Multichannel Power Line Communication

Von der Fakultät für Ingenieurwissenschaften,
Abteilung Elektrotechnik und Informationstechnik
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

von

Babak Nikfar
aus
Teheran, Iran

Gutachter: Prof. Dr. A. J. Han Vinck
Gutachter: Prof. Dr. Gerd Bumiller
Tag der mündlichen Prüfung: 16.11.2017

Abstract

Power line communication (PLC) is the technology in which the data signals of a communication system are transmitted through the conductors of a power delivery infrastructure. The unique environment of the PLC channels create specific challenges and requirements, which need to be modeled and analyzed properly in order to obtain a clear understanding of the communication system as well as attaining the ability to further improve the performance and reliability of the transmission. Moreover, the demand for increased data throughput as well as increased reliability and robustness of the transmission is of fundamental importance in any communication system as it is in PLC systems. In order to address these challenges and demands, the concept of multichannel PLC is studied and developed in this thesis. Multichannel PLC in this context is referred to the transmission of multiple information-carrying signals though the power line channel from one source to one destination.

We study multiple scenarios of multichannel data transmission in order to cover the diverse situations and requirements of a PLC transmission. One of the multichannel scenarios discussed in this thesis is the multiple-input multiple-output (MIMO) transmission, in which multiple data signals are transmitted via spatially separated PLC channels. Another scenario discussed in this thesis is the cooperative transmission between the source and destination of a PLC system by means of intermediate relay nodes in the network. Finally, the multiband transmission by utilizing different parts of the available PLC spectrum is studied. The core objective of this thesis is to develop and study novel algorithms and models to address the challenges and problems introduced in different scenarios of the multichannel PLC. These problems can be categorized as the channel selection problem for MIMO transmission, the relay selection problem for the cooperative communication, and the spectrum assignment problem for the multiband transmission. The basis of all these problems is a decision making problem, which can greatly influence the performance of the system.

To address these decision making problems, a powerful mathematical tool, namely the multi-armed bandit model, is used to model the different problems emerging in different scenarios of the multichannel PLC. This modeling approach is then used as a building block for developing machine learning algorithms in order to solve the aforementioned selection problems. Finally, novel machine learning algorithms are developed and their performances are analyzed and assessed. It is shown that the machine learning approach can considerably improve the performance of the multichannel PLC systems compared to the existing state of the art approaches, by enabling the selecting agent, i.e. the PLC transmitter, to perform intelligent decisions which improves the overall performance.

Zusammenfassung

Die Power-Line-Communication (PLC) ist die Technologie, bei der die Datensignale eines Kommunikationssystems über die Leiter einer Energieversorgungsinfrastruktur übertragen werden. Die einzigartige Umgebung der PLC-Kanäle stellt konkrete Herausforderungen und Anforderungen dar, die modelliert und analysiert werden müssen, um ein klares Verständnis des Kommunikationssystems zu erhalten und die Fähigkeit zur Verbesserung der Leistung und Zuverlässigkeit der Übertragung zu erreichen. Darüber hinaus ist in Kommunikationssystem die Nachfrage nach erhöhtem Datendurchsatz, sowie erhöhter Zuverlässigkeit und Robustheit der Übertragung von grundlegender Bedeutung. Um diesen Herausforderungen und Anforderungen gerecht zu werden, wird in dieser Arbeit das Konzept der Mehrkanal-PLC untersucht und weiterentwickelt. Die Mehrkanal-PLC wird in diesem Zusammenhang auf die Übertragung mehrerer informationstragenden Signale über den PLC-Kanal von einer Quelle zu einem Ziel bezogen.

Wir untersuchen mehrere Szenarien der Mehrkanal-Datenübertragung, um die vielfältigen Anforderungen einer PLC-Übertragung zu behandeln. Eines der in dieser Arbeit besprochenen Mehrkanal-Szenarien ist die Multiple-Input-Multiple-Output-Übertragung (MIMO), bei der mehrere Datensignale über räumlich getrennte PLC-Kanäle übertragen werden. Ein weiteres Szenario, das in dieser Arbeit diskutiert wird, ist die kooperative Übertragung zwischen der Quelle und dem Ziel eines PLC-Systems mittels Zwischenrelais als Knoten im Netzwerk. Schließlich wird die Multiband-Übertragung unter Verwendung unterschiedlicher Teile des verfügbaren PLC-Spektrums untersucht. Das Kernziel dieser Arbeit ist es, neuartige Algorithmen und Modelle zu entwickeln und zu untersuchen, um die Herausforderungen und Probleme zu lösen, die in verschiedenen Szenarien der Mehrkanal-PLC existieren. Diese Probleme sind als das Kanalauswahlproblem für die MIMO-Übertragung, das Relaiauswahlproblem für die kooperative Kommunikation und das Spektrum-Zuweisungsproblem für die Multibandübertragung kategorisiert werden. Die Basis all dieser Probleme ist ein Entscheidungsproblem, das die Leistungsfähigkeit des Systems stark beeinflussen kann.

Um diese Probleme lösen zu können, wird ein mathematisches Werkzeug, nämlich das mehrarmige Bandit-Modell, verwendet, um die verschiedenen Probleme zu modellieren, die sich in verschiedenen Szenarien der Mehrkanal-PLC ergeben. Dieser Modellierungsansatz wird als Baustein für die Entwicklung von maschinellen Lernalgorithmen verwendet, um die zuvor beschriebenen Auswahlprobleme zu lösen. Schließlich werden neuartige maschinelle Lernalgorithmen entwickelt und ihre Leistungen analysiert sowie bewertet. Es zeigt sich, dass der maschinelle Lernansatz die Leistungsfähigkeit der Mehrkanal-PLC-Systeme im Vergleich zu den bestehenden Ansätzen des Standes der Technik erheblich verbessern kann, indem es dem Auswahlagenten, d.h. dem PLC-Sender, ermöglicht, intelligente Entscheidungen durchzuführen, die die Gesamtleistung verbessern.

Acknowledgements

Finally, my PhD studies are finished and I have managed to write it all down. But before getting to more technical stuff, there are some people I would like to thank, who without them this work would not be possible.

First and foremost, I would like to thank my family for their love and support throughout my entire life. Their incredible support through every step of my life gave me the courage to move forward and provided me with the love and solace which made my every endeavor significantly easier. My appreciation toward my family is endless. I would also like to thank the newest member of our family, Paris, for all the joy she has provided us all.

I would also like to express my greatest appreciation to my supervisor, Prof. Dr. Han Vinck, who has been a great teacher to me, and not just for the scientific matters. He was always there when I needed him and helped me along the way as much as he could. I really enjoyed working with him.

I would also like to thank Prof. Dr. Gerd Bumiller for his co-supervision for the last two years. I certainly learned a lot from him and his passion and enthusiasm in what he does taught me and motivated me a great deal in my own endeavors. I hope the cooperation continues in the future.

I would also like to thank Chao, Danica, George, Thokozani, Victor, and Yanlin for all the good times that we spent together. Their friendship certainly made every activity more enjoyable. I would also like to thank my coworkers at the Hochschule Ruhr West who made a warm and welcoming environment for me during the last two years. And a special thanks to Sven for proofreading my abstract in German.

Last but certainly not least, I would like to declare my wholehearted appreciation and love toward Setare for her relentless love and support throughout our incredible journey in life. She supported me every step of the way and was there whenever I needed her, in sickness and in health, in hard times and in pleasant times. She made me a better person and the person who I am today, which makes her an inseparable part of my very own existence. I am eternally grateful to her.

Contents

List of Figures

List of Tables

List of Abbreviations

ARIB	**A**ssociation of **R**adio **I**ndustries and **B**usinesses
AWGN	**A**dditive **W**hite **G**aussian **N**oise
BER	**B**it **E**rror **R**ate
BPSK	**B**inary **P**hase **S**hift **K**eying
CD-UCB	**C**yclic **D**iscount **U**pper **C**onfidence **B**ound
CENELEC	**C**omité **E**uropéen de **N**ormalisation **Élec**trotechnique
CM	**C**ommon **M**ode
CSI	**C**hannel **S**tate **I**nformation
CW-UCB	**C**yclic **W**indow **U**pper **C**onfidence **B**ound
DFT	**D**iscrete **F**ourier **T**ransform
DM	**D**ifferential **M**ode
D-UCB	**D**iscounted **U**pper **C**onfidence **B**ound
EN	**E**uropean **N**orm
FCC	**F**ederal **C**ommunications **C**ommission
FFT	**F**ast **F**ourier **T**ransform
FSK	**F**requency **S**hift **K**eying
GB	**G**rouped **B**ands
HV	**H**igh **V**oltage
IDFT	**I**nverse **D**iscrete **F**ourier **T**ransform
IEEE	**I**nstitute of **E**lectrical and **E**lectronic **E**ngineers
IEEE-SA	**I**nstitute of **E**lectrical and **E**lectronic **E**ngineers - **S**tandards **A**ssociation
IFFT	**I**nverse **F**ast **F**ourier **T**ransform
ITU	**I**nternational **T**elecommunication **U**nion
LV	**L**ow **V**oltage
MAB	**M**ulti-**A**rmed **B**andit
MAC	**M**edia **A**ccess **C**ontrol
MCU	**M**aster **C**ontrol **U**nit
MIMO	**M**ultiple-**I**nput **M**ultiple-**O**utput
MISO	**M**ultiple-**I**nput **S**ingle-**O**utput
MRC	**M**aximal **R**atio **C**ombining
MV	**M**edium **V**oltage
OFDM	**O**rthogonal **F**requency **D**ivision **M**ultiplexing
OSTBC	**O**rthogonal **S**pace-**T**ime **B**lock **C**ode
PLC	**P**ower **L**ine **C**ommunication

PMT	**P**robability **M**atching **T**echnique
PRIME	**P**owerline **R**elated **I**ntelligent **M**etering **E**volution
PSD	**P**ower **S**pectral **D**ensity
PSK	**P**hase **S**hift **K**eying
p.u.l	**p**er **u**nit **l**ength
QAM	**Q**uadrature **A**mplitude **M**odulation
QoS	**Q**uality **o**f **S**ervice
QPSK	**Q**uadrature **P**hase **S**hift **K**eying
S-FSK	**S**pread-**F**requency **S**hift **K**eying
SDO	**S**tandard **D**eveloping **O**rganization
SGB	**S**imilar **G**rouped **B**ands
SIMO	**S**ingle-**I**nput **M**ultiple-**O**utput
SISO	**S**ingle-**I**nput **S**ingle-**O**utput
SNR	**S**ignal to **N**oise **R**atio
STBC	**S**pace-**T**ime **B**lock **C**ode
SVD	**S**ingular **V**alue **D**ecomposition
SW-UCB	**S**liding Window **U**pper **C**onfidence **B**ound
TEM	**T**ransverse **E**lectro**M**agnetic
TL	**T**ransmission **L**ine
UCB	**U**pper **C**onfidence **B**ound

To the memory of my father
and
to my mother.

Chapter 1

Introduction

1.1 State of the Art

The use of electrical power delivery wires to provide data transmission capabilities to power delivery networks, is known as power line communication (PLC). In a PLC system, the data signal is transmitted as a differential voltage signal between any two conductors of the power line infrastructure. This differential signal propagates through the power lines from the data source to the destination, providing a PLC communication channel which its features depend on the transmission line characteristics as well as the connected loads to the power line network.

Figure 1.1 depicts an example of a power line network with three wires denoted as Line (L), Neutral (N), and Protective Earth (PE). In this figure, three differential voltage signals, $\Delta v_{i,in}, i \in \{1, 2, 3\}$, are shown which indicate the signals which are produced between phase, neutral, and protective earth wires at the transmitter. These signals travel through the power line network and can be detected at the receiver as $\Delta v_{j,out}, j \in \{1, 2, 3\}$ signals. However, in a real PLC system, due to the Kirchhoff's law, one of the signals are always dependent to the other two signals and hence cannot contain information. The transmitted signal is subject to channel fading, noise, and interference as well as frequency and phase offsets. The main purpose of a PLC system design is to combat these obstacles and secure a reliable communication.

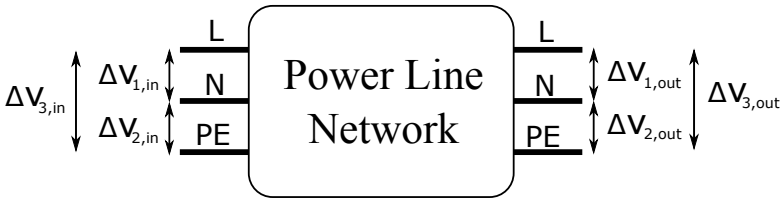


FIGURE 1.1: An example of PLC input and output voltage signals.

The concept of PLC is almost as old as the power delivery networks themselves. However, the amount of research in this area was relatively small until the late 1990s, when the interest in PLC research grew exponentially ever since. The wide range of applications in which PLC can prove useful and the number of associated challenges, resulted in gathering a substantial attention from the research community as well as industry in the past two decades. The PLC usage covers a wide range of applications such as in-home PLC in association with smart homes, communication in smart grids, communication in transport systems for instance in-vehicle communication, etc. The main advantages of using PLC for the aforementioned applications are twofold. Firstly, the cost of installation and infrastructures are considerably lower than other means of communication. The reason for that is the existing wires and power delivery infrastructures in most of the places where other means of communication are not available. Secondly, PLC can facilitate communication between some nodes in a network which are otherwise not connected to each other for data transmission. For instance, PLC can make communication through certain obstacles possible whereas other means of communication are proved incompetent.

Since power lines were not originally designed for data transmission, the PLC channel is proved to be a harsh environment for communication. The PLC channel is subject to frequency-selective fading, background noise, time-variant synchronous as well as asynchronous impulsive noise, and narrowband interference from other signals operating in the same frequency range as the transmitted signal. One of the major challenges in the PLC research is the appropriate modeling of the channel fading and noise of the PLC channel. The existing models for other communication media (for instance wireless communication) have proven inadequate for the PLC channel due to the unique characteristics of the PLC channel. Therefore, channel and noise modeling for PLC has attracted the attention of researchers and a few attempts have been made in this regard. Generally, the channel modeling approaches in PLC applications are either based on the transmission line theory or the multipath nature of the signal propagation in transmission lines.

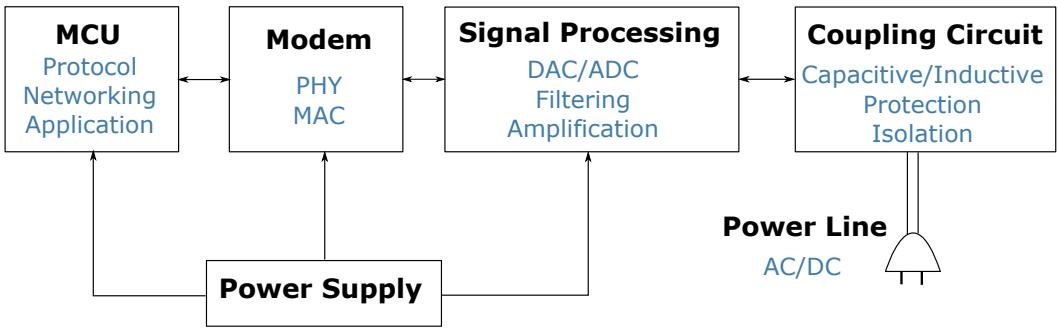


FIGURE 1.2: Architecture of a PLC transceiver.

Furthermore, the existence of multiple conductor in power line infrastructure gave rise to the idea of multiple-input multiple-output (MIMO) transmission in PLC. Today, different MIMO processing options, with the purpose of increasing data rates

and communication reliability, are in operation in major wireless cellular systems as well as wireless local area networks. This has been the motivation for the research in MIMO PLC with hopes of increased data rates (in case of spatial multiplexing) or improved transmission reliability (in case of spatial diversity and space-time coding). However, utilization of MIMO signal processing concepts in PLC applications is not without its own challenges. Namely, the existence of power line conductors in the vicinity of each other throughout the length of the PLC channel, results in a spatial correlation between the MIMO channels which may degrade the performance of the transmission. Moreover, other uses of multiple channel transmission such as cooperative transmission and multiband transmission has been proposed in the PLC literature. In all of these schemes the main focus of the system design is to increase the reliability and performance of the communication.

The structure of a PLC transceiver including all the higher layers of communication is illustrated in Figure 1.2 [1]. The master control unit (MCU) consists of the protocol, networking and the application layers, whereas the PLC modem consists of the physical and MAC layers. After performing appropriate signal processing, such as amplification, filtering, and digital to analog (or analog to digital) conversion, a coupling circuit is used in order to couple the data signals into the power line conductors.

1.2 Scope and Contributions of the Thesis

In this thesis, the multichannel power line communication is studied and further contributions has been made to this field. By multichannel PLC, it is implied that the communication link between the data source and destination consists of more than one unique communication channel. This is normally performed in order to introduce some notion of diversity into the communication system and through this introduced diversity, the performance and reliability of the communication link can be improved. We focus on three multichannel PLC scheme, which their descriptions follows.

- MIMO transmission: data signal as multiple data streams are transmitted simultaneously through the power line medium, constructing multiple spatially separated MIMO channels, which either results in higher data rates or increases the reliability and performance of the transmission.
- Multihop transmission: the data is transmitted from source to the destination via intermediate relay nodes, constructing multiple hops of the transmission each with its own channel properties. This relay-aided transmission provides cooperative diversity, which can in turn, increase the reliability of the channel as well as enabling the communication between long distances.
- Multiband transmission: separate parts of the spectrum are allocated for data transmission. This transmission can be simultaneous transmission in different frequency bands or use different frequency bands in different times. This may

help the communication system to avoid deep fades, narrowband interferences, tone masks, etc., which can result in increased reliability and performance.

Moreover, in order to introduce selection diversity to MIMO PLC, the problem of channel selection is studied. Similarly, in order to introduce cooperative diversity to multihop PLC transmission, the problem of relay selection is studied. Finally, in order to increase the spectral efficiency of the transmission, the problem of spectrum assignment is studied. In all of these problems, a decision has to be made at the PLC transmitter in form of a certain selection. The decision making process at the transmitter requires appropriate information about the PLC channels in order to optimize the decision. However, frequency-selective and time-variant nature of the PLC channels makes it a challenging task for the PLC transmitter to obtain channel state information for all communication links, at all time, and all frequencies. Therefore, it is a fair assumption to consider the channel state information at PLC transmitter unknown. We proposed a class of reinforcement learning algorithms which provides a strong decision making tool to the PLC transmitter. The selection policies based on these algorithms can be applied to the problems under study which enables the PLC transmitter to perform proper decisions without any prior knowledge of the channel.

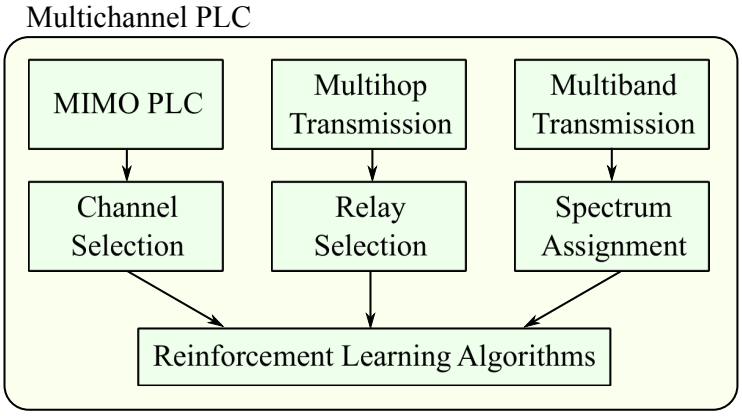


FIGURE 1.3: Scope of the thesis.

A visualization of the topics covered in this thesis is depicted in Figure 1.3. A list of the contributions to the multichannel PLC research which has been presented in this thesis follows.

- The MIMO PLC system is analyzed.
- The MIMO PLC channel capacity is derived for different levels of information availability at the transmitter and the receiver.
- The problem of channel selection in MIMO PLC is analyzed.
- The problem of relay selection in cooperative PLC is analyzed.
- The problem of spectrum assignment in multiband PLC is analyzed.
- The above-mentioned problems are modeled by multi-armed bandit problem formulation.

– Reinforcement learning approach is introduced to PLC applications in order to solve the above-mentioned problems.

– Novel reinforcement learning algorithms have been designed in order to incorporate the PLC characteristics and improve the performance of the seminal algorithms.

1.3 List of Publications

The content of this thesis has been published in high quality international academic papers in the form of journal papers and conference papers. The following is a list of these publications with a reference to the part of thesis where their contents have been discussed.

1.3.1 Journal Papers

[J2] **Babak Nikfar**, Gerd Bumiller, A. J. Han Vinck, *Dynamic Spectrum Assignment for Narrowband Power Line Communication*, in preparation for submission.
The contents of this paper can be found in Chapters 4 and 5.

[J1] **Babak Nikfar**, A. J. Han Vinck, *Relay Selection in Cooperative Power Line Communication: A Multi-Armed Bandit Approach*, to appear in IEEE/KICS Journal of Communications and Networks (accepted on 29.07.2016).
The contents of this paper can be found in Chapters 4 and 5.

1.3.2 Conference Papers

[C8] **B. Nikfar** and G. Bumiller, *Real-time Synchronization and Multiband Detection for Narrowband Power Line Communication*, IEEE International Symposium on Power Line Communications and Its Applications (ISPLC), Madrid, Spain, April 2017.
Some of the contents of this paper can be found in Chapter 4.

[C7] **B. Nikfar**, G. Bumiller, and A. J. H. Vinck, *An Adaptive Pursuit Strategy for Dynamic Spectrum Assignment in Narrowband PLC*, IEEE International Symposium on Power Line Communications and Its Applications (ISPLC), Madrid, Spain, April 2017.
The contents of this paper can be found in Chapters 4 and 5.

[C6] **B. Nikfar**, G. Bumiller, and A. J. H. Vinck, *Spectrum Assignment in Narrowband Power Line Communication*, Workshop on Power Line Communication (WSPLC), Paris, France, October 2016.
The contents of this paper can be found in Chapters 4 and 5.

[C5] **B. Nikfar**, S. Maghsudi, and A. J. H. Vinck, *Multi-armed Bandit Channel Selection for Power Line Communication*, IEEE International Conference on Smart Grid Communications, Miami, USA, November, 2015.
The contents of this paper can be found in Chapters 4 and 5.

[C4] **B. Nikfar** and A. J. H. Vinck, *Analysis of the MTL configuration for MIMO PLC with Mutual Coupling*, Workshop on Power Line Communication (WSPLC), Bottrop, Germany, September 2014.
The contents of this paper can be found in Chapters 2 and 3.

[C3] **B. Nikfar**, T. Akbudak, and A. J. H. Vinck, *MIMO Capacity of Class A Impulsive Noise Channel for Different Levels of Information Availability at Transmitter*, IEEE International Symposium on Power Line Communications and Its Applications (ISPLC), Glasgow, Scotland, April 2014.
The contents of this paper can be found in Chapter 3.

[C2] **B. Nikfar** and A. J. H. Vinck, *Combining Techniques Performance Analysis in Spatially Correlated MIMO-PLC Systems*, IEEE International Symposium on Power Line Communications and Its Applications (ISPLC), Johannesburg, South Africa, March 2013.
The contents of this paper can be found in Chapter 3.

[C1] **B. Nikfar** and A. J. H. Vinck, *Space Diversity in MIMO Power Line Channels with Independent Impulsive Noise*, Workshop on Power Line Communication (WSPLC), Rome, Italy, September 2012.
The contents of this paper can be found in Chapters 2 and 3.

1.3.3 Copyright Information

1.4 The Structure of the Thesis

The content of this thesis is structured in six chapters. The contributions and the simulation results are presented throughout each chapter. An overview of the contents as well as a description of the main focus, challenges, and contributions of each of the rest of the chapters chapters follows.

– **Chapter 2: Power Line Communication System Model**

In this chapter an overview of the PLC system model is presented. The system model presented in this chapter is used in later chapters as the ground model of the PLC system used in numerical simulations. In this chapter, the PLC

standardizations and regulations for both narrowband and broadband PLC are summarized and referenced. Moreover, the common PLC channel models are presented and discussed. Channel noise, as a significant component of the PLC channels, has been presented and for narrowband interference, background noise, as well as impulsive noise. Finally, the structure of the PLC transmitter and receiver is presented and discussed in detail. The multicarrier transmission technique of orthogonal frequency division multiplexing (OFDM), which has been accepted for PLC applications is described in this chapter as well.

– **Chapter 3: MIMO PLC**

In this chapter, a special multichannel PLC transmission scheme, namely the multiple-input multiple-output (MIMO) PLC transmission scheme is introduced and discussed. The utilization of MIMO transmission in PLC applications requires specific coupling methods, which are described in this chapter. Subsequently, the MIMO system model and the MIMO channel description, as well as the two most common MIMO transmission techniques, namely the spatial multiplexing and the spatial diversity, are presented. Spatial correlation in MIMO PLC systems, as a consequence of the natural characteristics of the power delivery structures, and its effect on the performance of the MIMO transmission ate described as well. Finally, the Channel capacity of MIMO PLC system in all cases of the availability of the channel state information at the PLC transmitter and receiver are derived in this chapter.

– **Chapter 4: Resource Allocation in Multichannel PLC**

In this chapter, the problem of resource allocation in multichannel PLC is discussed. More specifically, three forms of multichannel transmission is considered: first the MIMO transmission in multiple spatial PLC channels, second the cooperative transmission in multiple PLC channels constructed by multihop relaying, and finally the multiband transmission in separate frequency bands of the PLC spectrum. Moreover, for each of the aforementioned multichannel PLC scenarios a selection problem is introduced and discussed: channel selection problem in order to acquire selection diversity for MIMO PLC transmission, relay selection problem in order to acquire cooperative diversity for multihop transmission in PLC, and finally the spectrum assignment problem in order to acquire performance improvements in multiband PLC. Finally, all these selection problems are formulated as a multi-armed bandit problem to be used in the next chapter as the building blocks of the proposed reinforcement learning algorithms.

– **Chapter 5: Reinforcement Learning Applications in Multichannel PLC**

In this chapter, reinforcement learning algorithms are introduced in order to solve the multi-armed bandit problems formulated in the previous chapters for the selection problems described before. More specifically, the upper-confidence bound algorithms, probability matching techniques, and greedy algorithms as the most common reinforcement learning algorithms are described and analyzed. Moreover, four new algorithms based on the existing algorithms are proposed which exploits

the PLC characteristics in order to improve the existing algorithms. Simulation results which are presented throughout this chapter demonstrates the improvement achieved by the proposed algorithms.

– **Chapter 6: Conclusion**

Finally, this chapter concludes the thesis.

Chapter 2

System Model

2.1 Chapter Overview

The first step to understand and model any communication system, including power line communication systems, is through modeling and characterization of the channel. There is a considerable body of work in the literature for PLC channel and noise modeling, for instance in [2]–[6]. In general, PLC channels are frequency-selective, multipath propagating channels with cyclic short-term variations and abrupt long-term variations [7]. The channel noise in PLC systems, although additive, is a harsh noise which not only depends on the frequency and the time of operation, but also depends on the load impedance of the power line network. The PLC noise is normally divided into three main categories, namely colored background noise, impulsive noise (both synchronous and asynchronous to the line frequency), and narrowband interference [8], which will be discussed later on this chapter. The frequency in which the PLC systems operate can be divided into different ranges, which has been standardized by different standardization institutes as well as industry alliances. Figure 2.1 demonstrates the most common usable frequency ranges of PLC in practice. Broadband PLC utilizes the frequency band between 1.8 MHz up to 100 MHz, mediumband PLC is a new concept which operates in frequencies between 500 kHz and 1.8 Mhz, narrowband PLC utilizes the frequency ranges between 3 KHz and 500 kHz, and ultra-narrowband PLC utilizes the frequency ranges between 125 Hz and 3 kHz.

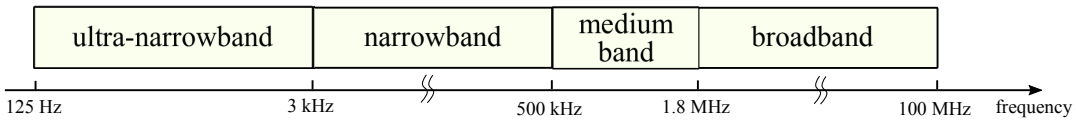


FIGURE 2.1: Ultra-narrowband, narrowband and broadband frequencies of PLC.

In this chapter, the OFDM-based PLC system model which will be used in later chapters, is presented. The PLC system model consists of the frequency specifications for OFDM transmission, channel and noise models, as well as the transmitter and receiver design. The models and system characteristics which are presented in this chapter make the basis of the contributions and simulations presented in the next chapters. The rest of the chapter is structured as follows. In Section 2.2, the spectral specifications as well as a brief review of the standardization of PLC systems, are presented. A review of the existing PLC channel modeling approaches as well as two of the most common models, namely multipath PLC channel model and PLC channel model based on the transmission line theory, are presented in Section 2.3. The noise in PLC channels and the mathematical models through which the noise is characterized will be discussed in Section 2.4. Subsequently, the transmitter and receiver structure of a PLC system based on orthogonal frequency division multiplexing (OFDM) transmission is presented in Section 2.5. Finally, Section 2.6 concludes the chapter.

2.2 PLC Regulations and Frequency Specifications

2.2.1 Regulation Activities

The growing number of PLC applications in the past few decades, caused the emergence of a number of frequency and system specifications. The need for unified specifications which leads to a harmonic cooperation of different PLC devices, was the starting point of the PLC standardization by standards developing organizations (SDOs). These standards have adopted many of the previously existing specifications and incorporated them into new and improved standards [9]. The main concern of a PLC standardization effort is to provide the coexistence of the PLC system with other communication systems operating in the same frequency range. Furthermore, the regulations ought to limit the strength of the signals which are coupled into the power lines. Although, even with the limited signal power, electromagnetic radiation occur in the PLC systems which needs to be addressed properly. However, since the ratio of the wavelength of the PLC signal to the length of the power lines is considerably shorter in broadband PLC compared to that of the narrowband PLC, the electromagnetic radiation is more important in broadband PLC applications. Therefore, the regulation constraints are different for broadband and narrowband PLC applications. In the following, a brief summary of the existing standardization and regulations for both narrowband and broadband PLC is presented.

Year	Standardization Activity
2008	PRIME specifications released
2009	PRIME Alliance established
2009	G3-PLC specifications released
2010	ITU G.hnem start of project
2010	IEEE 1901.2 authorization of project
2010	IEEE 1901.2 first draft released
2011	G3-PLC Alliance established
2011	ITU-T G.9955 released (G3-PLC and PRIME as annexes)
2012	ITU-T G.9902 (G.hnem), G.9903 (G3-PLC), G.9904 (PRIME) released
2013	IEEE 1901.2 released

TABLE 2.1: Time-line for the development of narrowband PLC standardization [10], [11].

Year	Standardization Activity
2001	HomePlug 1.0 released
2005	IEEE 1901 authorization of project
2005	HomePLug AV released
2006	ITU G.hn start of project
2008	ITU-T G.9960 consented
2010	HomePlug Green PHY released
2010	ITU-T G.9960 released
2010	IEEE 1901 released
2011	ITU-T G.9963 released
2012	HomePlug AV2 released

TABLE 2.2: Time-line for the development of broadband PLC standardization [12]–[14].

The standardization efforts in PLC applications have been achieved by a few industry alliances as well as standardization institutes. Among main institutes and alliances, one can name HomePlug Powerline Alliance, Powerline Related Intelligent Metering Evolution (PRIME), G3-PLC Alliance, International Telecommunication Union, Telecommunication Standardization Sector (ITU-T), and Institute of Electrical and Electronics Engineers Standards Association (IEEE-SA). The most common use of the PLC occurs in narrobwnd and broadband PLC regions. The mediumband of power line communication has recently gained interest after switching off of the AM radio in some parts of the world. The IEEE P1901.1 Working Group

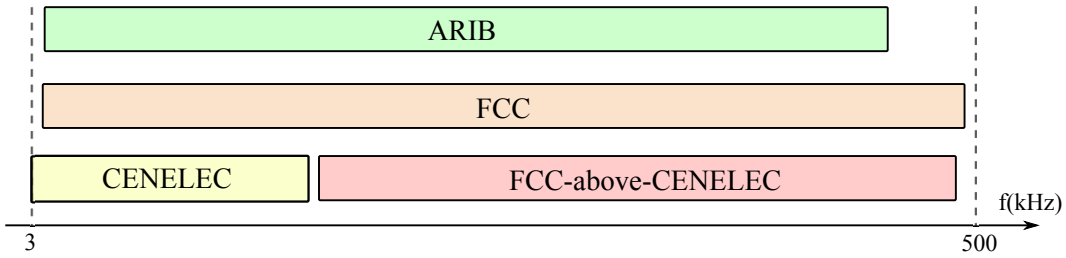


FIGURE 2.2: Frequency regulation of narrowband PLC.

has been established and held their first meeting in China in 2016. However, in this thesis we focus our work on narrowband and broadband PLC. Table 2.1 presents a brief overview of the time-line of the standardization process for narrowband PLC and Table 2.2 presents a brief overview of the time-line of the standardization process for broadband PLC, respectively. The regulations and standardizations of the PLC applications are aptly summarized in [15].

2.2.2 Narrowband Frequencies

One of the major markets for PLC equipments is Europe. European harmonized standards have been developed by recognized European SDOs. One of such standards for narrowband PLC is the European Norm (EN). EN 50065, which was first published by CENELEC[1] in 1992 delivers specifications for narrowband PLC systems for home and industry automation and for utility use such as smart metering [16]. This standard issues four distinct frequency bands which are referred to as CENELEC-A (3 – 95 kHz), CENELEC-B (95 – 125 kHz), CENELEC-C (125 – 140 kHz), and CENELEC-D (140 – 148.5 kHz). Moreover, the regulations mandate certain utilizations by these frequency bands. For instance, CENELEC-A band is reserved for power utilities, whereas CENELEC-B–D bands can only be used by consumer installations. There is no harmonized standard for the frequencies between 150 kHz and 500 kHz of the narrowband PLC spectrum, however the IEEE 1901.2 standard provides guidelines for using these frequencies in Europe as well [11], which is referred to as the FCC-above-CENELC frequency band (154.6875 – 487.5 kHz). In the US, PLC frequency bands are regulated through the FCC[2]. The FCC regulations allow the use of PLC applications in the frequency range of 9–490 kHz. Moreover, in Japan, the regulatory restrictions of ARIB[3], permits the use of PLC applications in the frequency range of 10–450 kHz. Figure 2.2 summarizes the frequency regulations of narrowband PLC.

[1]European committee for electrotechnical standardization.
[2]Federal communications commission.
[3]Association of radio industries and businesses.

