $\psi_{kn} \in (0, 1)$; (iii) $\hat{\eta}_{kn} = 0$; (iv) $\check{\eta}_{kn} = 0$; and (v) $\xi_{kn} = 0$.[13]

(C2) $\forall k > M_n^\star$, the following conditions must be satisfied: (i) $p_{kn} = 0$; (ii) $\psi_{kn} = 0$; (iii) $\check{\eta}_{kn} = 0$; and (iv) $\mu_{kn} = 0$.[14]

Going back to (5.7), the radiated power of the streams $k \leq M_n^\star$ at the channel access cutoff water level is $p_{kn} = \psi_{kn}\left(\hat{W}_n(M_n^\star) - h_{kn}^{-1}\right)$, where we have used that $W_j = \hat{W}_n(M_n^\star) = (T_s \sum_{\ell=j}^{J} \lambda_\ell)^{-1}$. Plugging this into the KKT condition in (5.6b), we have that all the streams that contribute to the n-th channel access activation, i.e., $k \leq M_n^\star$, must satisfy:

$$\frac{\partial \tilde{\mathcal{L}}}{\partial \psi_{kn}} = \log(h_{kn}\hat{W}_n(M_n^\star)) - 1 + \frac{1}{\hat{W}_n(M_n^\star)h_{kn}} - \frac{\beta}{\hat{W}_n(M_n^\star)} - \mu_{kn} = 0, \quad \forall k \leq M_n^\star. \tag{5.12}$$

Note that we cannot isolate $\hat{W}_n(M_n^\star)$ in the previous equation due to the dependence on the Lagrange multiplier μ_{kn}. To get rid of this dependence, we can use the KKT condition in (5.6c) evaluated at the cutoff water level, i.e.,

$$\sum_{k=1}^{K} \mu_{kn} = \sum_{k=1}^{M_n^\star} \mu_{kn} = \frac{\alpha}{\hat{W}_n(M_n^\star)}, \tag{5.13}$$

where we have used that $\mu_{kn} = 0$,$\forall k > M_n^\star$, which follows from (C2).

[13] Where (i) and (ii) follow from the fact that the stream must contribute to the channel access activation and (iii), (iv), and (v) follow from the slackness conditions (5.6f), (5.6g), and (5.6e), respectively.

[14] Where (i) and (ii) follow from the fact that the stream must not contribute to the channel access activation and (iii) and (iv) follow from the slackness conditions (5.6g) and (5.6h).

With this, we can obtain an equation that does not depend on μ_{kn} by adding the equations $\frac{\partial \tilde{\mathcal{L}}}{\partial \psi_{kn}} = 0, \forall k \leq M_n^\star$. Thus, the n-th channel access cutoff water level, $\hat{W}_n(M_n^\star)$, is obtained by solving $\sum_{k=1}^{M_n^\star} \frac{\partial \tilde{\mathcal{L}}}{\partial \psi_{kn}} = 0$, which is performed in Appendix 5.A.7 (set $M := M_n^\star$, $\hat{W} := \hat{W}_n(M_n^\star)$, $H_k := h_{kn}$, and $P_c := M_n^\star \beta + \alpha$), and is the one given in (5.8).

Up to now, we have shown that if $W_j = \hat{W}_n(M_n^\star)$, the channel access is "partially on", i.e., $\rho_n \in (0,1)$, $\psi_{kn} \in (0,1), \forall k \leq M_n^\star$ and $\psi_{kn} = 0, \forall k > M_n^\star$. Note that if $W_j = \hat{W}_n(M_n^\star) + \epsilon$, with $\epsilon > 0$, then we have that $\log(h_{kn} W_j) - 1 + \frac{1}{W_j h_{kn}} - \frac{\beta}{W_j} - \mu_{kn} > 0$. Thus, in order to satisfy (5.6b), we must have $\check{\eta}_{kn} > 0, \forall k \leq M_n^\star$. Then, from the slackness condition (5.6g), we know that $\psi_{kn} = 1, \forall k \leq M_n^\star$, and hence $\rho_n = 1$. A similar approach can be used to show that if $W_j < \hat{W}_n(M_n^\star)$, then $\psi_{kn} = 0, \forall k$, and hence $\rho_n = 0$. ■

5.A.2 Proof of Lemma 5.2

Now, we derive the expression of the k-th *stream cutoff water level*, $\bar{W}_{kn}$, i.e., the water level at which the k-th stream becomes partially active, where now $k > M_n^\star$. Similarly as in the proof of Lemma 5.1, in the k-th stream cutoff water level, we must have that $p_{kn} > 0$, $\psi_{kn} \in (0,1)$, $\rho_n = 1$, and from the slackness conditions we know that $\check{\eta}_{kn} = 0$, $\hat{\eta}_{kn} = 0$, $\xi_{kn} = 0$, $\mu_{kn} = 0$. The k-th stream cutoff water level is obtained by solving the equation obtained from the KKT condition in (5.6b) for $\bar{W}_{kn}$ (set $M := 1$, $\hat{W} := \bar{W}_{kn}$, $H_1 := h_{kn}$ and $P_c := \beta$ in Appendix 5.A.7) and is the one given in (5.9). Following the same procedure as in the last paragraph of the proof of Lemma 5.1, it can be shown that if $W_j > \bar{W}_{kn}$, $k \in (M_n^\star, K]$, then $\psi_{kn} = 1$. Alternatively, if $W_j < \bar{W}_{kn}$, $k \in (M_n^\star, K]$, then $\psi_{kn} = 0$. ■

5.A.3 Proof of Proposition 5.1

In this appendix, we prove that $M_n^\star$ can be found by performing a forward search over M_n and that it is the smallest M_n that satisfies the condition $\hat{W}_n(M_n) < \bar{W}_{(M_n+1)n}$. To prove Proposition 5.1, we need to make use of the following two lemmas:

Lemma 5.3. *If $\hat{W}_n(M_n) = \bar{W}_{(M_n+1)n}$, then $\hat{W}_n(M_n+1) = \hat{W}_n(M_n)$.*

Proof: See Appendix 5.A.4. ■

Lemma 5.4. *If $\hat{W}_n(M_n) \geq \bar{W}_{(M_n+1)n}$, then: (a) $\hat{W}_n(M_n+1) \geq \bar{W}_{(M_n+1)n}$; (b) the function $\hat{W}_n(M_n+1)$ is monotonically decreasing with $h_{(M_n+1)n}$.*

Proof: See Appendix 5.A.5. ■

Observe that $\bar{W}_{(M_n+1)n}$ is a function of $h_{(M_n+1)n}$ in opposition to $\hat{W}_n(M_n)$ that does not depend on $h_{(M_n+1)n}$. In this context, let $\tilde{h}_{(M_n+1)n}$ be the specific value of the channel gain of the stream $k = M_n + 1$ that satisfies $\bar{W}_{(M_n+1)n} = \hat{W}_n(M_n)$.

We first show that any M_n smaller than $M_n^\star$ is suboptimal. We know that $\hat{W}_n(M_n) \geq \bar{W}_{(M_n+1)n}, \forall M_n < M_n^\star$, since the condition $\hat{W}_n(M_n) < \bar{W}_{(M_n+1)n}$ is not satisfied until $M_n := M_n^\star$. Since $\bar{W}_{(M_n+1)n}$ is decreasing with the channel gain (see Remark 5.4), the condition $\hat{W}_n(M_n) \geq \bar{W}_{(M_n+1)n}$ implies $h_{(M_n+1)n} \geq \tilde{h}_{(M_n+1)n}$. From Lemma 5.3, we have that at $\tilde{h}_{(M_n+1)n}$, $\hat{W}_n(M_n+1) = \hat{W}_n(M_n)$ and from Lemma 5.4.b if $h_{(M_n+1)n} \geq \tilde{h}_{(M_n+1)n}$, then $\hat{W}_n(M_n+1) \leq \hat{W}_n(M_n), \forall M_n < M_n^\star$ or, equivalently, $\hat{W}_n(1) \geq \hat{W}_n(2) \geq \cdots \geq \hat{W}_n(M_n^\star)$. Therefore, any M_n in [1, $M_n^\star$) is suboptimal.

Now, we prove the suboptimality of any M_n greater than $M_n^\star$. From the structure of the forward search, the following relationship is satisfied $\hat{W}_n(M_n) < \bar{W}_{(M_n+1)n}, \forall M_n \in [M_n^\star, K]$. Thus, the streams $k > M_n^\star$ cannot contribute to the channel access activation since these streams are not active at the cutoff water level.

Finally, we must show that the streams $k \leq M_n^\star$ are active in the channel access cutoff water level, i.e., $\bar{W}_{k,n} \leq \hat{W}_n(M_n^\star)$, which is verified as proved in Lemma 5.4.a. ■

5.A.4 Proof of Lemma 5.3

Let the channel access cutoff water level for a given M_n be expressed as

$$\hat{W}_n(M_n) = \frac{X_{M_n}}{\frac{1}{e}(\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}})\mathcal{W}_0(X_{M_n})},$$

where X_{M_n} is the argument of the Lambert function, i.e.,

$$X_{M_n} = \frac{\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}}}{eM_n}(\alpha + M_n\beta - \sum_{k=1}^{M_n} h_{kn}^{-1}). \tag{5.14}$$

In the following lines, we impose the condition $\bar{W}_{(M_n+1)n} = \hat{W}_n(M_n)$, and, after some algebra, we obtain the argument of the Lambert function in $\hat{W}_n(M_n+1)$, i.e., $\tilde{X}_{M_n+1}$, where the tilde denotes that $\bar{W}_{(M_n+1)n} = \hat{W}_n(M_n)$:

$$\frac{(\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}})\bar{W}_{(M_n+1)n}}{e}\log\left(\frac{(\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}})\bar{W}_{(M_n+1)n}}{e}\right)=\frac{\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}}}{eM_n}(\alpha+M_n\beta-\sum_{k=1}^{M_n}h_{kn}^{-1})\Rightarrow \quad (5.15)$$

$$\tilde{X}_{M_n+1}=\frac{\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}}}{e(M_n+1)}\left[M_n\bar{W}_{(M_n+1)n}\log\left(\frac{(\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}})\bar{W}_{(M_n+1)n}}{e}\right)+\beta-h_{(M_n+1)n}^{-1}\right]\Rightarrow \quad (5.16)$$

$$\tilde{X}_{M_n+1}=\frac{(\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}})M_n\bar{W}_{(M_n+1)n}}{e(M_n+1)}\left[\log\left(\frac{(\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}})\bar{W}_{(M_n+1)n}}{e}\right)+\frac{\mathcal{W}_0(\frac{\beta h_{(M_n+1)n}-1}{e})}{M_n}\right]\Rightarrow \quad (5.17)$$

$$\tilde{X}_{M_n+1}=\frac{(\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}})\bar{W}_{(M_n+1)n}}{e(M_n+1)}\log\left(\frac{(\prod_{k=1}^{M_n} h_{kn})\bar{W}_{(M_n+1)n}^{M_n}}{e^{M_n}}e^{\mathcal{W}_0(\frac{\beta h_{(M_n+1)n}-1}{e})}\right)\Rightarrow \quad (5.18)$$

$$\tilde{X}_{M_n+1}=\frac{(\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}})\bar{W}_{(M_n+1)n}}{e}\log\left(\frac{(\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}})\bar{W}_{(M_n+1)n}}{e}\right) \quad (5.19)$$

To obtain (5.16), we have multiplied both sides in (5.15) by $\frac{M_n\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}}}{\prod_{k=1}^{M_n} h_{kn}^{\frac{1}{M_n}}}$ and used the definition of $\tilde{X}_{M_n+1}$, which follows from (5.14). In (5.19), we have used that $e^{\mathcal{W}(z)}=\frac{z}{\mathcal{W}(z)}$, which directly follows from the definition of the Lambert function.

The cutoff water level for M_n+1 active streams is $\hat{W}_n(M_n+1)=\frac{\tilde{X}_{M_n+1}}{\frac{1}{e}(\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}})\mathcal{W}_0(\tilde{X}_{M_n+1})}$, from where

$$\tilde{X}_{M_n+1}=\frac{\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}}}{e}\hat{W}_n(M_n+1)\log\left(\frac{\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}}}{e}\hat{W}_n(M_n+1)\right).$$

By comparing this expression with (5.19), we have that $\hat{W}_n(M_n+1)$ must be equal to $\bar{W}_{(M_n+1)n}$ and, thus, we have $\hat{W}_n(M_n+1)=\hat{W}_n(M_n)$. ■

5.A.5 Proof of Lemma 5.4

5.A.5.1 Proof of Lemma 5.4.a

Following similar steps as in (5.15)-(5.19), it is easy to show that if $\hat{W}_n(M_n)\geq\bar{W}_{(M_n+1)n}$, then $X_{M_n+1}\geq\tilde{X}_{M_n+1}$. From where, it follows $\frac{X_{M_n+1}}{\mathcal{W}_0(X_{M_n+1})}\geq\frac{(\prod_{k=1}^{M_n+1} h_{kn}^{\frac{1}{M_n+1}})\bar{W}_{(M_n+1)n}}{e}$ and, therefore, we also have $\hat{W}_n(M_n+1)\geq\bar{W}_{(M_n+1)n}$. ■

5.A.5.2 Proof of Lemma 5.4.b

In the following lines, we demonstrate that $\frac{\partial \hat{W}_n(M_n+1)}{\partial h_{(M_n+1)n}} \leq 0$ for $\hat{W}_n(M_n) \geq \bar{W}_{(M_n+1)n}$ or, equivalently, for $X_{M_n+1} \geq \tilde{X}_{M_n+1}$. Thus, we look at the sign of the derivative, $\text{sign}\left[\frac{\partial \hat{W}_n(M_n+1)}{\partial h_{(M_n+1)n}}\right]$, i.e.,

$$\text{sign}\left[\frac{h_{(M_n+1)n}^{-2}}{M_n+1}\mathcal{W}_0(X_{M_n+1}) - \frac{1}{M_n+1}\left((M_n+1)\beta+\alpha-\sum_{k=1}^{M_n+1}h_{kn}^{-1}\right)\frac{\mathrm{d}\mathcal{W}_0(X_{M_n+1})}{\mathrm{d}X_{M_n+1}}\frac{\partial X_{M_n+1}}{\partial h_{(M_n+1)n}}\right] \tag{5.20}$$

$$= \text{sign}\left[\mathcal{W}_0(X_{M_n+1}) - X_{M_n+1}\left(1+\frac{eh_{(M_n+1)n}X_{M_n+1}}{\prod_{k=1}^{M_n+1}h_{kn}^{\frac{1}{M_n+1}}}\right)\frac{\mathrm{d}\mathcal{W}_0(X_{M_n+1})}{\mathrm{d}X_{M_n+1}}\right] \tag{5.21}$$

$$= \text{sign}\left[\mathcal{W}_0(X_{M_n+1}) - X_{M_n+1}\left(1+mX_{M_n+1}\right)\frac{\mathcal{W}_0(X_{M_n+1})}{X_{M_n+1}(1+\mathcal{W}_0(X_{M_n+1}))}\right] \tag{5.22}$$

$$= \text{sign}\Big[\mathcal{W}_0(X_{M_n+1})\left(\mathcal{W}_0(X_{M_n+1}) - mX_{M_n+1}\right)\Big] \tag{5.23}$$

In (5.21), we have used that

$$\frac{\partial X_{M_n+1}}{\partial h_{(M_n+1)n}} = \frac{X_{M_n+1}}{(M_n+1)h_{(M_n+1)n}} + \frac{\prod_{k=1}^{M_n+1}h_{kn}^{\frac{1}{M_n+1}}}{e(M_n+1)h_{(M_n+1)n}^2}.$$

In (5.22), we have defined $m \triangleq eh_{(M_n+1)n}\left(\prod_{k=1}^{M_n+1}h_{kn}^{\frac{1}{M_n+1}}\right)^{-1}$ and evaluated the derivative of the Lambert function, i.e.,

$$\frac{\mathrm{d}\mathcal{W}_0(X_{M_n+1})}{\mathrm{d}X_{M_n+1}} = \frac{\mathcal{W}_0(X_{M_n+1})}{X_{M_n+1}(1+\mathcal{W}_0(X_{M_n+1}))}.$$

From (5.23), we see that the sign of the derivative depends on the product of the Lambert function (which is positive for $X_{M_n+1} > 0$ and negative for $X_{M_n+1} < 0$) and the difference between the Lambert function and a line with slope m. To demonstrate that $\hat{W}_n(M_n+1)$ is monotonically decreasing with $h_{(M_n+1)n}$ for $X_{M_n+1} \geq \tilde{X}_{M_n+1}$, we must show that

$$\mathcal{W}_0(X_{M_n+1}) \overset{(a)}{<} \bar{m}X_{M_n+1} \overset{(b)}{\leq} mX_{M_n+1}, \qquad \forall X_{M_n+1} > \max\{0, \tilde{X}_{M_n+1}\}, \tag{5.24}$$

$$\mathcal{W}_0(X_{M_n+1}) \overset{(a)}{>} \bar{m}X_{M_n+1} \overset{(b)}{\geq} mX_{M_n+1}, \qquad \forall \tilde{X}_{M_n+1} < X_{M_n+1} < 0, \tag{5.25}$$

where $\bar{m}$ is an arbitrary constant. Observe that when $\bar{m} \leq m$, inequalities (b) in (5.24) and (5.25) are satisfied. In the following lines, we propose a specific $\bar{m}$ that allows us to prove inequalities in (a) and thus to demonstrate that $\hat{W}_n(M_n+1)$ is monotonically decreasing with

$h_{(M_n+1)n}$.