2.2.3 Broadband Frequencies

The electromagnetic radiation is a bigger concern for broadband PLC, because of the higher frequencies. Therefore, the main concern of broadband regulations is to limit the electromagnetic emissions in order to prevent interferences. In Europe, the EN 50561-1 [17], has been developed to apply to in-home PLC systems operating in the 1.6–30 MHz frequency range. These regulations differentiate between a power port (only for power supply), a telecommunication port (only for communication signals), and a PLC port (for power supply as well as communication). Furthermore, power adaptation and spectrum notching is considered in order to avoid harmful interference to radio services. Further standards, namely EN 50561-2 for access networks and EN 50561-3 for frequencies above 30 MHz are under development. In the US, the 47 CFR §15 regulates the broadband PLC systems for the frequency band of 1.8–80 MHz. As is EN 50561-1, power adaptation and notching is intended for this frequency band.

2.3 PLC Channel Modeling

PLC channels are characterized as frequency-selective and time-variant with multipath propagation of the transmitted signal via transmission line conductors. The multipath fading feature of the PLC channel is due to inhomogeneities in the transmission line and the connected load impedances in the PLC network. These inconsistencies cause the signal to reflect and therefore multiple versions of the transmitted signal can be received at the receiver via multiple paths. In general, the power profile of the received signal can be obtained by convolving the power profile of the transmitted signal with the impulse response of the channel. Convolution in time domain is equivalent to multiplication in the frequency domain. Therefore, the time-variant transmitted signal $X(f,\tau)$ represented in frequency domain, after propagation through the channel becomes

$$Y(f,\tau) = H(f,\tau)X(f,\tau) + N(f,\tau), \tag{2.1}$$

where $H(f,\tau)$ is the channel response , and $N(f,\tau)$ is the additive noise term. τ represents the dependency to time. The purpose of the channel modeling is to model the channel response $H(f,\tau)$ to correspond to the real scenarios. In this section, we briefly describe the different channel modeling approaches and discuss the most common channel models available in the literature, which are also used in later chapters.

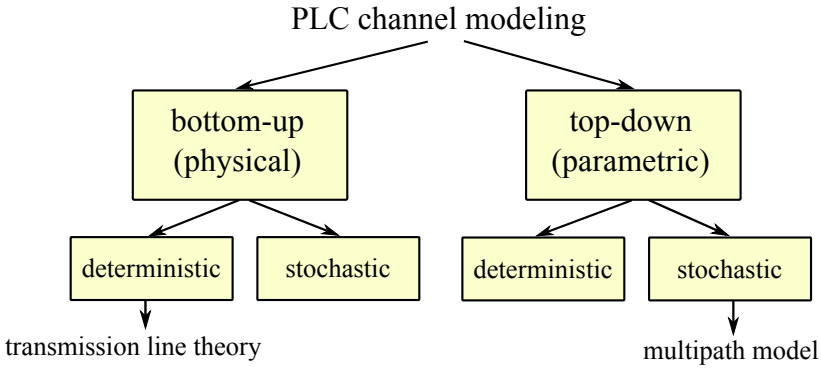


FIGURE 2.3: PLC channel modeling approaches and two of the most common channel models.

Feature	Physical Deterministic	Physical Stochastic	Parametric Deterministic	Parametric Stochastic
Modeling principle	electromag. theory	electromag. theory and topology	playback of experimental measurement	statistical fit to experimental measurement
Measurement requirements	none	none	large data base	large data base
Topology knowledge	detailed	detailed	none	none

TABLE 2.3: Comparison of PLC Channel Modeling Approaches.

2.3.1 Channel Modeling Approaches

The channel modeling approaches available in the PLC literature, can be generally divided into two main categories. The first category is the physical or bottom-up approach, which is based on the electrical properties of the transmission line and electromagnetic theory of signal propagation in transmission lines. For this approach the knowledge of the topology of PLC system is usually needed. The other category is the parametric or top-down approach in channel modeling, which is mainly based on the channel characteristics and impulse responses acquired through measurements. Both of these approaches can be further divided into two subcategories of deterministic and stochastic approaches, as illustrated in Figure 2.3.

The physical deterministic approach uses the transmission line theory to establish the channel characteristics via ABCD-parameters or S-parameters [18]. Whereas, in physical stochastic models, the channel transfer function is derived from the network topology with stochastic generation of realistic electrical elements, e.g. in

[19], [20]. On the other hand, in parametric deterministic channel modeling, experimental measurements are used to characterize the channel. This results in models which are closer to the realistic scenarios, however, it needs a large data base to be able to model the channel. Parametric stochastic models characterize the impulse response of the channel via the collected data. PLC channel models developed in [2] and [3] are well-known examples of this approach. Table 2.3 summarizes the different channel modeling approaches and their features, such as modeling principles, measurement requirements, and the required level of topology knowledge [21].

In the following, we present two of the most well-known channel models in PLC literature, which we use later in this thesis. Firstly, the multipath PLC channel model which is an attempt to model the PLC channel in a parametric stochastic manner, which provides a reasonable model of the real PLC channel, and is used widely among the researchers. The second model is the ABCD-parameters of the channel based on the transmission line theory, which provides a physical deterministic approach of channel modeling.

2.3.2 The Multipath Model

The concept of multipath signal propagation in a PLC channel was first introduced in [3]. It has been shown in [18] that wave propagation on transmission lines is very similar to the uniform plane electromagnetic waves. this means that at any point in the space, the electric and the magnetic field intensity vectors lie in a plane, and the planes at any two different points are parallel to each other. Furthermore, the electric and magnetic field vectors are perpendicular to each other and both are perpendicular to the direction of wave propagation at any point in space. Such a field structure is known as transverse electromagnetic (TEM) field structure. However, inhomogeneities in the dielectric surrounding the transmission lines, results in a quasi-TEM mode of propagation. Nevertheless, the results are almost the same and the differences are negligible [18].

We assume an indoor power line communication system with multiple branches in the power lines from the PLC transmitter to the PLC receiver. Line junctions (branches) and unmatched terminals, including the open outlets in an indoor environment, are considered as line discontinuities in the system. On a line discontinuity, the propagated signal is partially reflected toward the transmitter and partially transmitted over the discontinuity. Partial reflection and transmission can be modeled by reflection coefficient $\rho(f)$, and transmission coefficient $\Gamma(f)$, respectively. It can be shown, that these coefficients depend on the frequency, on which the signal is propagating. Figure 2.4 illustrates the signal propagation in a PLC system with one branch. All the reflected and transmitted signals, generate multiple versions of the originally transmitted signal with different attenuations and interferences. Thus, the receiver collects different versions of the transmitted signal, each one corresponding to its own propagated path. The behavior of the electromagnetic propagated wave along the transmission line can be characterized

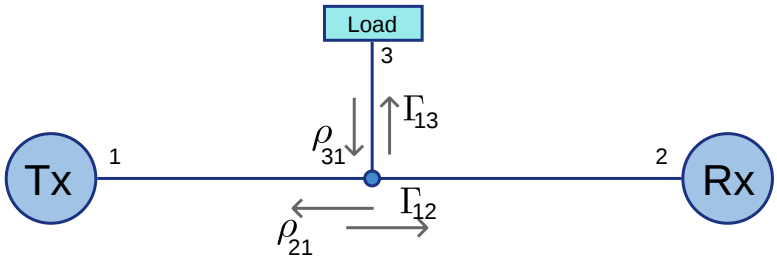


FIGURE 2.4: Multipath signal propagation in a PLC system with one branch.

by the complex-valued propagation constant $\gamma(f)$, which is defined as

$$\gamma(f) = \alpha(f) + j\beta(f), \tag{2.2}$$

where $\alpha(f)$ and $\beta(f)$ are the frequency dependent attenuation constant and phase constant, respectively. The attenuation constant $\alpha(f,\tau)$ in the multipath model is denoted by $\alpha_{MP}(f)$ and it is defined as

$$\alpha_{MP}(f) = a_0 + a_1 f^K, \tag{2.3}$$

where a_0, a_1, and K are the attenuation parameters of the transmission line and depend on the physical characteristics of the line and can be evaluated with extensive measurements. The phase constant $\beta(f)$ in the multipath model is denoted by $\beta_{MP}(f)$ and it is defined as

$$\beta_{MP}(f) = \frac{2\pi f}{\nu}, \tag{2.4}$$

where ν is the phase velocity. Therefore, the propagation constant in the multipath model can be expressed as

$$\begin{aligned} \gamma_{MP}(f) &= \alpha_{MP}(f) + j\beta_{MP}(f) \\ &= a_0 + a_1 f^K + j\frac{2\pi f}{\nu}. \end{aligned} \tag{2.5}$$

As the signal propagates through the power line conductors, at each line discontinuity, the reflected and the transmitted signal create new paths; therefore, multiple copies of the transmitted signal travel in multiple paths to the receiver. Therefore, based on (2.1) the received signal at the receiver can be modeled as [3]

$$\begin{aligned} Y(f,\tau) &= H_{MP}(f,\tau)X(f,\tau) + N(f,\tau) \\ &= \sum_{i=1}^{N_p} \left(\prod_{m=1}^{R_i} \rho_m(f) \prod_{n=1}^{T_i} \tau_n(f) \right) e^{-\gamma_{MP}(f)\ell_i} X(f,\tau) + N(f,\tau), \end{aligned} \tag{2.6}$$

where N_p is the number of paths that the transmitted signal is traveling before reaching the receiver, R_i and T_i are the number of reflection and transmission coefficients in the ith path respectively, and ℓ_i is the length of the ith path. Therefore,

the frequency response of the multipath model can be expressed as [3]

$$H_{MP}(f,\tau) = \sum_{i=1}^{N_p} \left(\prod_{m=1}^{R_i} \rho_m(f) \prod_{n=1}^{T_i} \tau_n(f) \right) e^{-\gamma_{MP}(f)\ell_i}. \tag{2.7}$$

In order to clarify the representation and divide the attenuation and delaying parts of the H_{MP}, (2.7) can be re-written as

$$H_{MP}(f,\tau) = \sum_{i=1}^{N_p} \underbrace{g_i}_{\text{weighting factor}} \cdot \underbrace{e^{-(a_0+a_1 f^K)\ell_i}}_{\text{attenuation factor}} \cdot \underbrace{e^{-j2\pi f(\ell_i/\nu)}}_{\text{delay portion}} \tag{2.8}$$

where g_i is the product of the reflection and transmission coefficients in and is known as the path gain, and is expressed as

$$g_i = \prod_{m=1}^{R_i} \rho_m(f) \prod_{n=1}^{T_i} \tau_n(f). \tag{2.9}$$

Since the statistical distribution of the product of a large number of uniform random variables approaches log-normality, we can model the path gain and consequently the impulse response of the channel as a complex log-normal distributed random variable [7]. The statistical properties of the channel gains are however time-variant. In particular, the average channel gains vary during transmission, which is a consequence of varying loads to the power line network.

2.3.3 The Transmission Line Model

Another well-known and heavily used model for PLC channels is the bottom-up deterministic approach based on the transmission line (TL) theory model. In indoor environments, the conductors which carry the power as well as the data signals are approximated to be ideal conductors and the dielectrics which contain them are approximated as uniform [18]. Therefore, the transverse electromagnetic (TEM) or quasi-TEM mode assumption is valid, and thus, the electrical quantities along the power line can be considered as scalars. The transmission line theory and the PLC channel model based on TL theory is described in [5], [6] among other resources. We briefly demonstrate the ABCD-parameter channel model based on the TL theory. In this model the power delivery cable is considered to be a series of infinitesimal parts which are connected to each other. Each part consists of resistance of the power delivery cables, conductance between different cables, capacitance between different cables, and finally self-inductance of each cable as well as mutual inductance between different cables. The resistance, conductance, capacitance and inductance parameters are called the primary parameters of the cable. These primary parameters can be expressed for each of the infinitesimal sections of the cable, as a per unit length (p.u.l) parameter.

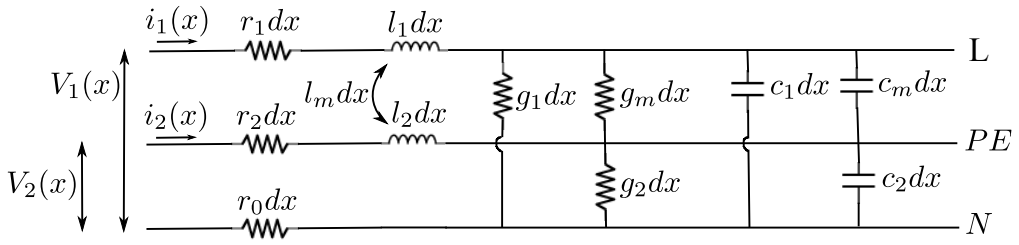


FIGURE 2.5: Equivalent p.u.l. representation of a three conductor PLC channel.

Figure 2.5 illustrates the equivalent representation of the p.u.l. parameters of the PLC channel. In this figure, L, PE, and N represent the Line, Protective Earth and Neutral cables of the power line system, respectively. $V_1(x)$ and $V_2(x)$ are the differential voltages between the conductors which contain the data signals and the reference conductor, whereas $i_1(x)$ and $i_2(x)$ are the currents. As mentioned before, the Kirchhoff's law prohibits the existence of more than two independent data signals when three conductors are available. The neutral conductor in the picture is considered to be the reference conductor. This has been chosen arbitrarily, and any other conductor could play the role of the reference conductor. r, g, c, and l represent the p.u.l. resistance, conductance, capacitance, and inductance, respectively, for each conductor.

The resistance, conductance, capacitance, and inductance matrices of the whole PLC channel can be composed from the p.u.l elements as

$$\mathbf{R} = \begin{bmatrix} r_0 & 0 & 0 \\ 0 & r_1 & 0 \\ 0 & 0 & r_2 \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} 0 & g_1 & g_2 \\ g_1 & 0 & g_m \\ g_2 & g_m & 0 \end{bmatrix}$$
$$\mathbf{C} = \begin{bmatrix} 0 & c_1 & c_2 \\ c_1 & 0 & c_m \\ c_2 & c_m & 0 \end{bmatrix} \quad \mathbf{L} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & l_1 & l_m \\ 0 & l_m & l_2 \end{bmatrix} \tag{2.10}$$

where each element of the matrix is the self or mutual parameter between different conductors. Indexes are 0 for reference conductor, 1 and 2 for the two data carrying conductors, and m for the mutual parameter between the conductors. From the cable parameter matrices, two important characteristics of the PLC channel can be derived, namely the characteristic impedance of the line and the propagation constant.

The characteristic impedance , Z_0, of a uniform transmission line is the ratio of the amplitudes of voltage and current of a single wave propagating along the line; and can be calculated as

$$Z_0 = \sqrt{\frac{\mathbf{R} + j2\pi f\mathbf{L}}{\mathbf{G} + j2\pi f\mathbf{C}}}. \tag{2.11}$$

The propagation constant, γ of the (2.2), which describes the behavior of the electromagnetic wave along a transmission line, can be calculated as

$$\gamma_{TL} = \sqrt{(\mathbf{R} + j2\pi f\mathbf{L})(\mathbf{G} + j2\pi f\mathbf{C})}. \tag{2.12}$$

We use the above-mentioned parameters of the transmission line to derive a PLC channel model, which is based on the ABCD-parameters of the transmission line. We are able to model each two nodes of the PLC system with input and output voltages $V_i(f)$ and $V_j(f)$, respectively, and corresponding currents $I_i(f)$ and $I_j(f)$, with ABCD-parameters as follows [22], [23]. Note that both voltages and currents are frequency dependent.

$$\begin{bmatrix} V_i(f) \\ I_i(f) \end{bmatrix} = \begin{bmatrix} A(f) & B(f) \\ C(f) & D(f) \end{bmatrix} \begin{bmatrix} V_j(f) \\ I_j(f) \end{bmatrix}. \tag{2.13}$$

The corresponding transfer function between these two nodes can be derived as

$$H_{TL}(f) \triangleq \frac{V_j(f)}{V_i(f)} = \frac{Z_j(f)}{A(f)Z_j(f) + B(f)}, \tag{2.14}$$

where $Z_j(f) = V_j(f)/I_j(f)$ is the impedance of node j.

The ABCD-parameters or transmission line parameters depend on the power line characteristics as well as the length ℓ of the transmission line, and can be calculated as

$$\begin{bmatrix} A(f) & B(f) \\ C(f) & D(f) \end{bmatrix} = \begin{bmatrix} \cosh(\gamma_{TL}\ell) & Z_0 \sinh(\gamma_{TL}\ell) \\ \frac{1}{Z_0}\sinh(\gamma_{TL}\ell) & \cosh(\gamma_{TL}\ell) \end{bmatrix}, \tag{2.15}$$

where Z_0 is the characteristic impedance of the transmission line per unit length as described in (2.11), and γ is the propagation constant as described in (2.12). The primary cable parameters depend on the physical cable characteristics, and are derived for a PLC channel in [24].

2.3.4 State of the Art in Channel Modeling

In addition to the two fundamental approaches to PLC channel modeling described above, namely the multipath model and the transmission line model, there exists numerous efforts in the literature in order to accurately model the PLC channels in different scenarios. These models can be categorized as Low Voltage (LV, 230/400 V), Medium Voltage (MV, between 1 and 35 kV), and High Voltage (HV, 110 kV and above) channel models. In the following we briefly describe the state of the art in LV, MV, and HV channel modelings.

The LV channel models are used for the access domain, which indicates the low voltage power distribution grid between the transformer stations and home connections, as well as the in-home communications. The echo-based channel model is developed for outdoor LV channels and verified in numerous applications [3], [25], [26]. This model is similar to the multipath model and is based on the reflections in

power supply networks, which causes the transmitted signal to travel different paths with different delays. For indoor LV channels different models have been proposed. The most important feature of the indoor LV channel is the strong sensitivity in frequency domain due to the impedance mismatch problems and their time variation. A deterministic bottom-up approach has been introduced in [27] and [28], which presents a model based on the transmission line model in terms of cascaded two-port networks. Another bottom-up approach is the statistical models, which define the channel parameters from the physical network features [19], [20], [29]–[31]. Other proposed models for the indoor channel adopt a top-down approach, which use measurement campaigns in order to model the channel in such way, that it corresponds to the real measurements [2], [32]–[36].

The MV channel models are used for the part of the power line, which connects the distribution substations that terminate the HV transmission network to the LV distribution transformers. The channel impairments encountered in MV lines are usually intermediate between those impairments encountered in LV and HV lines. Measurement-based characterization of the MV channels is done in two different levels: the component level and the network level. Component-level characterization generally takes place in a test lab [37], [38], while the network-level characterization generally takes place in a real MV distribution network [39]–[42]. Another approach for MV channel modeling is the theory-based characterization. In [43] the authors obtained a solution for single wire PLC and the case of multiwire was solved in [44] and [45]. In [46] this model was combined by the network topology in order to accurately predict the channel response. The HV channel models are used for the part of the power line between the generating power plants and the remote electrical substations. Some efforts for HV channel modeling can be found in [47]–[50]. Table 2.4 summarizes the state of the art in PLC channel modeling.

2.4 PLC Noise Modeling

2.4.1 Overview of PLC Noise

The PLC channel suffers from various types of noise and interference. Since power lines were not originally manufactured for the purpose of data transmission, admittedly, the noise which is available in the PLC channel is relatively harsh and needs to be properly modeled in order to provide a realistic understanding of the PLC system. The works dealing with noise characterization and modeling of PLC channels can be found extensively in the PLC literature, for instance in [51]–[55]. Generally, there are three kinds of noise in a PLC channel, which are described as follows.

1. White or colored background noise with a relatively low power spectral density (PSD), which is caused by summation of numerous noise sources of low power,
2. Narrow band interference, mostly amplitude modulated sinusoidal signals caused by other sources operating on the same frequencies,

Usage	Description	Literature
LV (outdoor)	echo-based model based on multipath model	[3], [25], [26]
LV (indoor)	deterministic bottom-up model based on the transmission line model	[27], [28]
LV (indoor)	statistical bottom-up model based on physical network features	[19], [20], [29]–[31]
LV (indoor)	top-down model based on measurements	[2], [32]–[36]
MV	measurement-based model component level	[37], [38]
MV	measurement-based model network level	[39]–[42]
MV	theory-based model	[43]–[46]
HV	echo-based models based on multipath model	[47]–[50]

TABLE 2.4: State of the art in PLC channel modeling.

3. Impulsive noise which can be either asynchronous or synchronous to the mains frequency, which is mostly caused by switched-mode power supplies.

Figure 2.6 illustrates the three kinds of PLC noise. All three kinds of noise are additive, which means they are added to the transmitted signal. In the following, we briefly demonstrate the modeling approaches for each kind of PLC noise.

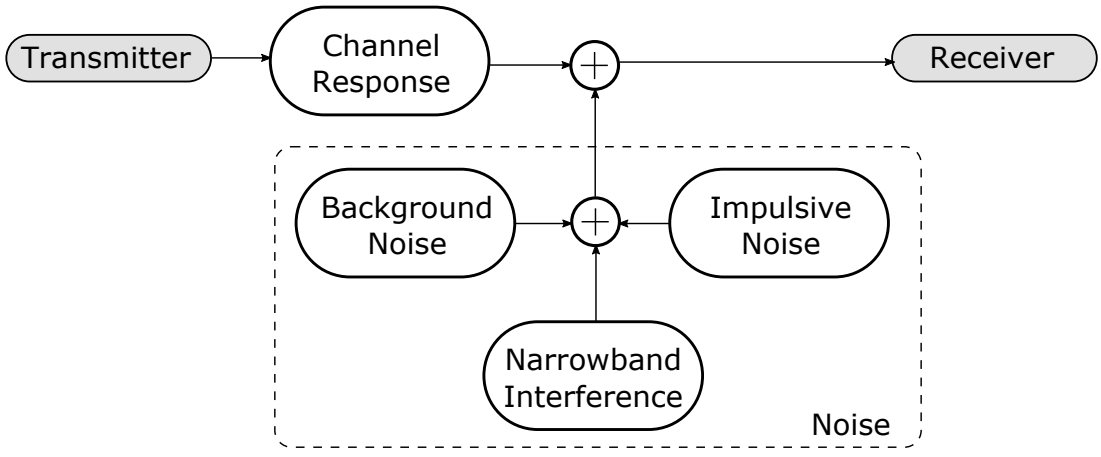


FIGURE 2.6: PLC noise.

2.4.2 Background Noise

The term background noise is given to a range of usually relatively low energy noise which is present in the PLC system at all times. The source of this noise can vary based on the environmental factors such as thermal noise of the system, but mostly it is due to the contribution of multiple noise sources of unknown origin. Some of them can be located even outside the considered premises and are coupled via radiation or via conduction [51].

The background noise consists of three main parts, which can be categorized as

1. Time-invariant continuous noise, which has a constant envelope for a long period of time, and its power is constant in all frequencies;
2. Time-variant continuous noise, which has an envelope that changes synchronously to the absolute voltage of the line, and its power is constant in all frequencies;
3. Colored noise, which has larger power in lower frequency regions.

The background noise with constant power in all frequencies is usually modeled as an additive white Gaussian noise (AWGN) process. This model is additive, that is, it is added to the signal, as well as it is white, that is, it has uniform power across the frequency band of the PLC system. The probability distribution of this noise in time domain is normal or Gaussian distribution. Let us denote the AWGN component of the PLC noise $N(f, \tau)$ in (2.1) as N_{AWGN}. The probability density function of N_{AWGN} can then be expressed as

$$f_{AWGN}(x) = \frac{1}{\sigma_{AWGN}\sqrt{2\pi}} \exp\left(-\frac{(x-\mu_{AWGN})^2}{2\sigma_{AWGN}^2}\right), \tag{2.16}$$

where μ_{AWGN} is the mean value and σ_{AWGN} is the standard deviation of the AWGN process. An exemplary sample values of PLC background noise with $\mu_x = 0$ and $\sigma^2 = 4$, and the corresponding probability density function is depicted in Figure 2.7.

The colored noise has larger power in lower frequency regions and hence cannot be modeled by an AWGN process. The larger power in lower frequency regions is due to the larger attenuation in higher frequencies of the propagation between each noise source and a receiver. Moreover, many noise sources have more power in lower frequency regions. This noise is usually modeled by pink noise process. The PSD of an exemplary white and an exemplary pink noise is illustrated in Figure 2.8.

2.4.3 Narrowband Interference

The narrowband interference in PLC channels is mostly formed by sinusoidal or modulated signals with different origins such as broadcast stations, disturbances caused by electrical appliances with a transmitter or a receiver, etc. The intensity of this noise usually varies by the time of the day and the environment in which the PLC system is operating. Let us denote the narrowband interference component of

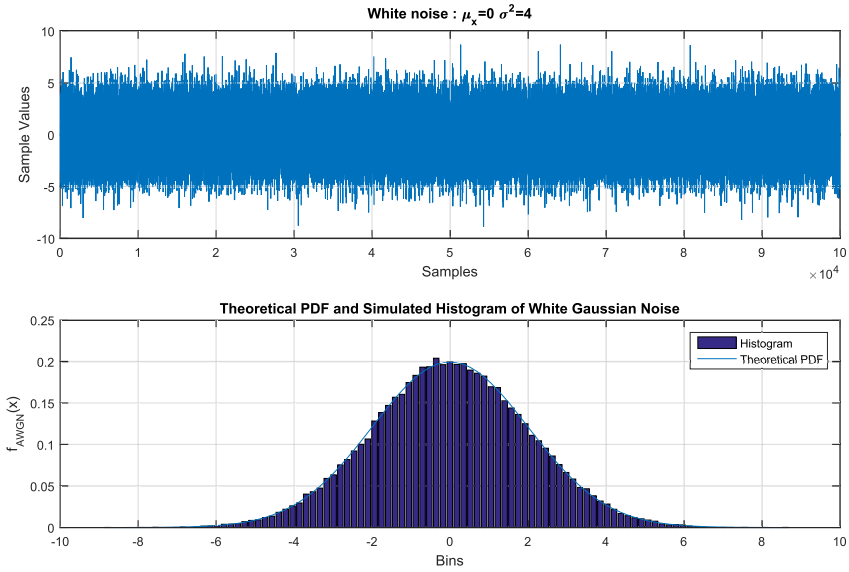


FIGURE 2.7: Noise samples and the probability density function of PLC background noise.

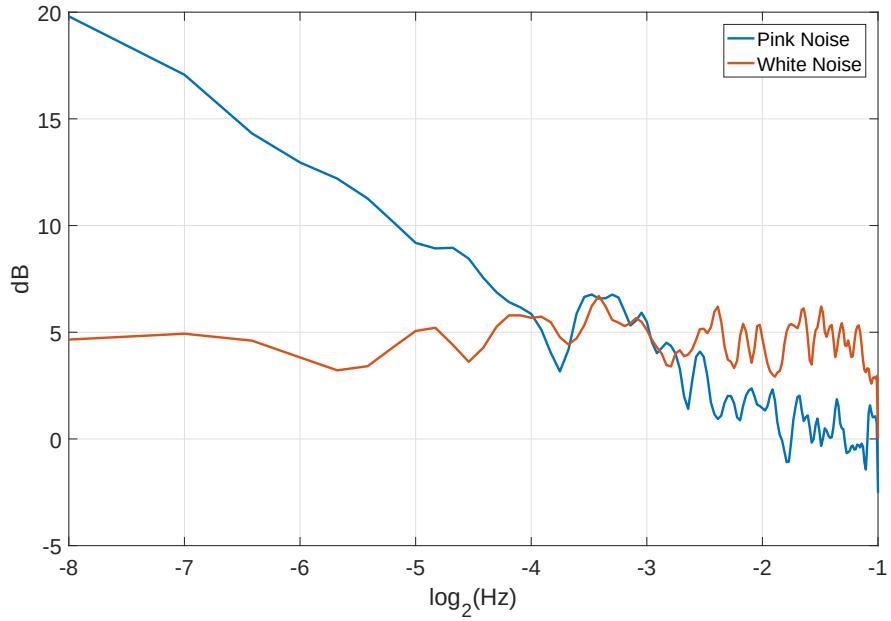


FIGURE 2.8: Power spectral density of exemplary white and colored (pink) background noise.

the PLC noise $N(f,\tau)$ in (2.1) as N_{NB}. Then, N_{NB} can be modeled as a modulated sinusoidal signal with interference frequency f_{NB} which can be expressed as

$$N_{NB} = A_{NB}\cos(2\pi f_{NB}t + \phi_{NB}), \tag{2.17}$$

where A_{NB} is the amplitude of the narrowband interference and ϕ_{NB} is the phase. Narrowband interference can cause degradation in the performance of the PLC and is usually considered as a serious problem for data transmission. Coding techniques as well as proper notching of the frequency spectrum of PLC can help to mitigate the narrowband interference.

2.4.4 Impulsive Noise

The PLC channel suffers from a class of noise known as the impulsive noise. This kind of noise has a relative high amplitude but occurs is a very short period of time. Generally there are three kinds of impulsive noise available in the PLC channel [56]:

1. Periodic impulsive noise synchronous with the mains: It is a cyclostationary noise, synchronous with the mains and with a frequency of 50 Hz in Europe.
2. Periodic impulsive noise asynchronous with the mains: This noise although periodic, does not show synchronization behavior with the frequency of the mains.
3. Asynchronous impulsive noise: This noise has an unpredictable nature, with no regular occurrence and is mainly due to transients caused by the connection and disconnection of electrical devices.

The impulsive noise and its behavior can be expressed in closed-form equations. One of the most utilized models for impulsive noise is the Middleton's model developed for the electromagnetic interference in communication channels [57]. Three general classes of interference is categorized by Middleton which the first class i.e. class A is often used to model the impulsive interference and shows characteristics which is close to that of a real PLC channel. The Middleton class A model is described by its amplitude probability density function which is defined as

$$f_{IN}(x) = e^{-A}\sum_{m=0}^{\infty}\frac{A^m}{m!\sqrt{2\pi\sigma_{IN}^2}}\exp\left(-\frac{x^2}{2\sigma_{IN}^2}\right), \tag{2.18}$$

where A is the *impulsive index* which is the product of the average number of impulses per unit time and the mean duration of the emitted impulses and σ_{IN}^2 is the noise variance and is defined as

$$\sigma_{IN}^2 = \left(\sigma_G^2 + \sigma_I^2\right)\frac{m/A + \Gamma}{1+\Gamma}, \tag{2.19}$$

where σ_G^2 and σ_I^2 are the Gaussian and impulsive variance respectively and $\Gamma = \sigma_G^2/\sigma_I^2$ is the Gaussian to impulsive noise power ratio.

Middleton's description of impulsive noise models the impulses from different interference sources as Poisson distributed random variables. It is noteworthy that (2.18) includes both the impacts of impulsive noise as well as the thermal noise in the communication system. The parameters A and Γ define the extent of impulsive noise in the system. For a small A (close to zero), the probability density function becomes very close to a Gaussian distribution and impulsive noise has a very low impact. Moreover, for A and Γ close to one the probability density function is very impulsive and the impulsive noise has a great impact on the communication system [58].

The Middleton class A interference channel can also be interpreted as a stationary two-state model, with infinite number of memoryless channels with common input and output alphabets [7]. The channel is considered to be either in state s_0 or s_1 with corresponding variances σ_0^2 and σ_1^2. The variance of the two states can be calculated by (2.19) as

$$\sigma_0^2 = \sigma_G^2 \tag{2.20a}$$

$$\sigma_1^2 = \sigma_G^2 + \frac{\sigma_I^2}{A}. \tag{2.20b}$$

The two-state representation of the channel is demonstrated in Figure 2.9. In a channel with infinite states the probability π_m of each state can be written as

$$\pi_m = e^{-A} \cdot \frac{A^m}{m!}, m \geq 0. \tag{2.21}$$

In a two-state channel, the probability of each state can be normalized by the factor of $e^{-A}(1+A)$ resulting in the two probabilities as

$$\pi_0 = \frac{1}{1+A} = \frac{1-A}{1-A^2} = 1 - A \tag{2.22a}$$

$$\pi_1 = \frac{A}{1+A} = \frac{A-A^2}{1-A^2} = A, \tag{2.22b}$$

where the approximation has been made due to the fact that $A^2 \ll 1$ and $A^2 \ll A$. Therefore, for each symbol there is a probability of being in either of two states with different variances, where the first state is purely Gaussian and the second state is a combination of Gaussian noise and impulsive noise.

Moreover, the periodic impulsive noise which is synchronous to the mains is considered to be a *cyclostationary* process and the periodic instantaneous noise power is derived in [59] as

$$\sigma_N^2(t) = \sum_{l=0}^{L-1} A_l \left| \sin\left(\frac{2\pi t}{T_{AC}} + \theta_l\right) \right|^{n_l}, \tag{2.23}$$

where L represents the number of noise classes (for narrowband PLC, $L = 3$ [59]), A_l, θ_l, and n_l are different characteristic parameters of the l-th noise class, and

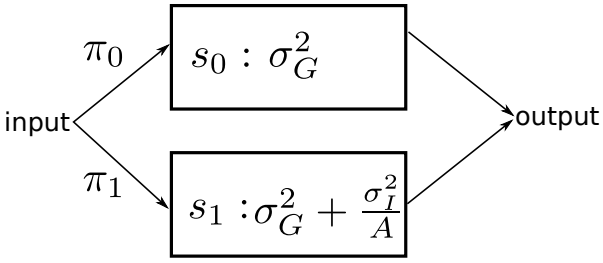


FIGURE 2.9: A two-state Markov chain demonstration of Middleton class A model

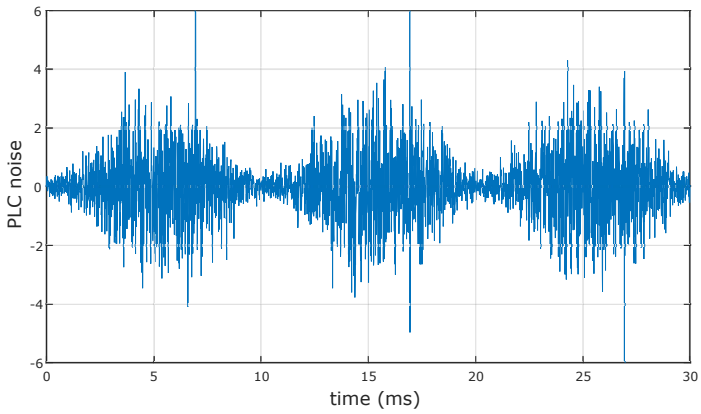


FIGURE 2.10: Periodic impulsive noise synchronous with the mains in time domain.

T_{AC} is the period of the mains voltage. Figure 2.10 illustrates the simulated PLC cyclostationary noise in time domain. The periodic nature of the noise can be clearly seen in this figure.

2.4.5 State of the Art in Noise Modeling

There is a body of work available in the PLC literature, which aims at modeling the PLC noise in different environment and scenarios. In this section, we briefly describe the different approached in PLC noise modeling. There are two main approaches in PLC noise modeling, which are referred to as the statistical-physical and the empirical noise modeling approaches. The statistical-physical modeling approach utilizes the physical characteristics of the PLC channel in order to derive the statistics of the interference at the receiver. The Middleton Class A model [57], [60], as described above, is a common statistical-physical model of the PLC noise. In [61] the authors have derived a canonical statistical-physical model of the instantaneous statistics of asynchronous noise based on the physical properties of the PLC network. A statistical PLC noise model suitable for low voltage networks has been derived in [62].