Note that, by replacing m in (5.19), we obtain

$$\tilde{X}_{M_n+1} = \frac{\bar{W}_{(M_n+1)n} h_{(M_n+1)n}}{m} \log\left(\frac{\bar{W}_{(M_n+1)n} h_{(M_n+1)n}}{m}\right).$$

From this relation, we can express m as a function of $\tilde{X}_{M_n+1}$, i.e.,

$$m = \bar{W}_{(M_n+1)n} h_{(M_n+1)n} \frac{\mathcal{W}_0(\tilde{X}_{M_n+1})}{\tilde{X}_{M_n+1}}.$$

Note that $\bar{W}_{(M_n+1)n} h_{(M_n+1)n} \geq 1$ because $\bar{W}_{(M_n+1)n} \in [h^{-1}_{(M_n+1)n}, \infty)$ as pointed out in Remark 5.4. Thus, $m \geq \frac{\mathcal{W}_0(\tilde{X}_{M_n+1})}{\tilde{X}_{M_n+1}}$. Let $\bar{m}$ denote the minimum slope, i.e., $\bar{m} = \frac{\mathcal{W}_0(\tilde{X}_{M_n+1})}{\tilde{X}_{M_n+1}}$. Since $\bar{m} \leq m$ the inequalities (b) in (5.24) and (5.25) are satisfied. Note that the Lambert function, $\mathcal{W}_0(X_{M_n+1})$, and the line $\bar{m}X_{M_n+1}$ cross both at the origin ($X_{M_n+1} = 0$) and at the point $X_{M_n+1} = \tilde{X}_{M_n+1}$, i.e., $\mathcal{W}_0(\tilde{X}_{M_n+1}) = \bar{m}\tilde{X}_{M_n+1}$. Finally, by using the concavity of the positive branch of the Lambert function [80] and these two crossing points, it is straight forward to show that the inequalities (a) in (5.24) and (5.25) are satisfied. Thus, $\hat{W}_n(M_n+1)$ is monotonically decreasing with $h_{(M_n+1)n}$. ■

5.A.6 Proof of Proposition 5.2

Let $\mathbf{E}_1$ and $\mathbf{E}_2$ be two different energy harvesting profiles, where the energy packet arrival instants are the same but the amount of energy in the packets is different, and let $\left\{\mathbf{P}^{\star}_{\mathbf{E}_1}, \boldsymbol{\rho}^{\star}_{\mathbf{E}_1}, \boldsymbol{\Psi}^{\star}_{\mathbf{E}_1}\right\}$ and $\left\{\mathbf{P}^{\star}_{\mathbf{E}_2}, \boldsymbol{\rho}^{\star}_{\mathbf{E}_2}, \boldsymbol{\Psi}^{\star}_{\mathbf{E}_2}\right\}$ be the associated optimal solutions to (5.4), respectively. Therefore, showing that the time-sharing condition is fulfilled is equivalent to demonstrating that

$$\begin{aligned} &I(\mathbf{P}^{\star}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\rho}^{\star}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\Psi}^{\star}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}) \geq \\ &\quad I(\mathbf{P}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\rho}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\Psi}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}) \geq \\ &\quad\quad \theta I(\mathbf{P}^{\star}_{\mathbf{E}_1}, \boldsymbol{\rho}^{\star}_{\mathbf{E}_1}, \boldsymbol{\Psi}^{\star}_{\mathbf{E}_1}) + (1-\theta) I(\mathbf{P}^{\star}_{\mathbf{E}_2}, \boldsymbol{\rho}^{\star}_{\mathbf{E}_2}, \boldsymbol{\Psi}^{\star}_{\mathbf{E}_2}) \end{aligned}$$

where $\theta \in [0,1]$, $\left\{\mathbf{P}^{\star}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\rho}^{\star}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\Psi}^{\star}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}\right\}$ is the optimal resource allocation for an energy harvesting profile equal to $\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2$ and $\left\{\mathbf{P}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\rho}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\Psi}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}\right\}$ is any feasible resource allocation.

In the following lines, we construct a feasible resource allocation,

$$\left\{\mathbf{P}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\rho}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\Psi}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}\right\},$$

that satisfies the time sharing condition in an epoch by epoch basis. In this context, since the procedure is the same for all the epochs, we just explain how to obtain the resource allocation of the streams contained in a generic epoch τ_j. Let $u = 1, \ldots, \mathcal{U}$ be an index used to indicate the different channel realizations observed within the streams in τ_j. Thus, $\mathcal{U}$ is the number of different channel realizations within the epoch. Let the set $\mathbb{S}_u$ contain all the streams in τ_j that have channel gain equal to $\bar{h}_u$, i.e., $\mathbb{S}_u = \{\{k, n\} | h_{kn} = \bar{h}_u, n \in \tau_j\}$. K_u denotes the cardinality of $\mathbb{S}_u$, which is a large number since, in the proposition statement, we have considered that within each epoch every channel realization is observed a sufficiently large number of times. In the following lines, we explain how to construct the resource allocation of the channel accesses in $\mathbb{S}_u$. Since K_u is large, θ can be approximated as $\theta \approx \frac{\bar{N}_u}{K_u}$, where $\bar{N}_u$ is an integer in the interval $[0, K_u]$. Due to the nature of DWF, the power allocated by the optimal solution given some energy harvesting profile is equal for all the streams in $\mathbb{S}_u$ because when a certain channel access is active, the transmission power in (2.21) only depends on the epoch water level and on the channel gain. We construct the resource allocation in the streams in $\mathbb{S}_u$ by assigning the resource allocation in $\{[\mathbf{P}^\star_{\mathbf{E}_1}]_{k'n'}, [\boldsymbol{\rho}^\star_{\mathbf{E}_1}]_{n'}, [\boldsymbol{\Psi}^\star_{\mathbf{E}_1}]_{k'n'}\}$ to $\bar{N}_u$ streams of $\mathbb{S}_u$ and the resource allocation in $\{[\mathbf{P}^\star_{\mathbf{E}_2}]_{k'n'}, [\boldsymbol{\rho}^\star_{\mathbf{E}_2}]_{n'}, [\boldsymbol{\Psi}^\star_{\mathbf{E}_2}]_{k'n'}\}$ to the remaining $K_u - \bar{N}_u$ streams, where $\{k', n'\}$ is any stream contained in $\mathbb{S}_u$. This procedure is repeated for the different channel realizations to obtain the resource allocation in all the streams of τ_j, then the total power consumption in τ_j is equivalent to θ times the power consumption in τ_j given by $\left\{\mathbf{P}^\star_{\mathbf{E}_1}, \boldsymbol{\rho}^\star_{\mathbf{E}_1}, \boldsymbol{\Psi}^\star_{\mathbf{E}_1}\right\}$ and $(1-\theta)$ times the power consumption in τ_j given by $\left\{\mathbf{P}^\star_{\mathbf{E}_2}, \boldsymbol{\rho}^\star_{\mathbf{E}_2}, \boldsymbol{\Psi}^\star_{\mathbf{E}_2}\right\}$. After repeating this process for all the epochs, the constructed resource allocation $\left\{\mathbf{P}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\rho}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}, \boldsymbol{\Psi}_{\theta\mathbf{E}_1+(1-\theta)\mathbf{E}_2}\right\}$ is a feasible solution as the ECCs are satisfied and the obtained mutual information is $\theta I(\mathbf{P}^\star_{\mathbf{E}_1}, \boldsymbol{\rho}^\star_{\mathbf{E}_1}, \boldsymbol{\Psi}^\star_{\mathbf{E}_1}) + (1-\theta) I(\mathbf{P}^\star_{\mathbf{E}_2}, \boldsymbol{\rho}^\star_{\mathbf{E}_2}, \boldsymbol{\Psi}^\star_{\mathbf{E}_2})$. Therefore, we have shown that the *time-sharing condition* is satisfied. ■

5.A.7 Derivation of the cutoff water level

In this appendix, we use a generic notation that serves us to derive both $\hat{W}_n(M_n)$ and $\bar{W}_{kn}$. Thus, $\hat{W}$ denotes the cutoff water level for the case of considering M active streams at the cutoff region with gains H_k, $k \in [1, M]$, and when the circuitry power consumption is P_c. Specifically, $\hat{W}$ is obtained by solving the equation $\sum_{k=1}^{M} \left[\log(\hat{W} H_k) - 1 + \frac{1}{\hat{W} H_k}\right] - \frac{P_c}{\hat{W}} = 0$ as follows:

$$\hat{W}\left(\prod_{k=1}^{M} H_k^{\frac{1}{M}}\right)\left[\log\left(\hat{W}\prod_{k=1}^{M} H_k^{\frac{1}{M}}\right) - \log e\right] + \frac{1}{M}\left(\prod_{k=1}^{M} H_k^{\frac{1}{M}}\right)\left(\sum_{k=1}^{M} H_k^{-1}\right) - \frac{P_c}{M}\left(\prod_{k=1}^{M} H_k^{\frac{1}{M}}\right) = 0 \Rightarrow \quad (5.26)$$

$$\frac{\hat{W}\prod_{k=1}^{M} H_k^{\frac{1}{M}}}{e}\log\left(\frac{\hat{W}\prod_{k=1}^{M} H_k^{\frac{1}{M}}}{e}\right) = \frac{\prod_{k=1}^{M} H_k^{\frac{1}{M}}}{Me}(P_c - \sum_{k=1}^{M} H_k^{-1}) \Rightarrow \quad (5.27)$$

$$\hat{W} = \frac{P_c - \sum_{k=1}^{M} H_k^{-1}}{M\mathcal{W}_0\left(\frac{\prod_{k=1}^{M} H_k^{\frac{1}{M}}}{Me}(P_c - \sum_{k=1}^{M} H_k^{-1})\right)} \quad (5.28)$$

In (5.26), we have multiplied by $\frac{\hat{W}}{M}\left(\prod_{k=1}^{M} H_k^{\frac{1}{M}}\right)$. In (5.28), we have used that $b \log b = a \Leftrightarrow b = \frac{a}{\mathcal{W}(a)}$, which follows from the definition of the Lambert function [80]. Moreover, since $\hat{W} > H_k^{-1}, \forall k = [1, M]$, so that the streams are active in the cutoff region, the term b is always greater than e^{-1} and, thus, the positive branch of the Lambert function, which is denoted by $\mathcal{W}_0(\cdot)$, is used. ■

Chapter 6

Sum-rate maximization in multiuser communications with WEHNs considering a generalized power consumption model

6.1 Introduction

In the previous chapter, we have seen that it is key that the power consumption model accounts not only for the transmission radiated power but also for the remaining energy sinks, which can be modeled, in general, by step functions. As a result of considering the channel access and stream activation costs, the (asymptotically) optimal transmission strategy alternates between "off" and "on" cycles.

The power consumption model considered in the previous chapter, $\mathcal{C}^2$ in (5.1), is a good first step towards the design of more efficient transmission strategies; however, still many sources of energy consumption escape from the model, which must be progressively included in the design. For example, "off-on" transitions also entail energy consumption due to the startup time of the transceiver [51, 121] that can be taken into account by refining the power

consumption model $\mathcal{C}^2$ as:

$$\mathcal{C}_n^3(\mathbf{p}_n, \mathbf{p}_{n+1}) = \begin{cases} \mathcal{C}^2(\mathbf{p}_n) & \text{if } n = N, \\ \mathcal{C}^2(\mathbf{p}_n) + \underbrace{\gamma \left(1 - \mathcal{H}^\ell \left(\sum_{k=1}^{K} p_{kn}\right)\right) \mathcal{H}^\ell \left(\sum_{k=1}^{K} p_{k(n+1)}\right)}_{\text{Startup power consumption}} & \text{if } n = 1, \ldots, N-1, \end{cases} \tag{6.1}$$

where $\gamma \geq 0$ is the mean startup power consumption of the transceiver; $\left(1 - \mathcal{H}^\ell \left(\sum_{k=1}^{K} p_{kn}\right)\right)$ is one when the n-th channel access is "off", $\mathcal{H}^\ell \left(\sum_{k=1}^{K} p_{k(n+1)}\right)$ is one when the $(n+1)$-th channel access is "on", and their product takes value one when an "off-on" transition occurs. Note that now the power consumption model depends on the channel access index, n, since the last channel access cannot incur on startup power consumption.

Additionally, in the previous chapter, we have seen that when the power consumption model includes step functions, the linear transmitter design is a nonconvex nonsmooth problem since the ECCs are no longer linear (as occurs when solely the transmission radiated power is considered, e.g., in DWF, cf. §2.3.1). Therefore, the optimal transmission strategy cannot be found with classical convex optimization techniques. To overcome the nonconvexity of the problem, in the previous chapter we have replaced the step functions by additional optimization variables, named IVs, which has allowed us to derive an asymptotically optimal solution of the problem.

The main advantages of reformulating the problem with IVs are:

(A1) Although the resulting problem is nonconvex (due to the presence of the binary IVs), by integer relaxation, dual decomposition, or other techniques, one might obtain a convex problem that performs close to the optimal solution as occurs in Chapter 5.

(A2) When the resulting problem is analytically tractable, one can obtain an intuition of the structure of the solution as it has been shown with the *Boxed Water-Flowing*.

Nevertheless, the approach based on IVs also has the following inconveniences:

(I1) The computational complexity of the problem increases due to the additional number of optimization variables.

(I2) In many transmitter designs, as the one explored in this chapter, even after performing integer relaxation of the IVs, the problem is still nonconvex; accordingly, for such

problems, the use of IVs does not seem the best approach to obtain a good transmission strategy.

(I3) It is generally difficult to deal with products of step functions in the power consumption model since products of optimization variables are nonconvex even after integer relaxation, which leads to (I2); accordingly, it would be difficult to derive good solutions given the power consumption model $\mathcal{C}_n^3$ in (6.1).

(I4) It is not robust to variations on the power consumption model, i.e., one must formulate and solve an optimization problem for every single variation of the power consumption model and, as we have seen in Chapter 5, finding the solution to the problem might be indeed involved.

In this context, in the current chapter we propose an alternative approach for solving power allocation problems that contain step functions; the proposed approach performs a successive approximation of the step functions by smooth functions in order to derive a smooth nonconvex optimization problem that can be solved by SCA [67]. The main advantages of this approach w.r.t. the use of IVs are:

- It applies to nonconvex optimization problems.
- It accepts products of step functions, which allows us to take into account, among others, the startup power consumption.
- It can be easily adapted to include new sources of energy consumption.
- The problem can be solved in a distributed way under very mild assumptions.
- In some applications, it provides a faster convergence due to the lower dimensionality of the problem.

More precisely, in this chapter, we consider a network of WEHNs transmitting simultaneously in a Gaussian interference channel and investigate a distributed power allocation algorithm that maximizes the sum-rate by considering a general power consumption model, which is composed of step functions and permits to take into account the different sources of energy consumption at each transmitter (e.g., the energy spent in the different components of the RF chain when the transmitter is "on" or the startup power consumption).

To the best of our knowledge, a few works have considered the interference channel for WEHNs: the authors of [114] assumed that the radiated power is the unique energy sink and derived a policy that maximizes the sum-rate; whereas, the two-user Gaussian interference channel was studied in [115] by considering the cost of having the transmitter "on".

The major contributions of this chapter are:

- Proposing a smooth approximation of the step function in (2.6) that can be easily convexified by state of the art SCA algorithms.
- Studying the power allocation strategy that maximizes the sum-rate in a network of energy harvesting nodes, which transmit simultaneously over a shared channel, by considering a generalized power consumption model that accepts products of step functions.
- Proposing the Iterative Smooth and Convex approximation Algorithm (ISCA), which distributedly computes a stationary solution of the aforementioned problem.

The remainder of the chapter is structured as follows. In §6.2, the sum-rate maximization problem for a network of WEHN is formulated. In §6.3, the smooth and convex approximations of the step functions are given. In §6.4, the ISCA is presented for a general power consumption model, which is particularized in §6.5 to the power consumption model $\mathcal{C}^2$ in (5.1). The performance of the ISCA in terms of achieved rate and computational complexity is numerically evaluated in §6.6. Finally, the chapter is concluded in §6.7.

6.2 System model and problem formulation

In the remaining of the chapter, we use the same notation as in the previous chapters with the only difference that we use the subindex t to differentiate between the different transmitters.

We consider a Gaussian interference channel composed of $\mathcal{T}$ transmitter and receiver pairs sharing the same band over SISO frequency-selective links composed of K parallel subcarriers. Transmission takes place during N time slots of duration T_s where, at the n-th slot, the channel power gain from transmitter t to receiver r at the k-th subcarrier is denoted by $h_{tr}(k,n)$. We do not consider interference cancellation techniques in order to avoid the need of having a centralized control or coordination in the network. Accordingly, we treat the MUI as additive colored noise. Thus, assuming that Gaussian signaling is used, the rate of user t depends on

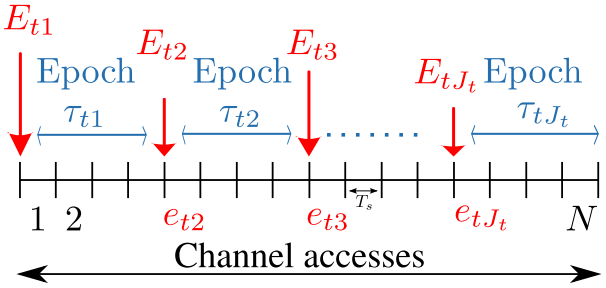

Figure 6.1: Representation of the energy harvesting process at the t-th transmitter.

its radiated power, $\mathbf{p}_t \in \mathbb{R}_+^{KN}$, $\mathbf{p}_t = (\mathbf{p}_{tn})_{n=1}^{N}$, $\mathbf{p}_{tn} = (p_t(k,n))_{k=1}^{K}$, and on the transmission power of all the other transmitters, i.e., $\mathbf{p}_{-t} = (\mathbf{p}_{t'})_{t' \neq t=1}^{\mathcal{T}}$, and reads as

$$r_t(\mathbf{p}_t, \mathbf{p}_{-t}) = \sum_{n=1}^{N} \sum_{k=1}^{K} \log \left(1 + \frac{p_t(k,n) h_{tt}(k,n)}{\sigma_t^2(k) + \sum\limits_{t' \neq t} p_t(k,n) h_{t't}(k,n)} \right), \tag{6.2}$$

where $\sigma_t^2(k)$ denotes the noise power spectral density at the t-th receiver and k-th subcarrier.[1]

We consider that the transmitters can harvest energy from the environment to recharge their batteries. As argued in §2.1.1, we model the energy harvesting process at each transmitter with a packetized model, which is depicted in Figure 6.1, where the subindex t is used to differentiate between transmitters.

The aim of this chapter is not to propose the ultimate power consumption model but to derive an algorithmic framework that is able to compute a distributed power allocation strategy for a very general power consumption model, which is composed of summations and products of step functions. This framework allows the use of a different power consumption model for each transmitter, which can be specified when the network is being deployed and one knows the different energy sinks of each node. Note that the solutions in [10–12, 49–51, 93, 115] are not robust to variations of the power consumption model, but a new optimization problem must be formulated and solved. In this context, we consider that the power consumption model has the following general form

$$\mathcal{C}_{tn}^{4}(\mathbf{p}_t) = \underbrace{\sum_{k=1}^{K} p_t(k,n)}_{\text{Transmission power}} + \underbrace{\sum_{s=1}^{S_{tn}} w_{tns} \prod_{q=1}^{Q_{tns}} \mathcal{H}^{\ell}\left(\phi_{tnsq}(\mathbf{p}_t)\right)}_{\text{Remaining energy sinks}}, \tag{6.3}$$

[1] We consider that each channel varies sufficiently slowly, so that the information theoretical results are meaningful.

where the subindex t and n denote a specific transmitter and channel access, respectively; $w_{tns} \neq 0$ is a given set of weights; S_{tn} and $Q_{tns} \in \mathbb{N}$; and $\phi_{tnsq} : \mathbb{R}_+^K \to \mathbb{R}_+$ must be concave, Lipschitz continuous, and continuously differentiable.

Since the different transmitters may have different architectures, they may also experience different consumptions for the channel access/slot activation costs and startup of the transceiver. Accordingly, we index the power consumption models $\mathcal{C}^2$ in (5.1) and $\mathcal{C}_n^3$ in (6.1) with an additional index to distinguish between transmitters. Thus, we have

$$\mathcal{C}_t^2(\mathbf{p}_{tn}) = \underbrace{\sum_{k=1}^{K} p_t(k,n)}_{\text{Transmission power}} + \underbrace{\alpha_t \mathcal{H}^\ell \left(\sum_{k=1}^{K} p_t(k,n) \right)}_{\text{Power consumption per active slot}} + \sum_{k=1}^{K} \underbrace{\beta_t \mathcal{H}^\ell \left(p_t(k,n) \right)}_{\text{Power consumption per active stream}}, \tag{6.4}$$

and $\mathcal{C}_{tn}^3(\mathbf{p}_{tn}, \mathbf{p}_{t(n+1)}) =$

$$\begin{cases} \mathcal{C}_t^2(\mathbf{p}_{tn}) & \text{if } n = N, \\ \mathcal{C}_t^2(\mathbf{p}_{tn}) + \underbrace{\gamma_t \left(1 - \mathcal{H}^\ell \left(\sum_{k=1}^{K} p_t(k,n) \right) \right) \mathcal{H}^\ell \left(\sum_{k=1}^{K} p_t(k,n+1) \right)}_{\text{Startup power consumption}} & \text{if } n = 1, \ldots, N-1, \end{cases} \tag{6.5}$$

respectively.

Note that $\mathcal{C}_t^2$ and $\mathcal{C}_{tn}^3$ are particular cases of $\mathcal{C}_{tn}^4$ in (6.3); the associated weights and inner functions, $\phi_{tnsq}(\mathbf{p}_t)$, are given in Table 6.1.

Our objective is to design a distributed power allocation strategy that maximizes the sum-rate under an offline approach (cf. §2.1.3), i.e., by assuming that we have non-causal knowledge of the channel gains and energy harvesting process. As argued in §2.1.3, this knowledge is available when the channel is time-static and the power harvesting profile is controllable or predictable, which happens, for example, with a solar harvester. Note that if a distributed power allocation strategy is derived, then the nodes can recompute the solution when a substantial change in the power profile prediction is observed. In non-static scenarios, the offline power allocation is still useful because it can be used as a benchmark for the design and evaluation of online transmission strategies, which does not require future knowledge of the harvested energy and channel state.

In this context, the offline sum-rate maximization problem is

$\mathcal{C}_t^2$ and $\mathcal{C}_{tN}^3$.

s	$1,\ldots,K$	$K+1$
w_{tns}	β_t	α_t
$\phi_{tns1}(\mathbf{p}_t)$	$p_t(s,n)$	$\sum_{k=1}^K p_t(k,n)$

$\mathcal{C}_{tn}^3$ for $n=1,\ldots,N-1$.

s	$1,\ldots,K$	$K+1$	$K+2$	$K+3$
w_{tns}	β_t	α_t	γ_t	$-\gamma_t$
$\phi_{tns1}(\mathbf{p}_t)$	$p_t(s,n)$	$\sum_{k=1}^K p_t(k,n)$	$\sum_{k=1}^K p_t(k,n+1)$	$\sum_{k=1}^K p_t(k,n)$
$\phi_{tns2}(\mathbf{p}_t)$	-	-	-	$\sum_{k=1}^K p_t(k,n+1)$

Table 6.1: Weights and inner functions to rewrite $\mathcal{C}_t^2$ and $\mathcal{C}_{tn}^3$ as $\mathcal{C}_{tn}^4$.

$$(\hat{P}): \quad \underset{\mathbf{p}}{\text{maximize}} \quad \sum_{t=1}^{\mathcal{T}} r_t(\mathbf{p}_t,\mathbf{p}_{-t}) \tag{6.6a}$$

$$\left.\begin{aligned} \text{subject to} \quad & \boldsymbol{\mathcal{B}}_t(\mathbf{p}_t)\succeq \mathbf{0}_{J_t}, \quad \forall t=1,\ldots,\mathcal{T}, \\ & \mathbf{p}_t\in\mathcal{P}_t, \qquad\quad \forall t=1,\ldots,\mathcal{T}, \end{aligned}\right\}\triangleq\hat{\mathcal{P}} \tag{6.6b, 6.6c}$$

where the transmission strategy $\mathbf{p}=(\mathbf{p}_t)_{t=1}^{\mathcal{T}}$ must satisfy the ECCs in (6.6b), which impose that the battery level must be nonnegative or, equivalently, that the energy cumulatively expended by the end of the ℓ-th epoch is not greater than the energy cumulatively harvested, i.e.,

$$[\boldsymbol{\mathcal{B}}_t(\mathbf{p}_t)]_\ell \triangleq \underbrace{\sum_{j=1}^{\ell} E_{tj}}_{\text{Harvested energy}} - \underbrace{T_s\sum_{j=1}^{\ell}\sum_{n\in\tau_{tj}}\mathcal{C}_{tn}^4(\mathbf{p}_t)}_{\text{Expended energy}} \geq 0, \qquad \ell=1,\ldots,J_t, \tag{6.7}$$

where $\mathcal{C}_{tn}^4(\mathbf{p}_t)$ is given in (6.3). Additionally, the transmission strategy $\mathbf{p}$ must fulfill the temporal-spectral mask constraints (cf. §2.1.5) in (6.6c) that limit the maximum transmit power, where $\mathcal{P}_t=\{\mathbf{p}_t:\mathbf{p}_t\succeq\mathbf{0}_{KN},\ \mathbf{p}_t\preceq\mathbf{p}_t^{\max}\}$ with $\mathbf{p}_t^{\max}\triangleq((p_t^{\max}(k,n))_{k=1}^K)_{n=1}^N$.

The problem in (6.6) has the following major difficulties: first, it is nonsmooth, nonconvex and NP-hard, which was shown in [122] for the simpler scenario of not having energy harvesting nor step functions; and second, it is key that the solution can be computed distributedly

by the network nodes because then they can adjust their strategies when strong variations of the energy harvesting profile are observed. In this context, we propose the ISCA that is able to compute a distributed power allocation strategy that aims at maximizing the sum-rate. The details of the ISCA are presented in §6.4. By now, it is important to know that it is composed of two loops: the outer loop performs a SSA of the step functions to derive, at each iteration, a smooth nonconvex problem that approximates (6.6); then, the inner loop solves this smooth nonconvex problem by means of the SCA algorithm in [67].

In the following section, we introduce a smooth approximation of the step function that can be easily convexified by SCA algorithms.

6.3 Approximations of the step function

The objective of this section is two-fold: (i) to design a smooth approximation of the unit step function in (2.6), which is used in the outer loop of the ISCA; and (ii) to derive a convex approximation of the smooth approximation in (i) that can be handled by the SCA algorithm in the inner loop. The later approximation, (ii), has to satisfy some tight requirements in order to guarantee convergence [67], which are listed in [67, Assumption 3], that intrinsically couple the design of the approximations in (i) and (ii) because depending on the chosen smooth approximation, it might be either easy or extremely difficult to later find an accurate convex approximation. In this context, we can easily derive a convex approximation if the smooth approximation is:

(C1) Differentiable.

(C2) An addition of concave and convex functions.

6.3.1 Smooth approximation of the step function

In this section, we present a smooth approximation of the step function that satisfies (C1)-(C2); whereas, in §6.3.2, we present its convexification that satisfies [67, Assumption 3].

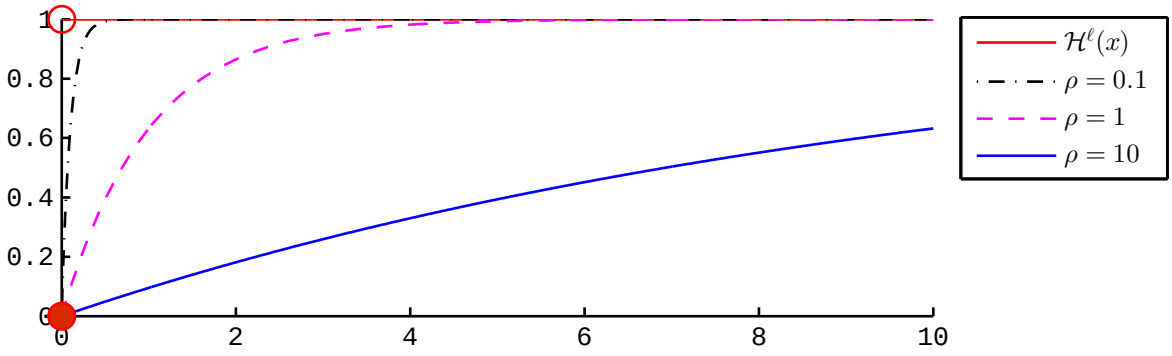

Figure 6.2: Representation of $\mathcal{H}_{\rho}(x)$ in (6.8) for different values of the approximation control parameter ρ.

6.3.1.1 Single step function

We first propose an approximation of the single step function in (2.6). Note that since we focus on power allocation problems, we limit the domain of the step function to the nonnegative real numbers. According to the requirements in (C1)-(C2), we approximate $\mathcal{H}^{\ell}(x)$ for $x \in \mathbb{R}_{+}$ with the function $\mathcal{H}_{\rho} : \mathbb{R}_{+} \to [0, 1]$ with

$$\mathcal{H}_{\rho}(x) = 1 - e^{-\frac{x}{\rho}}, \tag{6.8}$$

where $\rho > 0$ is a parameter that controls how good the approximation is (the smaller the value of ρ the better the approximation) as illustrated in Figure 6.2. Additionally, it can be easily shown that

$$\lim_{\rho \to 0} \mathcal{H}_{\rho}(x) = \mathcal{H}^{\ell}(x). \tag{6.9}$$

6.3.1.2 Product of step functions

In practice, it is also possible to encounter products of step functions as happens with the startup power consumption in $\mathcal{C}^{3}_{tn}$ (see (6.5)). For illustrative reasons, we first consider a single product of step functions, i.e., $\mathcal{H}^{\ell}(x_1)\,\mathcal{H}^{\ell}(x_2)$, and later, in Lemma 6.1, we present a smooth approximation of higher order products. We approximate the product of step functions $\mathcal{H}^{\ell}(x_1)\,\mathcal{H}^{\ell}(x_2)$, $x_1, x_2 \in \mathbb{R}_{+}$, with the function $\mathcal{H}_{\rho} : \mathbb{R}^{2}_{+} \to [0, 1]$ with

$$\mathcal{H}_{\rho}(x_1, x_2) \;=\; \mathcal{H}_{\rho}(x_1)\mathcal{H}_{\rho}(x_2) = \underbrace{1 + e^{-\frac{x_1 - x_2}{\rho}}}_{Convex} \underbrace{- e^{-\frac{x_1}{\rho}} - e^{-\frac{x_2}{\rho}}}_{Concave}. \tag{6.10}$$

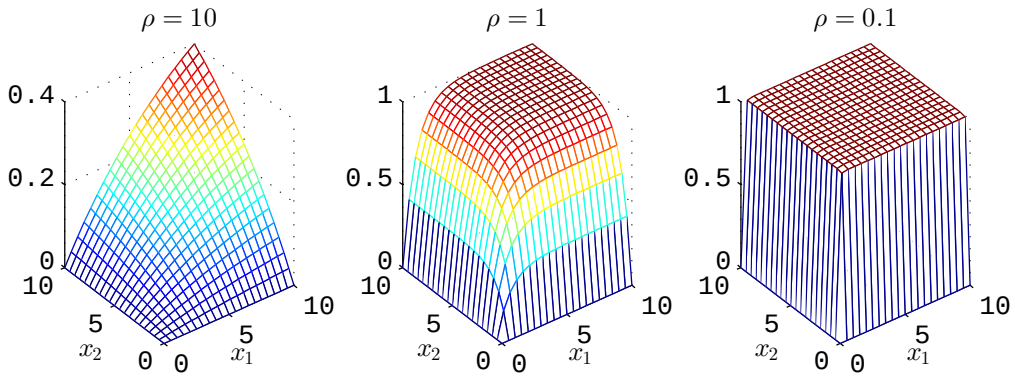

Figure 6.3: Representation of the smooth approximation of $\mathcal{H}^{\ell}(x_1)\,\mathcal{H}^{\ell}(x_2)$, i.e., $\mathcal{H}_{\rho}(x_1, x_2)$ in (6.10) for different values of the approximation control parameter ρ.

This approximation is depicted in Figure 6.3, where it is observed that the approximation improves when the control parameter ρ is reduced.[2]

Lemma 6.1. *Given a set of variables $x_q \in \mathbb{R}_+$, $q = 1, \ldots, Q$, and $\mathbf{x} = [x_1, \ldots, x_Q]^{\mathsf{T}}$, then the product of Q step functions, $\prod_{q=1}^{Q} \mathcal{H}^{\ell}(x_q)$, can be approximated by the differentiable function $\mathcal{H}_{\rho} : \mathbb{R}_+^Q \to [0, 1]$, with*

$$\mathcal{H}_{\rho}(\mathbf{x}) = \prod_{q=1}^{Q} \mathcal{H}_{\rho}(x_q) = \underbrace{1 + \sum_{i \in \mathcal{E}} \sum_{0<j_1 \cdots < j_i \leq Q} e^{-\frac{\sum_{k=1}^{i} x_{j_k}}{\rho}}}_{Convex} \underbrace{- \sum_{i \in \mathcal{O}} \sum_{0<j_1 \cdots < j_i \leq Q} e^{-\frac{\sum_{k=1}^{i} x_{j_k}}{\rho}}}_{Concave}, \quad (6.11)$$

where $\mathcal{E}$ and $\mathcal{O}$ are a partition of the set $\{1, \ldots, Q\}$ that take the even and odd elements, respectively; and $\rho > 0$ is the parameter that controls the approximation. Additionally,

$$\lim_{\rho \to 0} \mathcal{H}_{\rho}(\mathbf{x}) = \prod_{q=1}^{Q} \mathcal{H}^{\ell}(x_q)\,. \quad (6.12)$$

Proof: Note that the function $\mathcal{H}_{\rho}(\mathbf{x})$ in (6.11) is obtained by developing the product $\prod_{q=1}^{Q} \mathcal{H}_{\rho}(x_q)$. Since $e^{-x/\rho}$ is differentiable so it is $\mathcal{H}_{\rho}(\mathbf{x})$. The concavity and convexity of the different terms follows by noting that $e^{-x/\rho}$ is a convex function. Finally, $\lim_{\rho \to 0} \mathcal{H}_{\rho}(\mathbf{x}) = \lim_{\rho \to 0} \prod_{q=1}^{Q} \mathcal{H}_{\rho}(x_q) = \prod_{q=1}^{Q} \mathcal{H}^{\ell}(x_q)$. ■

Note that in the second and fourth sums, the values of j_i take all the possible combinations of i elements without repetition from the set $\{1, \ldots, Q\}$; accordingly, each of these sums con-

[2]Note that, for compactness in the notation, we use $\mathcal{H}_{\rho}$ to denote both the smooth approximation of the single step and the product of step functions. Throughout the chapter, we differentiate between them by the dimension of the argument.

tains $\frac{Q!}{i!(Q-i!)}$ summands. Observe that if we particularize the approximation in Lemma 6.1 to $Q = 1$ and $Q = 2$, then we recover the approximations of the single step function and the first order product given in (6.8) and (6.10), respectively.

6.3.2 Convex approximation of the smooth step function

In this section, we derive a convex approximation of the smooth step function, $\mathcal{H}_\rho(\boldsymbol{\phi}(\mathbf{x}))$, around the point $\mathbf{y}$ for a given transformation $\boldsymbol{\phi}$, whose component functions are all concave, denoted as $\breve{\mathcal{H}}_\rho(\mathbf{x}; \mathbf{y}, \boldsymbol{\phi})$, that satisfies the SCA requirements from [67, Assumption 3], which for completeness are given in Appendix 6.A.1.

6.3.2.1 Convex approximation of the single step function

We first consider the convex approximation of the single step function in (6.8), $\mathcal{H}_\rho(\phi(\mathbf{x}))$, which, from the rules of function composition, is a concave function [63]. Thanks to this concavity, it is easy to show that its linearization at the point $\mathbf{y}$, i.e., $\breve{\mathcal{H}}_\rho(\mathbf{x}; \mathbf{y}, \phi) = 1 + \xi_\rho(\mathbf{x}; \mathbf{y}, \phi)$, is a convex function that satisfies the requirements in [67, Assumption 3], where we have defined $\xi_\rho(\mathbf{x}; \mathbf{y}, \phi)$ as the linearization of $-e^{\frac{-\phi(\mathbf{x})}{\rho}}$ around the point $\mathbf{y}$, i.e.,

$$\xi_\rho(\mathbf{x}; \mathbf{y}, \phi) \triangleq \frac{\nabla_{\mathbf{x}}\phi(\mathbf{y})}{\rho} e^{-\frac{\phi(\mathbf{y})}{\rho}} (\mathbf{x} - \mathbf{y}) - e^{-\frac{\phi(\mathbf{y})}{\rho}}. \tag{6.13}$$

6.3.2.2 Convex approximation of products of step functions

Similarly, by linearizing the concave terms of the smooth product of step functions, we can obtain a convex approximation that satisfies the requirements in [67, Assumption 3].