Empirical noise modeling approach, on the other hand, uses the collected noise data from field measurements in order to propose models that match the characteristics

exhibited by the collected data. The PLC noise parameters have been experimentally investigated in [54], [63]–[65]. In [66] and [67] the authors fit the collected noise data to the Middleton Class A statistical model.

2.5 PLC Transmitter and Receiver

2.5.1 Physical Layer Overview

The transmitter and receiver structure of the physical (PHY) layer of a PLC system is standardized by several standardization entities such as IEEE [11] and ITU-T [13] among others. The main objective of such standardizations is to establish a secure and reliable communication between nodes of a PLC network. Generally, the transmitter is composed of a scrambler followed by an error-correction encoder module, followed by an interleaver, followed by a modulation mapper and finally an OFDM modulator module. The generated signal then, after applying appropriate transmit filter, is transmitted through the PLC channel. The block diagram of a PLC transmitter is illustrated in Figure 2.11. In this figure, the encoder module consists of three error-correction encoders, namely Reed-Solomon encoder, convolutional encoder and repetition encoder. The received signal from the PLC channel, is processed by the PLC receiver. The receiver of the PLC system is the corresponding component for each module of the transmitter in reverse order. Therefore, the receiver consists of an OFDM demodulator, followed by a demodulation module, followed by a deinterleaver, followed by a decoder module, and finally a descrambler. In the following, the design of the important parts of the PLC transmitter, as well as some simulation results are presented. In the rest of the thesis, the structure of the transmitter and receiver which is presented in this section, is used as the basis of our simulation results.

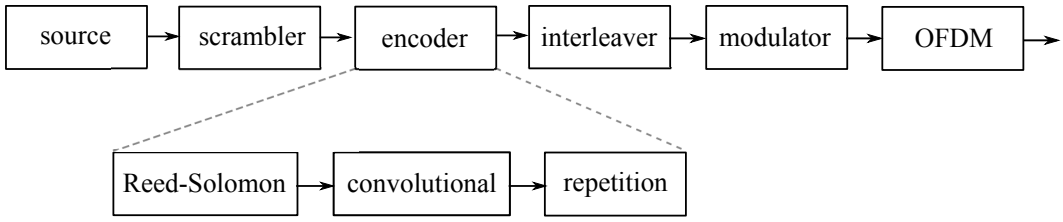
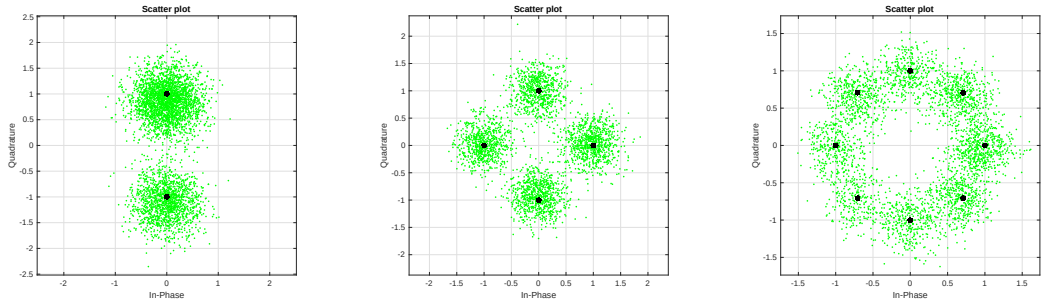


FIGURE 2.11: Block diagram of the PHY layer transmitter of a PLC system.

2.5.2 Modulation and Coding

With the help of modulation and coding design for PLC systems, a simple and robust communication system is achieved. Since the channel conditions are usually not known at the receiver, sophisticated demodulator features cannot be used effectively at the PLC receiver. The modulation and coding of the PLC system is



(A) BPSK (B) QPSK (C) 8-PSK

FIGURE 2.12: Constellation diagram of the received PLC signal with modulation schemes of BPSK, QPSK, and 8-PSK.

usually designed to comply with the local regulations, such as CENELEC, FCC, etc. Spread-frequency shift keying (S-FSK) as introduced by Shaub [68], is one of the robust modulations which has been used for PLC applications. Moreover, modulation schemes such as binary phase shift keying (BPSK) and differential BPSK can be used as well. Higher orders of phase shift keying (PSK), such as quadrature PSK (QPSK), 8-PSK, and their differential modes are also possible. The choice of the modulation scheme is based on the characterization of the channel and the quality of service (QoS)requirements. The differential mode, makes the design and implementation of the detector easier, since in the differential mode, the phase synchronization is not necessary and each symbol has a phase which only depends on the phase of the previous symbol. The coherent mode of modulation is also supported by the standards but the differential mode is used more commonly due to its simpler implementation in the system.

The constellation diagram of the received PLC signal with modulation schemes of BPSK, QPSK, and 8-PSK ate illustrated in Figure 2.12. It can be seen in this Figure that the distance between different received signals with 8-PSK modulation is smaller than that of BPSK/QPSK modulations, which results in a higher probability of error. However, since three bits are modulated as an 8-PSK symbol, two bits are modulated as a QAM symbol, and one bit is modulated as a BPSK symbol, the throughput which is achieved by the 8-PSK modulation is higher. Therefore, a trade-off between the data throughput and the BER performance of the communication system is to be considered, where achieving higher data throughputs results in a higher BER and consequently a lower performance. As a result, the transmitter can decide the modulation scheme based on the requirements of the transmission and the channel conditions.

Error correcting codes play an important role in communication systems with severe channel error conditions. The combination of modulation and coding provides a robust transmission scheme for data transmission in PLC channels. The PLC encoder module consists of a block of concatenated codes made of Reed-Solomon encoder and convolutional encoder, followed by an optional repetition encoder. The repetition encoder can secure a more reliable communication at the cost of lower

data throughput, since the repetition coding divides the data rate by a factor equal to the spreading factor of the encoder. Therefore, the repetition encoder is only used as an optional feature in severe situations and low signal to noise ratios.

2.5.3 OFDM

Multicarrier transmission is a commonly used communication technology, which achieves high rates by transmitting a set of parallel signals at low symbol rates [69]. Multicarrier transmission is able to achieve channel capacity through optimally allocating the available transmission power among subcarriers by means of the water-filling algorithm [70]. Multicarrier transmission can be implemented in different ways, such as vector coding [71], or orthogonal frequency division multiplexing (OFDM) [72]. These techniques are based on the same premise, which is dividing a wide channel into multiple parallel narrowband channels by means of orthogonal channel partitions. In the PLC systems, OFDM modulation and demodulation is applied in order to encode orthogonal signals of digital data in multiple subcarriers to be transmitted with different frequencies through the PLC channel. The combination of advanced channel coding techniques with OFDM modulation facilitates a very robust communication over the power line channel. The data signals before being modulated by the OFDM modulator are orthogonal to each other. This orthogonality prevents interference between subcarriers and facilitates the transmission of subcarriers at the same time.

Let us assume that the available PLC bandwidth is divided by N subcarriers with the same distance in the frequency from each other. We can denote the time-domain transmit signal in complex baseband as [73]

$$s(t) = \frac{1}{\sqrt{N}} \sum_{k} \sum_{n=0}^{N-1} a_n[k] g_n(t - kT_u), \tag{2.24}$$

where k is the time index, $a_n[k]$ is the k-th data symbol of the n-th subcarrier, $T_u = NT$ is the OFDM symbol duration, and $g_n(t)$ is the pulse shaping function of the n-th subcarrier, which in OFDM modulation is defined as

$$g_n(t) = \text{rect}\left(\frac{t}{T_u} - 0.5\right) \exp(j2\pi f_n t), \tag{2.25}$$

where $f_n = n/T_u$ is the center frequency of the n-th subcarrier. The exponential term in (2.25) corresponds to a shift of f_n is frequency domain. Therefore, the subcarrier spacing in frequency domain is $1/T_u$.

Every two subcarriers are said to be orthogonal. The orthogonality of subcarriers i and j is defined as [73]

$$\frac{1}{T_u} \int_0^{T_u} g_i(t) g_j^*(t) dt = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \tag{2.26}$$

Because of this orthogonality, the received data can be separated with a matched filter. Les us denote the received signal from the PLC channel as $r(t)$. Then, the output of the matched filter can be expressed as

$$y_n(t) = \frac{1}{T_u} \int_{-\infty}^{\infty} r(\tau) g_n^*(t-\tau) d\tau. \tag{2.27}$$

After sampling at $t = (k+1)T_u$, (2.28) becomes

$$y_n[k] = \frac{1}{T_u} \int_{kT_u}^{(k+1)T_u} r(\tau) \exp(-j2\pi(n/T_u)\tau) d\tau. \tag{2.28}$$

Because of the orthogonality between the subcarriers, no inter-symbol interference (ISI) is observed and the transmission can be regarded as N parallel transmissions.

The power spectral density of an OFDM transmit signal is the sum of the power spectral densities of each subcarrier. Therefore, the overall power spectral density can be expressed as [73]

$$P(f) = \sum_{n=0}^{N-1} |G(f - f_n)|^2, \tag{2.29}$$

where $G(f)$ is the Fourier transform of the rect function in (2.25). (2.29) can be further calculated as

$$P(f) = \sum_{n=0}^{N-1} \frac{\sin^2(\pi(f-f_n)T_u)}{(\pi(f-f_n)T_u)^2}. \tag{2.30}$$

Figure 2.13 illustrates the simulated power spectral density of a transmitted PLC signal in FCC-above-CENELEC frequency band with QPSK modulation, as an example. For the simulations, the frequency regulations of the IEEE 1901.2 standard for narrowband PLC has been applied and the FCC-above-CENELEC frequency band is chosen for data transmission. According to this standard, the FCC-above-CENELC operating band consists of 72 used subcarriers with a total amount of 256 subcarriers. The frequency specifications as well as the OFDM parameters of the used spectrum in this example is listed in Table 2.5.

The so-called side-lobes in Figure 2.13 are unwanted and are referred to as out of band radiation. This radiation can be reduced by proper windowing at the transmitter. The window function is applied to each OFDM symbol separately. This function can be chosen freely based on the out-of-band emission requirements. Figure 2.14 illustrates the application of a raised-cosine window (also known as the Hann window) to one OFDM symbol. Let us assume $N_t = N_{CP} + M$. The raised-cosine window is then defined as

$$w[n] = \begin{cases} 0.5\left(1 - \cos\left(\frac{\pi n}{N_w}\right)\right) & 0 \le n \le N_w \\ 1 & N_w + 1 \le n \le N_t - N_w - 1 \\ 0.5\left(1 - \cos\left(\frac{\pi(n-N_t)}{N_w}\right)\right) & N_t - N - w \le n \le N_t - 1 \end{cases} \tag{2.31}$$

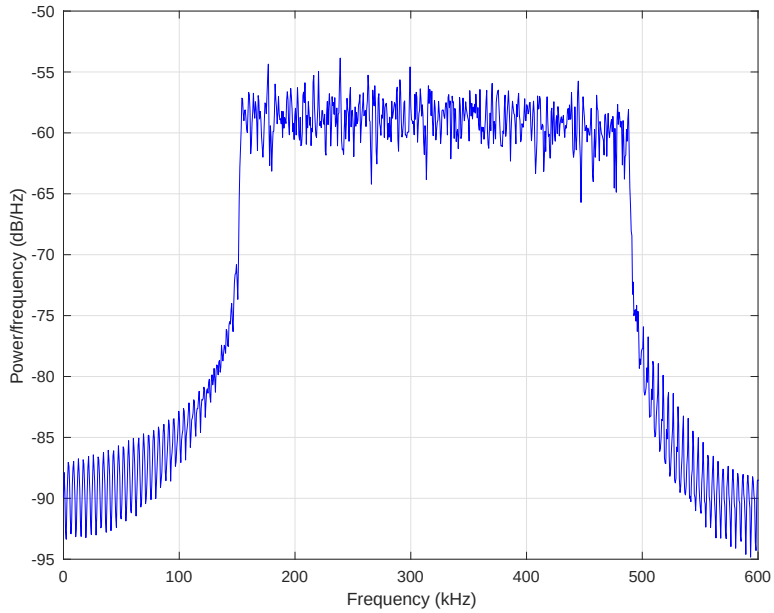


FIGURE 2.13: Spectrum of a transmitted PLC signal based on the specifications of Table 2.5.

Parameter	Value
First subcarrier	154.6875 kHz
Last subcarrier	487.5 kHz
Number of subcarriers	72
Inter-carrier spacing	4.6875 kHz
FFT-length	256
Number of cyclic prefix	30
Sampling frequency	1.2 Mhz
N_w	8

TABLE 2.5: An example of the OFDM parameters of a narrowband PLC system based on the IEEE 1901.2 standard.

Furthermore, it has been shown that the OFDM transmit signal of (2.24) can be produced with inverse discrete Fourier transform (IDFT), and (2.28) at the receiver can be realized by a discrete Fourier transform (DFT). IDFT and DFT can be implemented by inverse fast Fourier transform (IFFT) and fast Fourier transform (FFT), respectively, which have a considerably faster performance.

2.6 Chapter Conclusion

In this chapter, an overview of the PLC system and channel model has been presented. The PLC regulations and frequency specifications has been discussed for

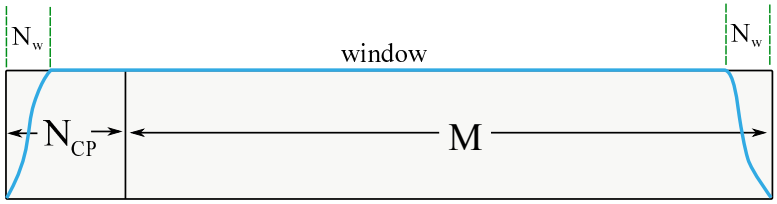


FIGURE 2.14: Applying a raised-cosine window to the OFDM symbol.

both narrowband and broadband PLC. In later chapters, both of the narrowband and broadband PLC and their corresponding frequency specifications are used for different multichannel transmission scenarios. Two of the most common channel models for PLC, namely the multipath PLC channel model and the channel model based on the transmission line theory, has been discussed in this chapter, which makes the basis of the channel models used in later chapters of the thesis. Moreover, the PLC channel noise and its models were introduced in this chapter. A complete PLC transmitter and receiver model has been presented and different parts of it has been discussed, which will be the basis of the system model for the simulations presented in later chapters of this thesis.

Chapter 3

MIMO PLC

3.1 Chapter Overview

Electric power lines normally consist of multiple conductors. In an in-home environment with one-phase power delivery structure, there are three conductors available, namely Line (L), Neutral (N), and Protective Earth (PE). Moreover, in a multi-phase power delivery structure, the number of phases in the system is more than one. The conventional PLC system, uses the phase and neutral conductors in order to construct a differential voltage between them to utilize for the purpose of data transmission. However, the existence of multiple conductors in power line infrastructures, gives rise to the idea of multiple spatial transmission channels. Therefore, multiple differential data signals between multiple conductors can be constructed, which can transmit several data signals at the same time. However, due to the Kirchhoff's law, with $N+1$ conductors, only N independent data signals can be constructed and transmitted at the same time. The availability of multiple transmission channels, enables the use of multiple-input multiple-output (MIMO) transmission in broadband PLC.

MIMO transmission has been heavily investigated in wireless communications literature and has shown significant performance improvement and increased data throughputs [74], [75]. These results encouraged researchers to investigate the benefits of MIMO transmission in the field of PLC as well. The first large scale public measurement results on MIMO power line channel and noise characteristics was published in 2008 [76]. MIMO PLC has been introduced in [77], and has been further investigated in numerous publications such as [76], [78]–[82] to name a few. The efforts in MIMO channel modeling and system analysis in the PLC literature, have been aptly summarized in [21]. It has been shown in these literatures, that MIMO PLC is capable of increasing the throughput of the transmission or improving the reliability and performance of the transmission. These, can be achieved by two separate methods:

1. Spatial multiplexing, which is a transmission technique to transmit independent and separately encoded data signals, through each of the MIMO PLC channels. Therefore, the space dimension is reused, and hence the throughput is increased.

2. Spatial diversity, which is a transmission technique to transmit multiple, redundant copies of the data signal, through each of the MIMO PLC channels. Therefore, the reliability of the transmission is increased which results in a better performance.

However, MIMO PLC is not without its challenges. Particularly, the capacitive and inductive coupling between the power line conductors would result in a spatial correlation between the MIMO channels, which in turn, would degrade the performance of the transmission. Another challenging aspect of the MIMO PLC is the derivation of the channel capacity in the presence of the impulsive noise which will be addressed later in this chapter.

In this chapter, MIMO PLC system and its challenges are analyzed and discussed. The rest of the chapter is organized as follows. Section 3.2 describes the coupling methods which enables the power line system to utilize MIMO transmission are described. The system model of a MIMO PLC transmission is described and analyzed in Section 3.3. Spatial correlation in MIMO PLC systems and its effect on the performance of the data transmission are described in 3.4. In Section 3.5 the capacity of a MIMO PLC system modem different levels of information availability at the PLC transmitter and receiver is derived. Finally, Section 3.6 concludes the chapter.

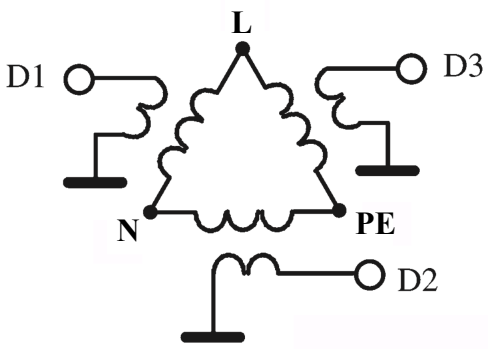


FIGURE 3.1: Delta-style coupling method.

3.2 MIMO PLC Coupling Methods

Before describing a MIMO PLC transmission, it is important to describe the MIMO PLC coupling methods. The idea of MIMO transmission is to transmit high frequency data signals over power-carrying transmission lines, which were originally

designed for electricity transmission at low frequencies. For this reason, couplers are used to connect the communication equipments to the power grid. Generally, there are two types of couplers: inductive and capacitive. Inductive couplers guarantee a balance between the lines, whereas capacitive couplers often introduce asymmetries due to component manufacturing tolerances [21]. PLC couplers designed for high voltage (HV, 110 kV to 380 kV) and medium voltage (MV, 10 kV to 30 kV) can be found in [83]. PLC couplers designed for low voltage (LV, 110 V to 400 V) single-input single-output (SISO) transmission can be found in [84]. In the following, we describe the MIMO couplers for LV PLC applications.

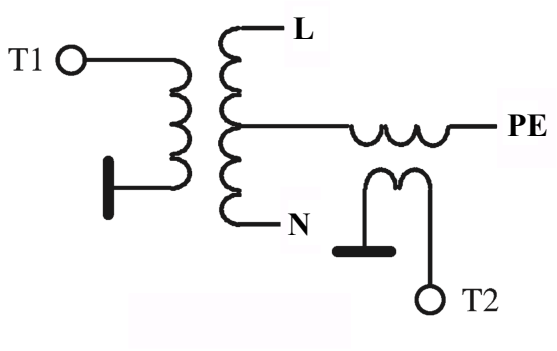


FIGURE 3.2: T-style coupling method.

In order to avoid the radiated emission in PLC transmission, the coupler design is preferred to be as symmetric as possible. This way, 180 degrees out of phase electric fields are generated that neutralize each other, resulting in reduced emission. This desired symmetrical way of propagation is also known as differential mode (DM). In case of asymmetries, some part of the injected signal turns into a current mode (CM) signal. Normally, the devices connected to the power line network are also able to produce asymmetries in the system [85].

Figure 3.1 illustrates the internal design of a delta-style coupler [86]. This design is also known as transversal probe. In delta-style design, the phase, neutral and protective Earth conductors at the location of signal injection are connected to each other by means of electrical baluns, which forms a triangle. The data signals are injected separately by means of additional baluns coupled with the othe baluns, which are denoted by D_1, D_2, and D_3 in Figure 3.1. However, the sum of the three injected voltages is zero and therefore, only two of them can be independent.

Figure 3.2 illustrates the internal design of a T-style coupler [87]. In this design, phase, neutral, and protective Earth conductors are each connected to one end of a balun, whereas the other end of the baluns are connected to each other ot one point. Two data signals are injected in a T-style coupler: one coupled with the baluns of the phase and neutral conductors, and the second one coupled to the balun of the protective Earth conductor. These injected signals are denoted by T_1 and T_2 in Figure 3.2.

Finally, the star-style coupler is illustrated in Figure 3.3 [86]. In this coupler design, the injected signals are coupled to inter-connected baluns in the form of a star.

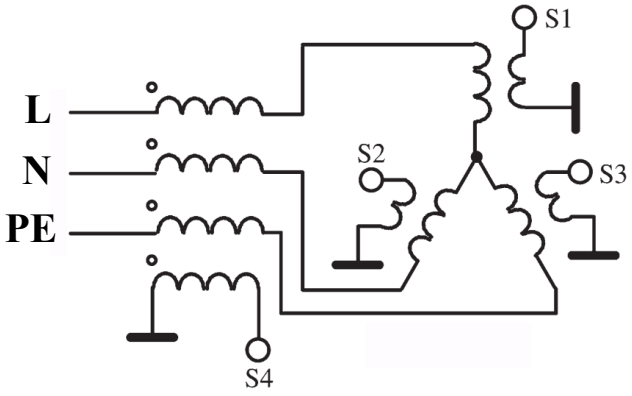


FIGURE 3.3: Star-style coupling method.

These data signals are denoted by S_1, S_2, and S_3 in Figure 3.3. Furthermore, those baluns are each connected to the phase, neutral, and protective Earth conductors, which together make a common mode signal, denoted by S_4 in Figure 3.3. The star-style coupling is normally used at the receiver of a PLC transmission system. Further details on the coupling methods and their advantages and disadvantages are presented in [88].

3.3 MIMO Transmission

3.3.1 MIMO System Model

MIMO transmission exploits the spatial diversity in order to improve the performance of the communication system, or rather improves the data rate with respect to conventional SISO transmission. In wireless communication, MIMO is enabled by the use of multiple antennas at both transmitter and receiver. MIMO can be also applied to PLC. Figure 3.4 illustrates a three wire installation of a PLC system as well as the transmitted or received data signals which are represented as differential voltages. A data signal can be constructed as the differential voltage values between two of the three available conductors with one of the conductors as a voltage reference. Therefore, with three conductors, three differential voltages of Δv_1, Δv_2, and Δv_3 can be produced, which are the differential voltages between phase and neutral, neutral and protective Earth, and phase and protective Earth, respectively. However, due to the Kirchhoff's law, only two of these three signals can be independent from each other and carry information data and the third one depends on the other two and cannot carry independent information.

MIMO PLC was first proposed to be used for multiple phase installations [89]. In this case, multiple channels were established between two uncoupled line wires and therefore the transmitter and the receiver must have access to all the three phase wires. Later, the use of MIMO PLC has been extended to one-phase installations,

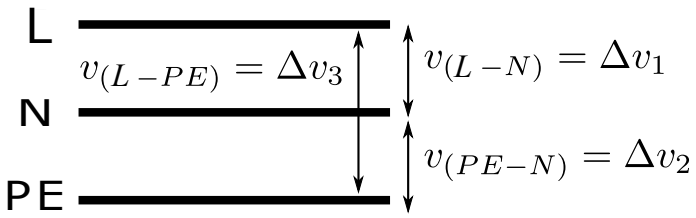


FIGURE 3.4: Three differential voltages of PLC constructing three different signals.

where the multiple channels were established between line, neutral, and protective earth wires. The latter is the basis of the MIMO system model in this thesis. In a one-phase power line installation, at the PLC transmitter, three transmit signals can be generated between the three available power line wires. However, at the PLC receiver, four receive signals can be received, with the fourth signal being the common mode (CM) signal between all the power line wires.

As an example of a MIMO PLC transmission, let us assume two transmit signals at the PLC transmitter and two received signals at the PLC receiver, which constructs a 2×2 MIMO transmission system. Figure 3.5 illustrates a MIMO PLC system model with two transmit signals as well as two received signals, constructing four MIMO transmission channels, where $\Delta v_{i,T_x}$ and $\Delta v_{i,R_x}$ are the differential voltages at the transmitter and at the receiver, respectively. The four MIMO channels h_{11}, h_{12}, h_{21}, and h_{22} match each transmitter port to the corresponding receiver port.

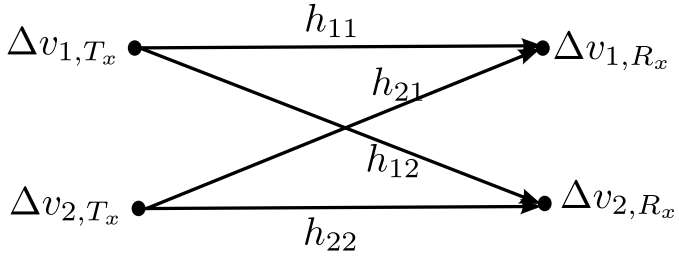


FIGURE 3.5: A MIMO PLC channel with two transmitters and two receivers.

In a general sense, let us assume a MIMO PLC system with n_T differential transmit signals and n_R differential received signals. Due to electromagnetic coupling between adjacent wires in a power line infrastructure, the transmitted signal from any of the signal feeds can be received and detected to different extents in all of the reception signal ports. Therefore, the i-th received signal from the i-th port at the receiver, when $i \in \{0, 1, \cdots, n_R - 1\}$, is the summation of all the transmitted signals with different weights and can be represented as

$$y_i(t) = \sum_{j=1}^{n_T} h_{ij}(t) x_j(t) + n_i(t), \tag{3.1}$$

where $h_{ij}(t)$ is the channel gain from the i-th received signal and the jth transmitted signal, $x_j(t)$ is the j-th transmitted signal from the j-th port at the transmitter,

when $j \in \{0, 1, \cdots, n_T - 1\}$, and $n_i(t)$ is the channel noise corresponding to the ith received signal.

In order to be able to demonstrate all the received signals, transmitted signals, channel gains, and noise terms as one equation, the matrix format representation of the MIMO system can be used which is shown bellow (time index is omitted due to readability)

$$\mathbf{Y} = \mathbf{HX} + \mathbf{N} \tag{3.2}$$

where $\mathbf{Y} \in \mathbb{C}^{n_R \times 1}$ is the vertical vector of received signals, $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ is the channel gain matrix, $\mathbf{X} \in \mathbb{C}^{n_T \times 1}$ is the vertical vector of transmitted signals and $\mathbf{N} \in \mathbb{C}^{n_R \times 1}$ is the vertical vector of received noise. The channel matrix $\mathbf{H}$ is deterministic and assumed to be constant at all times. The system model of such a MIMO system can be expressed as

$$\begin{bmatrix} y_1(t) \\ \vdots \\ y_{n_R}(t) \end{bmatrix} = \begin{bmatrix} h_{11}(t) & \cdots & h_{1n_T}(t) \\ \vdots & \ddots & \vdots \\ h_{n_R1}(t) & \cdots & h_{n_Rn_T}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ \vdots \\ x_{n_T}(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ \vdots \\ n_{n_R}(t) \end{bmatrix}. \tag{3.3}$$

Based on this system model, the transmission of data signals in PLC systems, can be categorized into four categories:

1. Single-input single-output (SISO) system, in which $n_T = 1$ and $n_R = 1$,
2. Single-input multiple-output (SIMO) system, in which $n_T = 1$ and $n_R > 1$,
3. Multiple-input single-output (MISO) system, in which $n_T > 1$ and $n_R = 1$, and
4. Multiple-input multiple-output (MIMO) system, in which $n_T > 1$ and $n_R > 1$.

3.3.2 MIMO PLC Channel Modeling

The MIMO PLC channel has been characterized from the results of measurement campaigns e.g. in [78], [90]–[93]. In these papers, a statistical analysis of the experiment data was performed and the MIMO channel were characterized according to this analysis. The modeling of the MIMO channel can be performed by top-down or bottom-up approaches. A frequency-domain top-down statistical MIMO PLC channel model was proposed in [90], which is based on the multipath propagation model described in the previous chapter. A time-domain top-down statistical MIMO PLC channel model was proposed in [91]. The bottom-up approach based on the transmission line theory is also used in order to characterize the MIMO PLC channel, e.g. in [94]. In this case the multi-conductor transmission line theory is used to characterize the pul parameters of the network and hence model the MIMO channel.

3.3.3 Spatial Multiplexing

In spatial multiplexing, each spatial MIMO channel carries independent data signals. Therefore, the data rate of the system is increased and more data can be transmitted in a certain amount of time. Note that, all these sub-channels function in the same allocated bandwidth and are just separated spatially from each other. This means that using spatial multiplexing does nor cost additional bandwidth and power for the communication system. The multiplexing gain is referred to as the degrees of freedom with reference to signal space constellation [95]. The degree of freedom can be calculated as

$$DOF = \min(n_T, n_R), \tag{3.4}$$

where n_T and n_R are the number of transmit and receive PLC ports, respectively.

In order to recover the transmitted data signals at the PLC receiver, it is necessary to perform a considerable amount of signal processing. First the MIMO system decoder must estimate the individual channel transfer functions in order to determine the channel transfer matrix. Once all of this has been estimated, then the matrix $\mathbf{H}$ has been produced and the transmitted data streams can be reconstructed by multiplying the received vector with the inverse of the transfer matrix.

3.3.4 Spatial Diversity

In spatial diversity techniques, same information is sent across independent PLC channels to combat fading. When multiple copies of the same data are sent across independently fading channels, the amount of fade suffered by each copy of the data will be different. This guarantees that at least one of the copies will suffer less fading compared to rest of the copies. Thus, the chance of properly receiving the transmitted data increases. In effect, this improves the reliability of the entire system. This also reduces the co-channel interference significantly.

Receive diversity is achieved by obtaining a transmitted signal at multiple receiver ports. This technique uses the maximum ratio combining (MRC) [73], which weighs the received signals y_n, $n \in \{1, \cdots, N\}$, by the square root of the SNRs γ_n $n \in \{1, \cdots, N\}$, such that

$$y = \sum_{n=1}^{N} \gamma_n y_n. \tag{3.5}$$

Therefore, the main contribution to the received signal comes from the channel with the highest SNR.

Transmit diversity is achieved by transmitting multiple versions of the same signal at the PLC transmitter. This scheme can be utilized by, for instance, space-time block coding (STBC) [96]–[98].

For a MIMO PLC system with n_T transmitters and n_R receivers, the maximum number of diversity paths is $n_T \times n_R$. Note that, the spatial diversity technique does

not increase the transmission throughput, since the data is repeated in transmission. However, this technique is able to increase the reliability of the transmission compared to SISO transmission, with the same amount of data throughput. Figure 3.6 illustrates a MIMO transmission example and demonstrates the difference between spatial multiplexing and spatial diversity.

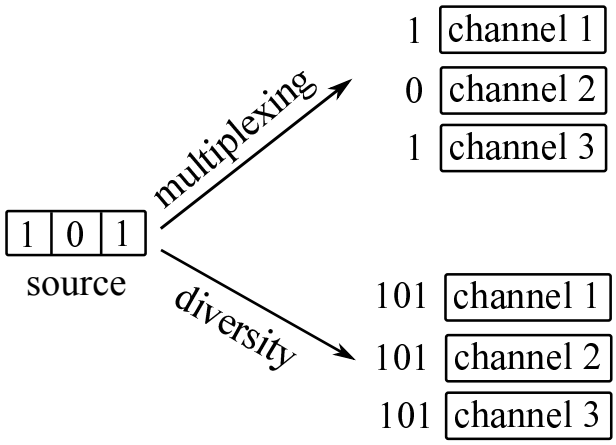


FIGURE 3.6: Spatial multiplexing and transmit spatial diversity for a MIMO system with three channels and a source of three bits.

3.4 Spatial Correlation in MIMO PLC

In this section we address one key aspect of MIMO PLC channels, namely the spatial correlation, where the term spatial refers to the multiple input and output ports. Spatial correlation is the phenomenon when the changes in voltage in one conductor can be detected in another conductor which lies in the vicinity of the first conductor. This can be the result of capacitive and inductive coupling between conductors which are in each others proximity. The inductive and capacitive coupling cause a change in voltage of one conductor affect the change of voltage in another conductor. Since the MIMO channels in PLC are based on the changes in voltage, capacitive and inductive coupling can cause the MIMO channels to influence each other which is known as the spatial correlation.

Spatial correlation can directly influence the capacity of the MIMO channel. The more the correlation is between the MIMO channels, the more the degradation of the capacity. Therefore, it is important to include this aspect of the PLC MIMO channels in the channel modeling approach. There has been some mentions of the spatial correlation for MIMO PLC in the literature [80], [99].

In order to mathematically analyze the spatial correlation and its effects on the MIMO PLC channel, we have to first evaluate the singular values of the channel matrix in frequency domain and then derive the covariance matrix of the channel. On any given subcarrier, we can decompose the channel matrix $\mathbf{H}$ to its singular

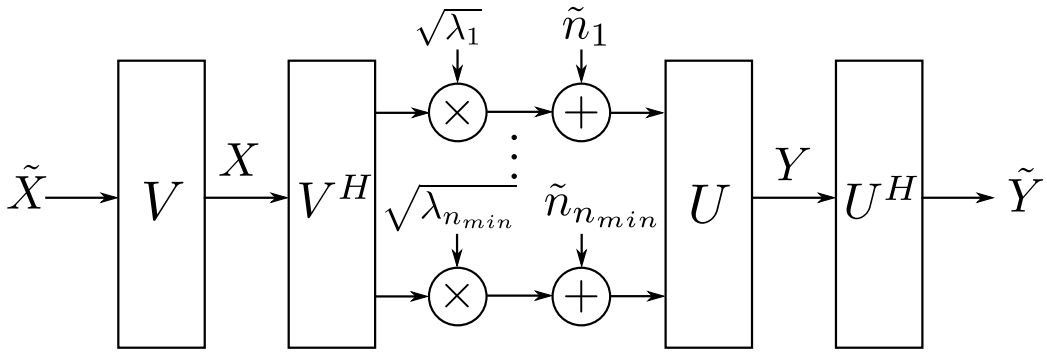


FIGURE 3.7: Singular value decomposition of the MIMO channel.

values with singular value decomposition (SVD) defined as [99]

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{3.6}$$

where the operator $(.)^H$ represents the Hermitian transpose or complex conjugate transpose. Matrices $\mathbf{U}$ and $\mathbf{V}$ are two unity matrices which are defined as $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$ and $\mathbf{V}\mathbf{V}^H = \mathbf{V}^H\mathbf{V} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix. Matrix $\mathbf{\Sigma}$ is a diagonal matrix which its diagonal elements are the singular values of $\mathbf{H}$ and are shown as $\lambda^{(i)}$, and its off-diagonal elements are zero. Figure 3.7 illustrates the singular value decomposition of the channel matrix which transforms the MIMO channel into several independent SISO channels.

The spatial correlation makes the available MIMO channels to be dependent to some extent on each other. This dependency can be demonstrated in the channel *covariance matrix* $\mathbf{R}_h$ which is defined as [99]

$$\mathbf{R}_h = E\left[\text{vect}(\mathbf{H})\text{vect}(\mathbf{H})^H\right], \tag{3.7}$$

where the operator $E(.)$ is the mathematical expectation and the operator vect(.) aligns all the columns of a given matrix with size $n_R \times n_T$ to form a vertical column of size $n_R n_T \times 1$. We use the most common correlation model, namely the Kronecker model to decompose the channel matrix into three different matrices as follows. The only assumption for the Kronecker channel decomposition is that the MIMO channels are only correlated at the two link extremities and not in between. Under these circumstances, the MIMO channel can be decomposed as [99]

$$\mathbf{H} = K \cdot \mathbf{R}_r^{1/2}\tilde{\mathbf{H}}\mathbf{R}_t^{1/2}, \tag{3.8}$$

where K is a constant introducing an overall channel gain, $\tilde{\mathbf{H}}$ is the channel matrix of a correlation-free channel with elements $\tilde{h}_{ij} \sim \mathcal{CN}(0,1)$ which have complex Gaussian distribution with zero mean and unit variance and are completely independent from each other. The matrices $\mathbf{R}_r^{1/2}$ and $\mathbf{R}_t^{1/2}$ are called the receiver and transmitter correlation matrices, respectively, and can be arbitrarily scaled. In case

we chose the normalization [99]

$$\begin{aligned} tr(\mathbf{R}_t^{1/2}) &= n_T \\ tr(\mathbf{R}_r^{1/2}) &= n_R, \end{aligned} \tag{3.9}$$

the channel covariance matrix of (3.7) can be re-written as

$$\mathbf{R}_h = K^2 \cdot \mathbf{R}_t \otimes \mathbf{R}_r, \tag{3.10}$$

where the operator $\otimes$ is the Kronecker product. Therefore, in order to model the covariance matrix of the channel and consequently the spatial correlation of the MIMO channel, we only need to model the transmitter and receiver correlation matrices $\mathbf{R}_t$ and $\mathbf{R}_r$. The elements of $\mathbf{R}_t$ and $\mathbf{R}_r$ can be directly modeled based on the measurements performed in a real situation or they can be modeled mathematically to generate results which are similar to those observed in the real measurements.

3.5 MIMO PLC Channel Capacity Analysis

3.5.1 PLC Channel Capacity

One of the fundamental characteristics of a communication channel is the capacity of the channel. The performance of any communication system is impaired by interference, or noise. One of the measures of calculating the extent of this impairment is the capacity of the channel. Calculating the capacity of the channel is a well studied subject in the literature and capacity formulations are available for both SISO and MIMO systems in wireless communications [73]. Unfortunately, the same can not be said about power line communications. The presence of impulsive noise in the PLC channel and the complicated models of the impulsive channel makes this task a challenging endeavor. The capacity of SISO PLC systems with Middleton class A impulsive noise has been studied in [100] and an equation for the channel capacity has been derived. Capacity of MIMO channels in wireless communication has been derived as well and published in the literature such as [101].

In this section, we derive the capacity of a MIMO PLC channel with Middleton class A impulsive noise. We consider two cases of channel state information availability at transmitter, i.e. the case when the channel state information is available at the transmitter as well as the case when the channel state information is not available at the transmitter. For both cases we assume that the channel information is available at the receiver. We will investigate the effect of possessing the knowledge about channel state information at transmitter on the channel capacity. The impulsive MIMO PLC channel is modeled by Middleton class A model as a two state channel as described before in this chapter. It is shown that in such channels, the overall capacity is the summation of the weighted capacity of the channel in each of the states, and a general term for the capacity of MIMO PLC is derived [81].

In this section we demonstrate the channel capacity of a SISO PLC system as described in [100]. Let us assume a finite state system with the capacity of channel in state m defined as C_m and the probability of the channel being in that state as π_m. We can obtain the average channel capacity C as [102]

$$C = \sum_{m=0}^{M} \pi_m C_m, \tag{3.11}$$

where M is the number of states. The PLC channel, as shown previously, can be regarded as a two-state model as illustrated in the previous chapter. The capacity of the AWGN channel in state m is defined as [100]

$$C_m = B \cdot \log_2 \left(1 + \frac{S}{N}\right) \quad [bits/sec] \tag{3.12}$$

where B is the bandwidth, S is the total signal power over the bandwidth and $N = 2\sigma_m^2$ is the total noise power over the bandwidth in state m.

The channel capacity of a SISO PLC system with impulsive noise and channel state information available at transmitter has been derived in [100]. It has been shown using the equation (3.11) that the capacity of SISO channel can be written as

$$\begin{aligned} C_{SISO} &= \sum_{m=0}^{M} \pi_m \cdot B \cdot \log_2 \left(1 + \frac{P}{\sigma_m^2}\right) \\ &= B \cdot e^{-A} \cdot \sum_{m=0}^{M} \frac{A^m}{m!} \cdot \log_2 \left(1 + \frac{P}{\sigma_m^2}\right), \end{aligned} \tag{3.13}$$

where B is the bandwidth, P is the transmitted signal power and σ_m^2 is the noise power in state m. The derivation of the capacity is supported by the fact that Middleton class A model can be viewed as a Markov chain, which gives a simple theoretical expression of the channel capacity of a SISO PLC channel.

3.5.2 Decoupling MIMO Channels and Waterfilling Algorithm

In order to calculate the capacity of a MIMO PLC channel, first we decompose the channel into a set of parallel, independent sub-channels using the singular value decomposition (SVD) method. The SVD of the channel matrix $\mathbf{H}$ can be written as [73]

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{3.14}$$

where $\mathbf{U} \in \mathbb{C}^{n_R \times n_R}$ and $\mathbf{V} \in \mathbb{C}^{n_T \times n_T}$ are unitary matrices and $\mathbf{\Sigma} \in \mathbb{C}^{n_R \times n_T}$ is a rectangular matrix whose diagonal elements are non-negative real numbers and whose off-diagonal elements are zero.

The diagonal elements of $\mathbf{\Sigma}$ are the singular values of the channel matrix $\mathbf{H}$ and can be shown as $\{\sigma_1, \sigma_2, \cdots, \sigma_{n_{min}}\}$ or equivalently $\{\sigma_i\}_{i=1}^{n_{min}}$ where $n_{min} = \min(n_T, n_R)$. Since

$$\mathbf{H}\mathbf{H}^H = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^H\mathbf{U}^H = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H, \tag{3.15}$$

where $\mathbf{\Lambda} \in \mathbb{C}^{n_R \times n_R}$ is a diagonal matrix which its diagonal elements are the eigenvalues $\{\lambda_i\}_{i=1}^{n_R}$ of the matrix $\mathbf{H}\mathbf{H}^H$ or equivalently the squared singular values σ_i^2 of the matrix $\mathbf{H}$.

When channel state information (CSI) is available at the transmitter side, channel decomposition to a set of parallel sub-channels can be performed as shown in Figure 3.7. For decomposing the channel into a set of parallel sub-channels, we will first define

$$\tilde{\mathbf{X}} \stackrel{\text{def}}{=} \mathbf{V}^H\mathbf{X}, \tag{3.16a}$$

$$\tilde{\mathbf{Y}} \stackrel{\text{def}}{=} \mathbf{U}^H\mathbf{Y}, \tag{3.16b}$$

$$\tilde{\mathbf{N}} \stackrel{\text{def}}{=} \mathbf{U}^H\mathbf{N}, \tag{3.16c}$$

then we can rewrite the channel (3.2) as:

$$\tilde{\mathbf{Y}} = \mathbf{\Sigma}\tilde{\mathbf{X}} + \tilde{\mathbf{N}}, \tag{3.17}$$

where $\tilde{\mathbf{N}}$ has the same distribution as $\mathbf{N}$. Therefore, we have the equivalent representation of MIMO channel as a parallel set of virtual SISO sub-channels as shown in Figure 3.7.

The power allocation at the PLC transmitter, which results in the optimal capacity is provided by the water filling algorithm [103]. The water filling algorithm calculates the inverse channel SNRs prior to the power allocation and determines a water level μ, such that the area under μ is equal to the power budget P_0. The difference between the inverse channel SNR and the water level μ is the allocated transmit power. Channels whose inverse SNR is not reached by the water level μ are deselected from transmission and are not allocated any transmit power. The largest transmit power is allocated to channels with higher SNRs.

3.5.3 MIMO Capacity with Different Information Availability

3.5.3.1 CSI Known at Transmitter

In this section we derive the channel capacity of a MIMO PLC system when the channel state information (CSI) is available at the transmitter.

In order to calculate the MIMO channel capacity, first we have to calculate the auto-correlation matrix of $\mathbf{Y}$ as

$$\begin{aligned} \mathbf{R}_{yy} &= E\left\{\mathbf{Y}\mathbf{Y}^H\right\} \\ &= E\left\{(\mathbf{H}\mathbf{X}+\mathbf{N})\left(\mathbf{H}^H\mathbf{X}^H+\mathbf{N}^H\right)\right\} \\ &= \mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H+\sigma_m^2\mathbf{I}_{n_R} \end{aligned} \tag{3.18}$$

where $\mathbf{R}_{xx} = E\{\mathbf{X}\mathbf{X}^H\}$ is the auto-correlation matrix of the transmitted signal vector and $E\{\mathbf{N}\mathbf{N}^H\} = \sigma_m^2$ is the power spectral density of the additive impulsive noise and $\mathbf{I}$ is the identity matrix.

It can be proven [73] that the mutual information between the input and output signals can be expressed as

$$I(\mathbf{X};\mathbf{Y}) = \log_2 \det\left(\mathbf{I}_{n_R} + \frac{1}{\sigma_m^2}\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H\right). \tag{3.19}$$

From (3.18) and (3.19), the capacity of the MIMO channel can be written as

$$C = \max_{\{\mathbf{R}_{xx}\}} \log_2 \det\left(\mathbf{I}_{n_R} + \frac{1}{\sigma_m^2}\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H\right). \tag{3.20}$$

The channel capacity of the equivalent parallel sub-channels can be then calculated [73] using equation (3.20). The capacity of the ith virtual SISO sub-channel in state m can be written as

$$C_{m,i} = B \cdot \log_2\left(1 + \frac{P_i\lambda_i}{\sigma_m^2}\right), \tag{3.21}$$

where P_i is the the power allocated to the ith channel and must satisfy the total power constraint $\sum_{i=1}^{n_T} P_i = n_T$. The MIMO channel capacity in state m is then the sum of the capacities of the virtual SISO channels as

$$C_m = \sum_{i=1}^{n_{min}} C_{m,i} = B \cdot \sum_{i=1}^{n_{min}} \log_2\left(1 + \frac{P_i\lambda_i}{\sigma_m^2}\right). \tag{3.22}$$

The capacity of the MIMO channel can be derived using the equations (3.11) and (3.22) as

$$C_{MIMO} = B \cdot e^{-A} \cdot \sum_{m=0}^{M}\sum_{i=1}^{n_{min}} \frac{A^m}{m!} \cdot \log_2\left(1 + \frac{P_i\lambda_i}{\sigma_m^2}\right). \tag{3.23}$$

The capacity in (3.23) can be maximized by solving a power allocation problem and using the water-filling algorithm [73]. The channel capacity of a MIMO PLC channel when CSI is available at the transmitter side as a function of SNR is given in Figure 3.9. It can be seen that with increasing the A parameter (impulsive index), the channel capacity will reduce. In the simulation of this case, we have used the

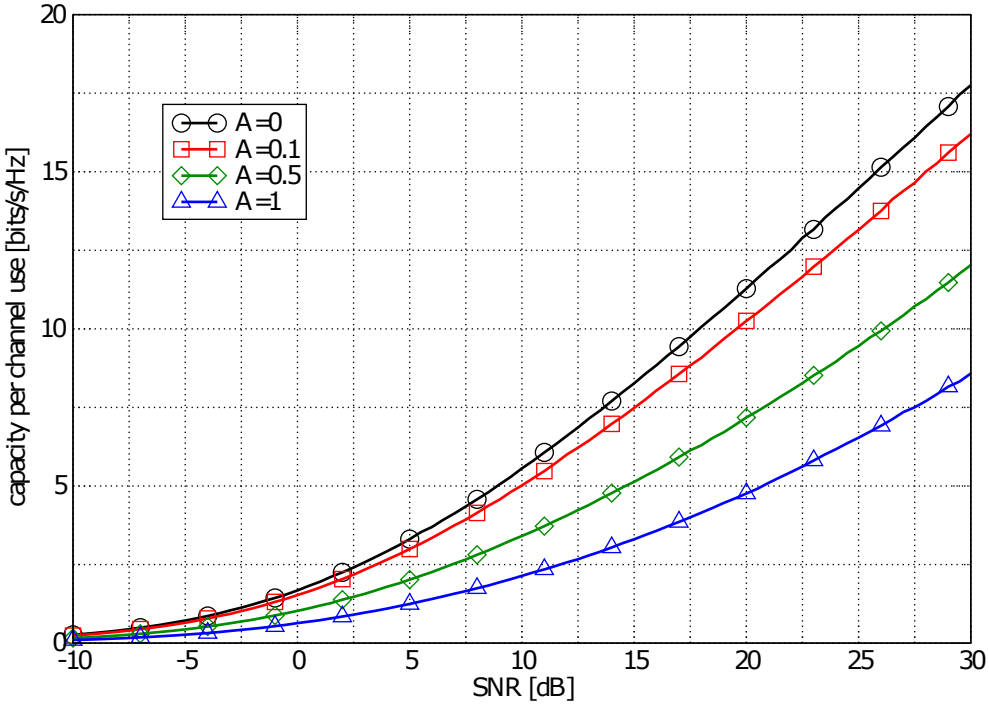


FIGURE 3.8: MIMO PLC channel capacity when CSI is not available at the transmitter with different A parameters and $\Gamma = 0.01$

water-pouring power allocation algorithm. Therefore, more power is allocated to the modes with higher SNR.

3.5.3.2 CSI Unknown at Transmitter

When CSI is known at both the transmitter and the receiver, each sub-channel has its own corresponding transmitted power which depends on the SNR of that channel. However, in the case where CSI is not known at the transmitter side, the transmitted power has to be distributed equally in all of the sub-channels. Therefore the auto-correlation function of the transmitted signals is given as

$$\mathbf{R}_{xx} = \mathbf{I}_{n_T}. \tag{3.24}$$

This might not be the optimal distribution of the input correlation matrix, since the channel gains are not Rayleigh distributed and the additive noise is not purely Gaussian. However, without the CSI it is reasonable to distribute the power equally in all the sub-channels. Consequently, the channel capacity of (3.20) can be rewritten as

$$C = \log_2 \det\left(\mathbf{I}_{n_R} + \frac{1}{\sigma_m^2}\mathbf{H}\mathbf{H}^H\right). \tag{3.25}$$

Using the decomposition of $\mathbf{H}\mathbf{H}^H = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ and the identity $\det(\mathrm{I}_m + \mathbf{A}\mathbf{B}) = \det(\mathrm{I}_n + \mathbf{B}\mathbf{A})$, where $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $\mathbf{B} \in \mathbb{C}^{n\times m}$, the channel capacity can be

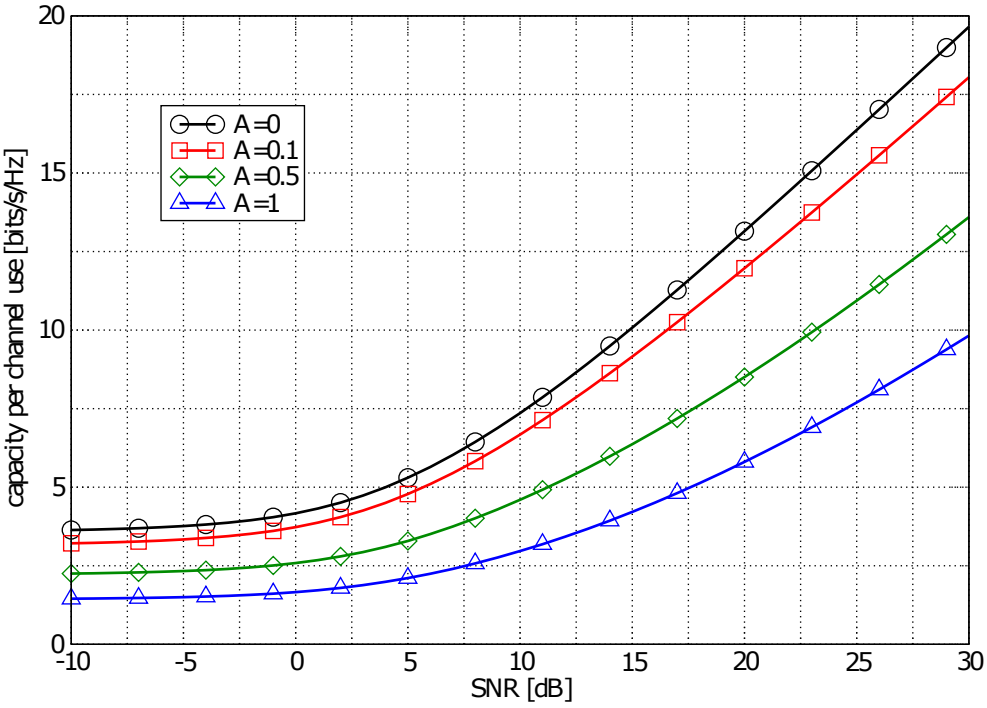


FIGURE 3.9: MIMO PLC channel capacity when CSI is available at transmitter with different A parameters and $\Gamma = 0.01$

expressed as

$$C = \log_2 \det \left(\mathbf{I}_{n_R} + \frac{1}{\sigma_m^2} \mathbf{\Lambda} \right). \tag{3.26}$$

Therefore, the channel capacity in the state m can be written

$$C_m = B \cdot \sum_{i=1}^{n_{min}} \log_2 \left(1 + \frac{\lambda_i}{\sigma_m^2} \right), \tag{3.27}$$

and consequently the capacity of the MIMO channel can be represented as

$$C_{MIMO} = B \cdot e^{-A} \cdot \sum_{m=0}^{M} \sum_{i=1}^{n_{min}} \frac{A^m}{m!} \cdot \log_2 \left(1 + \frac{\lambda_i}{\sigma_m^2} \right). \tag{3.28}$$

It can be seen that in this case, the MIMO channel has been converted into n_{min} virtual SISO channels with the channel gain λ_i for the ith SISO channel. The channel capacity in (3.28) is the special case of the capacity in (3.23) where CSI is not available at the transmitter and the power is equally allocated to all the channels hence $P_i = 1$ for $i = 1, 2, \cdots, n_{min}$.

The channel capacity of the MIMO PLC channel when CSI is not available at the transmitter as a function of SNR is given is in Figure 3.8 and Figure 3.10. Figure 3.8 shows the variation in channel capacity as varying the parameter A in the Middleton model, and Figure 3.10 shows the variation in channel capacity as varying

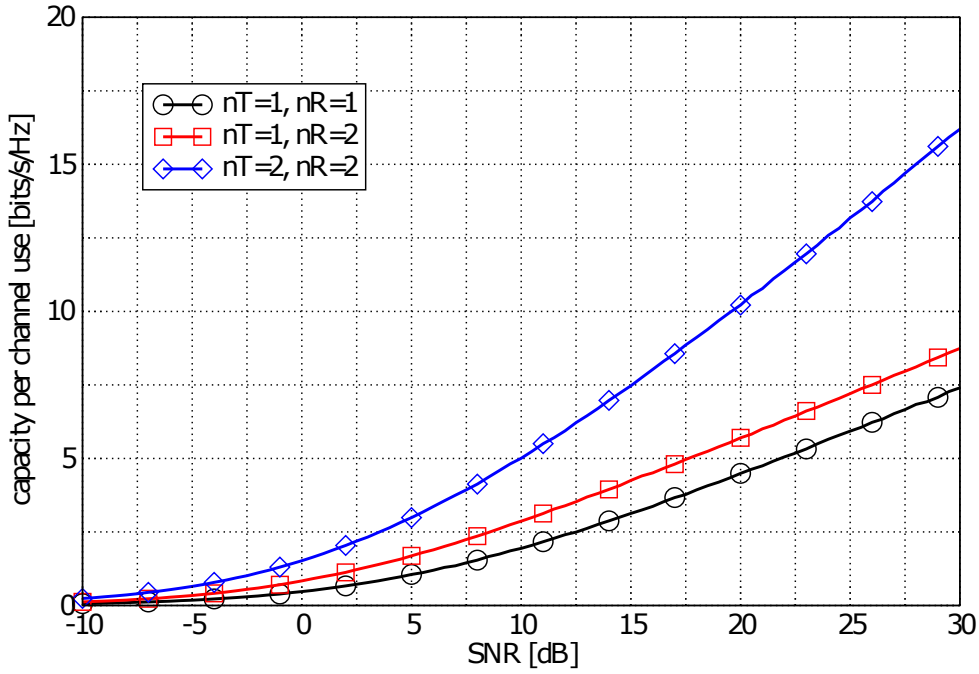


FIGURE 3.10: MIMO PLC channel capacity when CSI is not available at the transmitter with different numbers of transmitters and receivers.

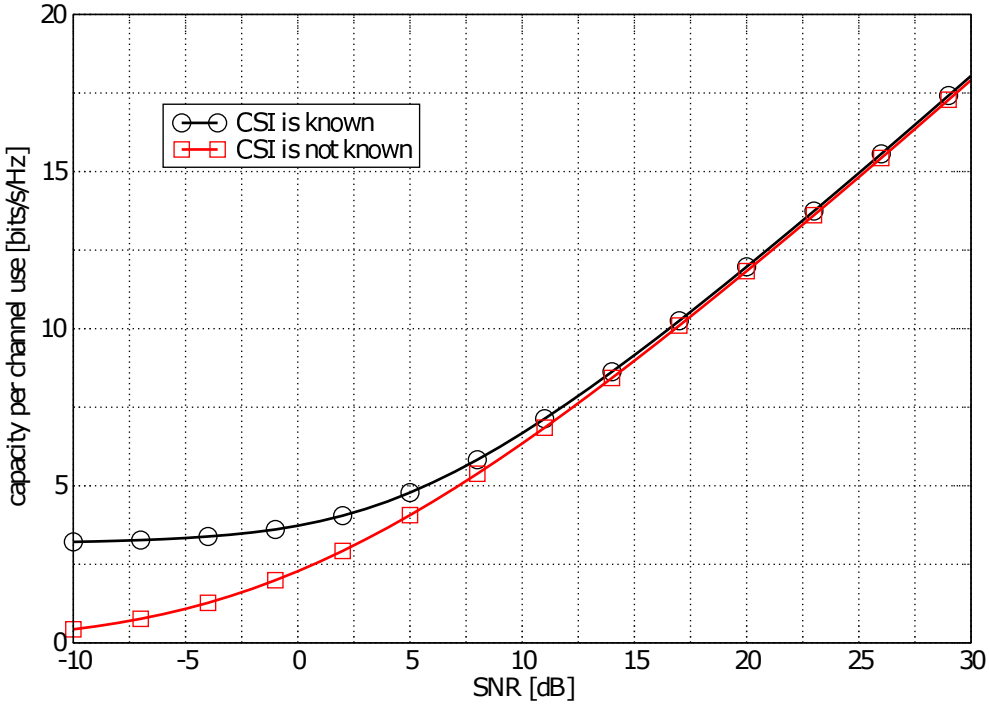


FIGURE 3.11: MIMO PLC channel capacity comparison between having CSI and not having CSI.

the number of transmit and receive feeds of PLC. It can be seen that increasing A results in a decreased channel capacity for the channel is more impulsive. On the other hand increasing the number of channels will increase the MIMO capacity as expected.

Figure 3.11 demonstrates the comparison between the case with known CSI and the unknown CSI. It can be seen that the availability of CSI at transmitter provides us with the needed knowledge for applying water-filling algorithm and hence maximum capacity is increased, specially in low SNR regions.

3.6 Chapter Conclusion

In this chapter, MIMO PLC as a major technique for multichannel transmission in PLC applications has been discussed. Spatial multiplexing and spatial diversity as two MIMO techniques have been described, where the former increases the data throughput of the communication system and the latter improves the performance and reliability of the transmission. Moreover, spatial correlation as the most important impairment in MIMO PLC systems has been described and it impact on the performance of the transmission has been analyzed. Finally, the MIMO PLC channel capacity in the presence of the impulsive noise has been derived for both the cases when the channel state information is available and is not available at the PLC transmitter, and the novel mathematical expressions for these capacities have been presented.

Chapter 4

Resource Allocation in Multichannel PLC

4.1 Chapter Overview

In any communication system, wireless or wired, there are circumstances in which the transmitter or the receiver has to make a decision in order to intelligently allocate the available resources between the available options. This decision can directly affect the overall performance of the system such as the bit error rate performance. For instance in a MIMO system where more than one transmission channel are available, the transmitter has the choice to transmit in all the available channels, as well as for some reasons which will be discussed in this chapter, the choice to transmit in only a few of the available channels. The choice of the selected channels, if done correctly, can further improve the performance of the MIMO system and is generally known as the channel selection problem. In such case, the transmitter for instance can use the channel state information and base its selection policy on how well the available channels are performing. By selecting the channels which suffer the least from deep fading, noise and interference, the transmitter can only transmit to the channels with better responses and therefore reduce the probability of losing information and increase the probability of maintaining a reliable communication.

As another example of the so-called *selection problems* in communication systems, we can name the relay selection problem in cooperative communication. In a communication network where several communication nodes and links are available, the transmission of data between two nodes which are far from each other in the network is a challenging task and may result in loss of data and unreliable communication. The noise and attenuation in the system makes it very difficult to establish a reliable communication between nodes of great distances. To overcome this problem, the cooperative communication is introduced where some intermediate nodes, known as relay nodes, help the transmitter in order to transmit the information to the

Multichannel scheme	Introduced diversity	Selection problem
MIMO transmission	spatial diversity	channel selection
Multihop transmission	cooperative diversity	relay selection
Multiband transmission	selection diversity	band selection

TABLE 4.1: Multichannel PLC and the resource allocation problems.

receiver. As a special case, we consider in this chapter a *two-hop* cooperative communication where one relay node helps the process of data transmission between transmitter and receiver by retransmitting the received data from data source to the data destination. However, when a number of different relays are available, the problem of selecting the right relay arises. Different relays result in different intermediate channels between transmitter and receiver and can greatly affect the overall performance. A proper relay selection policy can help the transmitter in order to select the best relay from a range of available relays in order to achieve a reliable communication.

Finally, another selection problem which frequently occurs in communication systems is the problem of selecting the right frequency band of transmission. In an OFDM-based communication system, selecting and prioritizing the most favorable subcarriers can be an important task which helps to avoid the faded parts of the spectrum or avoid the narrowband interference which may be available only in certain parts of the spectrum. This process is known as spectrum assignment or equivalently spectrum allocation. Spectrum allocation can be performed statically, that is with fixed pre-allocated data subcarriers, as well as dynamically which changes the data subcarriers throughout the transmission. The benefit of dynamic spectrum allocation is when the communication channel has a time-variant nature and the best subcarriers change position in the course of time. The selection policy in which the appropriate data subcarriers are selected is very important since the condition of the subcarriers directly affects the overall performance of the system.

All the problems mentioned above can be considered as a selection problem and the search for a good selection policy based on some pre-defined criteria can be applied to all selection problems. In the following we discuss and formulate in detail all the three above-mentioned problems and try to mathematically formulate the selection problem in each case. Table 4.1 summarizes the multichannel schemes and the corresponding selection problems discussed in this chapter.

The rest of the chapter is organizes as follows. Section 4.2 describes the channel selection problem and the corresponding spatial diversity in MIMO PLC systems. Section 4.3 described the multihop transmission and the process of relay selection, as well as the corresponding cooperative diversity in cooperative PLC systems. Section 4.4 describes the multiband transmission and the corresponding dynamic spectrum assignment problem. Moreover, in section 4.6, all the aforementioned selection problems will be modeled as a multi-armed bandit problem, in order to

be used by the reinforcement learning algorithms proposed in the next chapter. Finally, section 4.7 concludes the chapter.

4.2 Channel Selection Problem

4.2.1 Selection Diversity

As described in Chapter 2, Power line communication systems exploit the existence of the power delivery infrastructures to transmit data signals through power line conductors. The existence of multi-conductor power outlets makes it possible to develop the idea of multichannel communication in power lines. In general, however, power line channels exhibit time-variant statistical characteristics, which makes it challenging to develop performance optimizing transmission schemes. As a particular instance, consider a scenario where only one channel has to be selected among several available channels for data transmission in a multichannel PL C system. It is known that time-variability of PLC channels as well as the impulsive noise of the channel could yield sudden deep fades and thus dramatic data loss; hence, it is of utmost importance to choose a suitable channel that, with high probability and in an average sense, results in the highest data rate and is least affected by power line noise and narrowband interference. An appropriate channel selection, however, requires channel state information, which is very costly to acquire at the transmitter with respect to feedback and signaling overhead. The problem of information acquisition becomes even more challenging for statistically time-variant channels like PLC channels, where pilot signals have to be transmitted throughout the entire transmission. As a result, it becomes imperative to search for new solutions which are able to cope with the channel selection problem in a general multichannel PLC model, where channels are time-variant and information is scarce.

The channel selection problem has been studied in detail in the literature, for instance in [104]–[106]. The benefits of channel selection and the selection diversity in MIMO systems are aptly summarized and discussed in [107]–[109]. In the power line communication literature, the problem of channel selection in a MIMO PLC system has been addressed in [110] and solved by multi-armed bandit problem modeling and machine learning algorithmic solutions. In the following, we firstly demonstrate the selection diversity gained by the process of channel selection in a multichannel PLC system where a single channel is to be selected for data transmission. Then, we formulate the channel selection problem in MIMO PLC and indicate the challenges which we have to face.

Let us assume a MIMO power line communication system with N transmit ports at transmitter and M receive ports at receiver. The MIMO PLC system can use all the $M \times N$ available channels and establish a full MIMO transmission. However, because of the frequency selective and time-variant nature of the PLC channel and also the spatial correlation between the MIMO channels, as described in the previous chapter, the MIMO channels may perform at a sub-optimal level and

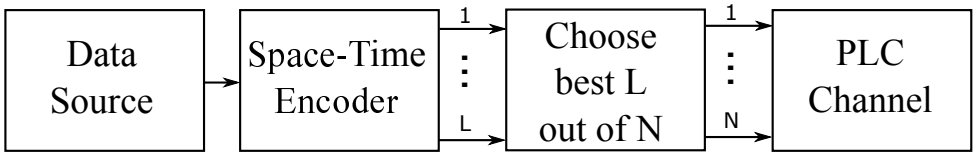


FIGURE 4.1: MIMO PLC system with channel selection at transmitter.

decrease the performance and reliability of the channel. To overcome this problem *transmit selection diversity* is introduced, where one or more transmit ports from all the available ports are selected for data transmission. Similarly, *receive selection diversity* refers to the selection of one or more receive ports from which the data is received. This selection is based on the performance of the corresponding channel, that is the channels which result in a better communication are to be selected for data transmission. In this thesis, we consider the case of transmit selection diversity where $1 \leq L \leq N$ transmit ports are selected for data transmission as depicted in Figure 4.1.

The selection process at the transmitter depends on the availability of the channel state information at the transmitter. When no channel state information is available at the transmitter, transmit diversity can be obtained through Alamouti space-time diversity scheme, or the extended orthogonal space-time block code (OSTBC) [111]. In this case, due to the absence of channel state information at transmitter, the transmitted power is distributed equally among transmit ports to maximize the channel capacity. For a single receive port when an OSTBC is employed, the received SNR is given by [111]

$$S_r = \frac{s}{N} \sum_{i=1}^{N} |h_i|^2, \tag{4.1}$$

where h_i denotes the MIMO channel between the i-th transmit port and the receiver, and s is the corresponding SNR of that particular channel. Therefore, the received signal to noise ratio approaches the average signal to noise ratio for transit diversity when N increases.

When the channel state information, through a feedback from the receiver, is available at the transmitter, the best L channels can be selected for data transmission. Let us assume the special case when a single transmit port is selected and therefore $L = 1$. This can significantly reduce the hardware cost and complexity. The use of the best channel of the available transmit channels also reduces the chance of spatial correlation and the corresponding channel degradation. When a single transmit port is selected, the average SNR is given by [112]

$$S_{r,CSI} = \bar{s} \sum_{i=1}^{N} \frac{1}{i}, \tag{4.2}$$

where which $\bar{s}$ is the average SNR.

The most important aspect of the transmit selection diversity is the ability to select the best channel from the available channels in the MIMO PLC system. When channel state information is available at transmitter, the aforementioned selection can be performed without any problem using this information. However, acquiring the channel state information and feeding this information back to the transmitter is not an easy task and requires a lot of overhead and increases complexity of the system and implementations. Therefore, it would be very convenient if the channel selection procedure were to be performed without any prior knowledge of the channel. In this case, the selection policy can be based on random selection of the transmitting channel or based on a fixed pre-defined selected channel. Neither of these methods can result in an optimum selection policy which gives the best performance. In the next chapter a set of machine learning algorithms are being introduced, which help the transmitter to perform the channel selection without any prior knowledge of the channel state information and have a near optimum performance at the same time.

4.2.2 Channel Selection at MIMO PLC Transmitter

Let us formally introduce the channel selection problem at the MIMO PLC transmitter. We consider a scenario in which a set $\mathcal{K}$ of K PLC channels are available in a MIMO PLC system.[1] At the transmitter, at each time slot t_j, a channel $k \in \mathcal{K}$ is to be selected for data transmission in that time slot based on some known selection policy. We do not transmit at the same time in all available channels as done in MIMO PLC transmission, whereas we are interested in a scenario in which only one channel is to be selected among numerous available channels for transmission. The most important criterion for our decision is the transmission performance in the selected channel. The transmission performance is measured in terms of data rate (here, also referred to as *utility*), which can be expressed as a function of the channel matrix H [7]

$$u_t(k) \sim H. \tag{4.3}$$

Clearly, as channel gains exhibit time-variant statistical characteristics, and in particular, time-variant mean, the utility function given by (4.3) can also be considered as non-stationary. However, it is a strictly increasing function of channel gains, and, therefore, higher channel gains yield higher utility. Now, assume that the transmission is performed in T trials, and the selected channel at time t is denoted by a_t. Intuitively, the transmitter desires to maximize its accumulated utility over the transmission time. Formally,

$$\underset{a_t \in \mathcal{K}}{\text{maximize}} \quad \sum_{t=1}^{T} u_t(a_t). \tag{4.4}$$

The maximization in (4.4) is however not feasible, since without prior channel knowledge, the objective function is not available. Therefore, at each time t, the

[1] Note that to satisfy the Kirchhoff's law and for simplicity we assume $K = 4$ in the simulations. However the results are valid for any value of K.

transmitter might opt to select the best channel in an average sense. That is, at time t, it tries to select channel k_t^* with

$$\mu_{k^*,t} = \mu_t^* = \max_{k \in \mathcal{K}} \ \mu_{k,t}, \tag{4.5}$$

so that the average utility is also maximized in an average sense. With the channel knowledge at transmitter and using an exhaustive search through the available channels, this can be achieved and the best channel can be selected for every transmission. However, accessing the channel knowledge requires the transmission of pilot signal throughout the whole transmission and that introduces a tremendous amount of undesirable overhead in the system which makes the implementation more difficult and increases the costs. Given no prior information about the channel, however, this is not a trivial task to accomplish. As described in the next chapter, machine learning algorithm can help to achieve such a task without the undesirable overhead.