Lemma 6.2. *Given the smooth approximation of the product of step functions* $\mathcal{H}_\rho(\boldsymbol{\phi}(\mathbf{x})) \approx \prod_{q=1}^{Q} \mathcal{H}^\ell(\phi_q(\mathbf{x}))$ *with* $\boldsymbol{\phi} \triangleq [\phi_1, \ldots, \phi_Q]^\mathsf{T}$*, where* ϕ_i *is a concave, Lipschitz continuous, and continuously differentiable function,* $\forall i = 1, \ldots, Q$*, then the function*

$$\breve{\mathcal{H}}_\rho(\mathbf{x}; \mathbf{y}, \boldsymbol{\phi}) = \underbrace{1 + \sum_{i \in \mathcal{E}} \sum_{0 < j_1 \cdots < j_i \leq Q} e^{-\frac{\sum_{k=1}^{i} \phi_{j_k}(\mathbf{x})}{\rho}}}_{Convex} + \underbrace{\sum_{i \in \mathcal{O}} \sum_{0 < j_1 \cdots < j_i \leq Q} \xi_\rho\left(\mathbf{x}; \mathbf{y}, \sum_{k=1}^{i} \phi_{j_k}\right)}_{Linear} \tag{6.14}$$

is a convex approximation of $\mathcal{H}_\rho(\boldsymbol{\phi}(\mathbf{x}))$ *around the point* $\mathbf{y}$ *that satisfies the required conditions in [67, Assumption 3], where* ξ_ρ *is given in* (6.13)*.*

Proof: See Appendix 6.A.1. ■

Remark 6.1. Using smooth approximations of step functions to deal with their discontinuity is not a new concept in the literature; it has been used before to approximate the cardinality operator (or ℓ^0 norm), e.g., see [123] and [124]. However, to the best of our knowledge this is the first work that derives an approximation for the product of step functions that takes into account the SCA requirements in [67, Assumption 3]. Note that the smooth approximation is not unique, e.g., one could use a logarithmic approximation. The key properties of the proposed approximation are: (i) it easily generalizes to products of step functions (products of logarithms are no longer concave and the approximation would be difficult to convexify); and (ii) the smooth approximation is accurate for not so small values of ρ (logarithmic approximations need to reduce much more the parameter ρ to have an accurate approximation, which leads to numerical problems due to the finite precision of the solvers).

These approximations are used by the ISCA, which is presented in the following section.

6.4 The Iterative Smooth and Convex approximation Algorithm (ISCA)

One natural approach to deal with the problem in (6.6) is to approximate the step functions in the ECCs with the approximation derived in Lemma 6.1 and then apply existing smooth nonconvex optimization algorithms (e.g., the algorithm in [67]) to find a locally optimal solution of the resulting problem. However, there exists a tradeoff in the selection of the approximation control parameter ρ. If we select a very small value of ρ, i.e., $\rho \to 0$, the approximation of the step functions is really accurate, but the nonconvex optimization problem has many local maxima and the SCA algorithm is highly influenced by the provided initial point. In opposition, if we set ρ to a large value, the nonconvex optimization algorithm is not affected that much by the initial point (because the approximation of the step functions is smoother), but then the resulting problem does not sufficiently resemble the original problem in (6.6).

In this context, we propose the ISCA that is composed of two loops as shown in Figure 6.4. The outer loop indexed by ς, performs a SSA of (6.6), deriving, at each iteration, a nonconvex smooth problem $(\tilde{P}^{\varsigma})$. This problem is obtained by using the approximation of the step functions derived in Lemma 6.1 given the approximation control parameter ρ^{ς}, which is initially set

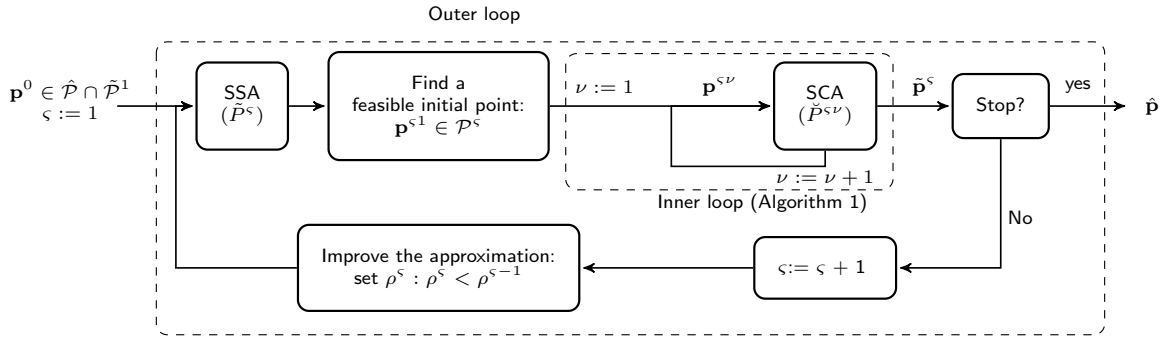

Figure 6.4: Block diagram of the ISCA.

to a large value $\rho^1 >> 0$ so that the ISCA is not influenced by the initial point, $\mathbf{p}^0$. Then, given a certain smooth problem $(\tilde{P}^\varsigma)$, the inner loop, which is based on SCA [67], determines a locally optimal solution. To do so, the inner loop, which is indexed by ν, sequentially convexifies $(\tilde{P}^\varsigma)$ around the current iterate, $\mathbf{p}^{\varsigma\nu}$, and solves the resulting strongly convex problem, $(\breve{P}^{\varsigma\nu})$, to determine the best update direction. When the inner loop converges to a local optimum of $(\tilde{P}^\varsigma)$, denoted as $\tilde{\mathbf{p}}^\varsigma$, a termination condition is checked: if it is satisfied, which implies that the approximation of the step functions is tight enough, the ISCA concludes that a good solution of the original problem (6.6) is $\hat{\mathbf{p}} = \tilde{\mathbf{p}}^\varsigma$; otherwise, a new outer iteration starts by reducing the approximation control parameter, which tightens the approximation of the step functions. Then, a feasible initial point for the inner loop is obtained that resembles the locally optimal solution of the previous outer iteration. Hence, by doing this double loop procedure, we hope to escape from locally optimal solutions that achieve a sum-rate that is far from the optimal one (in a similar way to graduated nonconvexity does [125]).

Given the smooth approximation of the step functions in Lemma 6.1, we can now approximate the original nonsmooth problem (6.6) by a smooth one, $(\tilde{P}^\varsigma)$, that, in spite of being nonconvex, has an inherently high degree of convexity, which is exploited by the inner loop to efficiently find its locally optimal solution. Then, the smooth problem that has to be solved at the ς-th outer iteration is:

$$(\tilde{P}^\varsigma): \quad \underset{\mathbf{p}}{\text{maximize}} \quad \sum_{t=1}^{\mathcal{T}} r_t(\mathbf{p}_t, \mathbf{p}_{-t}) \tag{6.15a}$$

$$\left.\begin{aligned} \text{subject to} \quad & \boldsymbol{\mathcal{B}}_t^\varsigma(\mathbf{p}_t) \succeq \mathbf{0}_{J_t}, \quad \forall t = 1, \ldots, \mathcal{T} \\ & \mathbf{p}_t \in \mathcal{P}_t, \qquad\quad \forall t = 1, \ldots, \mathcal{T} \end{aligned}\right\} \triangleq \tilde{\mathcal{P}}^\varsigma \tag{6.15b, 6.15c}$$

where

$$[\boldsymbol{\mathcal{B}}_t^{\varsigma}(\mathbf{p}_t)]_\ell \triangleq \sum_{j=1}^{\ell} E_{tj} - T_s \sum_{j=1}^{\ell} \sum_{n\in\tau_{tj}} \sum_{k=1}^{K} p_t(k,n) + \sum_{s=1}^{S_{tn}} w_{tns} \mathcal{H}_{\rho^{\varsigma}}(\boldsymbol{\phi}_{tns}(\mathbf{p}_t)), \quad \ell = 1,\ldots,J_t; \tag{6.16}$$

$\mathcal{H}_\rho$ is given in (6.11); and $\boldsymbol{\phi}_{tns}(\mathbf{p}_t)$ is a vector function defined as

$$\boldsymbol{\phi}_{tns}(\mathbf{p}_t) = [\phi_{tns1}(\mathbf{p}_t), \ldots, \phi_{tnsQ_{tns}}(\mathbf{p}_t)]^{\mathsf{T}}.$$

Accordingly, we have obtained a smooth nonconvex problem $(\tilde{P}^{\varsigma})$ in which all the terms associated with the step functions are either convex or concave. This simplifies the SCA inner loop, which is presented in the following section.

6.4.1 The inner loop: Nonconvex optimization of smooth problems with SCA

Among the algorithms that converge to stationary solutions of smooth nonconvex problems (e.g., gradient-based descend schemes [64], SCA algorithms [65–67], feasible sequential quadratic programming [68], parallel variable distribution [69], etc.), we have selected the algorithm in [67] for the inner loop because it has the following main advantages:

- It accepts nonconvex constraints.
- It exploits any degree of convexity present in the problem, which results in a much faster convergence.
- The problem can be generally solved in a distributed way under very mild assumptions.
- It includes as special cases SCA-based algorithms, such as (proximal) gradient or Newton type method, block coordinate (parallel) descent schemes and difference of convex functions methods.

The algorithm proposed in [67] is based on SCA and consists on solving a sequence of strongly convex inner approximations of the nonconvex smooth problem. Under some structural assumptions, the algorithm converges to a locally optimal solution. These assumptions enforce a specific structure of: (i) the original nonconvex smooth problem [67, Assumption 1];

(ii) the convex approximation of the objective function [67, Assumption 2]; and (iii) the convex approximation of the constraints [67, Assumption 3]. In order to use the algorithm in [67] in our inner loop, we need to satisfy these structural requirements.

It can be easily shown that the smooth problem in (6.15) satisfies (i).

Since the objective function is nonconvex, we need to derive a proper convex approximation. To do so, we exploit the "partial" concavity of the rate of a certain user, $r_t(\mathbf{p}_t, \mathbf{p}_{-t})$, w.r.t. its own transmission power $\mathbf{p}_t$. Hence, we approximate the objective function in (6.15a) around the current iterate $\mathbf{p}^{\varsigma\nu} = (((p_t^{\varsigma\nu}(k,n))_{k=1}^K)_{n=1}^N)_{t=1}^{\mathcal{T}}$ as

$$\sum_{t=1}^{\mathcal{T}} \breve{r}_t(\mathbf{p}_t, \mathbf{p}^{\varsigma\nu}), \quad \text{where} \quad \breve{r}_t(\mathbf{p}_t, \mathbf{p}^{\varsigma\nu}) = r_t(\mathbf{p}_t, \mathbf{p}_{-t}^{\varsigma\nu}) + \boldsymbol{\pi}_t^{\varsigma\nu\mathsf{T}}(\mathbf{p}_t - \mathbf{p}_t^{\varsigma\nu}) - \frac{b_t}{2}||\mathbf{p}_t - \mathbf{p}_t^{\varsigma\nu}||^2.$$

The term $\boldsymbol{\pi}_t^{\varsigma\nu}$ linearizes the rate functions of the users $t' \neq t$ w.r.t. $\mathbf{p}_t$, i.e.,

$$\boldsymbol{\pi}_t^{\varsigma\nu} \triangleq ((\pi_{tkn}^{\varsigma\nu})_{k=1}^K)_{n=1}^N = \sum_{t' \neq t} \nabla_{\mathbf{p}_t} r_{t'}(\mathbf{p}_{t'}, \mathbf{p}_{-t'})\Big|_{\mathbf{p}^{\varsigma\nu}} \text{with}$$

$$\pi_{tkn}^{\varsigma\nu} = \sum_{t' \neq t} \frac{-\text{snr}_{t'}^{\varsigma\nu}(k,n) h_{tt'}(k,n)}{\text{mui}_{t'}^{\varsigma\nu}(k,n)(1 + \text{snr}_{t'}^{\varsigma\nu}(k,n))}; \tag{6.17}$$

$\text{snr}_t^{\varsigma\nu}(k,n) \triangleq \frac{h_{tt}(k,n) p_t^{\varsigma\nu}(k,n)}{\text{mui}_t^{\varsigma\nu}(k,n)}$ and $\text{mui}_t^{\varsigma\nu}(k,n) \triangleq \sigma_t^2(k) + \sum_{t' \neq t} p_{t'}^{\varsigma\nu}(k,n) h_{t't}(k,n)$ are the SINR and the multiuser interference-plus-noise power experienced by user t given the power profile $\mathbf{p}^{\varsigma\nu}$. The term $\frac{b_t}{2}||\mathbf{p}_t - \mathbf{p}_t^{\varsigma\nu}||^2$ with $b_t \geq 0$ is a proximal regularization term that relaxes the convergence conditions of the inner loop algorithm or enhances the convergence speed [66].

Finally, we use the convex approximation of the step functions derived in Lemma 6.2 to approximate the constraints of the smooth problem.

Accordingly, the strongly convex problem that has to be solved in the ν-th inner loop iteration, which approximates the smooth problem $(\tilde{P}^{\varsigma})$ around the current iterate, $\mathbf{p}^{\varsigma\nu}$, is

$$(\breve{P}^{\varsigma\nu}): \quad \underset{\mathbf{p}}{\text{maximize}} \quad \sum_{t=1}^{\mathcal{T}} \breve{r}_t(\mathbf{p}_t, \mathbf{p}^{\varsigma\nu}) \tag{6.18a}$$

$$\left.\begin{array}{lll} \text{subject to} & \boldsymbol{\mathcal{B}}_t^{\varsigma\nu}(\mathbf{p}_t; \mathbf{p}_t^{\varsigma\nu}) \succeq \mathbf{0}_{J_t}, & \forall t = 1, \ldots, \mathcal{T} \\ & \mathbf{p}_t \in \mathcal{P}_t, & \forall t = 1, \ldots, \mathcal{T} \end{array}\right\} \triangleq \breve{\mathcal{P}}^{\varsigma\nu} \tag{6.18b}$$

$$\tag{6.18c}$$

with $[\boldsymbol{\mathcal{B}}_t^{\varsigma\nu}(\mathbf{p}_t;\mathbf{p}_t^{\varsigma\nu})]_\ell \triangleq$

$$\sum_{j=1}^{\ell} E_{tj} - T_s \sum_{j=1}^{\ell} \sum_{n\in\tau_{tj}} \sum_{k=1}^{K} p_t(k,n) + \sum_{s\in\mathcal{S}_{tn}^+} w_{tns}\breve{\mathcal{H}}_{\rho^\varsigma}(\mathbf{p}_t;\mathbf{p}_t^{\varsigma\nu},\boldsymbol{\phi}_{tns}) + \sum_{s\in\mathcal{S}_{tn}^-} w_{tns}\breve{\mathcal{H}}^-_{\rho^\varsigma}(\mathbf{p}_t;\mathbf{p}_t^{\varsigma\nu},\boldsymbol{\phi}_{tns}), \tag{6.19}$$

where $\mathcal{S}_{tn}^+ = \{s \in [1, S_{tn}] : w_{tns} > 0\}$ and $\mathcal{S}_{tn}^- = \{s \in [1, S_{tn}] : w_{tns} < 0\}$; $\breve{\mathcal{H}}_\rho$ is given in Lemma 6.2; and $\breve{\mathcal{H}}_\rho^-$ is defined as $\breve{\mathcal{H}}_\rho$ but swapping the odd and even sets. Note that we have defined $\breve{\mathcal{H}}_\rho^-$ because the negative weights invert the concavity or convexity of the terms of the smooth step function.

Additionally, since the objective function and constraints of the different transmitters are decoupled, we can apply primal decomposition [126] and solve $\mathcal{T}$ parallel problems, one for each user, which leads to a distributed resource allocation strategy that requires very limited feedback as presented later. Accordingly, each transmitter must solve the following problem at each inner loop iteration:

$$\left.\begin{aligned}
(\breve{P}_t^{\varsigma\nu}): \quad \underset{\mathbf{p}_t}{\text{maximize}} \quad & \breve{r}_t(\mathbf{p}_t,\mathbf{p}^{\varsigma\nu}) && \text{(6.20a)}\\
\text{subject to} \quad & \boldsymbol{\mathcal{B}}_t^{\varsigma\nu}(\mathbf{p}_t;\mathbf{p}_t^{\varsigma\nu}) \succeq \mathbf{0}_{J_t}, && \text{(6.20b)}\\
& \mathbf{p}_t \in \mathcal{P}_t. && \text{(6.20c)}
\end{aligned}\right\} \triangleq \breve{\mathcal{P}}_t^{\varsigma\nu}$$

Since $(\breve{P}_t^{\varsigma\nu})$ is a strongly convex problem, its unique solution, $\breve{\mathbf{p}}_t^{\varsigma\nu}$, can be easily determined by classical convex optimization algorithms, e.g., interior point methods [63]. However, since the solution to $(\breve{P}_t^{\varsigma\nu})$ has to be computed at each inner loop iteration, it is key to derive (if possible) a closed form solution in order to reduce the computational complexity of the ISCA.

The SCA-based inner loop algorithm that converges to a locally optimal solution of the smooth nonconvex problem $(\tilde{P}^\varsigma)$ is given in Algorithm 6.1 [67]. Basically, the algorithm uses the unique optimal solution to $(\breve{P}_t^{\varsigma\nu})$, $\breve{\mathbf{p}}_t^{\varsigma\nu}$, to determine the initial point of the following iteration, $\mathbf{p}_t^{\varsigma(\nu+1)}$, which is computed as a convex combination of $\breve{\mathbf{p}}_t^{\varsigma\nu}$ and the previous iterate $\mathbf{p}_t^{\varsigma\nu}$.

Theorem 6.1 ([67]). *Given the smooth nonconvex problem $(\tilde{P}^\varsigma)$, suppose that one of the two following conditions holds:*

a) The step size $a^{\varsigma\nu}$ is such that $0 < \inf_\nu a^{\varsigma\nu} \leq \sup_\nu a^{\varsigma\nu} \leq a^{\max} \leq 1$ and $2c_{\breve{r}} \geq a^{\max} L_{\nabla r}$, where $c_{\breve{r}}$ is the constant of uniform strong convexity of $\sum_{t=1}^{\mathcal{T}} \breve{r}_t(\mathbf{p}_t,\mathbf{p}^{\varsigma\nu})$ and $L_{\nabla r}$ is Lipschitz

Algorithm 6.1 The inner loop: SCA of $(\tilde{P}^{\varsigma})$

Input: $\mathbf{p}^{\varsigma 1} \in \tilde{\mathcal{P}}^{\varsigma}$, $a^{\varsigma\nu} > 0$.
Initialization: Set $\nu := 1$.
Step 1: If a termination condition is satisfied: STOP.
Step 2: For every user $t \in [1, \mathcal{T}]$, find $\breve{\mathbf{p}}_t^{\varsigma\nu}$ that is the unique optimal solution of the strongly convex problem $(\breve{P}_t^{\varsigma\nu})$.
Step 3: Update the iterate: $\mathbf{p}_t^{\varsigma(\nu+1)} = \mathbf{p}_t^{\varsigma\nu} + a^{\varsigma\nu}\left(\breve{\mathbf{p}}_t^{\varsigma\nu} - \mathbf{p}_t^{\varsigma\nu}\right), \forall t$.
Step 4: $\nu := \nu + 1$ and go to Step 1;

continuity constant of $\sum_{t=1}^{\mathcal{T}} r_t(\mathbf{p}_t, \mathbf{p}_{-t})$.

b) (i) $\tilde{\mathcal{P}}^{\varsigma}$ *is compact; (ii)* $\breve{\mathbf{p}}^{\varsigma\nu}$ *is regular for every possible initial point* $\mathbf{p}^{\varsigma 1} \in \tilde{\mathcal{P}}^{\varsigma}$*; and (iii) the step size* $a^{\varsigma\nu}$ *is such that* $a^{\varsigma\nu} \in (0, 1]$, $a^{\varsigma\nu} \to 0$, *and* $\sum_{\nu} a^{\varsigma\nu} = +\infty$.