4.3 Relay Selection Problem

4.3.1 Cooperative PLC

In this section we introduce another selection problem that can emerge in PLC systems, namely the relay selection problem. In power line communication networks, multiple network nodes are interconnected via the transmission lines and the data signals flow between different nodes of the network. If the distance between source and destination in a communication scenario is long, the limitation of the transmission range of the PLC node prevents the establishment of a reliable communication. This limitation is due to the harsh environment of the PLC channel, for instance frequency-selective fading, colored and impulsive noise, narrowband interference, and low receiver sensitivity. To overcome this problem, cooperative communication is used to transmit the data signals from source to destination with help of one or more intermediate nodes. In this case, the source node can communicate directly with nodes within its transmission range, and these nodes, in turn, can forward the message to the destination node. The intermediate nodes are called relays and the process of transmitting signals with the help of relays is referred to as multi-hop communication or relaying.

The application of relaying in cooperative wireless communication has been studied to a great extent. The use of relaying in cooperative power line communication has been mentioned and studied for certain communication scenarios as well. For example in [113] and [114], PLC relaying based on single frequency networking has been introduced and its performance has been shown and discussed. Distributed space-time coding for multi-hop transmission and decode-and-forward relaying has been studied in [115] and [116], respectively. The concept of cooperative multi-hop communication for PLC has been first introduced and discussed in [117], [22] and [23]. Existence of many intermediate nodes between source and destination, results

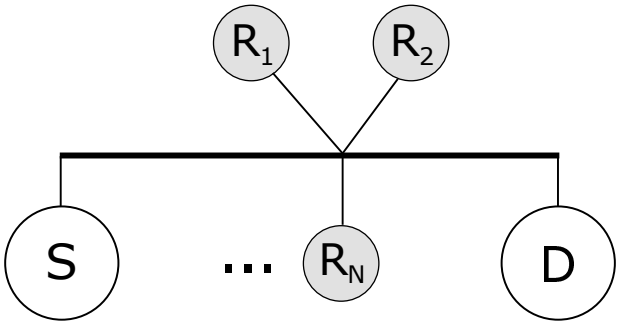


FIGURE 4.2: A PLC network with multiple relay nodes.

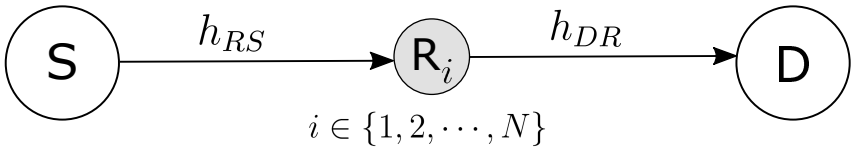


FIGURE 4.3: A two-hop transmission in cooperative PLC.

in the existence of many optional paths or routes to follow. This situation gives rise to the problem of proper relay selection, where the challenge is to pick the optimal path that satisfies the needed performance requirements. In [23] the relay selection problem has been discussed and the link rates has been introduced as a figure of merit for different relay selection criteria.

We consider the two-hop cooperative communication scenario, in which the transmitted signal from source travels to an intermediate relay before reaching the destination. We consider N available intermediate relay nodes, from which one of them as the relaying node is to be selected, as depicted in Figure 4.2. However, the proper relay selection policy requires the availability of channel side information at the transmitter, which in turn requires an increased complexity of signal processing and introduces a lot of overhead in the system. In order to avoid this problem, we introduce a class of machine learning algorithms in the next chapter to solve the relay selection problem without the channel state information at transmitter.

The idea of cooperative communication has been well investigated in wireless communication, e.g. in [118], [119]. The principle of this idea is to realize spatial diversity without the use of multiple antennas. In this case, cooperative users generate a virtual antenna array to achieve the desired cooperative diversity. This concept has been extended to relay networks with multi-hop transmission between source and destination, e.g. in [120]. It has been shown in [23] that the PLC relay channel consists of two keyhole channels, and thus a diversity gain as observed for wireless relaying can not be achieved for PLC. However, despite the lack of the cooperative diversity advantage, cooperative multi-hop transmission can provide significant power gains. In this paper, we assume a two-hop transmission, that is, a source node transmits the message to a destination node through a relay node between them as depicted in Figure 4.3. A generalization to a multi-hop transmission is straightforward.

4.3.2 Relay Selection at PLC Transmitter

We briefly describe and formulate the relay selection problem in the PLC transmitter and describe the different methods of selecting the relay for a two-hop transmission. In a two-hop cooperative PLC, as depicted in Figure 4.3, three nodes are available, namely source node S, selected relay node R_i, which is selected from a sequence of all the available relays $i \in \{1, 2, \cdots, N\}$, and destination node D. We assign a number n to each node so that $\{S, R_i, D\} \sim \{1, 2, 3\}$. The cooperative transmission is assumed to follow a time division protocol, which means at each time instant, only one node (either the source or the relay) can transmit data with a fixed transmit power. As a figure of merit, we consider the end-to-end achievable rate, also known as the *end-to-end capacity*. The end-to-end capacity in a two-hop transmission for a link between node n and node $n+1$ is expressed as $C_{n,n+1}$, which is a strictly increasing function of the transmitted signal-to-noise ratio (SNR) [7]. Therefore, less noise power results in a better link in terms of a higher end-to-end capacity of the link.

We use the conventional strategy of *fixed-rate two-hop* transmission [121]. In a fixed-rate cooperative transmission scheme, the relay node between the source and destination nodes, re-transmits the received message which it receives from source node, using the same transmission scheme and thus link rate, and then transmits the data stream towards the destination node. This fixed-rate strategy is applied over both hops with data rate over each hop denoted as $R_1 = R_2 = R$, for some fixed value of R. In order to ensure reliable communication,

$$R \leq C_{n,n+1} \tag{4.6}$$

must be satisfied over all hops. The selection of the fixed rate value of R is important it can directly influence the amount of the end-to-end capacity of the transmission. By selecting the proper value of R, the end-to-end capacity can be increased which directly results in an overall better performance of the system. It is proved in [121], that the maximum end-to-end capacity can be achieved by choosing the value of R as the minimum value of the end-to-end capacity of each hop. Formally, the fixed rate of transmission can be expressed as

$$R = \min_{n=1,2} C_{n,n+1}. \tag{4.7}$$

Therefore, the maximum rate that can be achieved over the entire transmission route is determined by the minimum of the rates achievable on the individual links. In the fixed-rate strategy of cooperative communication, the relay node retransmits the received message using the same transmission scheme and thus link rate. Hence, the maximum rate that can be achieved over this route is determined by the minimum of the rates achievable on the individual links. Therefore, the end-to-end capacity for a fixed-rate two-hop transmission can be expressed as

$$C_{total} = \frac{1}{2} \min_{n=1,2} C_{n,n+1}. \tag{4.8}$$

For fixed-rate cooperative transmission scheme, the receiver is assumed to have full channel state information, whereas the transmitter is assumed not to have the channel state information. This assumption is due to the fact that acquiring the channel state information at transmitter for PLC systems causes undesirable overhead in the system. Furthermore, the time-variant nature of the PLC channels, as described in the previous section, demands acquiring of the channel state information throughout the transmission, which results in an inefficient transmission scheme. Without access to the channel state information at the transmitter, however, selecting a relay for two-hop transmission among N available relays, is not an easy task. We wish to have a relay selection strategy which results in a high end-to-end capacity C_{total}, or equivalently results in the selection of a relay which provides the best link quality for both transmission hops. Formally, we aim to solve the following maximization problem;

$$\underset{R_i \in \{R_1, R_2, \cdots, R_N\}}{\text{maximize}} \quad C_{total}. \tag{4.9}$$

Let us assume the scenario in which the channel conditions are unknown at the transmitter. We consider three relay selection strategies to maximize the total end-to-end capacity without any information about the conditions of the channel at transmitter.

1. Fixed selection: in this strategy, a fixed relay node is assigned for cooperative transmission between the source and destination, regardless of the instantaneous channel conditions. This strategy neglects the variations is the PLC channel and therefore is not an optimal method of relay selection.

2. Random selection: in this strategy, at each transmission time interval a random relay is selected from the sequence of N arrays. This method neglects the variations of the channel over time as well, however, the randomness of the relay selection may decrease the probability of selecting a bad relay compared to the fixed selection method.

3. Learning algorithms: we propose learning algorithms based on the *multi-armed bandit (MAB)* model in machine learning. In our approach we consider the variations of the channel over time which occurs in a cyclostationary manner as described before, and try to adapt the relay selection with these variations without the knowledge of channel state information at the transmitter.approaches.

4.4 Spectrum Assignment Problem

4.4.1 Multi-Band Transmission

As described in Chapter 2, the PLC channel is time-variant and frequency-selective which may also suffer from narrowband interference. Narrowband interference is mainly caused by other PLC devices or other applications operating in the same

frequency range as the device of interest. Narrowband interference can severely degrade the transmission performance and increase the bit error rate which eventually leads to an unreliable communication. In order to establish a reliable communication and fulfill a certain quality of service (QoS) requirements, the narrowband interference and the parts of the spectrum in which the PLC channel is in deep fade should be avoided and the parts of the spectrum in which the PLC channel suffers the least from fading and interference should be selected for data transmission. Such a frequency band selection can be performed by evaluating the channel conditions on the available subcarriers and prioritizing the available spectrum accordingly. After evaluating the spectrum, the best frequency band is to be selected for the data transmission. This process is referred to as *spectrum assignment* problem.

The use of PLC spectrum is regulated through standardizations, for instance in [122], [123] for narrowband PLC. In these standards the spectrum in which the data is supposed to be transmitted is determined and the OFDM parameters of the data transmission has been established. Due to the time-variant nature of the PLC channel and the random occurrence of the narrowband interference, some parts of this spectrum might be in deep fade or suffer from narrowband interference. Therefore, it has been suggested in the standards. that instead of the entire spectrum, only one portion of the spectrum be used for data transmission. The transmitter may decide the frequency band in which the data is going to be transmitted, although the used sub-carriers have to be consecutive. This process is referred to as band selection or spectrum assignment, and the selection policy is based on the behavior of the channel. Therefore, the knowledge of channel state information is necessary to perform the spectrum assignment and without the channel state information this assignment may not lead to an optimal band selection. The problem of spectrum assignment without the channel state information at transmitter is addressed in the next chapter.

The problem of spectrum assignment has been studied extensively in wireless communication, for instance in [124]–[126]. Dynamic spectrum assignment has been introduced to PLC applications in [127] where the authors study a low-bandwidth PLC system where the spectrum allocation is performed based on the available channel state information at transmitter. In [110], the authors model the channel selection problem as a MAB and perform channel selection with unknown channels at transmitter. In this thesis we consider the problem of dynamic spectrum assignment with unknown channel at transmitter.

The spectrum assignment process can be performed as a fixed frequency assignment technique or dynamically throughout the data transmission. In the fixed spectrum assignment, the portions of the spectrum which result in the best bit error rates are selected before the transmission begins and stay the same throughout the transmission. This frequency assignment technique is applicable for stationary scenarios where for instance the location and intensity of the narrowband interference remains constant during the transmission. The time-variant nature of the PLC channels makes this method non-efficient due to the changes of the PLC channel throughout the transmission. For instance, the best frequency band at the start of

Parameter	Value
Start frequency	154.6875 kHz
End frequency	487.5 kHz
Total number of subcarriers	256
Total number of used subcarriers	72, 36, 18
Number of cyclic prefix samples	30
Subcarrier spacing	4.6875 kHz
Sampling Frequency	1.2 MHz

TABLE 4.2: Parameters of FCC-above-CENELEC frequency band.

Band	Start Frequency [kHz]	End Frequency [kHz]	Subcarriers
1	154.6875	487.5	72
2	154.6875	318.75	36
3	323.4375	487.5	36
4	154.6875	234.375	18
5	239.0625	318.75	18
6	323.4375	403.125	18
7	407.8125	487.5	18

TABLE 4.3: Spectrum division in FCC-above-CENELEC.

the transmission might not remain the best band and suffer from narrowband interference after some time has passed. Therefore, the monitoring and selecting the available spectrum have to be performed throughout the transmission in pre-defined intervals. In the next section, we briefly define the problem of dynamic spectrum allocation at PLC transmitters.

According to the narrowband PLC standards, the frequencies between 154.6875 kHz and 487.5 KHz have been dedicated to the FCC-above-CENELEC band. This frequency band is dedicated to an OFDM-based transmission which is divided into a total of 72 data subcarriers. The frequency and OFDM parameters of the FCC-above-CENELEC frequency band in narrowband PLC is summarized in Table 4.2. The spectrum division properties and the corresponding seven available sub-bands for spectrum assignment and their specifications are presented in Table 4.3.

Moreover, this frequency range can be further divided into two frequency bands of 36 subcarriers with the same inter-subcarrier spacing, and the data can be transmitted in either of the two 36-subcarrier-long sub-bands. Each frequency band with 36 subcarriers can, in turn, be further divided into two frequency sub-bands of 18 subcarriers with the same inter-subcarrier spacing as before, making four 18-subcarrier-long frequency sub-bands, and the data can be transmitted in either of

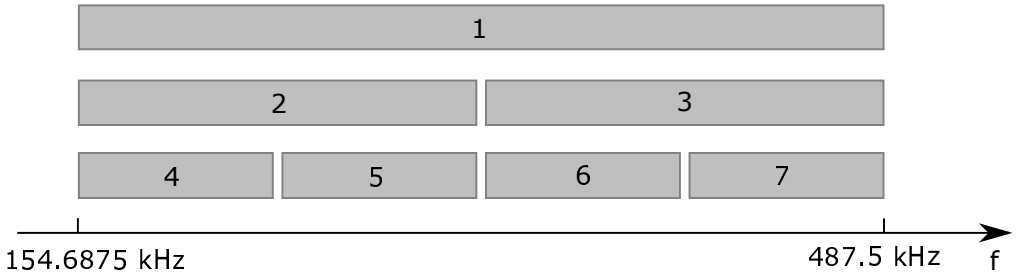


FIGURE 4.4: FCC-above-CENELEC spectrum partitions with seven different frequency bands available.

the four 18-subcarrier-long sub-bands. The different sub-bands and their relative frequency ranges have been illustrated in Figure 4.4. As shown in this Figure, there are seven different frequency bands available which can be selected for data transmission, each with its specific number of data subcarriers. The reason behind choosing a different band and even moving to a frequency band with less data subcarriers can be the fact that some parts of the whole FCC-above-CENELEC band can be affected with deep fade and severe narrowband interference. Therefore, selecting another available band even though may reduce the achievable throughput, however can significantly increase the reliability and performance of the transmission.

The OFDM parameters for each of the sub-bands do not change during a band transition. In other words, each of the six sub-bands with 36 or 18 data subcarriers as well as the main frequency band with 72 data subcarriers have the same frequency parameters such as inter-subcarrier spacing, number of FFT samples, sampling frequency, etc. This helps to have the same transmitter and receiver filters and the need for designing and implementing new filters is eliminated. On the other hand, if the frequency sub-bands have been shifted or have changed the frequency parameters, each sub-band would need a new transmitter and receiver filter redesigned and reimplemented specifically for each frequency band transition. In order to be able to have the same frequency parameters and consequently use the same filters at transmitter and receiver when we bisect the spectrum, any consecutive sub-matrix of the OFDM modulation matrix with 18 or 36 subcarriers must have an inverse or equivalently have a full rank. It can be shown that any $k \times k$ consecutive sub-matrix of an IDFT matrix has indeed full rank. Note that the sub-matrix has to consist of consecutive subcarriers in order to fulfill the full rank requirement.

4.4.2 Dynamic Spectrum Assignment at PLC Transmitter

The spectrum assignment process can be performed as a fixed frequency assignment technique or dynamically throughout the data transmission. In the fixed spectrum assignment, the parts of the spectrum which result in the best bit error rates are selected before the transmission begins and stay the same throughout the transmission. This frequency assignment technique is applicable for stationary scenarios

where for instance the location and intensity of the narrowband interference remains constant during the transmission. The time-variant nature of the PLC channels makes this method non-efficient due to the changes of the PLC channel during the transmission. Therefore, monitoring and selecting the available spectrum band have to be performed throughout the transmission in pre-defined intervals. Such sub-carrier evaluation introduces tremendous amount of overhead and complexity into the system and requires to be performed constantly throughout the transmission due to the time-variant nature of the PLC channel. To avoid this problem, we propose reinforcement learning algorithms which can in a reasonable time find the best part of the spectrum without any a priori knowledge of the channel conditions at transmitter.

Formally, let us assume $B_i,\ 1 \leq i \leq 7$, to be the selected band as depicted in Figure 4.4. As a figure of merit we use the average transmitting power $\bar{P}_i(Q)$ required to maintain a certain QoS of Q. We select the transmitting band at the beginning of each data frame transmission, here and thereafter known as *episodes*, based on some selection policy π. The selection policy should solve the following problem

$$\underset{B_i}{\text{minimize}}\ \bar{P}_i(Q). \tag{4.10}$$

Solving this problem yields i and consequently the corresponding frequency band B_i which is to be selected for data transmission. Solving this problem and finding the best band, requires the knowledge of the channel which is not a suitable solution in PLC as described in the previous section.

4.5 General Resource Allocation Problem

In the previous subsections, we briefly discussed different problems emerging in the multichannel PLC. All these problems have decision making as their core concept. In these problems, the PLC transmitter is faced with multiple options and it is trying to allocate its resources, in this case power, into the right options. In a real PLC system, all three resource allocation problems can occur at the same time. In other words, the PLC transmitter can choose the proper transmitting port at the transmitter, choose the proper intermediate relay node for a cooperative communication, and finally, choose the portion of the spectrum in which the data is transmitted. Each resource allocation strategy can in turn improve the performance of the system by a proper decision making policy. This causes the transmitter to transmit the data packets throughout the transmission time, via different transmitting ports, through different relay nodes and in different parts of the spectrum. Therefore, by proper decision making algorithms, a significant improvement in the overall system performance can be expected. However, given specific circumstances, the PLC transmitter may encounter any subset of these problems. In the following, we focus on each problem individually. Before further investigating the proposed solutions to these problems, we use a powerful mathematical model in order to model these problems to use them as building blocks of machine learning algorithms.

4.6 Multi-armed Bandit Problem Modeling

4.6.1 Multi-armed Bandit Model

In the previous chapter, we discussed several selection problems which may occur in the power line communication. At each of these problems, the PLC transmitter faces the problem of selecting a parameter or a set of parameters between several available options. The outcome of each selection would directly influence the fundamental properties of the communication system and the reliability of the transmission. Therefore, a proper selection policy can greatly influence the overall performance of the communication system and is of utmost importance. This selection policy should be based on certain criteria which helps the selecting agent, i.e. the PLC transmitter, to perform a knowledge based selection in order to obtain an optimal performance. The selection criteria vary in each case of selection problems. For instance, for the selection problems discussed in Chapter 3, the following criteria may be considered:

- The channel selection process at the transmitter of a MIMO PLC system is based on the conditions of the different channels which are formed by the MIMO system between the transmitter and the receiver.
- The relay selection process at the transmitter of a cooperative PLC system in a PLC network is based on the conditions of the channels which are formed at each hop of the transmission between source, relay, and the data destination.
- The spectrum selection process at the transmitter of a narrowband PLC system is based on the conditions of each subcarrier at the time of transmission.

It can be observed that in order to perform a proper selection policy, the decision has to be eventually based on the conditions of the PLC channel. Therefore, obtaining the channel conditions at the transmitter is necessary in order to be able to make a knowing decision based on the information obtained from the channel state. To acquire the channel state information at the transmitter, the transmitter has to transmit pilot signals at pre-defined locations in time and frequency.

The receiver then receives the pilot signals and performs channel estimation based on the received pilot signals to obtain the channel state information. The channel state information is then fed back to the transmitter to be used in the process of selection. However, as described in Chapter 2, the PLC channel is time variant and frequency selective. This means that in order to obtain the channel state information, pilot signals have to be transmitted regularly throughout the transmission as well as all the subcarriers. Moreover, even when only a single selection has to be made, due to the time-variant nature of the PLC channel, all the other options need to be regularly monitored in order to follow the changes in the environment. Therefore, the channel estimation should be performed not only in time and frequency domains, but also in different locations, links and channels. This makes a tremendous amount of undesired overhead in the system as well as reducing the data throughput.

Due to the difficulties in acquiring the channel state information as described above, it is a fair assumption to consider the channel conditions as unknown at the transmitter. However, the channel state information is needed at the transmitter in order to perform an informed decision in any type of described selection problems. To solve this problem we propose the use of a class of problem modeling approaches known as the multi-armed bandit (MAB) problem modeling. MAB is a class of decision making problems introduced in machine learning, where a selecting agent consecutively selects an *arm* (or *action*) from a set of predefined actions, and receives a *reward* drawn from some a priori unknown distribution. Only the reward of played arm is observable at each trial. As a result of lack of information, at each trial, the player may choose an inferior arm in terms of average reward, yielding a *regret* that is quantified by the difference between the reward that would have been achieved if the agent would have selected the best arm with the highest reward and the actual achieved reward. The solution to the MAB problem aims to find the decision policy which results in the highest reward, or equivalently, results in the lowest regret. In other words, the goal is to perform a decision without any a priori knowledge of the channel, which results in the best performance.

MAB problems have been extensively used in wireless communication to model the problem of balancing exploitation and exploration (see for instance [128]–[131]). The exploitation refers to the use of the existing knowledge of the environment in order to receive the highest reward and exploration refers to the searching for a new decision and reacting to the changes in the environment. The key feature of the MAB problem modeling is the establishment of a trade-off between exploitation and exploration. The applications of MAB in wireless communication is aptly summarized in [132]. The MAB problem modeling was first introduced to the PLC systems for the channel selection problems in [110]. It has been shown that the MAB problem modeling in combination with machine learning algorithms provides a near optimal solution to the single channel selection problem in MIMO PLC systems without the channel state information at the transmitter.

In this chapter, the MAB problem modeling approach in both stationary and non-stationary systems is briefly introduced. Furthermore, the formal formulation of the channel selection problem in MIMO PLC, the relay selection problem in cooperative PLC networks, and the spectrum assignment in narrowband PLC systems, based on the MAB problem modeling are introduced.

Let us consider the problem of making a decision in an environment of incomplete information. The selecting agent (in out case the PLC transmitter), is constantly faced with the problem of maximizing its profit based on its current knowledge of the environment and trying to learn more about the environment in order to improve the quality of its decisions. This problem is widely modeled by the MAB problem [133], [134]. The selecting agent is repeatedly faced with a decision between N different choices, here and thereafter referred to as actions or arms. Depending on the selected arm, a reward is observed. The reward in only observable after the selection has been made and the reward of the other choices cannot be revealed to the selecting agent. The goal of the selecting agent is to maximize the average

reward over the time of operation. The MAB problem can be divided into three separate problems as described in the following.

1. Stationary bandits: in these problems the environment in which the decision making process is taking place does not change over time. Therefore, as time of the operation increases the knowledge of the environment at the selecting agent increases and better decisions are made.
2. Non-stationary bandits: in these problems the environment changes during time. Therefore, the selecting agent has to react to the changes in the environment in an appropriate amount of time. Consequently, the recent rewards are more relevant to the selecting agent than the older ones.
3. Adversarial bandits: these problems were introduced in [135]. In this problem the rewards collected by the agent do not follow some probability distribution but are generated by an adversary in order to make things more complicated for the agent. Since this scenario is out of the scope of this thesis and does not apply to the PLC systems, we do not discuss it further.

Let us consider a general selection problem in a PLC system. Let us assume a set $\mathcal{K}$ of available options from which the selecting agent is able to select, also known as actions or arms. Each frame of data is transmitted after selecting an arm $k \in \mathcal{K}$, $k = \{1, 2, \cdots, N\}$, where N is the number of available arms. The selected arm results in a particular total profit, also known as a reward. We define an episode as a slot in which a frame of data is transmitted according to a selected arm by the selecting agent. At each episode, an action a_t is selected, yielding the instantaneous reward $X_t(a_t)$. The rewards $\{X_t(k)\}_{t\geq 1}$ for each arm $k \in \{1, 2, \cdots, N\}$ is calculated according to the received signal and this information is fed back to the transmitter via a robust mode of transmission, which is normally chosen as the acknowledgment (ACK) packet which is already being fed back to the transmitter.

The selecting agent chooses an arm at each episode according to a certain selecting policy π. Let us denote the expectation of the reward $X_t(k)$ by $\mu_{k,t}$. Let k_t^* denote the optimal arm at time t, with expected reward $\mu_{k^*,t}$, where by definition

$$\mu_{k^*,t} = \mu_t^* = \max_{k\in\mathcal{K}} \ \mu_{k,t}. \tag{4.11}$$

We define the instantaneous regret at time t as the difference between mean rewards of the selected arm and the optimal arm. The expected regret of a decision making policy π after T trials, therefore, can be expressed as

$$R_{\pi,T} = \mathbb{E}_\pi \left[\sum_{t=1}^{T} (\mu_t^* - \mu_{a_t,t}) \right], \tag{4.12}$$

where $\mathbb{E}$ represents the mathematical expectation. The goal of a good policy is to select the optimal arm at each trial, which results in a minimum expected regret over all trials. Therefore, the goal of the MAB problem is to minimize the expected

regret with a certain decision making policy π, or equivalently

$$\underset{\pi}{\text{minimize}}\ R_{\pi,T}. \tag{4.13}$$

The best decision policy of a MAB problem is one that yields the best results for the equation (4.13).

4.6.2 Stationary and Non-stationary Bandits

MAB problem modeling of the communication system problems are generally divided into stationary and non-stationary bandits. In stationary bandit problems, the environment in which the selecting agent performs its decisions is considered to be stationary and not changing over time. At each episode, the selecting agent chooses an arm based a certain selection policy which is, in turn, based on the current knowledge of the environment. This knowledge at the beginning of the operation is limited and may not correspond the actual conditions of the environment. However, over time the selecting agent obtains more and more information about the environment in the form of observed rewards. As the information about the environment at the side of selecting agent increases, the quality of the decisions will increase as well. Since the environment is stationary, the decision making needs to explore the environment less as the time of operation increases and the confidence on the decisions becomes stronger as well. This means that in a stationary environment the exploration of the environment becomes less frequent in order to exploit the existing information. Figure 4.5 illustrates the trade-off between exploitation and exploration in a stationary bandit problem. As depicted in this figure, as the time of the operation increases less explorations are necessary and the average reward can be increased by increasing the number of exploiting episodes.

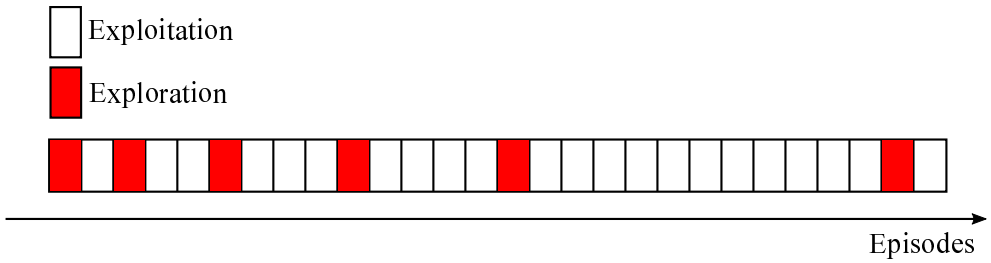


FIGURE 4.5: Exploitation-exploration trade-off in a stationary bandit.

On the other hand, in a non-stationary bandit, the environment is changing over time. This can be seen as a more realistic scenario for PLC problems, due to the time-variant nature of the PLC channel. In this case, the selecting agent faces the problem of adapting itself to the changes of the environment and reacting to these changes in an appropriate time. For this reason, the trade-off between exploitation and exploration is more than ever important. More exploitation results in a better performance and higher average rewards but results in a delayed reaction to the changes in the environment. On the other hand, more exploration results in a

faster reaction to the changes of the environment, however, it may also result in the selection of a sub-optimal arm which, in turn, results in a decrease in performance and lower average reward. Finding the optimal trade-off between a high average reward and a fast reaction to the changes of the environment is the main focus of a proper solution to the non-stationary bandit problem. Figure 4.6 depicts the exploitation-exploration trade-off in a non-stationary bandit in two scenarios. In (a), the explorations are more frequent which results in a faster reaction to the changes of the environment, whereas in (b) the exploitations are more frequent which results in a slower reaction to the changes in the environment but a higher average reward.

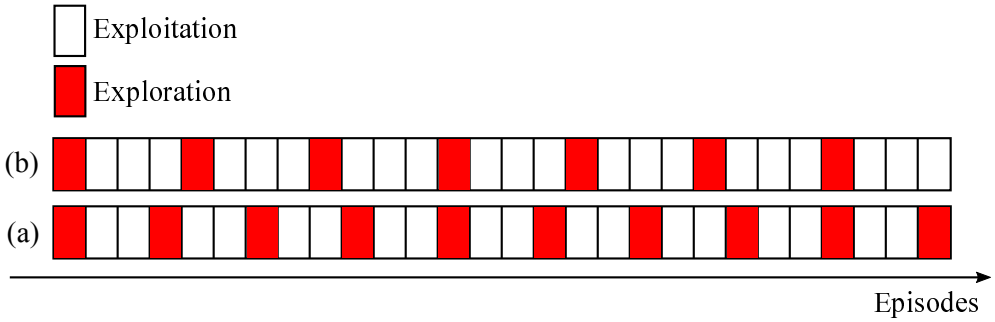


FIGURE 4.6: Exploitation-exploration trade-off in a non-stationary bandit with (a) more exploration and (b) more exploitation.

4.6.3 Channel Selection Problem as MAB

We consider a scenario in which a set $\mathcal{K}_c$ of K_c PLC channels are available in a MIMO PLC system.[2] At each episode, a single channel $k_C \in \mathcal{K}_c$ is selected for data transmission. This can be considered as a single transmit diversity in a MIMO system as described in Chapter 3. In other words, we do not transmit at the same time in all channels as done in a conventional MIMO transmission, whereas we are interested in a scenario that only one channel is to be selected among numerous available channels for transmission in order to exploit the single transmit diversity of the system. The most important criterion for our decision is the transmission performance in the selected channel. The transmission performance is measured in terms of data rate (here, also referred to as *utility*). Clearly, as channel gains exhibit time-variant statistical characteristics, and in particular, time-variant mean, the utility function given above is also a non-stationary function. However, it is a strictly increasing function of channel gains, and, therefore, higher channel gains yield higher utility. Now, assume that the transmission is performed in T trials, and the selected channel at time t is denoted by $a_t^{(c)}$. Intuitively, the transmitter

[2]Note that since the PLC signals are formed on a set of conductors, they must satisfy the Kirchhoff's circuit laws. For simplicity and the ability to conduct simulations, we assume $K_c = 4$. However the results are valid for any value of K_c.

desires to maximize its accumulated utility over the transmission time. Formally,

$$\underset{a_t^{(c)} \in \mathcal{K}_c}{\text{maximize}} \quad \sum_{t=1}^{T} u_t^{(c)} \left(a_t^{(c)} \right). \tag{4.14}$$

The maximization in (4.14) is however not feasible, since without prior channel knowledge, the objective function is not available. Therefore, at each time t, the transmitter might opt to select the best channel in an average sense. That is, at time t, it tries to select channel k_t^* with

$$u_t^{(c)}(k^*) = \max_{k \in \mathcal{K}_c} \; u_t^{(c)}(k), \tag{4.15}$$

so that the average utility function is also maximized in an average sense. With the channel knowledge available at the PLC transmitter and using an exhaustive search through the available channels, this can be achieved and the best channel can be selected for every transmission based on the available knowledge. However, accessing the channel knowledge requires the transmission of pilot signal throughout the whole transmission and that introduces a tremendous amount of undesirable overhead in the system. Given no prior information, however, this is not a trivial task to accomplish. The expected regret of some decision making policy π after T rounds (episodes), denoted by $R_{\pi,T}^{(c)}$, yields

$$R_{\pi,T}^{(c)} = \mathbb{E}_\pi \left[\sum_{t=1}^{T} \left(u_t^{(c)}(k^*) - u_t^{(c)} \left(a_t^{(c)} \right) \right) \right], \tag{4.16}$$

Let $\mathcal{P}$ be the set of all possible decision making policies. The agent aims at minimizing its average regret over the horizon by choosing the optimal policy, which guides the agent for selecting an arm at each trial, so that the average regret is minimized. This goal can be formally stated as

$$\underset{\pi \in \mathcal{P}}{\text{minimize}} \; R_{\pi,T}^{(c)}. \tag{4.17}$$

Now recalling from the problem formulation, in multi-channel PLC, our goal is to select the best channel k_t^* with the utility function of $u_t^{(c)}(k^*)$, so that the average utility is maximized. As channel characteristics are unknown a priori, and after transmitting in any channel, only the performance of that specific channel can be observed, the problem resembles a MAB, where arms are mapped to PLC channels. Furthermore, as described previously, PLC channels are time-variant, in the sense that although channel gains can be attributed to a log-normal distribution, the statistical characteristics do not remain constant over time. Therefore, it can be assumed that the problem is a non-stationary bandit problem.

4.6.4 Relay Selection Problem as MAB

We consider the two-hop cooperative communication scenario, in which the transmitted signal from source travels to an intermediate relay before reaching the destination. We consider N_r available intermediate relay nodes, from which one of them as the relaying node is to be selected for our two-hop fixed-rate transmission scheme. In the following, we model the PLC single relay selection problem as a MAB problem. Let us assume a set $\mathcal{K}_r$ of available relay nodes, known as actions or arms. Each frame of data is transmitted in one episode by selecting a relay node

$$k_r \in \mathcal{K}_r, \quad k_r = \{1, 2, \cdots, N_r\}, \tag{4.18}$$

resulting in a particular total end-to-end capacity as described in the previous chapter, here and thereafter known as reward. At each episode t (corresponding to one frame of data), an action $a_t^{(r)}$ is selected, yielding the instantaneous reward $X_t(a_t^{(r)})$. The rewards $\{X_t(k_r)\}_{t\geq 1}$ for each arm $k_r \in \{1, 2, \cdots, N_r\}$ is calculated according to the received signal and this information is fed back to the transmitter via a robust mode of transmission and the transmitter chooses an arm at each trial according to a policy π. The only difference between the available relay nodes is the corresponding link rate, which in turn is dependent on the channel response as well as the noise power spectral density of the PLC channel. Therefore, we use the link rates as a figure of merit for the modeling of the problem. Given the frequency response $H_{n,n+1}(f)$ for a link from node n to node $n+1$, the link rates are computed as [23]

$$C_{n,n+1} = \int_{f_1}^{f_2} \log_2 \left(1 + \frac{S_T |H_{n,n+1}(f)|^2}{S_N \Gamma}\right) \mathrm{d}f, \tag{4.19}$$

where S_T is the transmitter-side power spectral density, S_N is the receiver-side noise power spectral density, and Γ is the margin taking into account the gap between information theoretic capacity and achievable rate using practical coding and modulation schemes.