Then every regular limit point of $\{\breve{\mathbf{p}}^{\varsigma\nu}\}_{\nu=1}^{\infty}$ *is a stationary solution of* $(\tilde{P}^{\varsigma})$. *Furthermore, none of such points is a local maximum.*

6.4.2 Determining a feasible initial point for the inner loop.

The inner loop in Algorithm 6.1 requires a feasible initial point, i.e., $\mathbf{p}^{\varsigma 1} \in \tilde{\mathcal{P}}^{\varsigma}$, in order to determine the solution of $(\tilde{P}^{\varsigma})$. Ideally, we would like to use as initial point the solution to $(\tilde{P}^{\varsigma-1})$, denoted as $\tilde{\mathbf{p}}^{\varsigma-1}$; however, in most of the cases this point is not feasible, i.e., $\tilde{\mathbf{p}}^{\varsigma-1} \notin \tilde{\mathcal{P}}^{\varsigma}$, due to the reduction of the approximation control parameter, which is formally shown later in Lemma 6.3. Then, we could use as initial point the projection of $\tilde{\mathbf{p}}^{\varsigma-1}$ to the feasible set $\tilde{\mathcal{P}}^{\varsigma}$. However, since $\tilde{\mathcal{P}}^{\varsigma}$ is nonconvex, computing the projection would suppose solving a nonconvex problem, which is not practical because we need something simple and fast. There are many heuristic approaches to find the initial feasible point. The simplest and most general option, which in practice works well, is to move from $\tilde{\mathbf{p}}^{\varsigma-1}$ towards the ISCA initial point, $\mathbf{p}^0$, which is required to belong to $\hat{\mathcal{P}} \cap \tilde{\mathcal{P}}^1$. It can be easily shown that if $\mathbf{p}^0 \in \hat{\mathcal{P}} \cap \tilde{\mathcal{P}}^1$, then there exists a step length, s^{ς}, such that an initial feasible point is obtained, i.e., $\mathbf{p}^{\varsigma 1} \triangleq \tilde{\mathbf{p}}^{\varsigma-1} + s^{\varsigma}(\mathbf{p}^0 - \tilde{\mathbf{p}}^{\varsigma-1}) \in \tilde{\mathcal{P}}^{\varsigma}$. Observe that given the power consumption models $\mathcal{C}_t^2$ and $\mathcal{C}_{tn}^3$ in (6.4) and (6.5), respectively, we can select $\mathbf{p}^0 = \mathbf{0}$.

Depending on the specific power consumption model one may find better ways to obtain the initial feasible point for the inner loop. For example, given the power consumption model $\mathcal{C}_t^2$, it can be shown that the steepest ascend direction of a given ECC is an ascend direction of the remaining ECCs. Hence, for such a power consumption model, we could find the feasible initial point by successively moving in the steepest ascend direction of the latest unfulfilled

ECC, i.e., $\nabla_{\mathbf{p}_t}[\boldsymbol{\mathcal{B}}_t^{\varsigma}(\tilde{\mathbf{p}}_t^{\varsigma-1})]_{j_t^{\varsigma}}$, with $j_t^{\varsigma} = \max_j\{[\boldsymbol{\mathcal{B}}_t^{\varsigma}(\tilde{\mathbf{p}}_t^{\varsigma-1})]_j < 0, j = 1,\ldots,J_t\}$, until the j_t^{ς} constraint is satisfied. This process must be repeated until all the constraints are fulfilled.

6.4.3 Convergence of the feasible sets and distributed implementation

Now, the details of all the building blocks of the ISCA have been introduced. The following lemma characterizes the relations between the feasible sets of the outer loop problems $(\tilde{P}^{\varsigma})$, $\tilde{\mathcal{P}}^{\varsigma}$, w.r.t. the feasible set of the original problem, $\hat{\mathcal{P}}$.[3]

Lemma 6.3. *(a) The sequence of feasible sets of the smooth problems $\{\tilde{\mathcal{P}}^{\varsigma}\}_{\varsigma=1}^{\infty}$ converges to $\hat{\mathcal{P}}$, i.e.,*

$$\lim_{\varsigma\to\infty} \tilde{\mathcal{P}}^{\varsigma} \to \hat{\mathcal{P}}. \tag{6.21}$$

(b) If the step function weights are all positive, $w_{tns} > 0$, $\forall t, s$, then we have that $\hat{\mathcal{P}} \subseteq \tilde{\mathcal{P}}^{\infty} \subset \cdots \subset \tilde{\mathcal{P}}^{\varsigma+1} \subset \tilde{\mathcal{P}}^{\varsigma} \subset \cdots \subset \tilde{\mathcal{P}}^{1}$.

(c) If the weights are all negative, $w_{tns} < 0$, $\forall t, s$, then $\tilde{\mathcal{P}}^{1} \subset \cdots \subset \tilde{\mathcal{P}}^{\varsigma} \subset \tilde{\mathcal{P}}^{\varsigma+1} \subset \ldots \tilde{\mathcal{P}}^{\infty} \subseteq \hat{\mathcal{P}}$.

Proof: See Appendix 6.A.2. ■

Note that when a step function has a positive weight, $w_{tns} > 0$, it is tightening the ECCs w.r.t. solely considering the radiated power in the ECCs. Conversely, the ECCs are loosed by negative weighted step functions. As shown in Table 6.1, the power consumption model $\mathcal{C}_t^2$ in (6.4) has all the weights positive; whereas, $\mathcal{C}_{tn}^3$ in (6.5) has positive and negative weights. Observe that, it would be estrange that a power consumption model had all the weights negative since it would imply that all the step functions reduce the power consumption.

In the near future we want to use the result on the convergence of the feasible sets in Lemma 6.3 in order to prove that the sequence of stationary solutions of the smooth problem, $(\tilde{P}^{\varsigma})$, obtained by Algorithm 6.1, converges to a stationary solution of the original nonconvex and nonsmooth problem $(\hat{P})$. We believe that by using the tools of variational analysis [127] we might be able to obtain the desired proof; however, further research is required.

As argued in §2.1.3, the solution to the offline maximization problem is specially meaningful when the channel and energy harvesting processes can be predicted within the optimization

[3] Lemma 6.3 uses the limit definition of a sequence of sets, which is given in [127, Definition 4.1]. Basically, the limit exists if the outer and inner limits (see definitions in [127, Definition 4.1]) are equal. In such a case, the limit is the set given by the inner and outer limits.

time horizon. When this happens, it is key that the power allocation strategy can be computed distributedly by the network nodes because then the nodes can adapt their transmission strategies when substantial changes on the energy harvesting process are observed.

In order to compute the solution in a distributed manner, the following signaling is required so that the remaining transmitters can compute the weights $\pi_{tkn}^{\varsigma\nu}$: (i) at each outer loop iteration, each transmitter has to broadcast the feasible initial point of the inner loop; and (ii) at each inner loop iteration, each transmitter t solely has to broadcast $\nabla_{\mathbf{p}_{-t}} r_t(\mathbf{p}_t, \mathbf{p}_{-t}^{\varsigma\nu})$, which can be computed with the local measurements of the SINR and the MUI.

Note that the fulfillment of the ECCs is not guaranteed until convergence of the ISCA. Accordingly, it is required that each of the nodes has a backup battery to be used for this transitory regime, which can be recharged with the harvested energy.

Remark 6.2. Note that if a different rate function applies, the ISCA can be used by deriving a proper convex approximation of the objective function.

In the following section, we focus on the power consumption model $\mathcal{C}_t^2$ and derive a closed form solution of the inner loop problem.

6.5 The ISCA algorithm for $\mathcal{C}_t^2$.

In this section, we use the ISCA to distributedly compute a locally optimal power allocation given the power consumption model $\mathcal{C}_t^2$ in (6.4), which is the power consumption model that we used in Chapter 5.

As it has been mentioned in §6.4, at each inner loop iteration, the t-th transmitter must solve (6.20) to obtain the update direction. Given the power consumption model $\mathcal{C}_t^2$, we have that

$$[\boldsymbol{\mathcal{B}}_t^{\varsigma\nu}(\mathbf{p}_t; \mathbf{p}_t^{\varsigma\nu})]_\ell = \sum_{j=1}^{\ell} \left[E_{tj} - \sum_{n\in\tau_j} \left(\varepsilon_t^{\varsigma\nu}(n) + \sum_{k=1}^{K} \varphi_t^{\varsigma\nu}(k,n) p_t(k,n) \right) \right], \quad \text{with} \tag{6.22}$$

$$\varphi_t^{\varsigma\nu}(k,n) = T_s \Big(1 + \frac{\alpha_t}{\rho^\varsigma} e^{\frac{-\sum_{k=1}^{K} p_t^{\varsigma\nu}(k,n)}{\rho^\varsigma}} + \frac{\beta_t}{\rho^\varsigma} e^{\frac{-p_t^{\varsigma\nu}(k,n)}{\rho^\varsigma}} \Big) \quad \text{and} \tag{6.23}$$

$$\varepsilon_t^{\varsigma\nu}(n) = T_s\alpha_t\left(1-\left(1+\frac{\sum_{k=1}^{K}p_t^{\varsigma\nu}(k,n)}{\rho^{\varsigma}}\right)e^{\frac{-1}{\rho^{\varsigma}}\sum_{k=1}^{K}p_t^{\varsigma\nu}(k,n)}\right)+ \tag{6.24}$$

$$T_s\beta_t\sum_{k=1}^{K}\left(1-\left(1+\frac{p_t^{\varsigma\nu}(k,n)}{\rho^{\varsigma}}\right)e^{\frac{-1}{\rho^{\varsigma}}p_t^{\varsigma\nu}(k,n)}\right), \tag{6.25}$$

where the constants $\varphi_t^{\varsigma\nu}(k,n)$ and $\varepsilon_t^{\varsigma\nu}(n)$ are obtained after linearizing the positive step functions at the current iterate, $\mathbf{p}_t^{\varsigma\nu}$.

Lemma 6.4. *Given the power consumption model $\mathcal{C}_t^2$, the optimal solution to* (6.20) *is obtained in closed form as*

$$\breve{p}_t^{\varsigma\nu}(k,n)=\left[\ \frac{1}{2}\left(p_t^{\varsigma\nu}(k,n)-\frac{\mathit{mui}_t^{\varsigma\nu}(k,n)}{h_{tt}(k,n)}\right)-\frac{1}{2b_t}\left(\frac{1}{W_t^{\varsigma\nu\star}(k,n)}-\right.\right. \tag{6.26}$$

$$\left.\left.\sqrt{\left[\frac{1}{W_t^{\varsigma\nu\star}(k,n)}-b_t\left(p_t^{\varsigma\nu}(k,n)+\frac{\mathit{mui}_t^{\varsigma\nu}(k,n)}{h_{tt}(k,n)}\right)\right]^2+4b_t}\right)\right]_0^{p_t^{\max}(k,n)}, \qquad n\in\tau_j,$$

where

$$W_t^{\varsigma\nu\star}(k,n)=\frac{1}{-\pi_{tkn}^{\varsigma\nu}+\bar{\lambda}_{tj}^{\varsigma\nu\star}\varphi_t^{\varsigma\nu}(k,n)}, \tag{6.27}$$

$\bar{\lambda}_{tj}^{\varsigma\nu\star}=\sum_{\ell=j}^{J_t}\lambda_{t\ell}^{\varsigma\nu\star}$ *with* $\{\lambda_{t\ell}^{\varsigma\nu\star}\}_{\ell=1}^{J_t}$ *being the optimal Lagrange multipliers associated to the ECCs in* (6.18b)*, which can be efficiently obtained similarly to the FSA algorithm in §4.5.2. Additionally, if we do not include the proximal regularization term, i.e.,* $b_t=0$*, we obtain the following iterative directional water-filling like solution:*

$$\breve{p}_t^{\varsigma\nu}(k,n)=\left[W_t^{\varsigma\nu\star}(k,n)-\frac{\mathit{mui}_t^{\varsigma\nu}(k,n)}{h_{tt}(k,n)}\right]_0^{p_t^{\max}(k,n)}. \tag{6.28}$$

Proof: See Appendix 6.A.3. ■

From the expression in (6.28), we can get some intuition on the solution. First, if the water level, $W_t^{\varsigma\nu\star}(k,n)$, is smaller than $\frac{\mathrm{mui}_t^{\varsigma\nu}(k,n)}{h_{tt}(k,n)}$, then it is preferable to turn off the (k,n)-th subchannel. Second, the water level, which is given in (6.27), decreases with the interference produced to other users, which is quantified in the term $-\pi_{tkn}^{\varsigma\nu}$. This implies that the users will try to reduce the interference as much as possible to increase the sum-rate. Third, the water level depends on $\varphi_t^{\varsigma\nu}(k,n)$ in (6.23), i.e., the partial derivative of the n-th slot smooth power consumption w.r.t. $p_t(k,n)$ evaluated at the current iterate $p_t^{\varsigma\nu}(k,n)$. Accordingly, if the power of a certain subchannel is small, $p_t^{\varsigma\nu}(k,n)\to 0$, the derivative of the smooth step functions

is large and the water level is penalized. In opposition, if the power of a certain subchannel is large, $p_t^{\varsigma\nu}(k,n) \to p_t^{\max}(k,n)$, the derivative of the smooth step function tends to zero and the water level is rewarded. Note that these penalizations or rewards are weak at the initial ISCA iterations, because the approximation of the step functions is smooth, but they gain in importance as the ISCA iterations go by. Finally, the water level is a function of the Lagrange multipliers that depend on the energy availability of the node in a similar way to the DWF solution [84].

It is worth mentioning that if the number of transmitter and receiver pairs in the network is set to one, $\mathcal{T} = 1$, then the ISCA computes the solution to the problem studied in Chapter 5.

Note that the ISCA can be applied to a broad class of problems. In the following remarks, we use the ISCA to derive power allocation strategies that, to the best of our knowledge, have not been yet derived in the literature.

Remark 6.3 (Transmission power only). Consider the sum-rate maximization problem of a network of energy harvesting nodes, where the unique source of energy consumption is the transmission power ($\alpha_t = 0$, $\beta_t = 0$, $\gamma_t = 0$). Then, the inner loop of the ISCA (or the algorithm in [67]) can be used to determine a distributed locally optimal solution, where the solution to the $\varsigma\nu$-th inner loop problem is

$$\breve{p}_t^{\varsigma\nu}(k,n) = \left[\frac{1}{-\pi_{tkn}^{\varsigma\nu} + \bar{\lambda}_{tj}^{\varsigma\nu\star} T_s} - \frac{\mathrm{mui}_t^{\varsigma\nu}(k,n)}{h_{tt}(k,n)}\right]_0^{p_t^{\max}(k,n)}, \quad n \in \tau_j. \tag{6.29}$$

Remark 6.4 (Power consumption model $\mathcal{C}_{tn}^3$ in (6.5)). Consider the problem of maximizing the sum-rate given the power consumption model $\mathcal{C}_{tn}^3$. Then, the ISCA determines a distributed power allocation strategy, where the inner loop problem must be solved by numerical methods since it does not accept a closed form solution. Additionally, the power allocation strategy in a point-to-point link is obtained by particularizing $\mathcal{T} = 1$, which implies that $\pi_{tkn}^{\varsigma\nu} = 0, \forall k, n$.

Remark 6.5 (Non-harvesting nodes). Finally, consider the sum-rate maximization problem of a network of non-harvesting nodes given any power consumption model of the form (6.3). Then, the ISCA determines a distributed power allocation strategy, where the strongly convex problem that has to be solved at each inner loop iteration is (6.20) given that $J_t = 1, \forall t$, which imposes a sum-power constraint.

In the following section, we evaluate the performance of the ISCA.

6.6 Results

In this section, we numerically evaluate the performance of the ISCA in terms of achieved rate and computational complexity of the algorithm.

As it has been argued in the state of the art in Chapter 2, there are few works that consider the different energy sinks at each transmitter; indeed, to the best of our knowledge, the most similar work to the problem studied in this chapter is the problem that we studied in Chapter 5, where we considered the mutual information maximization problem in a point-to-point link with the power consumption model $\mathcal{C}_t^2$.

Accordingly, in order to have some benchmark with which to compare the performance of the ISCA, we particularize the solution derived in the previous section to a point-to-point link by setting $\mathcal{T} = 1$. The remaining system parameters have been set as follows. We have considered $N = 50$ channel accesses of duration $T_s = 20$ ms in which symbols are transmitted through $K = 2$ parallel streams. The power consumption constants are set to $\alpha_t = \gamma_t = 150$ mW and $\beta_t = 10$ mW. A Rayleigh fading channel has been considered with unit mean channel power gain. The energy harvesting process is modeled as a compound Poisson process as done in [84], where the arrival instants follow a Poisson distribution with rate $\frac{1}{10}$ and the energy in the packets is drawn from a uniform distribution and normalized by the total harvested energy that varies along the x-axis of Figures 6.5-6.7. The initial point of the ISCA is set to zero, and the approximation control parameter is $\rho^{\varsigma} = 0.5\rho^{\varsigma-1}$ with $\rho^1 = 5$. We have not used the proximal regularization term, $b_t = 0$, and the step size of the inner loop is $a^{\varsigma\nu} = a^{\varsigma(\nu-1)}(1 - 10^{-3}a^{\varsigma(\nu-1)})$ with $a^{\varsigma 0} = 1, \forall\varsigma$.

We consider two classes of strategies:

- **Power consumption model $\mathcal{C}_t^2$:** On the one hand, we consider strategies that use the power consumption model $\mathcal{C}_t^2$ (i.e., strategies that disregard the startup power consumption, $\gamma_t = 0$), which can be based either on the use of IVs or on the use of the ISCA: (i) the IVs based strategies are the upper bound and feasible solution obtained through integer relaxation as it has been explained in §5.3.1, which we label here as *IV-UB ($\gamma_t = 0$)* and *IV ($\gamma_t = 0$)*; (ii) the ISCA based strategies are *ISCA-FSA ($\gamma_t = 0$)* that uses the closed form solution derived in Lemma 6.4, where the Lagrange multipliers are obtained similarly to the FSA algorithm (cf. §4.5.2), and *ISCA-BM ($\gamma_t = 0$)* that solves (6.20)

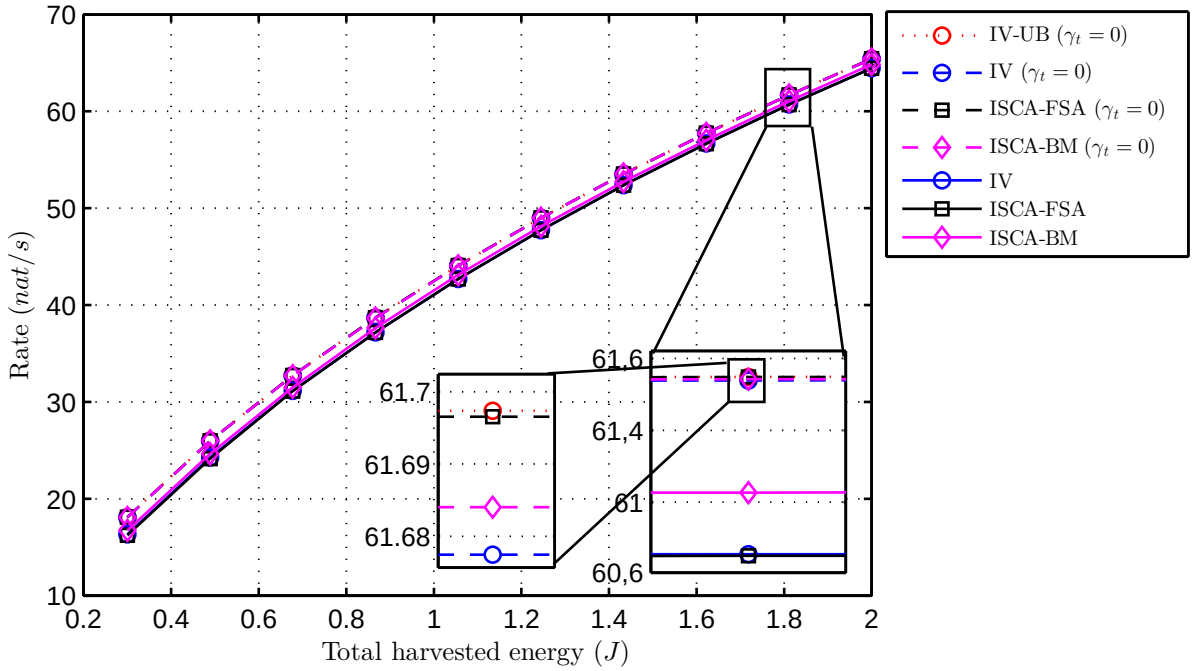

Figure 6.5: Achieved rate versus total harvested energy for the different considered strategies.

using the barrier method [63] given that $\gamma_t = 0$.

- **Power consumption model $\mathcal{C}_{tn}^3$:** On the other hand, we consider strategies that account for the startup power consumption (i.e., strategies that use the power consumption model $\mathcal{C}_{tn}^3$), *ISCA-BM*, which solves (6.20) using the barrier method, and the strategies *IV* and *ISCA-FSA*, which scale *IV* ($\gamma_t = 0$), and *ISCA-FSA* ($\gamma_t = 0$) until the ECCs with startup power consumption are satisfied.

In this setup, Figure 6.5 shows the achieved rate versus total harvested energy. First, we observe that the solution provided by the ISCA strategies is close to the global optimum since the gap with the upper bound, *IV-UB* ($\gamma_t = 0$), is small. Second, the ISCA based solutions perform slightly better than the feasible strategy *IV* ($\gamma_t = 0$). Finally, as expected, the strategies that consider the startup power consumption achieve a lower sum-rate, where the stationary solution *ISCA-BM* performs better than the other strategies that consider the startup power consumption.

Figure 6.6 shows the percentage of the total harvested energy that is expended in energy sinks different than the transmission power. The percentage of energy spent in the circuitry is much higher at low harvested energies, where the cost for turning on a subchannel is a

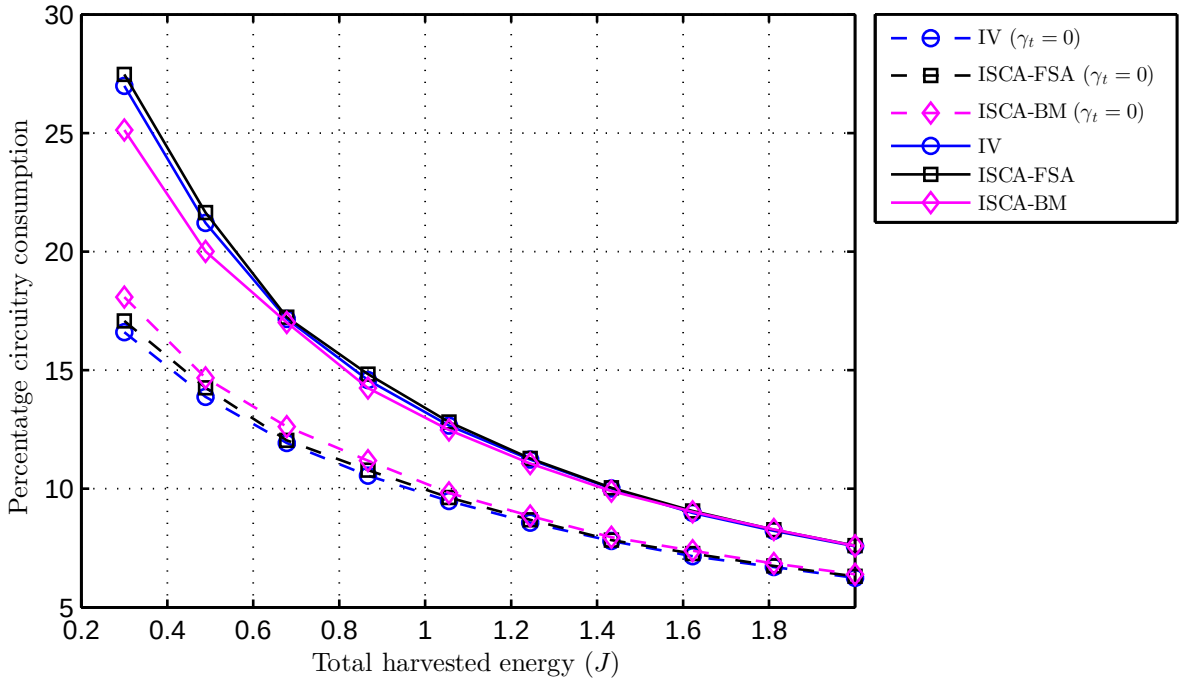

Figure 6.6: Percentage of the total harvested energy expended in the circuitry.

high fraction of the total available energy, and decreases in the high energy regime, where the transmission power in each subchannel increases. Additionally, at the high energy regime, the effect of disregarding the startup power consumption does not have a significant impact since most of the channel accesses are active and, accordingly, there are few off-on transitions.

Figure 6.7 evaluates the computational complexity of the different algorithms. It is observed that the worst performance is achieved by *ISCA-BM*, where most of the execution time is spent in the computation of the gradient and the Hessian required for the Newton method [63]; however, it also solves a more complex problem than the strategies that disregard the startup power consumption. Note that the performance of the ISCA with the barrier method improves when the startup power consumption is disregarded. Finally, it is important to mention that when a closed form solution of the inner loop problem is available, as happens with *ISCA-FSA*, the computational complexity of the ISCA is dramatically reduced outperforming the strategy based on IVs.

Finally, Figure 6.8 shows the violation of the original nonsmooth ECCs produced by the stationary solution of each smooth nonconvex problem, which is computed as the Euclidean norm of the nonfulfilled constraints. To obtain Figure 6.8, we have reduced the approximation

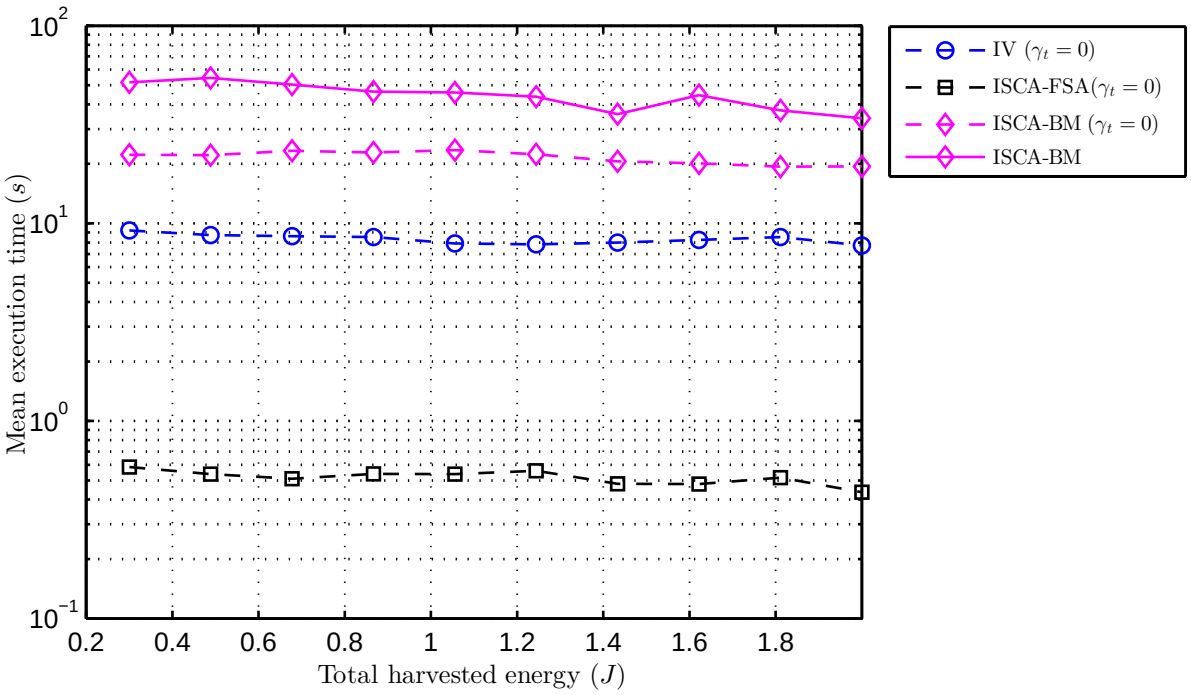

Figure 6.7: Mean execution time versus total harvested energy for the different considered strategies.

control parameter slightly slower to have more points, i.e., $\rho^{\varsigma} = 0.8\rho^{\varsigma-1}$ with $\rho^1 = 5$. As expected, when the outer loop iterations go by and the approximation control parameter is reduced, the violation of the original nonsmooth ECCs is reduced, reaching tolerable levels in EH applications.

6.7 Conclusions

In this chapter, we have studied the sum-rate maximization problem of a Gaussian interference channel composed of WEHNs by considering a general power consumption model that is composed of step functions. We have proposed the ISCA, a distributed power allocation algorithm that is based on SSA of the step functions to derive a sequence of smooth nonconvex problems that can be solved by means of SCA. By numerical simulation, we have shown that the sum-rate achieved by the ISCA is similar to the one of existing algorithms while it reduces the computational complexity when a closed form solution of the inner loop problem exists.

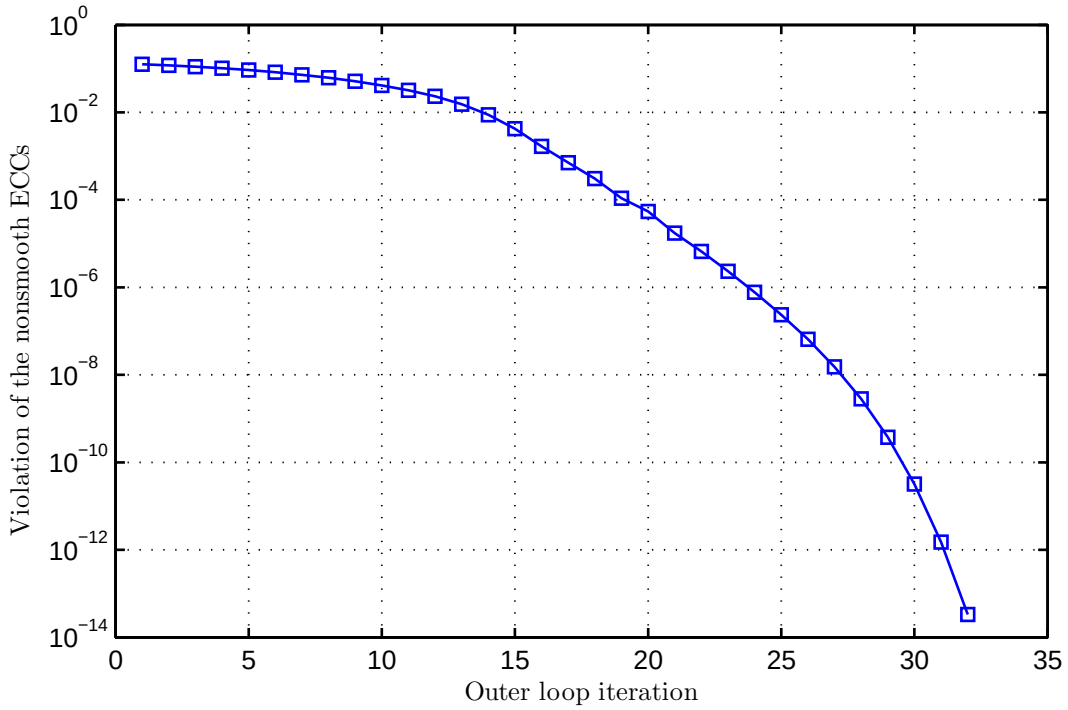

Figure 6.8: Violation of the original ECCs as a function of the ISCA outer iteration for the *ISCA-FSA ($\gamma_t = 0$)* power allocation strategy.

6.A Appendix

6.A.1 Proof of Lemma 6.2

In the following lines, we show that $\breve{\mathcal{H}}_\rho(\mathbf{x};\mathbf{y},\boldsymbol{\phi})$ satisfies all the requirements imposed in [67, Assumption 3] that are:

Assumption 1. *The function $\breve{\mathcal{H}}_\rho(\cdot;\cdot,\boldsymbol{\phi}) : \mathcal{X}\times\mathcal{Y}\to\mathbb{R}$ must satisfy [67, Assumption 3] for all $\boldsymbol{\phi}$:*

A1) $\breve{\mathcal{H}}_\rho(\cdot;\mathbf{y},\boldsymbol{\phi})$ is convex on $\mathcal{X}$ for all $\mathbf{y}\in\mathcal{Y}$;

A2) $\breve{\mathcal{H}}_\rho(\mathbf{x};\mathbf{x},\boldsymbol{\phi}) = \mathcal{H}_\rho(\boldsymbol{\phi}(\mathbf{x}))$, for all $\mathbf{x}\in\mathcal{Y}$;

A3) $\mathcal{H}_\rho(\boldsymbol{\phi}(\mathbf{x})) \leq \breve{\mathcal{H}}_\rho(\mathbf{x};\mathbf{y},\boldsymbol{\phi})$ for all $\mathbf{x}\in\mathcal{X}$ and $\mathbf{y}\in\mathcal{Y}$;

A4) $\breve{\mathcal{H}}_\rho(\cdot;\cdot,\boldsymbol{\phi})$ is Lipschitz continuous on $\mathcal{X}\times\mathcal{Y}$;

A5) $\nabla_{\mathbf{x}}\breve{\mathcal{H}}_\rho(\mathbf{y};\mathbf{y},\boldsymbol{\phi}) = \nabla_{\mathbf{x}}\mathcal{H}_\rho(\boldsymbol{\phi}(\mathbf{y}))$, for all $\mathbf{y}\in\mathcal{Y}$;

A6) $\nabla_{\mathbf{x}}\breve{\mathcal{H}}_\rho(\cdot;\cdot,\boldsymbol{\phi})$ is continuous on $\mathcal{X}\times\mathcal{Y}$;

where $\nabla_{\mathbf{x}}\breve{\mathcal{H}}_\rho(\mathbf{y};\mathbf{y},\boldsymbol{\phi})$ denotes the partial gradient of $\breve{\mathcal{H}}_\rho(\mathbf{x};\mathbf{y},\boldsymbol{\phi})$ w.r.t. $\mathbf{x}$ evaluated at $(\mathbf{y};\mathbf{y},\boldsymbol{\phi})$.

First note that since the component functions of $\boldsymbol{\phi}$ are all concave and $e^{\frac{-x}{\rho}}$ is convex and

decreasing, the function $e^{-\frac{\sum_{k=1}^{i}\phi_{j_k}(\mathbf{x})}{\rho}}$ is convex [63], which proves the convexity of the terms in (6.14). Accordingly, A1 is satisfied because $\breve{\mathcal{H}}_\rho$ is the addition of convex and affine terms. Since $\breve{\mathcal{H}}_\rho(\cdot;\mathbf{y},\boldsymbol{\phi})$ is obtained after linearizing the concave terms of $\mathcal{H}_\rho(\boldsymbol{\phi}(\cdot))$, it follows that $\breve{\mathcal{H}}_\rho(\cdot;\mathbf{y},\boldsymbol{\phi})$ is a global over estimator that has the same value and gradient at $\mathbf{y}$. Hence, conditions A2, A3 and A5 are also satisfied. Finally, since $e^{\frac{-x}{\rho}}$ is Lipschitz continuous, $\breve{\mathcal{H}}_\rho(\cdot;\cdot,\boldsymbol{\phi})$ is also Lipschitz continuous. ■

6.A.2 Proof of Lemma 6.3

Note that the difference between $\tilde{\mathcal{P}}^{\varsigma+1}$ and $\tilde{\mathcal{P}}^{\varsigma}$ is due to the reduction of the approximation control parameter ($\rho^{\varsigma+1} < \rho^{\varsigma}$) in the ECCs. It can be easily shown that: (i) $\mathcal{H}_\rho$ is strictly decreasing in ρ (for $x > 0$) and (ii) $\mathcal{H}_\rho(\mathbf{x}) \leq \prod_{q=1}^{Q}\mathcal{H}^\ell(x_q)$, $\forall \mathbf{x} \in \mathbb{R}_+^Q$, $\rho > 0$. Accordingly, when all the weights are positive, we have from (i) that the smooth ECCs are tightened when the approximation control parameter is reduced, i.e., $\tilde{\mathcal{P}}^{\varsigma+1} \subset \tilde{\mathcal{P}}^{\varsigma}, \forall\varsigma$. Additionally, from (ii), the ECCs are relaxed when using the smooth approximation, we have that $\hat{\mathcal{P}} \subseteq \tilde{\mathcal{P}}^{\varsigma}$. This proves (b), the proof of (c) follows similarly by noting that when all the weights are negative the original ECCs are tightened. In the following lines, we prove (a).