Let us denote the expectation of the reward $X_t(k_r)$ by $\mu_k^{(r)}(t)$. Let k_t^* denote the optimal arm at time t, with expected reward $u_t^{(r)}(k^*)$, where by definition

$$u_t^{(r)}(k^*) = \max_{k_r \in \mathcal{K}_r} \ \mu_k^{(r)}(t). \tag{4.20}$$

We define the *instantaneous regret* at time t as the difference between mean rewards of the selected arm and the optimal arm. The expected regret of a decision making policy π after T episodes, therefore, can be expressed as

$$R_{\pi,T}^{(r)} = \mathbb{E}_\pi \left[\sum_{t=1}^{T} \left(u_t^{(r)}(k^*) - u_t^{(r)}\left(a_t^{(r)}\right)\right)\right], \tag{4.21}$$

where $\mathbb{E}$ represents the mathematical expectation. The goal of a good policy is to select the optimal arm at each trial, which results in a minimum expected regret over all trials. Therefore, the goal of the MAB problem is to minimize the expected

regret with a certain decision making policy π, or equivalently

$$\underset{\pi}{\text{minimize}}\ R^{(r)}_{\pi,T}. \tag{4.22}$$

Now recalling from the problem formulation, in a PLC network, our goal is to select the best relay k^*_t so that the average reward is maximized or equivalently the regret is minimized. As link rates are unknown a priori, and after transmitting in any route, only the link rate of that specific route can be observed, the problem resembles a MAB, where arms are mapped to the selected relay. Furthermore, as described previously, PLC channels are time-variant, and consequently the statistical characteristics do not remain constant over time. Therefore, it can be assumed that the problem is a non-stationary bandit problem.

4.6.5 Spectrum Assignment Problem as MAB

In the spectrum assignment problem, as described in the previous chapter, the exploitation exploration problem is defined as exploring the available spectrum to find a profitable transmission band while taking the empirically best frequency band as often as possible. In the following, we model the PLC spectrum assignment problem as a MAB problem.

A set of seven available frequency bands B_i, $i \in \{1, 2, \cdots, 7\}$, as described in the previous chapter, are modeled as actions of the MAB problem. The selecting agent is the PLC transmitter and each frame of data is transmitted by selecting an action $a^{(b)}_k = B_i$ where k corresponds to that particular data frame. Data transmission through the selected arm results in a particular average transmitting power $\bar{P}(Q, a^{(b)}_k)$, required to reach a particular QoS of Q with an upper-bound of P_m which is defined as the maximum transmitting power of the PLC transmitter as

$$\bar{P}\left(Q, a^{(b)}_k\right) \leq P_m. \tag{4.23}$$

Based on the limit of transmission power at the PLC transmitter, the achieved reward at each trial $X_k(a^{(b)}_k)$ is defined as the difference between the maximum transmission power and the observed average transmitted power and can be written as

$$X_k\left(a^{(b)}_k\right) = P_m - \bar{P}\left(Q, a^{(b)}_k\right). \tag{4.24}$$

Let us denote the expectation of the reward $X_k(a^{(b)}_k)$ by $\mu_k(a^{(b)}_k)$. Let $a^{(b)}_{k^*}$ denote the optimal action (or equivalently the best frequency band) corresponding to the k-th data frame, with expected reward $\mu_k(a^{(b)}_{k^*})$, where by definition

$$\mu_k\left(a^{(b)}_{k^*}\right) = \max_{a^{(b)}_k}\ \mu_k\left(a^{(b)}_k\right). \tag{4.25}$$

We define the regret at trial k as the difference between mean rewards of the selected action and the optimal action. The expected value of the regret of a decision making

policy π after $\mathcal{K}_b$ trials, therefore, can be expressed as

$$R^{(b)}_{\pi,\mathcal{K}_b} = \mathbb{E}_\pi \left[\sum_{k=1}^{\mathcal{K}_b} \left(\mu_k \left(a^{(b)}_{k^*} \right) - \mu_k \left(a^{(b)}_k \right) \right) \right], \tag{4.26}$$

where $\mathbb{E}$ represents the mathematical expectation. The goal of a good policy is to select the optimal arm at each trial, which results in a minimum expected regret over all trials. Therefore, the goal of the MAB problem is to maximize the average reward over all trials or equivalently, to minimize the expected regret with a certain decision making policy π. Formally, we aim to solve the following problem

$$\underset{\pi}{\text{minimize}}\ R^{(b)}_{\pi,\mathcal{K}_b}. \tag{4.27}$$

It can be seen that equation (4.27) is the MAB model of the spectrum assignment problem formulated as equation (4.10). Recalling from the problem formulation, in a narrowband PLC system, our goal is to select the best spectrum portion of the available spectrum so that the average reward is maximized or equivalently the regret is minimized. As rewards are unknown a priori, and after transmitting in any spectrum, only the reward of that specific spectrum can be observed, the problem resembles a MAB, where arms are mapped to the frequency bands of the total spectrum. Furthermore, as described previously, PLC channels are time-variant, and consequently the statistical characteristics do not remain constant over time. Therefore, it can be assumed that the problem is a non-stationary bandit problem.

4.7 Chapter Conclusion

In this chapter, the multichannel PLC has been analyzed in explored in more detail. Three multichannel transmission scenarios, namely MIMO transmission, cooperative multihop transmission, and multiband transmission have been described and discussed in detail as the options of the PLC transmitter for a better communication via exploiting the multiple communication channels available inherently in PLC infrastructures. Furthermore, the corresponding diversity of the aforementioned multichannel schemes and the generated selection problems have been described. It has been shown that each selection problem can be solved by a proper decision making strategy, which in turn can significantly increase the performance and reliability of the communication system. However, it has been shown that in the state of the art solutions, the decision making policy is based on the channel state information at the PLC transmitter. Nonetheless, the channel state information is not accessible at the PLC transmitter due to time- and frequency selectivity as well as feedback delays of the channel estimation. In order to be able to propose a machine learning strategy for obtaining a decision making policy independent of the channel state information, these selection problems have been modeled by a powerful mathematical tool called the multi-armed bandit model. The MAB modeled problems will be used

in the next chapter in order to propose a class of solutions based on reinforcement learning.

Chapter 5

Reinforcement Learning Applications in Multichannel PLC

5.1 Chapter Overview

In chapter 3 we have presented three different selection problems which normally occur in a PLC system and discussed each selection problem in detail. We have formulated each selection problem as a MAB problem in Chapter 4 with accurately defined reward functions and regret functions for each case. The MAB problem for all the selection problems is a non-stationary bandit problem which concludes in a well-defined minimization (or maximization) problem with certain constraints. In this chapter we aim to solve these minimization problems with the assumption of an unknown channel at the transmitter. This assumption is a key aspect of the problem formulation, since as described in the previous chapters, obtaining the channel state information at the PLC transmitter is a challenging task and cannot be easily performed for all the times and all the frequencies. Therefore, we propose the use of reinforcement algorithms in order to solve the MAB problems without any a priori knowledge of the channel state information.

Reinforcement learning is an area of machine learning which involves in solving the problem of decision making of a selecting agent in an unknown environment in order to maximize some notion of cumulative reward. Therefore, reinforcement learning is a proper solution to the MAB problems and can provide the optimum or near optimum decisions in an unknown environment. Reinforcement learning has been extensively used in the selection problems of the communication theory (see for instance [136]–[139]). In this thesis we describe three of the most common families of reinforcement learning algorithms in order to solve the MAB problems of PLC. Each algorithm has been chosen based on the specific needs and system design of the corresponding MAB problem. Furthermore, we propose four new algorithms based on the reinforcement learning algorithms which are further adapted to the PLC

systems in order to achieve the optimal result. In the following we present each algorithm in detail and discuss the performance analyses. The reinforcement learning algorithms that are being discussed in this chapter are categorized as follows.

1. Upper confidence bound (UCB) algorithms
   (a) The UCB-1 algorithm
   (b) The discounted UCB algorithm
   (c) The sliding window UCB algorithm
   (d) The proposed cyclic discount UCB algorithm (novel)
   (e) The proposed cyclic window UCB algorithm (novel)
2. Probability matching technique
   (a) The seminal algorithm
   (b) The proposed algorithm (novel)
3. Greedy algorithms
   (a) The seminal algorithm
   (b) The ε-greedy algorithm
   (c) The proposed greedy algorithm (novel)

The rest of the chapter is structured as follows. Section 5.2 describes the upper confidence bound algorithms as well as the two proposed algorithms based on the upper confidence bound algorithms. Section 5.3 describes the probability matching technique and the proposed algorithm based on the probability matching technique. Section 5.4 describes the greedy algorithms and the proposed algorithm based on the greedy algorithms. Simulation results are presented for all the algorithms, which shows the performance improvement for the proposed algorithms. Finally, section 5.5 concludes the chapter.

5.2 Upper Confidence Bound Algorithms

5.2.1 UCB-1 Algorithm

The class of upper confidence bound (UCB) algorithms were first introduced to machine learning applications in [140] and further analyzed in [141]. These algorithms were investigated in the literature ever since in order to solve the problem of exploitation-exploration trade-off in an intuitive and efficient manner (see for instance [142], [143]). Though the stationary formulation of the MAB problem allows the UCB algorithms to address exploration versus exploitation trade-off, it may fail to be adequate to model an evolving environment where the reward

distributions undergo changes in time, same as a PLC channel. To model such situations, we need to consider non-stationary MAB problems, where distributions of rewards may change in time. Therefore, non-stationary algorithms were developed to address the problem of a changing environment [144]–[146]. Both stationary and non-stationary algorithms have been used in communication theory in order to solve the exploitation exploration trade-off problem in different scenarios (see for instance [147]–[149]). In this section the conventional UCB algorithm for stationary environments, namely the UCB-1 algorithms, as well as two of the most powerful non-stationary UCB algorithms, namely discounted UCB and sliding window UCB will be discussed in this section. Furthermore, simulation results are presented in order to evaluate the performance of the algorithms.

The first algorithm that we consider is the seminal UCB algorithm, which is known as the UCB-1 algorithm. This algorithm is mainly used in stationary scenarios where the physical characteristics of the environment does not change over time. Although the PLC channel cannot be considered as a stationary environment, but since UCB-1 is the underlying algorithm for the other non-stationary algorithms, we first investigate this algorithm.

Let us assume an environment with N_{arm} arms. The algorithm runs for N_{ep} times which is the number of episodes of data transmission. At each episode i, an arm a_i is selected by the algorithm and the corresponding reward $X_i^{(UCB-1)}(a_i)$ will be observed. In order to prioritize the arms, the concept of an *upper bound* is defined as follows.

Definition 1. The upper bound of a confidence interval of an arm, here and thereafter known as the upper bound function, is denoted as a well-defined function which presents the limits of the average reward which can be expected from that arm. This function includes the obtained knowledge of the environment as well as the accuracy of this knowledge.

Therefore, the upper bound function of each arm in the UCB-1 algorithm can be expressed as the summation of two other functions:

1. The average reward function, $\bar{X}_i^{(UCB-1)}(a_i)$, which denotes the empirical mean of the observed rewards,
2. The padding function, $c_i^{(UCB-1)}(a_i)$, which denotes the accuracy of the upper bound.

and hence, the upper confidence function of the arm a_i which is denoted by $U_i^{(UCB-1)}(a_i)$ can be written as [141]

$$U_i^{(UCB-1)}(a_i) = \bar{X}_i^{(UCB-1)}(a_i) + c_i^{(UCB-1)}(a_i). \tag{5.1}$$

The average reward function is constructed from the past observed rewards and represents the existing knowledge of the environment. Therefore, this function becomes a more accurately representation of the exact average reward of the arm as the number of the past episodes which use that arm increases. The average

reward of arm a_i is defined as [141]

$$\bar{X}_i^{(UCB-1)}(a_i) = \frac{1}{N_i(a_i)} \sum_{s=1}^{i} X_s^{(UCB-1)}(a_i) \mathbb{1}_{\{a_s=a_i\}}, \tag{5.2}$$

where $\mathbb{1}_{\{a_s=a_i\}}$ returns one when the condition $a_s = a_i$ is fulfilled and returns zero otherwise, and [141]

$$N_i(a_i) = \sum_{s=1}^{i} \mathbb{1}_{\{a_s=a_i\}} \tag{5.3}$$

denotes the number of times arm a_i has been played in the i first episodes. Equation (5.2) denotes the superposition of the observed rewards when the arm is played which is averaged over all the instances in which that arm was played. This returns the arm's reward in an average sense with increasing accuracies over time of operation.

The padding function is designed in a way to be able to control the inaccuracies of the average reward. In other words, the padding function is large when the arm has been played only a few amount of times and therefore the inaccuracy in the average reward is higher. After playing the arm more times, the average reward will approach its actual value and at the same time the padding function shrinks which indicates the accuracy of the average reward estimation. Hence, the bigger the padding function is, the less experienced is that arm and the more inaccuracies are involved in the estimation of the average reward. A standard choice for the padding function is [141]

$$c_i^{(UCB-1)}(a_i) = B\sqrt{\frac{\zeta \log(i)}{N_i(a_i)}}, \tag{5.4}$$

where $N_i(a_i)$ is defined as Equation (5.3), B is an upper bound on the rewards and is selected as the maximum value of the observed reward through many trials of the algorithm, and $\zeta > 0$ is some appropriate constant and is selected based on empirical applications of the algorithm. By assigning appropriate values of ζ, this function can tune the performance of the UCB-1 algorithm to better results in stationary environments.

The UCB-1 algorithm starts by the initialization phase, where all of the available arms are played successively in order to obtain the initial information of all the arms. After playing each arm, the returned reward of that arm is observed by the algorithm and the upper bound of that arm is updated according to Equation (5.1). After the initialization phase, all the arms have been played once and the upper bound of all the arms are updated. Although, due to the fact that each arm has been played only once, the padding functions are large and the average reward informations are still inaccurate.

After obtaining the initial information, the next arm is selected by algorithm. The selection policy is the selection of the arm with highest upper confidence bound among all the available arms which represents a higher reward in an average sense.

This selection policy can be formally expressed as

$$a_i = \underset{1 \leq n \leq N_{arm}}{\arg\max} U_i^{(UCB-1)}(a_n). \tag{5.5}$$

The signal transmission is then performed through the selected arm and at the end of the episode the returned reward of that arm is observed and once again the upper bound of the selected arm is updated according to Equation (5.1). This process repeats until the end of the transmission. Algorithm 1 summarizes the UCB-1 algorithm.

Algorithm 1 UCB-1

1: **for** i=1 **to** N_{arm} **do**
2: play arm $a_i = i$.
3: observe the reward $X_i^{(UCB-1)}(a_i)$.
4: calculate the average reward of a_i by Equation (5.2).
5: calculate the padding function of a_i by Equation (5.4).
6: update the upper bound of a_i by Equation (5.1).
7: **end for**
8: **for** i=$N_{arm} + 1$ **to** N_{ep} **do**
9: play arm a_i selected by Equation (5.5).
10: observe the reward $X_i^{(UCB-1)}(a_i)$.
11: calculate the average reward of a_i by Equation (5.2).
12: calculate the padding function of a_i by Equation (5.4).
13: update the upper bound of a_i by Equation (5.1).
14: **end for**

It has been proven [141], that the UCB-1 algorithm performs optimally in stationary environments. The following theorem formally this proposition.
Theorem 1. The UCB-1 algorithm is optimal in the sense that its expected regret matches the lower bound regret of all policies for stationary bandit problems [141].

5.2.2 Discounted UCB Algorithm

Theorem 1 denotes the optimality of the UCB-1 algorithm for stationary bandits. However, for non-stationary bandits the UCB-1 algorithm cannot be considered as an optimal policy. This is due to the fact that for non-stationary bandits, the environment is subject to the changes in the fundamental parameters of the environment, however the UCB-1 algorithm does not consider these changes and therefore is not able to react to them in a timely manner.

For non-stationary bandits the reward values are assumed to be from an unknown distribution which is also non-stationary, therefore the optimal selection policy must have the ability of adaptation to the changes of the statistical characteristics of the values of rewards. For this reason, *discounted upper confidence bound* (D-UCB) algorithm has been introduced in [145], where a discount factor $\gamma \in (0, 1)$ has been introduced to the UCB algorithm to mark the affects of the time in which

the selected arm has been played. This means the older selected actions do not have equal weights in calculating the empirical mean of the rewards, whereas the more recent selected actions weigh more in the calculation of the confidence bound function. The discount factor γ plays a major role in the outcome of the algorithm, therefore it is the most important parameter in this algorithm. Figure 5.1 illustrates the weight factor for UCB-1 algorithm and the D-UCB algorithm.

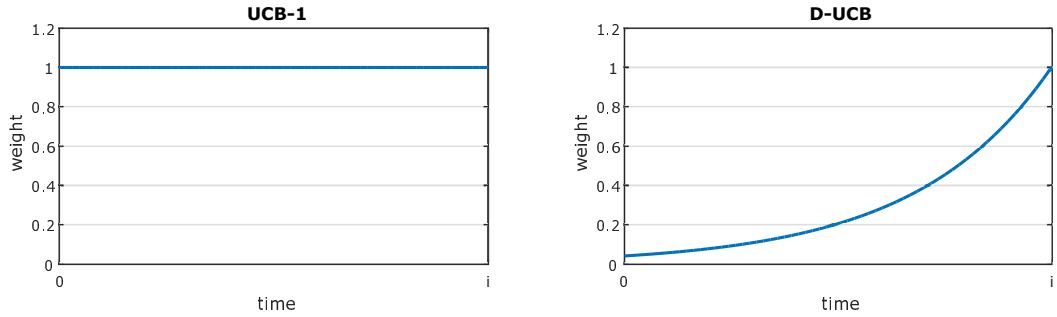


FIGURE 5.1: Weight factors in confidence bound calculations for UCB-1 and D-UCB algorithms.

In the D-UCB algorithm, at each episode i, an arm a_i is selected and the corresponding reward $X_i^{(D-UCB)}(a_i)$ will be observed. The average reward function is constructed from the past observed rewards with appropriate weights and represents the existing knowledge of the environment. Therefore, the average reward of arm a_i can be defined as

$$\bar{X}_i^{(D-UCB)}(a_i,\gamma) = \frac{1}{N_i(a_i,\gamma)} \sum_{s=1}^{i} \gamma^{i-s} X_s^{(D-UCB)}(a_i) \mathbb{1}_{\{a_s=a_i\}}, \tag{5.6}$$

where

$$N_i(a_i,\gamma) = \sum_{s=1}^{i} \gamma^{i-s} \mathbb{1}_{\{a_s=a_i\}}, \tag{5.7}$$

and

$$\mathbb{1}_{\{a_s=a_i\}} = \begin{cases} 1 & a_s = a_i \\ 0 & a_s \neq a_i \end{cases} \tag{5.8}$$

The term γ^{i-s} ensures that the recent results of the observed rewards become higher weights compared to the old ones. This weight reduces the influence of the outdated information and hence results in a quicker reaction of the algorithm to the changes of the environment.

The padding function on the other hand needs to be appropriately weighted so that the more recent rewards receiving a larger padding function. With a uniform weighting of the padding function, all the rewards contribute with the same weight and the information of the age of the reward is lost. In order to use this information, a proper padding function should be designed to increase the contribution of the most recent results, and at the same time, reduce the contribution of the older

rewards. A common choice of the padding function for this situation is

$$c_i^{(D-UCB)}(a_i,\gamma) = B\sqrt{\frac{\zeta \log(n_i(\gamma))}{N_i(a_i,\gamma)}}, \tag{5.9}$$

where

$$n_i(\gamma) = \sum_{i=1}^{N} N_i(\gamma, a_i). \tag{5.10}$$

Therefore, from Equation (5.6) and (5.9), the upper confidence function of the arm a_i which is denoted by $U_i^{(D-UCB)}(a_i)$ can be written as

$$U_i^{(D-UCB)}(a_i,\gamma) = \bar{X}_i^{(D-UCB)}(a_i,\gamma) + c_i^{(D-UCB)}(a_i,\gamma). \tag{5.11}$$

The D-UCB algorithm starts by the initialization phase same as the UCB-1 algorithm. In the initialization phase, all of the available arms are played successively in order to obtain the initial information of all the arms. After playing each arm, the returned reward of that arm is observed by the algorithm and the upper bound of that arm is updated according to Equation (5.11) which has been weighted accordingly. After the initialization phase, all the arms have been played once and the upper bound of all the arms are updated. After obtaining the initial information, the next arm is selected by the algorithm. The selection policy is the selection of the arm with highest upper confidence bound among all the available arms which represents a higher reward in an average sense. This selection policy can be formally expressed as

$$a_i = \underset{1\leq n\leq N_{arm}}{\arg\max}\, U_i^{(D-UCB)}(a_n,\gamma). \tag{5.12}$$

The signal transmission is then performed through the selected arm and at the end of the episode the returned reward of that arm is observed and once again the upper bound of the selected arm is updated according to Equation (5.11). This process repeats until the end of the transmission. The D-UCB algorithm is similar to the UCB-1 algorithm but with the aforementioned functions.

Theorem 2. Discounted upper-confidence bound algorithm is almost optimal in the sense that its expected regret matches the lower bound regret of all policies for non-stationary bandit problems [144], [145].

5.2.3 Sliding Window UCB Algorithm

Another algorithm which deals with the abrupt changes of the environment in a non-stationary bandit is the sliding window UCB (SW-UCB) algorithm. This algorithm was proposed in [144] and presents another solution for assigning different weights on the observed rewards based on the time of occurrence. The main difference between D-UCB and SW-UCB algorithms is that in D-UCB algorithm the weight factor is distributed as an exponential function over all the rewards so that old rewards are assigned with a low weight factors and recent rewards are assigned

with a high factor, as illustrated in Figure 5.1. However, in SW-UCB algorithm the weight factor function is denoted as a sliding window function where the old rewards do not contribute to the calculation of the average reward and only the recent rewards contribute to the calculation of the average reward. Moreover, the recent rewards have the same weight in the calculation of the average reward. This has been illustrated in Figure 5.2. It can be seen that in this algorithm the size of the window has an important role in the outcome of the algorithm.

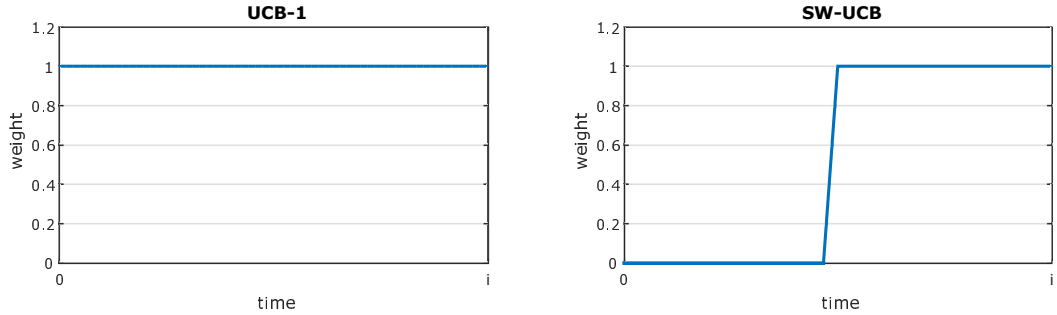


FIGURE 5.2: Weight factors in confidence bound calculations for UCB-1 and SW-UCB algorithms.

Let us assume the observed reward at the end of an episode where the action a_i has been selected, as $X_i^{(SW-UCB)}(a_i)$. The window size is set to be τ, which is selected by empirical calculations to tune the algorithm in order to retrieve best results. The average reward function in this case is constructed by averaging the past τ rewards of a_i and ignoring the older rewards. Therefore, the average reward of arm a_i can be defined as

$$\bar{X}_i^{(SW-UCB)}(a_i, \tau) = \frac{1}{N_i(a_i, \tau)} \sum_{s=i-\tau+1}^{i} X_s^{(SW-UCB)}(a_i) \mathbb{1}_{\{a_s=a_i\}}, \tag{5.13}$$

where

$$N_i(a_i, \tau) = \sum_{s=i-\tau+1}^{i} \mathbb{1}_{\{a_s=a_i\}}. \tag{5.14}$$

Note that the summation over the rewards is only considering the past τ elements, however with the same weight. This is as opposed to the weighting system in which all the past rewards have some weights increasing over time. This method of weighting helps to disregard the outdated information which may not be still valid due to changes of the environment.

The padding function of the algorithm follows the same logic and only the past τ elements are considered in the calculation of the padding function and the rest are disregarded. These past τ elements, however, have the same weights in the calculation of the padding function. This passing function also needs to be designed and a range of different functions can be applied to the SW-UCB algorithm. A

common choice of the padding function for the SW-UCB algorithm is

$$c_i^{(SW-UCB)}(a_i,\tau) = B\sqrt{\frac{\zeta \log(\min(i,\tau))}{N_i(a_i,\tau)}}, \tag{5.15}$$

where $\min(i,\tau)$ denotes the minimum of i and τ.

Therefore, from Equation (5.13) and (5.15), the upper confidence function of the arm a_i which is denoted by $U_i^{(SW-UCB)}(a_i)$ can be written as

$$U_i^{(SW-UCB)}(a_i,\tau) = \bar{X}_i^{(SW-UCB)}(a_i,\tau) + c_i^{(SW-UCB)}(a_i,\tau). \tag{5.16}$$

The SW-UCB algorithm starts by the initialization phase same as the UCB-1 and D-UCB algorithms. In the initialization phase, all of the available arms are played successively in order to obtain the initial information of all the arms. After playing each arm, the returned reward of that arm is observed by the algorithm and the upper bound of that arm is updated according to Equation (5.16) which only uses the past τ samples in the calculations. After the initialization phase, all the arms have been played once and the upper bound of all the arms are updated. After obtaining the initial information, the next arm is selected by the algorithm. The selection policy is the selection of the arm with highest upper confidence bound among all the available arms which represents a higher reward in an average sense. This selection policy can be formally expressed as

$$a_i = \underset{1\leq n\leq N_{arm}}{\arg\max}\, U_i^{(SW-UCB)}(a_n,\tau). \tag{5.17}$$

The signal transmission is then performed through the selected arm and at the end of the episode the returned reward of that arm is observed and once again the upper bound of the selected arm is updated according to Equation (5.16) with only the last τ samples contribute to the calculations. This process repeats until the end of the transmission. The SW-UCB algorithm is similar to the UCB-1 algorithm but with the aforementioned functions.

Theorem 3. Sliding window upper-confidence bound algorithm is almost optimal in the sense that its expected regret matches the lower bound regret of all policies for non-stationary bandit problems [144].

5.2.4 The Proposed UCB Algorithms: Introduction

As described in Chapter 2, the PLC noise and the channel transfer function can be modeled as a cyclostationary random process. In a cyclostationary process, the statistical characteristics of the process repeat periodically, which in the PLC case, this period is matched to the mains period. Therefore, in the calculation of the empirical mean and the padding function of the UCB-1 algorithm, considering all the previous actions with the same weight, is not an optimal policy and can be

further optimized by exploiting the cyclostationary nature of the channel. Furthermore, the mere consideration of the recent past actions as major contributors to the calculation of the confidence index, as in D-UCB and SW-UCB algorithms, may result in a sub-optimal policy as well, since the far past actions in the same cycle and hence with the same statistical characteristics are neglected due to their low discount weight or even completely disregarded if they lie outside of the weight window. To overcome this problem, we propose two novel algorithms, namely

1. The cyclic discounted upper-confidence bound (CD-UCB) algorithm, and
2. The cyclic window upper-confidence bound (CW-UCB) algorithm.

Furthermore, through simulation results, we show that for a cyclostationary system like PLC, these algorithms result in a better selection policy, and therefore an overall better performance. In the following we describe in detail the proposed algorithms.

5.2.5 The Proposed CD-UCB Algorithm

Let us assume the period of the alternating current (AC) waveform of the power lines as T_{AC} with the noise power and hence the rewards of each arm as a cyclostationary process with T_{AC} duration of each cycle. Up to the time index $t = T$, the total number of complete cycles, denoted by P, can be calculated as

$$P = \left\lfloor \frac{T}{T_{AC}} \right\rfloor. \tag{5.18}$$

Furthermore, the empirical mean value of each arm at time i, as well as the padding function, are to be calculated in a way that all the last cycles are included in the calculations. The weighing factor should be chosen in a way that it involves the cyclostationary behavior of the reward. For this purpose, let us consider a single period of time with duration T_{AC}. In this period, the first samples will have lower weights and the last samples will have higher weights. In other words, we use a similar discount factor as in the D-UCB algorithm for each period. Let us assume the observed reward of the arm a_i as $X_i^{(CD-UCB)}(a_i)$. For the last P complete periods, we calculate the empirical means of of arm a_i as

$$\begin{aligned} \Xi_2 &= \Xi_2(a_i, T_{Ac}, \gamma) \\ &= \sum_{p=1}^{P} \sum_{s=t-pT_{AC}}^{i-(p-1)T_{AC}} \gamma^{i-(p-1)T_{AC}-s} X_s^{(CD-UCB)}(a_i) \mathbb{1}_{\{a_s=a_i\}}, \end{aligned} \tag{5.19}$$

where the first summations goes through all the periods and in the second summation the term $\gamma^{t-(p-1)T_{AC}-s}$ applies a discounted factor for each period separately. The start of the period starts with a low weight factor and the weight factor increases as time reaches the end of the period.

Equation (5.19) only accounts for the complete periods, however, as depicted in Figure 5.3, the first period may be an incomplete period. For the incomplete period

at the beginning of the time index (see Figure 5.3), we calculate the corresponding portion of the empirical mean as

$$\begin{aligned}\Xi_1 &= \Xi_1(a_i, T_{Ac}, \gamma) \\ &= \sum_{s=1}^{t-PT_{AC}} \gamma^{t-PT_{AC}-s} X_s^{(CD-UCB)}(k) \mathbb{1}_{\{a_s=k\}}. \end{aligned} \tag{5.20}$$

This applies a discount factor to the incomplete period. Without doing so, the information available in this incomplete period is lost, and by using this the algorithm becomes a little stronger. Therefore, for the whole time $1 \leq t \leq T$, the average reward of the action a_i can be calculated as

$$\bar{X}_i^{(CD-UCB)}(a_i, T_{AC}, \gamma) = \frac{1}{N_i(a_i, T_{AC}, \gamma)}(\Xi_1 + \varkappa\, \Xi_2), \tag{5.21}$$

where $\varkappa = sign(P)$ returns one when the number of periods is more than one, and returns zero when the total time is less than a period. The term $N_t(a_i, T_{AC}, \gamma)$ is described as

$$\begin{aligned} N_i(a_i, T_{AC}, \gamma) &= \sum_{s=1}^{i-PT_{AC}} \gamma^{i-PT_{AC}-s} \mathbb{1}_{\{a_s=a_i\}} \\ &\quad + \varkappa \sum_{p=1}^{P} \sum_{s=i-pT_{AC}}^{i-(p-1)T_{AC}} \gamma^{i-(p-1)T_{AC}-s} \mathbb{1}_{\{a_s=a_i\}}. \end{aligned} \tag{5.22}$$

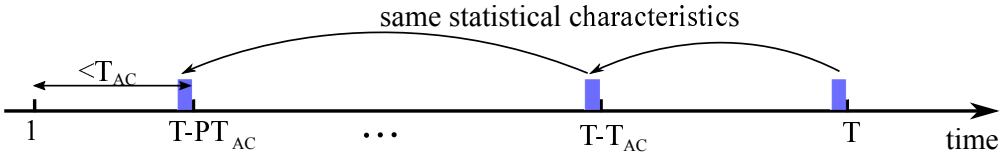


FIGURE 5.3: The cyclostationary behavior of the channels with mains period T_{AC}.

The padding function for the CD-UCB algorithm is calculated similar to the average reward calculations. This means that all the previous rewards contribute to the padding function, but with different weights and these weights are depending on the periods of the cyclostationary behavior of the noise. Therefore, each period receives separate weighing factors from the start of the episodes to the current time of operation. The padding of the proposed CD-UCB algorithm can be written as

$$c_i^{(CD-UCB)}(a_i, T_{AC}, \gamma) = B\sqrt{\frac{\zeta \log(n_i(T_{AC}, \gamma))}{N_i(a_i, T_{AC}, \gamma)}}, \tag{5.23}$$

where

$$n_i(T_{AC}, \gamma) = \sum_{i=1}^{N} N_i(a_i, T_{AC}, \gamma). \tag{5.24}$$

Therefore, from Equation (5.21) and (5.23), the upper confidence function of the arm a_i which is denoted by $U_i^{(CD-UCB)}(a_i)$ can be written as

$$U_i^{(CD-UCB)}(a_i, T_{AC}, \gamma) = \bar{X}_i^{(CD-UCB)}(a_i, T_{AC}, \gamma) + c_i^{(CD-UCB)}(a_i, T_{AC}, \gamma). \tag{5.25}$$

The proposed CD-UCB algorithm starts by the initialization phase same as the UCB-1, D-UCB, and SW-UCB algorithms. In the initialization phase, all of the available arms are played successively in order to obtain the initial information of all the arms. After playing each arm, the returned reward of that arm is observed by the algorithm and the upper bound of that arm is updated according to Equation (5.25) which uses the separate discounts for each period in the calculations. After the initialization phase, all the arms have been played once and the upper bound of all the arms are updated. After obtaining the initial information, the next arm is selected by the algorithm. The selection policy is the selection of the arm with highest upper confidence bound among all the available arms which represents a higher reward in an average sense. This selection policy can be formally expressed as

$$a_i = \underset{1 \leq n \leq N_{arm}}{\arg\max}\, U_i^{(CD-UCB)}(a_n, T_{AC}, \gamma). \tag{5.26}$$

The signal transmission is then performed through the selected arm and at the end of the episode the returned reward of that arm is observed and once again the upper bound of the selected arm is updated according to Equation (5.25) with discounted factors for all the periods accordingly. This process repeats until the end of the transmission. The proposed CD-UCB algorithm is similar to the UCB-1 algorithm but with the aforementioned functions.

5.2.6 The Proposed CW-UCB Algorithm

Similar to the approach in CD-UCB algorithm, we propose another algorithm based on UCB algorithm which its weight factor is periodic, hence adapted to the cyclo-stationary behavior of the channel. In CD-UCB algorithm, the statistical characteristics of the reward function at time $t = T$ and $t = T - T_{AC}$ are the same, and therefore the weight factor at these times is at its maximum. However, at time $t = T - T_{AC} + 1$, the reward function enters the next cycle and results in a low value of weight factor. Although, the statistical characteristics of the reward function at time $t = T - T_{AC}$ and $t = T - T_{AC} + 1$ are not that different. This problem can be addressed with the proposed cyclic window UCB algorithm.

In CW-UCB algorithm, the empirical mean and the padding function are not discounted as in the CD-UCB algorithm, but are windowed periodically. The windowing period is chosen to be matched with the periodic behavior of the reward

and hence matched with T_{AC}. Therefore, the most important parameter in this algorithm is the size of the window. The window size, W, is selected in a way to maximize the effect of windowing according to the period of operation. Figure 5.4 illustrates the weight factors in the calculations of average reward and the padding function in CD-UCB and CW-UCB algorithms.