Let us write $\tilde{\mathcal{P}}^{\varsigma} = \mathcal{P}_t \cap \mathbb{B}^{\varsigma}$ with $\mathbb{B}^{\varsigma} = \{\mathbf{p}_t : \boldsymbol{\mathcal{B}}_t^{\varsigma}(\mathbf{p}_t) \succeq \mathbf{0}_{J_t}, \forall t = 1,\ldots,\mathcal{T}\}$ and $\hat{\mathcal{P}} = \mathcal{P}_t \cap \hat{\mathbb{B}}$ with $\hat{\mathbb{B}} = \{\mathbf{p}_t : \boldsymbol{\mathcal{B}}_t(\mathbf{p}_t) \succeq \mathbf{0}_{J_t}, \forall t = 1,\ldots,\mathcal{T}\}$. Accordingly, we need to prove that $\mathbb{B}^{\varsigma}$ converges to $\hat{\mathbb{B}}$ as $\varsigma \to \infty$. To prove this, we define the following two sets:

$$\begin{aligned}
\mathbb{B}_+^{\varsigma} \triangleq \Big\{\mathbf{p}_t : \sum_{j=1}^{\ell} E_{tj} \;&-\; T_s \sum_{j=1}^{\ell}\sum_{n\in\tau_{tj}}\sum_{k=1}^{K} p_t(k,n) + \sum_{s\in\mathcal{S}_{tn}^+} w_{tns}\mathcal{H}_{\rho^{\varsigma}}(\boldsymbol{\phi}_{tns}(\mathbf{p}_t)) \\
&+\; \sum_{s\in\mathcal{S}_{tn}^-} w_{tns}\prod_{q=1}^{Q_{tns}}\mathcal{H}^\ell(\phi_{tnsq}(\mathbf{p}_t)) \leq 0, \forall \ell = 1,\ldots,J_t, \forall t = 1,\ldots,\mathcal{T}\Big\}, \\
\mathbb{B}_-^{\varsigma} \triangleq \Big\{\mathbf{p}_t : \sum_{j=1}^{\ell} E_{tj} \;&-\; T_s \sum_{j=1}^{\ell}\sum_{n\in\tau_{tj}}\sum_{k=1}^{K} p_t(k,n) + \sum_{s\in\mathcal{S}_{tn}^-} w_{tns}\mathcal{H}_{\rho^{\varsigma}}(\boldsymbol{\phi}_{tns}(\mathbf{p}_t)) \\
&+\; \sum_{s\in\mathcal{S}_{tn}^+} w_{tns}\prod_{q=1}^{Q_{tns}}\mathcal{H}^\ell(\phi_{tnsq}(\mathbf{p}_t)) \leq 0, \forall \ell = 1,\ldots,J_t, \forall t = 1,\ldots,\mathcal{T}\Big\},
\end{aligned}$$

where $\mathcal{S}_{tn}^+$ and $\mathcal{S}_{tn}^-$ are defined as in (6.19). Note that $\mathbb{B}_+^{\varsigma}$ approximates the positive step functions only and $\mathbb{B}_-^{\varsigma}$, the negative ones. Similarly to the proofs of (b) and (c), it follows that

$\mathbb{B}_{+}^{\varsigma+1} \subset \mathbb{B}_{+}^{\varsigma}, \forall \varsigma$ and $\mathbb{B}_{-}^{\varsigma} \subset \mathbb{B}_{-}^{\varsigma+1}, \forall \varsigma$. Additionally, following the same arguments, we have that $\mathbb{B}_{-}^{\varsigma} \subset \mathbb{B}^{\varsigma} \subset \mathbb{B}_{+}^{\varsigma}$. From [127, Exercise 4.3], we have that the limits of the sets $\mathbb{B}_{+}^{\varsigma}$ and $\mathbb{B}_{-}^{\varsigma}$ exists (the inner and outer limits are equal) and are $\mathbb{B}_{+}^{\varsigma} \to \hat{\mathbb{B}}$ and $\mathbb{B}_{-}^{\varsigma} \to \hat{\mathbb{B}}$. This leads to $\mathbb{B}^{\varsigma} \to \hat{\mathbb{B}}$, which proves (a). ■

6.A.3 Proof of Lemma 6.4

We prove Lemma 3 by using the KKT sufficient optimality conditions. The Lagrangian of the problem in (6.20) is $\mathcal{L}_t^{\varsigma\nu}(\mathbf{p}_t, \boldsymbol{\lambda}_t^{\varsigma\nu}) = r_t(\mathbf{p}_t, \mathbf{p}_{-t}^{\varsigma\nu}) + \boldsymbol{\pi}_t^{\varsigma\nu\mathsf{T}}(\mathbf{p}_t - \mathbf{p}_t^{\varsigma\nu}) - \frac{b_t}{2}||\mathbf{p}_t - \mathbf{p}_t^{\varsigma\nu}||^2 + \boldsymbol{\lambda}_t^{\varsigma\nu\mathsf{T}}\boldsymbol{\mathcal{B}}_t^{\varsigma\nu}(\mathbf{p}_t; \mathbf{p}_t^{\varsigma\nu})$, where $\boldsymbol{\lambda}_t^{\varsigma\nu}$ are the Lagrange multipliers associated to the ECCs in (6.20b). Taking the derivative of the Lagrangian w.r.t. $p_t(k,n)$, $n \in \tau_j$, and equating to zero, we have

$$\breve{p}_t^{\varsigma\nu}(k,n) = \begin{cases} 0 & \text{if } \frac{1}{W_t^{\varsigma\nu}(k,n)} \geq \frac{h_{tt}(k,n)}{\text{mui}_t^{\varsigma\nu}(k,n)} + b_t p_t^{\varsigma\nu}(k,n), \\ \bar{p}_t^{\varsigma\nu}(k,n) & \text{if } \frac{1}{W_t^{\varsigma\nu}(k,n)} < \frac{h_{tt}(k,n)}{\text{mui}_t^{\varsigma\nu}(k,n)} + b_t p_t^{\varsigma\nu}(k,n), \quad \text{and} \\ & \quad \frac{1}{W_t^{\varsigma\nu}(k,n)} > \frac{h_{tt}(k,n)}{h_{tt}(k,n)p_t^{\max}(k,n) + \text{mui}_t^{\varsigma\nu}(k,n)} - b_t(p_t^{\max}(k,n) - p_t^{\varsigma\nu}(k,n)), \\ p_t^{\max}(k,n) & \text{if } \frac{1}{W_t^{\varsigma\nu}(k,n)} \leq \frac{h_{tt}(k,n)}{h_{tt}(k,n)p_t^{\max}(k,n) + \text{mui}_t^{\varsigma\nu}(k,n)} - b_t(p_t^{\max}(k,n) - p_t^{\varsigma\nu}(k,n)), \end{cases}$$

where $W_t^{\varsigma\nu}(k,n) = \frac{1}{-\pi_{tkn}^{\varsigma\nu} + \bar{\lambda}_{tj}^{\varsigma\nu}\varphi_t^{\varsigma\nu}(k,n)}$ with $\bar{\lambda}_{tj}^{\varsigma\nu} = \sum_{\ell=j}^{J} \lambda_{t\ell}^{\varsigma\nu}$; and $\bar{p}_t^{\varsigma\nu}(k,n)$ is obtained as the solution of the following quadratic equation

$$\frac{h_{tt}(k,n)}{h_{tt}(k,n)\bar{p}_t^{\varsigma\nu}(k,n) + \text{mui}_t^{\varsigma\nu}(k,n)} = \frac{1}{W_t^{\varsigma\nu}(k,n)} + b_t(\bar{p}_t^{\varsigma\nu}(k,n) - p_t^{\varsigma\nu}(k,n)). \tag{6.30}$$

From [128, Lemma 35] (with $H_k := \frac{h_{tt}(k,n)}{\text{mui}_t^{\varsigma\nu}(k,n)}$, $\tau := b_t$, $c_k := p_t^{\varsigma\nu}(k,n)$, $\tilde{\mu}_k := \frac{1}{W_t^{\varsigma\nu}(k,n)}$), the previous equation has the following properties: (i) both roots are real, one root is always negative, and the other is nonnegative; (ii) both roots are increasing in $W_t^{\varsigma\nu}(k,n)$; and from the expression of the nonnegative root, we obtain (6.26).

Finally, note that if the proximal step is zero, i.e., $b_t = 0$, then (6.28) follows directly from the first order equation in (6.30). ■

Chapter 7

Thesis conclusions and future work

7.1 Conclusions

This dissertation has focused on the design of transmission strategies (linear precoding designs and resource allocation strategies) that take into account the specific characteristics of WEHNs. In particular, the dissertation has put an special emphasis on taking into account the energy availability variations and the different sources of energy consumption at the transmitter(s), which are not accounted for in classical transmission strategies. As a result, the proposed transmission strategies exploit more efficiently the available energy than classical transmission strategies improving the achieved performance given some figure of merit.

In Chapter 1, we have motivated the conducted research and presented the outline and research contributions of the dissertation.

Chapter 2 has presented the structure of a WEHN and the main characteristics that must be taken into account when designing transmission strategies for WEHNs. Additionally, the chapter has overviewed the state of the art on well-known transmission strategies.

Chapter 3 has considered an energy harvesting transmitter operating in a point-to-point link with a time-static channel that has a finite battery capacity, acquires the data packets over time, and has to fulfill some QoS constraint. This chapter has studied the (offline) data departure curve that minimizes the transmission completion time by satisfying the DCC, the ECC, and the QoS constraint. The optimal transmission strategy has been constructed by deriving structural properties of the solution: first, it has been shown that constant rate transmission is the strategy that requires less energy to transmit a certain amount of data if no energy is lost due

to battery overflows; afterwards, it has been shown that losing energy due to battery overflows is suboptimal as far as there is data to be transmitted and, accordingly, it is better to increase the rate before the overflow until either there is no overflow or the DCC is reached (i.e., there is no more data to transmit); finally, it has been demonstrated that the existence of the optimal solution depends both on the dynamics of the harvesting process and on the required QoS. According to these properties, an algorithm has been proposed that is able to determine whether the problem has a solution or not and, in case of having a solution, determines the optimal data departure curve.

Chapter 4 has been devoted to a WEHN transmitting arbitrarily distributed symbols in a discrete-time linear vector Gaussian channel. The chapter has investigated the offline linear precoding strategy that maximizes the mutual information by taking into account the causality constraints on the use of energy when there is an infinite backlog of data waiting to be transmitted. The chapter has proved that the optimal left singular vectors of the precoder matrix diagonalize the channel and argued that the derivation of the optimal right singular vectors is an involved problem that is left for future research. Then, the chapter has focused in the practical situation in which the right singular vectors are set to the identity matrix; in this setup, the optimal offline power allocation, the *MIMO Mercury Water-Flowing* solution, has been derived. The chapter has presented two different algorithms that compute the offline *MIMO Mercury Water-Flowing* solution whose computational complexity has been evaluated both analytically and experimentally; additionally, an online algorithm has been proposed. Finally, the performance of the different algorithms has been compared through numerical simulations showing a substantial increase in the achieved mutual information.

Chapter 5 has studied the power allocation strategy that maximizes the mutual information achieved by a WEHN along N independent channel accesses in which symbols are sent through K parallel streams. The main contribution of Chapter 5 w.r.t. previous works like DWF is that, not only the transmission radiated power has been accounted for in the ECCs, but also the cost associated with having a certain channel access or stream "on". The offline maximization problem has been shown to be nonsmooth and nonconvex due to the presence of step functions on the power consumption model. After replacing the step functions by additional optimization variables, named IVs, the chapter has explored the integer relaxation and dual problems to obtain two suboptimal solutions and an upper bound of the original offline maximization problem. Interestingly, we have shown that these feasible solutions are practi-

cally identical in the sense that a common condition for the activation of the channel access and streams is obtained, i.e., if the epoch water level is greater than the corresponding cutoff water level, and asymptotically tend to the optimal solution when the number of channel accesses or streams per epoch is large. Additionally, we have devised the *Boxed Water-Flowing*, an intuitive graphical representation of the asymptotically optimal offline resource allocation, which is based on the cutoff water level concept. Finally, a practical online algorithm has been proposed, which requires the lowest execution time compared to the offline algorithms and has a small performance loss.

Finally, Chapter 6 has focused on a multiuser communication system and has investigated the power allocation strategy in each transmitter that maximizes the network's sum-rate when a generalized power consumption model is considered in the ECCs, which is composed of products and summations of step functions. To address the problem, the ISCA has been proposed that is based on SSA of the indicator functions to derive a sequence of smooth nonconvex problems that can be solved by means of SCA. As it has been presented, one of the key features of the proposed algorithm is that the power allocation strategy can be computed in a distributed manner by the network nodes. By numerical simulation, we have shown that, under the setup of Chapter 5 (a point-to-point communication with channel access and stream activation costs), the mutual information achieved by the ISCA is similar to the asymptotically optimal *Boxed Water-Flowing* solution while it reduces the computational complexity.

7.2 Future Work

There are a myriad of possible research directions that can be considered to extend the results obtained in this work. In short, for the design of more efficient transmission strategies, it is key to develop better models of both the energy generation and consumption processes.

First of all, given the current available models of the energy generation and consumption processes, there are still many *offline transmission strategies* that should be investigated:

- This thesis has only considered the transmission completion time minimization and the mutual information maximization; accordingly, one possible research direction is the consideration of different figures of merit (for example, the ones discussed in §2.1.5).
- Regarding practical communication systems that use an arbitrary input distribution, the

following research lines can be considered:

- To determine the optimal right singular vectors of the precoding matrix under the setup studied in Chapter 4.
- To determine the optimal linear precoder with the generalized power consumption model $\mathcal{C}^4$ introduced in Chapter 6.
- To allow a dynamic selection of the modulation and coding schemes according to the node energy availability.

As briefly discussed in §2.1.1, few works have dealt with the modeling of the energy harvesting process given the different energy sources and transducers. Accordingly, the following research directions are required in the upcoming years:

- To develop measurement campaigns of the harvested energy given different energy sources and harvesting technologies (in a similar way than the measurements of the solar and kinetic energies obtained in [31, 35, 36]). These measurement campaigns must be broad enough to capture the energy variability due to environmental and climatic conditions as well as node's placement and mobility.
- To create statistical models of the acquired data in the previous point in order to design:
  - Low-complexity prediction algorithms of the harvested energy that can be used prior to the computation of the offline transmission strategy.
  - Online algorithms that exploit the available statistical knowledge of the harvested energy.

Additionally, for the design of accurate offline and online transmission strategies it is key to analyze the impact of not having complete CSI at the transmitter.

This work has proposed a generalized power consumption model, but still many sources of *power consumption* escape from the model and should be considered:

- *Energy consumption at the processing unit to perform the computation of the transmission strategy:*

  As we have seen throughout this dissertation, the computation of the transmission strategy entails a certain computational complexity, which in general grows with the number of channel accesses, N, (or optimization time horizon), the number of parallel data

streams, K, and the number of energy packets or ECCs, J. This computational complexity has an associated energy consumption which has been ignored in the designed transmission strategies (and, to the best of our knowledge, in all the related works in the literature).

On the one hand, when this source of energy consumption is neglected, a larger optimization time horizon is always preferred in terms of achievable performance; additionally, as it has been argued in §2.1.1, the packetized model of the energy harvesting process can be understood as a time sampling of the continuous power harvesting profile, where the smaller the sampling window (or the greater the value of J), the more the packetized model resembles to the real power harvesting profile, which also results in a better performance.

On the other hand, if this source of energy consumption is considered, there exists a tradeoff in the selection of the parameters N and J, which should be carefully analyzed and studied, since larger values of N and J imply a loss on the total available energy for transmission due to the higher energy consumption to compute the transmission strategy.

- *Energy consumption of the sensors to acquire data:*

  A wireless sensor node consumes a substantial amount of energy for sensing purposes (in the sensor itself and to perform the analog to digital conversion of the sensed data), which has been generally ignored when designing the node's transmission strategy. This source of consumption opens a new research paradigm since the sensing and transmission policies must be jointly designed to achieve the best performance given some figure of merit.

Bibliography

[1] European Comission, "Addressing the challenge of energy efficiency through information and communication technologies," May 2008.

[2] G. E. Moore, "Cramming more components onto integrated circuits," *IEEE Solid-State Circuits Society Newsletter*, vol. 11, no. 5, pp. 33–35, Sep. 2006.

[3] W. Vereecken, W. V. Heddeghem, D. Colle, M. Pickavet, and P. Demeester, "Overall ICT footprint and green communication technologies," in *proceedings of the 4th International Symposium on Communications, Control and Signal Processing*, Mar. 2010, pp. 1–6.

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

[5] Yole Développement, "Emerging energy harvesting devices," Nov. 2012.

[6] M. Gregori and M. Payaró, "Efficient data transmission for an energy harvesting node with battery capacity constraint," in *Proceedings of the IEEE Global Telecommunications Conference*, Dec. 2011, pp. 1–6.

[7] ——, "Energy-efficient transmission for wireless energy harvesting nodes," *IEEE Trans. on Wireless Communications*, vol. 12, no. 3, pp. 1244–1254, Mar. 2013.

[8] ——, "Optimal power allocation for a wireless multi-antenna energy harvesting node with arbitrary input distribution," in *Proceedings of the IEEE International Conference on Communications*, Jun. 2012, pp. 5794 –5798.

[9] ——, "On the precoder design of a wireless energy harvesting node in linear vector Gaussian channels with arbitrary input distribution," *IEEE Trans. on Communications*, vol. 61, no. 5, pp. 1868–1879, May 2013.

[10] ——, "Throughput maximization for a wireless energy harvesting node considering the circuitry power consumption," in *Proceedings of the IEEE Vehicular Technology Conference*, Sep. 2012, pp. 1–5.

[11] M. Gregori, A. Pascual-Iserte, and M. Payaró, "Mutual information maximization for a wireless energy harvesting node considering the circuitry power consumption," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, Apr. 2013, pp. 4238–4243.

[12] M. Gregori and M. Payaró, "On the optimal resource allocation for a wireless energy harvesting node considering the circuitry power consumption," *accepted in IEEE Trans. on Wireless Communications*, Jun. 2014.

[13] ——, "Multiuser communications with energy harvesting transmitters," in *Proceedings of the IEEE International Conference on Communications*, 2014.

[14] M. Payaró, M. Gregori, and D. P. Palomar, "Yet another entropy power inequality with an application," in *Proceedings of the IEEE International Conference on Wireless Communications and Signal Processing*, Nov. 2011, pp. 1–5.