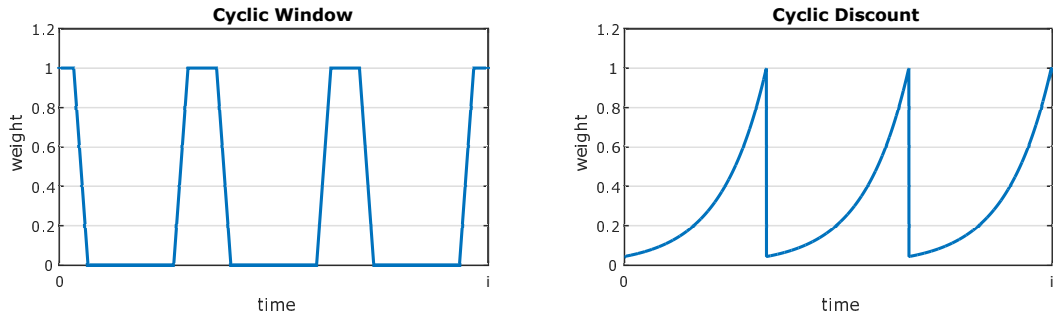


FIGURE 5.4: Weight factors in confidence bound calculations for the proposed CW-UCB and the proposed CD-UCB algorithms.

Let us assume the observed reward of arm a_i as $X_i(a_i)$. The average reward is windowed periodically. Formally, the average reward can be expressed as

$$\bar{X}_i^{(CW-UCB)}(a_i, W, T_{AC}) = \frac{1}{N_i(a_i, W, T_{AC})} \sum_{p=0}^{P} \sum_{s=1}^{i} w(s - i + pT_{AC}) X_s(a_i) \mathbb{1}_{\{a_s = a_i\}}, \tag{5.27}$$

where $w(s)$ is the window function and is defined as

$$w(s) = \begin{cases} 1 & |s| < \frac{W}{2} \\ 0 & \text{otherwise} \end{cases} \tag{5.28}$$

The term $w(s - t + pT_{AC})$ in (5.27) denotes that the windowing is performed at the current time in addition to the multiple times of the mains period before the current time (see Figure 5.4 for an illustration of this idea). This ensures that the periodicity of the PLC channel environment is considered in the calculation of the average reward. This enables the algorithm to adapt itself to the features of the PLC channel and perform better in these environments. The term $N_i(a_i, W, T_{AC})$ can be described as

$$N_i(a_i, W, T_{AC}) = \sum_{p=0}^{P} \sum_{s=1}^{i} w(s - i + pT_{AC}) \mathbb{1}_{\{a_s = a_i\}}. \tag{5.29}$$

The padding function for the CW-UCB algorithm is calculated similar to the average reward calculations. Therefore, each period receives separate weighing factors from the start of the episodes to the current time of operation. This is also similar to the calculations of the CD-UCB algorithm. The padding of the proposed CW-UCB

algorithm can be written as

$$c_i^{(CW-UCB)}(a_i, T_{AC}, W) = B\sqrt{\frac{\zeta \log(n_i(T_{AC}, W))}{N_i(a_i, T_{AC}, W)}}, \tag{5.30}$$

where

$$n_i(T_{AC}, W) = \sum_{i=1}^{N} N_i(a_i, T_{AC}, W). \tag{5.31}$$

Therefore, from Equation (5.27) and (5.30), the upper confidence function of the arm a_i which is denoted by $U_i^{(CW-UCB)}(a_i)$ can be written as

$$U_i^{(CW-UCB)}(a_i, T_{AC}, W) = \bar{X}_i^{(CW-UCB)}(a_i, T_{AC}, W) + c_i^{(CW-UCB)}(a_i, T_{AC}, W). \tag{5.32}$$

The proposed CW-UCB algorithm starts by the initialization phase same as the UCB-1, D-UCB, SW-UCB, and CD-UCB algorithms. In the initialization phase, all of the available arms are played successively in order to obtain the initial information of all the arms. After playing each arm, the returned reward of that arm is observed by the algorithm and the upper bound of that arm is updated according to Equation (5.32) which uses the separate windows for each period in the calculations. After the initialization phase, all the arms have been played once and the upper bound of all the arms are updated. After obtaining the initial information, the next arm is selected by the algorithm. The selection policy is the selection of the arm with highest upper confidence bound among all the available arms which represents a higher reward in an average sense. This selection policy can be formally expressed as

$$a_i = \underset{1 \le n \le N_{arm}}{\arg\max}\, U_i^{(CW-UCB)}(a_n, T_{AC}, W). \tag{5.33}$$

The signal transmission is then performed through the selected arm and at the end of the episode the returned reward of that arm is observed and once again the upper bound of the selected arm is updated according to Equation (5.32) with window factors for all the periods accordingly. This process repeats until the end of the transmission. The proposed CW-UCB algorithm is similar to the UCB-1 algorithm but with the aforementioned functions.

The proposed algorithms consist of an initialization phase which its duration is proportional to the number of relays. Then at each frame length the end-to-end capacity has to be obtained and instantaneous rewards are calculated at the transmitter. The calculation of reward for the selected arm consists of a linear calculation of the empirical mean in addition to the calculation of the padding function which consists of a square root and a logarithm function. These calculations are only done in one arm at each frame time. The algorithms are bounded in the sense that after a limited amount of time, the arm with the best rewards can be detected and selected for most of the consecutive selections. However, the amount of time needed for this convergence, denoted by $\tau_{alg}(N)$, is directly proportional to the number of relays N (see Figure 5.14). Let us denote the time in which the relay channels remain

in a particular state as τ_{ch}. In order to have a working algorithm, change of the channel in time must happen slower than the convergence time of the algorithm and $\tau_{alg}(N) < \tau_{ch}$ must hold. The exhaustive search method in the simulations, assumes the perfect CSI at transmitter and expectedly the returned reward is higher than learning algorithms. However, acquiring CSI is much more complicated than feeding back the observed reward. The reason for that is that the PLC channel is time-variant and frequency-selective. Therefore, pilot signals should be transmitted in all the subcarriers at pre-defined time intervals throughout the transmission. Moreover, the amount of overhead which this brings increases linearly by the number of available relays, since in order to react to the changes in the environment all the possible routs should be evaluated. On the other hand, feeding back the reward data can take place on the free bits of the ACK (acknowledgment) packet which is already being fed back to the transmitter and is not dependent on the number of available relays.

5.2.7 Performance Evaluation of UCB Algorithms

In this section we present numerical results to evaluate the performance of the above-mentioned UCB algorithms. For this reason we consider two selection problems in PLC, namely the PLC channel selection problem and the PLC relay selection problem as described in Chapter 4. In Chapter 4 these two selection problem have been modeled as a MAB problem with well-defined reward and regret functions. We use the reward and regret function of each case to apply the UCB algorithms. The simulation results show significant improvement when the machine learning algorithm is applied. In the following we present these simulation results as well as a discussion on the performance analysis.

5.2.7.1 UCB algorithms in PLC Channel Selection

We consider a PLC system with four channels as described in Chapter 2, however generalization to more than four channels is straight forward. Channel gains are considered to have log-normal distribution as discussed in Chapter 2. Channels are considered to be non-stationary, and we assume that the mean value of the log-normal distribution of each channel varies over time in a piece-wise stationary manner, as it is depicted in Figure 5.5. As illustrated in this figure, two of the channels are not changing over time whereas the other two change abruptly at time instances which are a priori unknown to the transmitter.

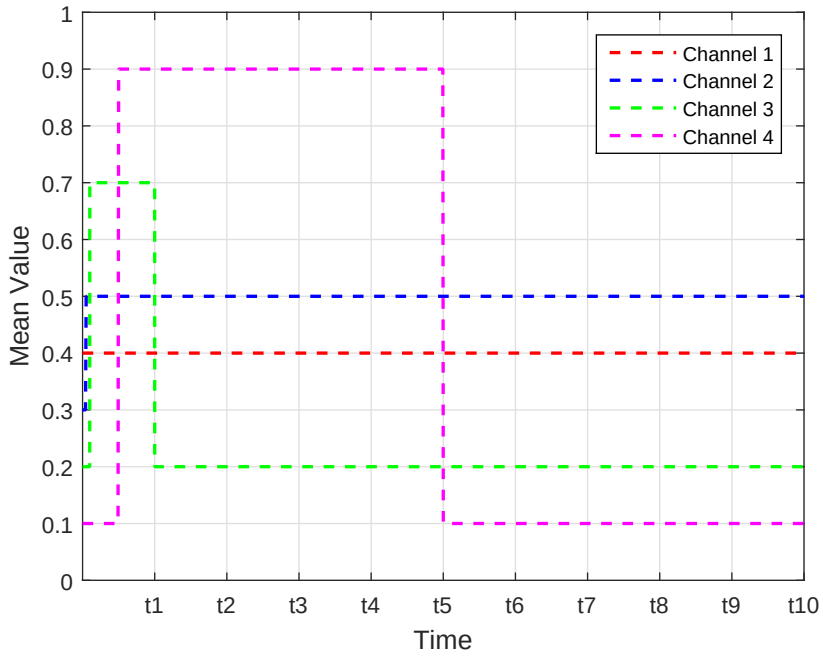


FIGURE 5.5: Exemplary time-variation in mean value for different PLC channels.

From the figure 5.5, it can be concluded that the optimal choice of channel on an average base yields:

$$a_i = \begin{cases} 1 & \text{if } t < t_1 \\ 2 & \text{if } t_1 \leq t < t_2 \\ 3 & \text{if } t_2 \leq t < t_3 \\ 4 & \text{if } t_3 \leq t < t_4 \\ 2 & \text{if } t \geq t_4 \end{cases}, \tag{5.34}$$

with $t_1 = 5 \times 10^2$, $t_2 = 10^3$, $t_3 = 5 \times 10^3$and $t_4 = 5 \times 10^4$. We evaluate the performance of the proposed model by using two selection strategies, namely the D-UCB algorithm and the SW-UCB algorithm as described previously. In doing so, we select $T = 10^5$, $\zeta = 0.55$, $\tau = 7 * 10^2$, $B = 1$, $\gamma = 0.9$ based on empirical results of the algorithms. These parameters has been chosen in regard to the criteria described in [144], in which the parameters are chosen so that the optimal result is achieved for piece-wise stationary channels.

For the sake of comparison, we use two other strategies as benchmarks for performance comparison:

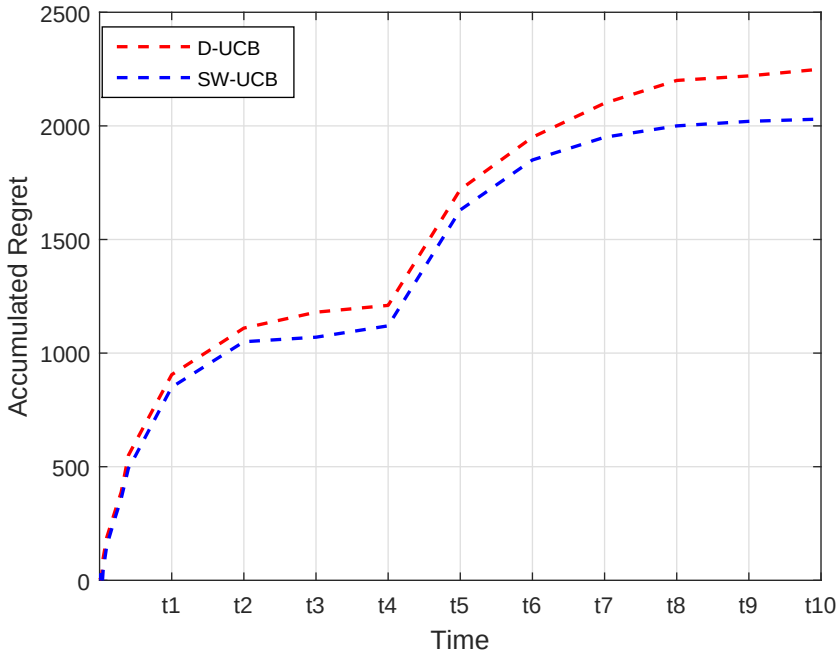


FIGURE 5.6: Regret growth of two algorithms of D-UCB and SW-UCB for the designed channel selection bandit model.

1. Optimal channel selection given full information, which finds the best channel with the largest mean value of channel gains by exhaustive search; and
2. Random channel selection, where at each time a channel is selected uniformly at random.

The regret growth of the two algorithms is shown in Figure 5.6. As it is clear from this figure, at change points, where the average gain of some channel, thereby the optimal choice of channel, changes, the regret increases very fast. The reason is that both algorithms require some time to detect the change and adapt their selection. However, after the best channel is selected, the regret growth slows down, thus the accumulated regret remains almost fixed. It should be also noted that if the changes occur in short time intervals, like the first and second change-points depicted in Figure 5.5, a phase, where the accumulated regret remains almost fixed cannot be observed, simply since the optimal channels changes again even before (or right after) the algorithm adapts its choice. Thus, it can be concluded that the proposed model and approach is more suitable for slowly-varying environments. Moreover, it can be observed that the performance of *SW-UCB* is slightly better than that of *D-UCB*.

Figure 5.7 depicts the average reward achieved by the proposed channel selection

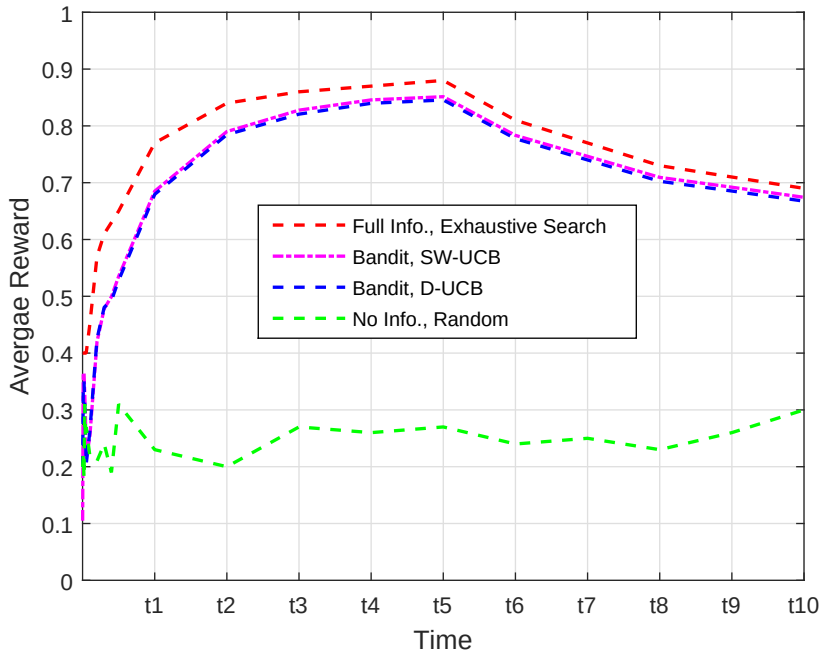


FIGURE 5.7: Performance of proposed bandit model and solution in comparison with that of exhaustive search given full information and uniformly random channel selection.

model in comparison with those of exhaustive search given full information and uniformly random selection. From this figure, it can be concluded that the model exhibits acceptable performance which only slightly differs from a fully informed channel selection. Furthermore it can be observed that the two MAB channel selection algorithms result in an approximately equal performance. Also note that the average utility increases in the time periods where the average gain of the selected (optimal) channel is large, for example in the time $t_3 \leq t < t_4$, where the optimal channel (Channel 4) has an average gain equal to $\mu_t = 0.9$. Shortly after, in time $t \geq t_4$, when Channel 2 with average $\mu_t = 0.5$ becomes optimal, average decreases, and the final average yields ≈ 0.7.

5.2.7.2 UCB Algorithms in PLC Relay Selection

We consider a two-hop cooperative communication with a PLC channel as described in Chapter 4. The number of available relays is considered to be 6 nodes, from which a single node is selected for transmission at each transmission instant. However the results for more and less number of relays are presented as well in order to observe the effects of the number of relays on the algorithms. The only difference between

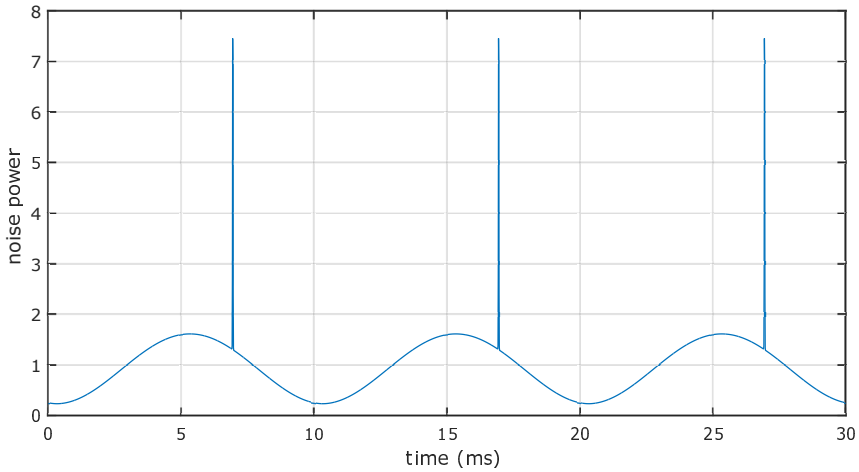


FIGURE 5.8: PLC noise power.

the available relay nodes is the corresponding link rate, which in turn is dependent on the channel response as well as the noise power spectral density of the PLC channel as described in the previous chapter.

The impulsive noise of the power line channel is considered to be periodic and synchronous to the frequency of the mains as described in Chapter 2, with a variance determined as cyclostationary function. Figure 5.8 demonstrates the periodic impulsive noise in the power lines which is synchronous to the frequency of the mains $f_m = 50Hz$.

Different relay selection policies are applied to the PLC system to acquire a quantitative comparison between the existing policies in the literature and the selection policies described in the proposed algorithms. Formally, the following approached have been examined:

1. In the first method, an exhaustive search is performed to find the best available relay node for transmission with the assumption of perfect CSI. This returns the highest reward at each time instance, or equivalently the maximum achievable reward for other algorithms which is a function of the channel conditions.

2. In the second method, a fixed relay node is selected at the beginning of the transmission in a random manner and the selected node is used for the entire transmission time, regardless of the changes in the channel.

3. In the third method the relay node is selected randomly for each instant of transmission. Therefore the all the available nodes have the chance of being selected, but the changes in channel does not play a role in the selection procedure.

4. The fourth and the fifth methods of relay selection are the basic UCB and Discounted UCB algorithms described previously.

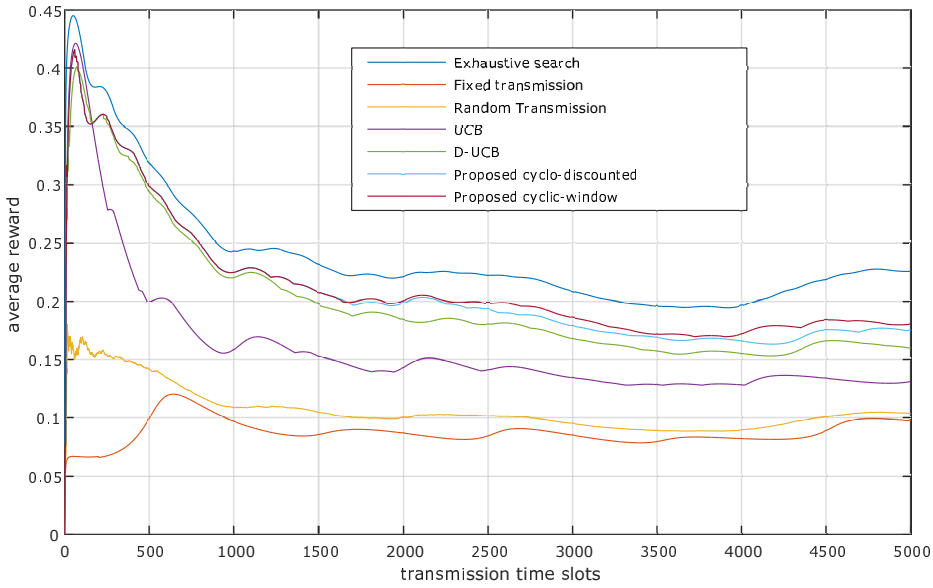


FIGURE 5.9: Average reward comparison between different relay selection policies.

5. Finally, the sixth and seventh selection policies are the proposed CD-UCB and CW-UCB algorithms.

Figure 5.9 and 5.10 show the average reward and the accumulated regret of the above mentioned relay selection policies, respectively. It can be seen that the best performance belongs to the exhaustive search method, as expected. The worst performance is from the random selection of the relay node and the fixed relay node, due to the fact that the changes in channel is disregarded in these methods. Basic UCB and discounted-UCB algorithms provide a significant performance gain in terms of average reward and demonstrate lower regrets throughout the transmission. It can be seen that the proposed algorithms can further improve the performance and obtain more reward and lower regret compared to the existing algorithms. Although, at the beginning of the operation, the padding functions are still large for all the arms and therefore a uniform weight as in UCB outperforms the discounted weights of the proposed algorithms.

Figure 5.11 illustrates the percentage of correct selection in different transmission policies. This percentage is calculated after the entire duration of the transmission, by comparing the selected relay nodes of each method to the selected relay nodes of the exhaustive search method as a measure of reference. It can be seen that the basic MAB algorithms result in a higher percentage of correct selection and the proposed MAB algorithms can achieve yet higher percentages. This visualization helps us to determine the accuracy and the level of success for each individual algorithm. It can be seen that both the proposed algorithms significantly improve the state of the art algorithms.

The first proposed algorithm, namely the cyclo-discounted UCB algorithm, uses

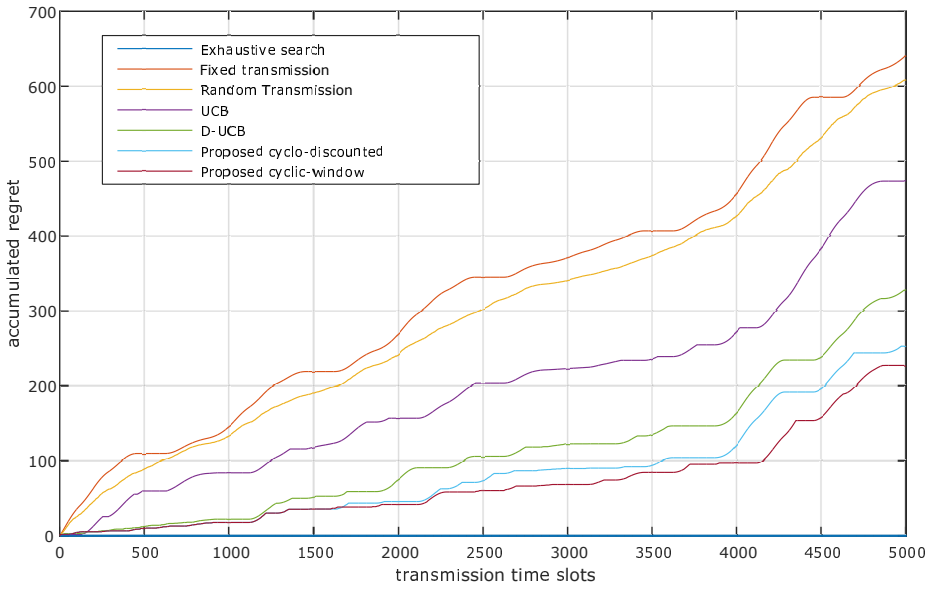


FIGURE 5.10: Accumulated regret comparison between different relay selection policies.

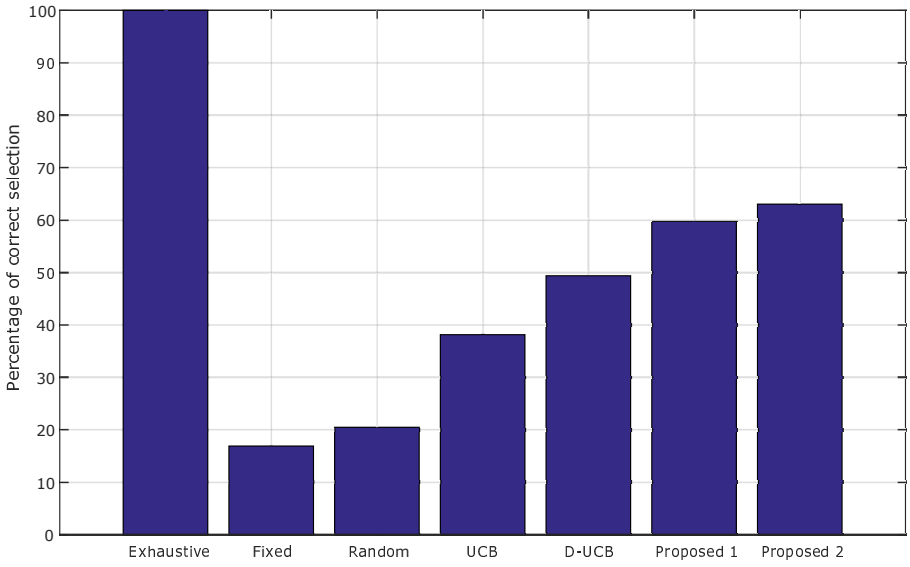


FIGURE 5.11: Percentage of correct selection of different transmission policies.

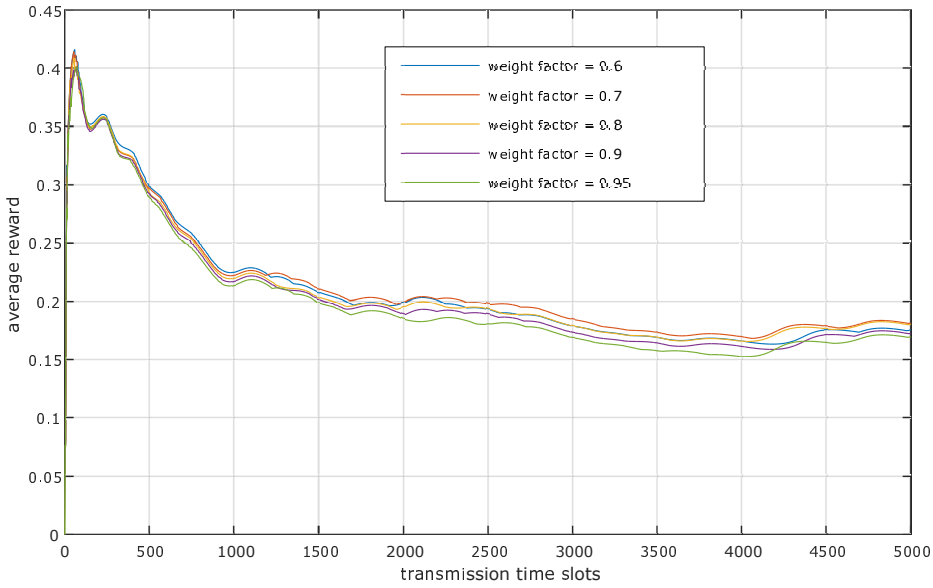


FIGURE 5.12: Impact of weight factor in the proposed cyclo-discounted UCB algorithm.

a weight factor as a parameter of cyclic weights for the calculation of the UCB index, whereas the second proposed algorithm, namely the cyclic-window UCB algorithm, uses a window size parameter for the calculation of the UCB index. Figure 5.12 and 5.13 demonstrates the impact of different weight factors and window sizes, respectively. Compared with Figure 5.9 we can realize that the change in window size or weight factor can relatively affect the performance of the algorithms. These parameters can be determined beforehand by empirical results which are gathered by running the algorithm in real environments. The proper value of each parameter, i.e. the window size and the weight factors can then be implemented in the algorithms to tune them into better and more accurate results.

Figure 5.14 demonstrates the effect of the number of relays on the cyclic-window algorithm. It can be seen that in the same frame of time, with higher the number of relays it takes longer for the algorithm to adapt itself to the environment and the accumulated regret grows by the number of available relays. This is due to the fact that with a higher number of relays between the source and the destination of the cooperative communication system, the algorithm has more options to evaluate at the transmitter. Each option, i.e. relay node, must then have the possibility of being selected, which makes the probing time longer and the probability of selecting a sub-optimal relay is also higher. Therefore, with more relay nodes the probability of regret is increased and the algorithm converges with a slightly lower rate.

We have provided results of a two-hop relay selection scheme in PLC networks. However, the use of learning algorithms can be further generalized to multi-hop relay selection problems as well. In the case of M-hop transmission, with N_i, $i \in \{1, \cdots, M-1\}$ available relays at each hop, the number of available options raise to $\prod_{i=1}^{M-1} N_i$ which makes the time needed for finding the best path longer.

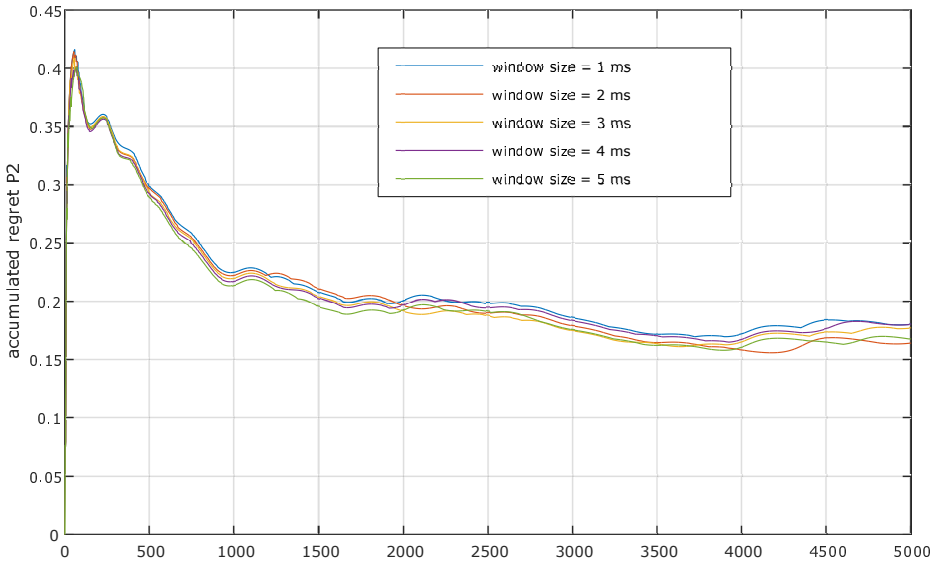


FIGURE 5.13: Impact of window size in the proposed cyclic-window UCB algorithm.

Finding the best strategy for multi-hop relay selection can be regarded as a future work for this problem. Moreover, a generalized version of the proposed algorithm can be developed to further increase the reliability of the transmission. In this case, a multi-relay selection approach based on the learning algorithm is applied and at the receiver an appropriate combining technique is performed to further exploit the combining diversity among the best chosen relays. This strategy can further improve the reliability and decrease the probability of transmission in a less optimal route.

5.3 Probability Matching Technique

5.3.1 Action Value Estimation

Another approach to deal with the selection problem in an environment of incomplete information with non-stationary conditions is the probability matching technique [150], [151]. The selecting agent, i.e. the PLC transmitter, is faced with the choice of maximizing the reward function based on its current knowledge of the channel, which eventually leads to a better performance, and trying to obtain more information about the PLC channel through exploration.

In this section, we consider the dynamic spectrum assignment problem, described in the previous chapter. The PLC transmitter is faced with the problem of dynamic spectrum assignment in an environment of incomplete information. The selection policy π is supposed to minimize its regret function based of its current knowledge

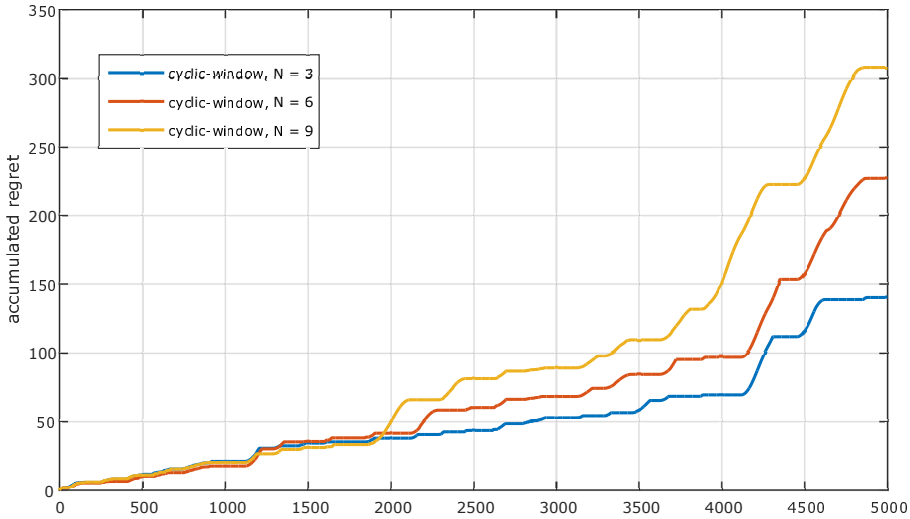


FIGURE 5.14: The proposed cyclic-window algorithm for different number of available relays.

of the environment and try to learn more about its environment in order to improve its decisions. This has been traditionally called the exploitation-exploration problem. Moreover, due to the time-variant nature of the PLC channel, the transmitter has to detect the changes of the environment throughout the transmission which is referred to as a non-stationary MAB problem. In this case, the best frequency band in some time of transmission does not necessarily remain the best band throughout the transmission and therefore the spectrum assignment must be performed dynamically.

In the non-stationary MAB problem, all the expected reward values $\bar{X}_e(a_e)$ are considered to be initially unknown and furthermore, at some unknown time during the transmission they tend to attain a new value. Therefore, the transmitter aims to estimate the actual expected reward values. This estimate is called an *action value estimate* and is denoted by $\hat{X}_q(a_e)$, where q is the number of times that action a_e has been selected by the transmitter. After the q-th instance of selection of the action a_e, the action value estimate of the selected arm is updated according to the following action value estimate update rule [146]

$$\hat{X}_q(B_n) = (1-\beta)^q \hat{X}_0 + \sum_{i=1}^{q} \beta(1-\beta)^{q-i} X_e(a_e = B_n), \tag{5.35}$$

where $\hat{X}_0$ is the initial estimation performed randomly, $X_e(a_e = B_n)$ is the observed expected reward when the action at episode e is set to be B_n, and β is a constant parameter which is obtained from empirical results of the algorithm and is tuned to result in the near optimal estimations. The weight given to the observed value of expected reward decreases exponentially as the age of the observation increases, making the more recent results play a bigger role in evaluating the value of action

value estimate, through which the algorithm is enabled to follow the changes which occur in the channel conditions.

In the following sections, we describe a reinforcement learning approach in order to define the selection policy π, namely the probability matching technique (PMT). Furthermore, we have modified the PMT algorithm to match the conditions of the PLC system model in order to further improve the performance of the algorithms.