[15] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 6, no. 4, p. 32, Sep. 2007.

[16] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 443–461, Third Quarter 2011.

[17] N. Wang, N. Zhang, and M. Wang, "Wireless sensors in agriculture and food industry—recent development and future perspective," *Computers and electronics in agriculture*, vol. 50, no. 1, pp. 1–14, Jan. 2006.

[18] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Computer networks*, vol. 51, no. 4, pp. 921–960, Mar. 2007.

[19] D. Gunduz, K. Stamatiou, N. Michelusi, and M. Zorzi, "Designing intelligent energy harvesting communication systems," *IEEE Commun. Mag.*, vol. 52, no. 1, pp. 210–216, Jan. 2014.

[20] D. P. Palomar, "A unified framework for communications through MIMO channels," Ph.D. dissertation, Technical University of Catalonia (UPC), Barcelona, Spain, 2003.

[21] M. Payaró, "Impact of channel state information on the analysis and design of multi-antenna communication systems," Ph.D. dissertation, Technical University of Catalonia (UPC), Barcelona, Spain, 2006.

[22] R. J. Vullers, R. Schaijk, H. J. Visser, J. Penders, and C. V. Hoof, "Energy harvesting for autonomous wireless sensor networks," *IEEE J. Solid-State Circuits*, vol. 2, no. 2, pp. 29–38, Spring 2010.

[23] "Cebekit: Windlab junior c-0200 technical characteristics." [Online]. Available: http://fadisel.cat/docs/c-0200-ing.pdf

[24] H. J. Visser, A. C. Reniers, and J. A. Theeuwes, "Ambient RF energy scavenging: GSM and WLAN power density measurements," in *Proceedings of the IEEE European Microwave Conference*, Oct. 2008, pp. 721–724.

[25] W. K. Seah, Z. A. Eu, and H.-P. Tan, "Wireless sensor networks powered by ambient energy harvesting (wsn-heap)-survey and challenges," in *Proceedings of the IEEE Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology*, May 2009, pp. 1–5.

[26] L. Mateu, C. Codrea, N. Lucas, M. Pollak, and P. Spies, "Energy harvesting for wireless communication systems using thermogenerators," in *Proceeding of the XXI Conference on Design of Circuits and Integrated Systems*, Nov. 2006.

[27] C. Park and P. H. Chou, "Ambimax: Autonomous energy harvesting platform for multi-supply wireless sensor nodes," in *Proceedings of the IEEE Sensor and Ad Hoc Communications and Networks*, vol. 1, Sep. 2006, pp. 168–177.

[28] Y. K. Tan and S. K. Panda, "Self-autonomous wireless sensor nodes with wind energy harvesting for remote sensing of wind-driven wildfire spread," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 4, pp. 1367–1377, Apr. 2011.

[29] S. J. Roundy, "Energy scavenging for wireless sensor nodes with a focus on vibration to electricity conversion," Ph.D. dissertation, University of California, 2003.

[30] S. P. Beeby, M. J. Tudor, and N. M. White, "Energy harvesting vibration sources for microsystems applications," *Measurement Science and Technology*, vol. 17, no. 12, p. R175, 2006.

[31] M. Gorlatova, J. Sarik, M. Cong, I. Kymissis, and G. Zussman, "Movers and shakers: Kinetic energy harvesting for the internet of things," *arXiv preprint:1307.0044*, 2013.

[32] J. Landt, "The history of RFID," *IEEE Potentials*, vol. 24, no. 4, pp. 8–11, Oct. 2005.

[33] S. A. Bhalerao, A. V. Chaudhary, R. B. Deshmukh, and R. M. Patrikar, "Powering wireless sensor nodes using ambient RF energy," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, Oct. 2006, pp. 2695–2700.

[34] D. Valerdi, Q. Zhu, K. Exadaktylos, S. Xia, M. Arranz, R. Liu, and D. Xu, "Intelligent energy managed service for green base stations," in *Proceedings of the IEEE GLOBECOM Workshops*, Dec. 2010, pp. 1453–1457.

[35] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: Measurements and algorithms," *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1853–1865, Sep. 2013.

[36] M. Gorlatova, M. Zapas, E. Xu, I. Kymissis, and G. Zussman, "Crawdad data set columbia/enhants (v. 2011-04-07)," 2011. [Online]. Available: http://crawdad.org/ crawdad/columbia/enhants/

[37] A. Kansal and M. Srivastava, "An environmental energy harvesting framework for sensor networks," in *Proceedings of the 2003 International Symposium on Low Power Electronics and Design*, Aug. 2003, pp. 481–486.

[38] K. Kinoshita, T. Okazaki, H. Tode, and K. Murakami, "A data gathering scheme for environmental energy-based wireless sensor networks," in *Proceedings of the IEEE Consumer Communications and Networking Conference*, Jan. 2008, pp. 719–723.

[39] S. Kaplanis and E. Kaplani, "A model to predict expected mean and stochastic hourly global solar radiation $I(h; nj)$ values," *Renewable energy*, vol. 32, no. 8, pp. 1414–1425, 2007.

[40] L. Wong and W. Chow, "Solar radiation model," *Applied Energy*, vol. 69, no. 3, pp. 191 – 224, Jul. 2001.

[41] R. Aguiar and M. Collares-Pereira, "Tag: A time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation," *Solar Energy*, vol. 49, no. 3, pp. 167–174, Sep. 1992.

[42] N. Sharma, J. Gummeson, D. Irwin, and P. Shenoy, "Cloudy computing: Leveraging weather forecasts in energy harvesting sensor systems," in *Proceedings of the IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks*, Jun. 2010, pp. 1–9.

[43] C. K. Ho, P. D. Khoa, and P. C. Ming, "Markovian models for harvested energy in wireless communications," in *Proceedings of the IEEE International Conference on Communication Systems*, Nov. 2010, pp. 311–315.

[44] J. R. Piorno, C. Bergonzini, D. Atienza, and T. S. Rosing, "Prediction and management in energy harvested wireless sensor nodes," in *Proceedings of the IEEE International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology*, May 2009, pp. 6–10.

[45] L. Berbakov, C. Antón-Haro, and J. Matamoros, "Optimal transmission policy for cooperative transmission with energy harvesting and battery operated sensor nodes," *Signal Processing*, vol. 93, no. 11, pp. 3159 – 3170, Nov. 2013.

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

[47] ——, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089 – 1098, Aug. 2004.

[48] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Trans. on Wireless Communications*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.

[49] J. Xu and R. Zhang, "Throughput optimal policies for energy harvesting wireless transmitters with non-ideal circuit power," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 322–332, Feb. 2014.

[50] P. Youssef-Massaad, M. Medard, and L. Zheng, "Impact of processing energy on the capacity of wireless channels," in *In Proceedings of International Symposium on Information Theory and its Applications*, Oct. 2004.

[51] P. Youssef-Massaad, L. Zheng, and M. Medard, "Bursty transmission and glue pouring: on wireless channels with overhead costs," *IEEE Trans. on Wireless Communications*, vol. 7, no. 12, pp. 5188 –5194, Dec. 2008.

[52] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-mse minimization," *IEEE Trans. on Signal Processing*, vol. 55, no. 11, pp. 5436–5446, Nov. 2007.

[53] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint tx-rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[54] B. Prabhakar, E. U. Biyikoglu, and A. E. Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *Proceedings of the IEEE International Conference on Computer Communications*, vol. 1, Apr. 2001, pp. 386–394.

[55] M. Zafer and E. Modiano, "A calculus approach to energy-efficient data transmission with quality-of-service constraints," *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 898–911, Jun. 2009.

[56] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[57] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. on Information Theory*, vol. 52, no. 7, pp. 3033 – 3051, Jul. 2006.

[58] M. Payaró and D. P. Palomar, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution," in *Proceedings of the IEEE International Symposium on Information Theory*, Jul. 2009, pp. 1085 –1089.

[59] M. Payaró, X. Mestre, and M.-A. Lagunas, "Optimum transmit architecture of a MIMO system under modulus channel knowledge at the transmitter," in *Proceedings of the IEEE Information Theory Workshop*, Oct. 2004, pp. 418–422.

[60] J. Yang and S. Ulukus, "Transmission completion time minimization in an energy harvesting system," in *Proceedings of the IEEE Information Sciences and Systems*, Mar. 2010, pp. 1–6.

[61] B. Devillers and D. Gunduz, "Energy harvesting communication system with battery constraint and leakage," in *Proceedings of the IEEE GLOBECOM Workshops*, Dec. 2011, pp. 383–388.

[62] N. Michelusi, L. Badia, R. Carli, L. Corradini, and M. Zorzi, "Impact of battery degradation on optimal management policies of harvesting-based wireless sensor devices," in *Proceedings of the IEEE International Conference on Computer Communications*, Apr. 2013, pp. 590–594.

[63] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ Pr, 2004.

[64] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: Numerical methods*. Old Tappan, NJ (USA); Prentice Hall Inc., 1989.

[65] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *Journal of Global Optimization*, vol. 47, no. 1, pp. 29–51, May 2010.

[66] G. Scutari, F. Facchinei, P. Song, D. Palomar, and J.-S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

[67] G. Scutari, F. Facchinei, L. Lampariello, and P. Song, "Parallel and distribiuted methods for nonconvex optimization," *ICASP*, 2013.

[68] C. T. Lawrence and A. L. Tits, "A computationally efficient feasible sequential quadratic programming algorithm," *Siam Journal on optimization*, vol. 11, no. 4, pp. 1092–1118, 2001.

[69] C. A. Sagastizábal and M. V. Solodov, "Parallel variable distribution for constrained optimization," *Computational Optimization and Applications*, vol. 22, no. 1, pp. 111–131, 2002.

[70] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[71] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[72] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York, NY, USA: Wiley-Interscience, 1991.

[73] G. D. Forney Jr and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. on Information Theory*, vol. 44, no. 6, pp. 2384–2415, Oct. 1998.

[74] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. on Information Theory*, vol. 51, no. 4, pp. 1261 –1282, Apr. 2005.

[75] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. on Information Theory*, vol. 52, no. 1, pp. 141–154, Jan. 2006.

[76] M. Payaró and D. P. Palomar, "Hessian and concavity of mutual information, differential entropy, and entropy power in linear vector Gaussian channels," *IEEE Trans. on Information Theory*, vol. 55, no. 8, pp. 3613 –3628, Aug. 2009.

[77] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *Proceedings of the IEEE International Symposium on Wireless Communication Systems*, Sep. 2009, pp. 26–30.

[78] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," *IEEE Trans. on Information Theory*, vol. 59, no. 7, pp. 3301–3314, Jul. 2011.

[79] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. on Information Theory*, vol. 36, no. 5, pp. 1019–1030, Sep. 1990.

[80] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the LambertW function," *Advances in Computational mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

[81] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 487–499, Aug. 2002.

[82] E. Uysal-Biyikoglu, A. El Gamal, and B. Prabhakar, "Adaptive transmission of variable-rate data over a fading channel for energy-efficiency," in *Proceedings of the IEEE Global Telecommunications Conference*, vol. 1, Nov. 2002, pp. 97–101.

[83] M. Zafer and E. Modiano, "Delay-constrained energy efficient data transmission over a wireless fading channel," in *Proceedings of the IEEE Information Theory and Applications Workshop*, Jan. 2007, pp. 289–298.

[84] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1732–1743, Sep. 2011.

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

[86] O. Ozel and S. Ulukus, "Information-theoretic analysis of an energy harvesting communication system," in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications Workshops*, Sep. 2010, pp. 330–335.

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

[88] M. Gorlatova, A. Bernstein, and G. Zussman, "Performance evaluation of resource allocation policies for energy harvesting devices," in *Proceedings of the IEEE Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, May 2011, pp. 189–196.

[89] P. Blasco, D. Gunduz, and M. Dohler, "A learning theoretic approach to energy harvesting communication system optimization," in *Proceedings of the IEEE GLOBECOM Workshops*, Dec. 2012, pp. 1657–1662.

[90] J. Ventura and K. Chowdhury, "Markov modeling of energy harvesting body sensor networks," in *Proceedings of the IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, Sep. 2011, pp. 2168–2172.

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

[92] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Trans. on Communications*, vol. 60, no. 1, pp. 220–230, Jan. 2012.

[93] O. Orhan, D. Gunduz, and E. Erkip, "Throughput maximization for an energy harvesting communication system with processing cost," in *Proceedings of the IEEE Information Theory Workshop*, Sep. 2012, pp. 84–88.

[94] J. Yang, O. Ozel, and S. Ulukus, "Broadcasting with an energy harvesting rechargeable transmitter," *IEEE Trans. on Wireless Communications*, vol. 11, no. 2, pp. 571 –583, Feb. 2012.

[95] M. A. Antepli, E. Uysal-Biyikoglu, and H. Erkal, "Optimal scheduling on an energy harvesting broadcast channel," in *Proceedings of the IEEE Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, May 2011, pp. 197–204.

[96] O. Ozel, J. Yang, and S. Ulukus, "Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery," *IEEE Trans. on Wireless Communications*, vol. 11, no. 6, pp. 2193–2203, Jun. 2012.

[97] ——, "Optimal transmission schemes for parallel and fading Gaussian broadcast channels with an energy harvesting rechargeable transmitter," *Computer Communications, Elsevier*, vol. 36, no. 12, pp. 1360 – 1372, Jul. 2013.

[98] R. Rajesh, P. Deekshith, and V. Sharma, "Capacity of a Gaussian MAC with energy harvesting transmit nodes," in *Proceedings of the IEEE Information Theory and Applications Workshop*, Feb. 2012, pp. 338–343.

[99] J. Yang and S. Ulukus, "Optimal packet scheduling in a multiple access channel with energy harvesting transmitters," *Journal of Communications and Networks*, vol. 14, no. 2, pp. 140–150, Apr. 2012.

[100] J. Jeon and A. Ephremides, "The stability region of random multiple access under stochastic energy harvesting," in *Proceedings of the IEEE International Symposium on Information Theory*, Jul. 2011, pp. 1796–1800.

[101] M. Khuzani and P. Mitran, "On online energy harvesting in multiple access communication systems," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1883–1898, Mar. 2014.

[102] P. Blasco, D. Gunduz, and M. Dohler, "Low-complexity scheduling policies for energy harvesting communication networks," in *Proceedings of the IEEE International Symposium on Information Theory Proceedings*, Jul. 2013, pp. 1601–1605.

[103] D. Gunduz and B. Devillers, "Two-hop communication with energy harvesting," in *Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Dec. 2011, pp. 201–204.

[104] B. Gurakan, O. Ozel, J. Yang, and S. Ulukus, "Energy cooperation in energy harvesting wireless communications," in *Proceedings of the IEEE International Symposium on Information Theory Proceedings*, Jul. 2012, pp. 965–969.

[105] O. Orhan and E. Erkip, "Optimal transmission policies for energy harvesting two-hop networks," in *Proceedings of the IEEE46th Annual Conference on Information Sciences and Systems*, Mar. 2012, pp. 1–6.

[106] ——, "Energy harvesting two-hop networks: Optimal policies for the multi-energy arrival case," in *Proceedings of the IEEE Sarnoff Symposium*, May 2012, pp. 1–6.

[107] Y. Luo, J. Zhang, and K. B. Letaief, "Optimal scheduling and power allocation for two-hop energy harvesting communication systems," *IEEE Trans. on Wireless Communications*, vol. 12, no. 9, pp. 4729–4741, Sep. 2013.

[108] A. Nasir, X. Zhou, S. Durrani, and R. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. on Wireless Communications*, vol. 12, no. 7, pp. 3622–3636, Jul. 2013.

[109] I. Ahmed, A. Ikhlef, R. Schober, and R. Mallik, "Power allocation for conventional and buffer-aided link adaptive relaying systems with energy harvesting nodes," *IEEE Trans. on Wireless Communications*, vol. 13, no. 3, pp. 1182–1195, Mar. 2014.

[110] M. Feghhi, A. Abbasfar, and M. Mirmohseni, "Optimal power and rate allocation in the degraded Gaussian relay channel with energy harvesting nodes," in *Proceedings of the IEEEIran Workshop on Communication and Information Theory*, May 2013, pp. 1–6.

[111] C. Huang, R. Zhang, and S. Cui, "Throughput maximization for the Gaussian relay channel with energy harvesting constraints," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1469–1479, Aug. 2013.

[112] B. Medepally and N. B. Mehta, "Voluntary energy harvesting relays and selection in cooperative wireless networks," *IEEE Trans. on Wireless Communications*, vol. 9, no. 11, pp. 3543–3553, Nov. 2010.

[113] L. Berbakov, C. Antón-Haro, and J. Matamoros, "Optimal transmission policy for collaborative beamforming with finite energy storage capacity," in *Proceedings of the IEEE Personal Indoor and Mobile Radio Communications*, Sep. 2013, pp. 559–563.

[114] K. Tutuncuoglu and A. Yener, "Sum-rate optimal power policies for energy harvesting transmitters in an interference channel," *Journal of Communications and Networks*, vol. 14, no. 2, pp. 151–161, Apr. 2012.

[115] X. Liu and E. Erkip, "Energy-efficient communication over Gaussian interference networks with processing energy cost," *CoRR*, vol. abs/1301.1661, 2013.

[116] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*, ser. Wiley series in probability and statistics. New York: Wiley, 1999.

[117] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. on Information Theory*, vol. 48, no. 6, pp. 1319 –1343, Jun. 2002.

[118] D. Guo, Y. Wu, S. Verdú *et al.*, "Estimation in Gaussian noise: Properties of the minimum mean-square error," *IEEE Trans. on Information Theory*, vol. 57, no. 4, pp. 2371–2385, Apr. 2011.

[119] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. on Communications*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.

[120] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," *lecture notes of EE392o, Stanford University, Autumn Quarter*, vol. 2004, 2003. [Online]. Available: http://www.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf

[121] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[122] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[123] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed ℓ^0 norm," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 289–301, Jan. 2009.

[124] X. Zheng, X. Sun, D. Li, and J. Sun, "Successive convex approximations to cardinality-constrained convex programs: a piecewise-linear DC approach," *Computational Optimization and Applications*, pp. 1–19, 2013.

[125] A. Blake and A. Zisserman, *Visual Reconstruction*. Cambridge MA: MIT Press, 1987.

[126] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[127] R. T. Rockafellar, R. J.-B. Wets, and M. Wets, *Variational analysis*. Springer, 1998, vol. 317.

[128] G. Scutari, F. Facchinei, J. Pang, and D. Palomar, "Real and complex monotone communication games," *IEEE Trans. Inf. Theory*, vol. PP, no. 99, pp. 1–1, 2014.