5.3.2 The PMT Algorithm

The seminal PMT algorithms tries to assign selection probabilities to the available actions. These probabilities change over time as the selecting agent gathers more information about the environment. The most intuitive way of assigning probabilities is to make the selection probability $P(B_n)$ of action B_n (the selected frequency band) proportional to its current action value estimate. Therefore, we assign selection probabilities to each of the seven available actions. and the selection probability of action B_n or equivalently the probability of the selection of the frequency band B_n can be expressed as [150]

$$P(B_n) = \frac{\hat{X}_q(B_n)}{\sum_n \hat{X}_q(B_n)}. \tag{5.36}$$

The conventional PMT algorithm starts with equiprobable actions. The PLC transmitter performs the selection based on the probabilities and an action will be selected. The reward of the selected action is observed and the corresponding action value estimate is updated according to (5.35). Afterwards, based on the updated action value estimate of the selected action, all the action probabilities will be updated according to the probability matching rule of (5.36). Next action will be selected with the new probabilities and the process will start again. This has been summarized in Algorithm 2.

Algorithm 2 Conventional PMT

1: assign equal probabilities and action value estimates to all actions.
2: **for** i=1 **to** number of episodes **do**
3: select a band B_i based on action probabilities.
4: observe the reward $X_e(a_e = B_i)$ of the selected action.
5: update the action value estimate of the selected action based on (5.35) to obtain $\hat{X}_q^{new}(B_i)$.
6: update all the action probabilities based on updated action value estimates and equation (5.36).
7: **end for**

5.3.3 The Proposed PMT Algorithm

In PMT algorithm, we assign selection probabilities to each of the seven available actions and this probability is proportional to its current action value estimate. However, there is a possibility that at some point in time the probability of a certain action reaches a near-zero value and makes this action unavailable for the rest of the algorithm. This is not desired in a PLC system, since the channel is time-variant and an action after a long time of idleness may become of interest at some point during the transmission. To overcome this problem we introduce a minimum probability P_{min} which replaces the actual probability of the action in case it falls bellow P_{min}. Formally, the replacement probability rule can be expressed as

$$Q(B_n) = \max(P_{min}, P(B_n)). \tag{5.37}$$

However, since the sum of all the probabilities must equal unity, we express the new probability matching rule as

$$P_{new}(B_n) = \frac{Q(B_n)}{\sum_n Q(B_n)}. \tag{5.38}$$

As described in the system model, some of the frequency bands include other smaller bands and vice versa, some of the frequency bands can be combined to make a larger band. This results in dependencies between the frequency bands and allows us to exploit these dependencies in order to improve the algorithm. For this reason, we propose two modifications to the conventional algorithm as follows;

5.3.3.1 Similarity Threshold

First we define *grouped bands* as follows.
Definition 2. Two frequency bands are grouped bands (GBs) if the combination of their spectrum constitutes another frequency band defined by the system, which is referred to as the *upper band*.

For instance, as described in the previous chapter, the GBs of B_4 and B_5 have the upper band of B_2. We define a similarity metric between two GBs of B_i and B_j as $|\hat{X}_q(B_i) - \hat{X}_q(B_j)|$ and we define the *similar grouped bands* as follows.
Definition 3. Two GBs of B_i and B_j are said to be *similar grouped bands* (SGBs) if their similarity metric falls bellow a certain threshold δ, or equivalently

$$|\hat{X}_q(B_i) - \hat{X}_q(B_j)| < \delta. \tag{5.39}$$

With a proper selection of the similarity threshold δ, the performances of two SGBs can be estimated to be quite close to each other and consequently the selection probabilities assigned to these two actions are also close enough for the selecting agent (PLC transmitter) to treat these frequency bands as equals. Therefore, at every E episodes, if the selected band is a SGB, we force the selecting agent to

select the upper band instead which its performance is again similar to its SGBs. Doing so, enables the transmitter to achieve higher throughputs with the same performance. Although, we cannot do this for every episode and have to perform this modification only at every E episode, because if at every episode an upper band is selected, the two SGBs will never get a chance to be updated.

5.3.3.2 Multiple Action Value Estimate Update

We propose to not only update the action value estimate of the selected frequency band, but to update the action value estimate of the upper and/or lower bands of the selected action as well. Formally, let us assume that B_i and B_j are two GBs with action value estimates of $\hat{X}_q(B_i)$ and $\hat{X}_q(B_j)$, respectively. We define B_{up} as the upper frequency band with action value estimate defined as the mean value of the action value estimates of its two GBs, which can be expressed as

$$\hat{X}_q(B_{up}) = \frac{1}{2}\left(\hat{X}_q(B_i) + \hat{X}_q(B_j)\right). \tag{5.40}$$

Note that equation (5.40) is based on the empirical results of the simulated system. Let us denote the selected band by B_{sel}. There are two cases to consider:

- If $B_{sel} = B_i$ which is a GB (all other bands except for B_1), the action value estimate of B_i is updated to $\hat{X}_q^{new}(B_i)$ based on (5.35) as well as the action value estimate of B_{up} which will be updated to $\hat{X}_q^{new}(B_{up})$ and can be calculated as

$$\hat{X}_q^{new}(B_{up}) = \frac{1}{2}\left(\hat{X}_q^{new}(B_i) + \hat{X}_q(B_j)\right). \tag{5.41}$$

- If $B_{sel} = B_{up}$ which is an upper action (bands B_1, B_2, and B_3), the action value estimate of B_{up} is updated to $\hat{X}_q^{new}(B_{up})$ based on (5.35) as well as the action value estimates of its two GBs, B_i and B_j. Since the changes in the amount of action value estimates for both B_i and B_j should be the same and there is no priority assigned to either one of them, we can claim

$$\hat{X}_q^{new}(B_i) - \hat{X}_q(B_i) = \hat{X}_q^{new}(B_j) - \hat{X}_q(B_j). \tag{5.42}$$

  Furthermore, from (5.40) we have

$$\hat{X}_q^{new}(B_{up}) = \frac{1}{2}\left(\hat{X}_q^{new}(B_i) + \hat{X}_q^{new}(B_j)\right). \tag{5.43}$$

  From (5.42) we substitute $\hat{X}_q^{new}(B_i) = \hat{X}_q^{new}(B_j) + \hat{X}_q(B_i) - \hat{X}_q(B_j)$ in (5.43) which results in

$$\hat{X}_q^{new}(B_j) = \hat{X}_q^{new}(B_{up}) + \frac{1}{2}(\hat{X}_q(B_j) - \hat{X}_q(B_i)). \tag{5.44}$$

Similarly,

$$\hat{X}_q^{new}(B_i) = \hat{X}_q^{new}(B_{up}) + \frac{1}{2}(\hat{X}_q(B_i) - \hat{X}_q(B_j)). \tag{5.45}$$

Note that bands B_2 and B_3 fall into both the above-mentioned cases. Therefore, depending on the selected band, multiple action value estimates are updated instead of only one, which may result in a faster convergence of the algorithm. Algorithm 3 summarizes the proposed modified PMT algorithm.

Algorithm 3 The Proposed PMT

```
assign equal probabilities and action value estimates to all actions.
for i=1 to number of episodes do
    select a band B_i based on action probabilities.
    for every E episodes do
        if B_i is a GB (with B_j in the same group) and |X̂_q(B_i) − X̂_q(B_j)| < δ then
            select the upper band B_up.
        end if
    end for
    observe the reward X_e(a_e = B_i) of the selected action.
    update the action value estimate of the selected action based on (5.35) to obtain X̂_q^new(B_i).
    if the selected action is a GB then
        update the action value estimate of the corresponding upper action based on (5.41).
    end if
    if the selected action is an upper action then
        update the action value estimates of the corresponding GBs based on (5.44) and (5.45).
    end if
    update all the action probabilities based on updated action value estimates.
end for
```

5.3.4 Performance Evaluation of PMT Algorithms

We define an episode of data transmission as the duration of a data frame in the OFDM system. For every episode the PLC transmitter has to select a transmission band. This selection is based on the selection policies described in Algorithms 3 and 5. As a frame of reference, we consider a random selection policy as well, where the transmission band is selected randomly with equal probabilities of the available bands. We assume that the entire duration of the transmission is 400 episodes and the channel conditions change abruptly after the 200th episode. Therefore, the first and the second halves of the transmission are performed with completely different channel conditions. Furthermore, we assume that for the first half of the transmission, the fourth band is the best transmission band and for the second half of the transmission the sixth band is the best transmission band. However, the PLC transmitter is not aware of this information.

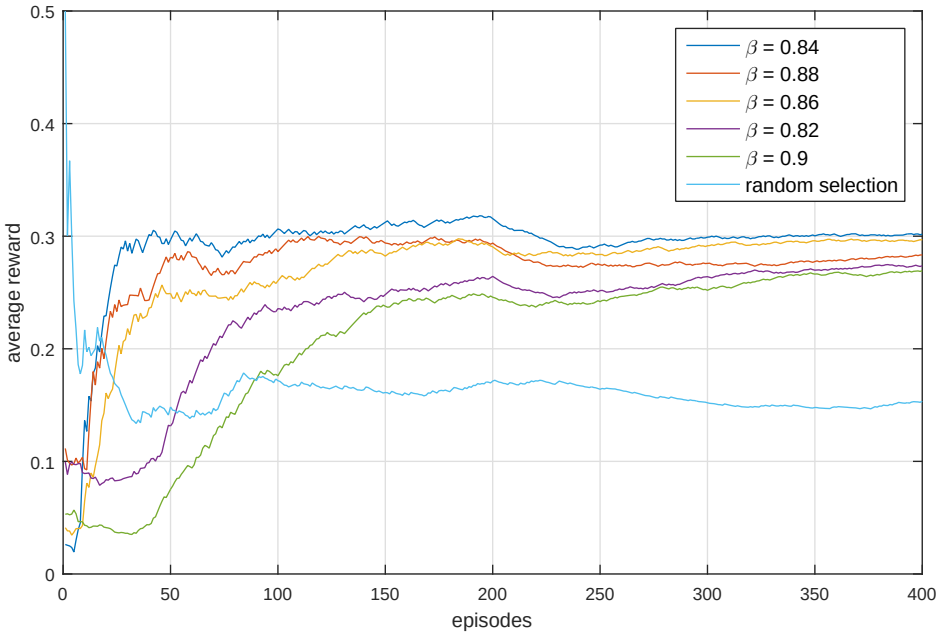


FIGURE 5.15: Average reward of the conventional PMT algorithm.

Figures 5.15 and 5.16 show the average reward obtained throughout the 400 episodes for conventional PMT and the modified PMT algorithms, respectively. Different values of the parameter β has been considered to evaluate the reinforcement learning process of the algorithms. Compared to the random channel selection at the transmitter, learning algorithms are clearly advantageous in terms of the average reward. It can be seen from these two figures that the algorithm requires some time at the beginning of the deployment in order to converge and assign the right probabilities to the different actions. This convergence time, as apparent form figures 5.15 and 5.16, are dependent on the value of the parameter β. This value is determined by empirical methods, which gather the results of finite experiments and evaluates the value of β to best fit the requirements of the system and acquire a beter performance of the algorithm.

By comparing Figure 5.15 and Figure 5.16 we can observe that the modifications which have been applied to the PMT algorithm have resulted in an overall better performance of the system. This means in the modified algorithm, the convergence of the algorithm is much faster for all the values of β which are depicted. In other words, in the modified algorithm the maximum performance is reached faster than the conventional PMT algorithm which gives the proposed algorithm the advantage of finding the best solution in a shorter times, thus increasing the overall performance of the communication system in the time of the operation. This is the direct result of the multiple action value estimate update, which was proposed in the modified version of the algorithm. Furthermore, although the abrupt change in the channel conditions did not affect the performance of the PMT dramatically as compared to the UCB algorithm family, but the performance of the modified PMT is nevertheless smoother at the time of the change. This means that the proposed

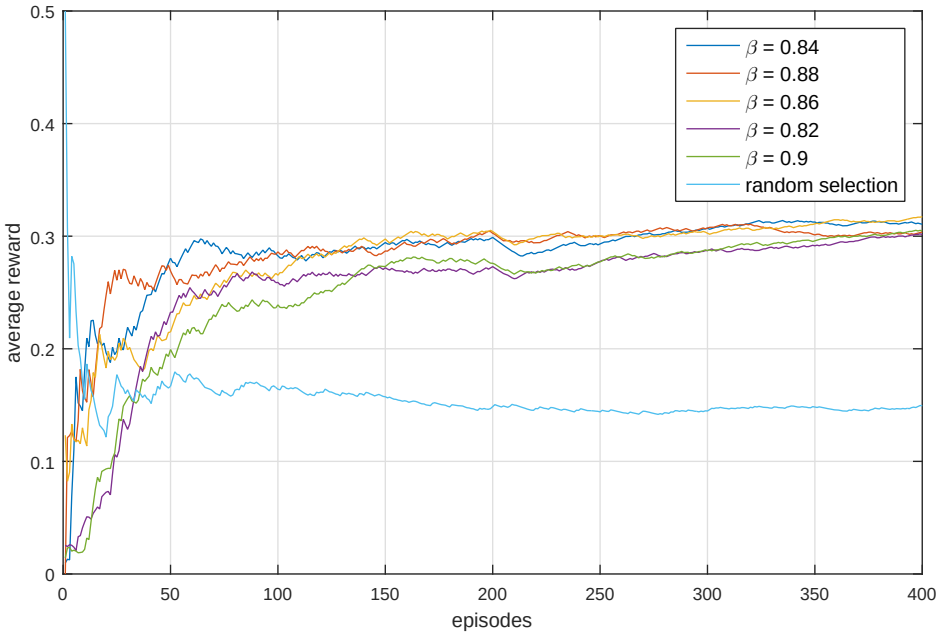


FIGURE 5.16: Average reward of the proposed modified PMT algorithm.

algorithm can adapt itself faster to the changes of the environment, should the environment change its characteristics during a transmission.

5.4 Greedy Algorithms

5.4.1 The ε-Greedy Algorithm

Greedy algorithms are a class of decision making algorithms which select the locally optimal choice at each stage with the hope of finding a global optimum reward. In the decision making problems of the multichannel PLC which are described previously, the greedy algorithms can provide a robust approach in establishing the selection policy at the PLC transmitters. However, since the greedy algorithms only consider the best option at the time of decision making, they are unable to follow the changes in the environment; i.e. the channel. Hence, a slight but powerful change is introduced in the greedy algorithm to enable it to follow the changes of the environment as the environment change through time. This algorithm is known as the ε-greedy algorithm

We describe and analyze the ε-greedy algorithm in order to solve the non-stationary MAB problem. The ε-greedy algorithm is an extension of the conventional greedy algorithm which has been developed for non-stationary problems. According to the greedy action selection rule the selecting agent selects the action with the highest action value estimate at each episode, which results in an overall convergence to

the best action in a stationary environment. However, the ε-greedy action selection scheme provides a method in which the selecting agent searches the sub-optimal actions in random time intervals in order to follow the changes in the channel. In other words, the ε-greedy agent:

- with a probability of $1-\varepsilon$, selects the action with highest action value estimate,
- with a probability of ε, selects an action at random.

The value of ε is normally selected in the space $\{0.01, 0.1\}$. Higher values of ε will force the selecting agent to make exploratory choices more often and as a result will prevent the selecting agent from concentrating its choices to the optimal action while giving it the ability to react rapidly to changes that take place in its environment. On the other hand, smaller values of ε provides the opportunity to exploit the existing knowledge of the environment in order to improve the transmission performance while it may cause slower reactions to the changes of the environment. Therefore, a proper value of ε is essential for a trade-off between exploitation and exploration. Algorithm 4 summarizes the ε-greedy algorithm.

Algorithm 4 The ε-greedy

1: assign equal action value estimates to all actions.
2: **for** i=1 **to** number of episodes **do**
3: with a probability of $1-\varepsilon$ return EXPLOIT = true otherwise return EXPLOIT = false.
4: **if** EXPLOIT **then**
5: select a band B_i which has the highest action value estimate.
6: **else**
7: select a random band B_i.
8: **end if**
9: observe the reward $X_e(a_e = B_i)$ of the selected action.
10: update the action value estimate of the selected action based on (5.35) to obtain $\hat{X}_q^{new}(B_i)$.
11: **end for**

5.4.2 The Proposed Greedy Algorithm

The ε-greedy algorithms shows promising results in a non-stationary system. However, we aim to further improve the ε-greedy algorithm by exploiting some of the unique characteristics of the multichannel PLC in order to develop a faster and better algorithm. In order to improve the ε-greedy algorithm, we have applied the same modifications described in modified PMT algorithm to the ε-greedy algorithm as well. These modifications utilize the same reasoning described in the modified PMT algorithm section with the objective of an improved performance. Algorithm 5 summarizes the proposed modified ε-greedy algorithm.

Algorithm 5 The proposed ε-greedy

assign equal action value estimates to all actions.
for i=1 **to** number of episodes **do**
  with a probability of $1-\varepsilon$ return EXPLOIT = true otherwise return EXPLOIT = false.
  if EXPLOIT **then**
    select a band B_i which has the highest action value estimate.
  else
    select a random band B_i.
  end if
  for every E episodes **do**
    if B_i is a GB (with B_j in the same group) **and** $|\hat{X}_q(B_i) - \hat{X}_q(B_j)| < \delta$ **then**
      select the upper band B_{up}.
    end if
  end for
  observe the reward $X_e(a_e = B_i)$ of the selected action.
  update the action value estimate of the selected action based on (5.35) to obtain $\hat{X}_q^{new}(B_i)$.
  if the selected action is a GB **then**
    update the action value estimate of the corresponding upper action based on (5.41).
  end if
  if the selected action is an upper action **then**
    update the action value estimates of the corresponding GBs based on (5.44) and (5.45).
  end if
end for

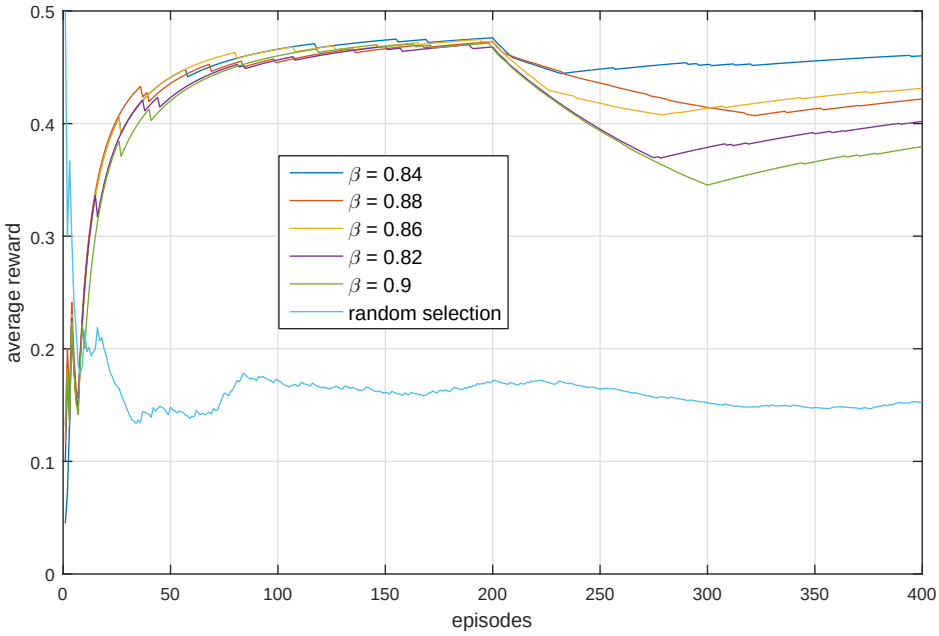


FIGURE 5.17: Average reward of the conventional ε-greedy algorithm.

5.4.3 Performance Evaluation of Greedy Algorithms

Figures 5.17 and 5.18 illustrate the average reward obtained for the conventional ε-greedy and the modified ε-greedy algorithms. By comparing Figure 5.17 with Figure 5.15, we can observe that the ε-greedy algorithm results in higher rewards since it exploits the highest action value estimate more often than the PMT algorithm. However, the ε-greedy algorithm is more sensitive to the abrupt changes in the channel condition.

As seen in Figure 5.17, the algorithm fails to maintain the average reward after the conditions of the channel change abruptly and needs more time to converge again compared to the PMT algorithm. The trade-off between higher rewards and faster response to the changes of the environment can be observed in these simulations. In other words, if a higher performance is needed, the convergence time is lower, and if a faster convergence is desired, a lower performance is observed. However, the modified ε-greedy algorithm solves this problem as can be seen in Figure 5.18. The modified ε-greedy algorithm can react faster to the changes of the environment and results in a better performance in non-stationary channels where the channel conditions are changing constantly. Therefore, with this proposed algorithm a fast convergence with a reasonably high performance is achievable. This is due to the fact that the proposed algorithm is using the unique features of the PLC channels and hence is able to reduce the regret generated by the seminal algorithm.

The effect of the similarity threshold is depicted in Figure 5.19. In this figure the number of OFDM symbols for every 100 episodes and for all the proposed

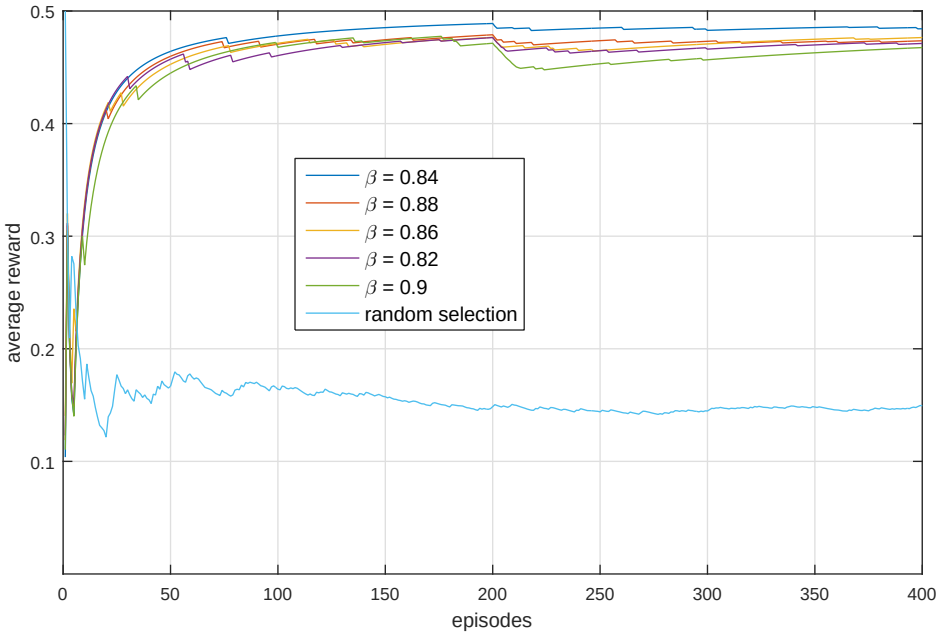


FIGURE 5.18: Average reward of the proposed modified ε-greedy algorithm.

algorithms is depicted. As mentioned before, the different parts of the spectrum correspond to different number of data subcarrier for every OFDM symbol (72, 36, or 18). Therefore, for the same amount of data, the bigger bands will result in lower number of OFDM symbols and consequently smaller time of transmission which can be interpreted as higher throughput in the same amount of time. With the similarity threshold modification of the algorithm, we have selected an upper band whenever the action value estimates were close. Therefore a lower number of OFDM symbols should be achieved with the same performance. As seen in Figure 5.19, in both cases of PMT and ε-greedy algorithm, the modified version resulted in lower number of OFDM symbols and consequently a higher throughput.

5.5 Chapter Conclusion

In this chapter, we proposed utilizing reinforcement learning algorithms and decision making strategies in order to solve the MAB problem for the selection problems in multichannel PLC without channel knowledge at the PLC transmitter. This enables us to solve the multichannel PLC problems, described in the previous chapter, in a manner that no channel state information is needed at the transmitter. Compared to other state of the art solutions, our proposed solution has the benefit of improving the performance and reliability of the communication and at the same time avoiding the difficulties and problems introduced by obtaining the channel state information at the transmitter. Furthermore, we proposed four novel algorithms based on the

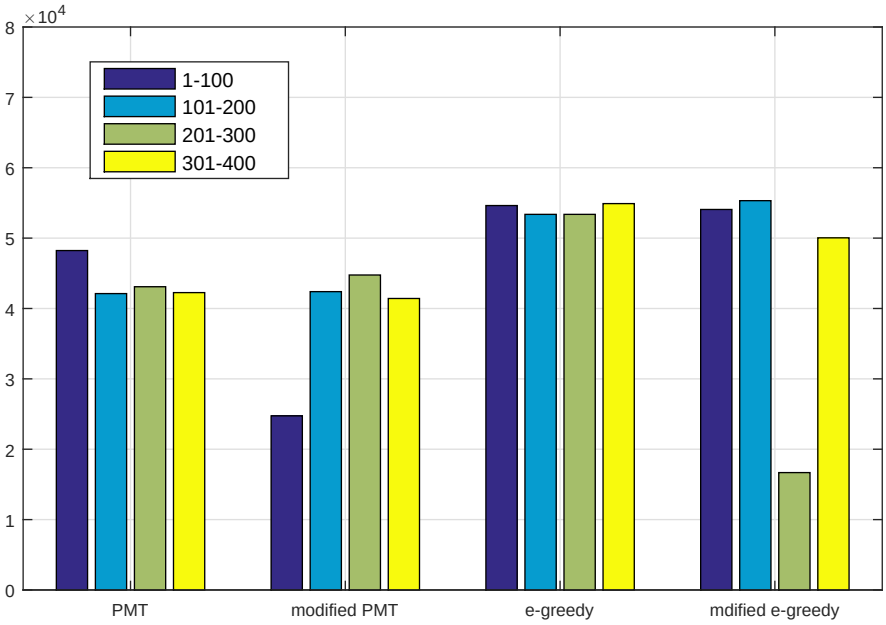


FIGURE 5.19: Number of OFDM symbols to be transmitted for each algorithm.

conventional reinforcement learning algorithms which accommodate the characteristics of the PLC channel. These algorithms are designed and developed to exploit the unique features of the multichannel PLC system model to further improve the performance of the communication. Simulation results have been presented which shows the improvements achieved by the proposed algorithms.

Chapter 6

Conclusion

In this thesis, we studied the multichannel power line communication and its challenges. The multichannel transmission has been categorized into three major transmission techniques as follows. The first transmission technique, namely the MIMO transmission, provides spatial diversity which can improve the performance of the transmission or the data throughput. The spatial diversity at the PLC transmitter provides the capability to select the transmission channel based on the available channel state information. This can further improve the performance of the transmission by transmitting to the channels with higher SNRs and is referred to as the transmit selection diversity. The problem of transmit selection can be defined as the problem of selecting the best transmit ports in order to optimize the selection diversity at the transmitter. However, obtaining the channel state information cause some problems to the communication system. The main problem, is the feedback delay in the PLC channel environment which is characterized as highly time-variant and frequency-selective, which leads to an out of date channel state information. Furthermore, the PLC channel is highly dependent on the connected loads of the power line networks, which any connection and disconnection dramatically changes the channel state information. Therefore, in order to nonetheless exploiting the performance improvement introduced by the transmit channel selection diversity, a method of performing intelligent decision at the transmitter without relying on the channel state information is needed.

Moreover, as another multichannel PLC technique, the cooperative multihop transmission in PLC applications has been discussed which provides cooperative diversity in the communication system. In this case, the data signal is transmitted in several channels between the source and destination by means of intermediate relay nodes in the network. These relay nodes must be selected in such way that results in the best transmission performance. Hence, the relay selection problem plays a major role in optimizing the cooperative diversity. The relay selection process is normally based on the channel state information of every hop which makes the entire communication system. For the same reasons stated above, a technique which can

provide a proper relay selection capabilities without relying on the channel state information is needed to be developed for PLC applications.

Another multichannel communication scenario is the multiband transmission in PLC applications. In this case, the available spectrum is divided into several transmission bands and the transmitter has the ability to choose from these transmission bands in order to maximize the performance of the transmission. This band selection or as referred to in this theses, spectrum assignment, must be based on the performance of the system in that particular band. The attenuation of the data signal can be different from band to band and the PLC noise can also be different in different bands. Moreover, tone masking and time-variant regulations of the use of spectrum may force the PLC transmitter to select the transmitting band constantly throughout the transmission. Selecting the proper transmit band requires the knowledge of the entire spectrum at all time. Obtaining this knowledge at the PLC transmitter requires a tremendous effort which decreases the throughput and efficiency of the communication system. Therefore, intelligent and effective spectrum assignment without relying on the knowledge of the whole spectrum is needed for PLC applications.

All the above-mentioned multichannel problems are commonly used in the state of the art PLC applications. However, to the best knowledge of the author, there solutions available in the PLC literature heavily rely on the availability of the channel state information at the PLC transmitter. This makes those solutions not implementable in real PLC applications or prevents them to fully achieve a high performance and reliable communication. The demand for solutions which do not rely on the channel state information is of utmost importance in the PLC systems. In this thesis, we contributed to the state of the art and pushed the limits of the PLC technology further, to enable it to perform high quality and reliable multichannel techniques. Our contribution to the state of the art, enables the PLC application to implement algorithms, which significantly improves the state of the art performance. These proposed algorithms are based on the machine learning principles.

Our approach to solve the aforementioned problems, is inspired by machine learning applications. We proposed a series of reinforcement learning solutions to be used in the selection problems of the multichannel PLC, which enables the PLC transmitter to perform intelligent decision in an environment of incomplete information, hence freeing the PLC transmitter from the need of acquiring the channel state information at all times and frequencies. These reinforcement learning algorithms enable the PLC transmitter to perform intelligent decision making policies, which are uniquely adapted to the PLC characteristics of the communication channel. In order to design these reinforcement learning algorithms, firstly we modeled each problem by a powerful mathematical tool known as the multi-armed bandit problem modeling approach. Then we evaluated selection policy algorithms and proposed new algorithms in order to solve the selection problems in each case with unknown channel at the transmitter. The proposed novel algorithms were designed by exploiting some specific features of the PLC channel, in order to further improve the performance of the seminal algorithms. The simulation results were provided which

indicates the improvement in performance of the transmission in all the proposed algorithms.

Bibliography

[1] Renesas Electronics America Inc., "Narrowband power line communication: Applications and challenges", 2012.

[2] H. Philipps, "Modelling of powerline communication channels", in *Conf. IEEE Int. Symp. on Power Line Communications and Its Applications*, 1999, pp. 14–21.

[3] M. Zimmermann and K. Dostert, "A multipath model for the powerline channel", *Communications, IEEE Transactions on*, vol. 50, no. 4, pp. 553–559, 2002.

[4] J. Anatory, M. Kissaka, and N. Mvungi, "Channel model for broadband power-line communication", *Power Delivery, IEEE Transactions on*, vol. 22, no. 1, pp. 135–141, 2007.

[5] J. Anatory, N. Theethayi, and R. Thottappillil, "Power-line communication channel model for interconnected networks part i: Two-conductor system", *Power Delivery, IEEE Transactions on*, vol. 24, no. 1, pp. 118–123, 2009.

[6] ——, "Power-line communication channel model for interconnected networks part ii: Multiconductor system", *Power Delivery, IEEE Transactions on*, vol. 24, no. 1, pp. 124–128, 2009.

[7] H. Ferreira, L. Lampe, J. Newbury, and T. Swart, *Power Line Communications: Theory and Applications for Narrowband and Broadband Communications over Power Lines*. Wiley, 2010, ISBN: 9780470661284.

[8] M. Zimmermann and K. Dostert, "An analysis of the broadband noise scenario in powerline networks", in *International Symposium on Powerline Communications and its Applications (ISPLC2000)*, 2000, pp. 5–7.

[9] S Galli, M Koch, H. Latchman, S Lee, and V Oksman, "Industrial and international standards on plc-based networking technologies", *Power Line Communications: Theory and Applications for Narrowband and Broadband Communications over Power Lines*, pp. 363–412, 2010.

[10] V. Oksman and J. Zhang, "G.HNEM: The new ITU-t standard on narrowband PLC technology", *IEEE Communications Magazine*, vol. 49, no. 12, pp. 36–44, 2011.

[11] *IEEE standard for low-frequency (less than 500 kHz) narrowband power line communications for smart grid applications*, IEEE 1901.2 Narrowband.

[12] L. Yonge, J. Abad, K. Afkhamie, L. Guerrieri, S. Katar, H. Lioe, P. Pagani, R. Riva, D. M. Schneider, and A. Schwager, "An overview of the HomePlug

AV2 technology", *Journal of Electrical and Computer Engineering*, vol. 2013, pp. 1–20, 2013.

[13] *Unified high-speed wire-line based home networking transceivers–foundation.*

[14] *IEEE standard for broadband over power line networks: Medium access control and physical layer specifications.*

[15] C. Cano, A. Pittolo, D. Malone, L. Lampe, A. M. Tonello, and A. G. Dabak, "State of the art in power line communications: From the applications to the medium", *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 7, pp. 1935–1952, Jul. 2016.

[16] *Signalling on low-voltage electrical installations in the frequency range 3 kHz to 148.5 kHz. low voltage decoupling filters.*

[17] *Power line communication apparatus used in low-voltage installations. radio disturbance characteristics. limits and methods of measurement.*

[18] C. R. Paul, *Analysis of multiconductor transmission lines*. John Wiley & Sons, 2008.

[19] A. M. Tonello and F. Versolatto, "Bottom-up statistical plc channel modeling part i: Random topology model and efficient transfer function computation", *Power Delivery, IEEE Transactions on*, vol. 26, no. 2, pp. 891–898, 2011.

[20] ——, "Bottom-up statistical plc channel modeling part ii: Inferring the statistics", *Power Delivery, IEEE Transactions on*, vol. 25, no. 4, pp. 2356–2363, 2010.

[21] L. T. Berger, A. Schwager, P. Pagani, and D. M. Schneider, "Mimo power line communications", *Communications Surveys & Tutorials, IEEE*, vol. 17, no. 1, pp. 106–124, 2014.

[22] L. Lampe and A. H. Vinck, "On cooperative coding for narrow band plc networks", *AEU-International Journal of Electronics and Communications*, vol. 65, no. 8, pp. 681–687, 2011.

[23] ——, "Cooperative multihop power line communications", in *Power Line Communications and Its Applications (ISPLC), 2012 16th IEEE International Symposium on*, IEEE, 2012, pp. 1–6.

[24] F. Gruber and L. Lampe, "On plc channel emulation via transmission line theory", in *Power Line Communications and its Applications (ISPLC), 2015 International Symposium on*, 2015, pp. 178–183.

[25] K. Dostert and S. Galli, "Keynote II: Modeling of electrical power supply systems as communication channels", in *International Symposium on Power Line Communications and Its Applications, 2005.*, Institute of Electrical and Electronics Engineers (IEEE), 2005.

[26] M. Gotz, M. Rapp, and K. Dostert, "Power line channel characteristics and their effect on communication system design", *IEEE Communications Magazine*, vol. 42, no. 4, pp. 78–86, 2004.

[27] S. Galli and T. Banwell, "A deterministic frequency-domain model for the indoor power line transfer function", *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 7, pp. 1304–1316, 2006.

[28] T. Sartenaer and P. Delogne, "Deterministic modeling of the (shielded) outdoor power line channel based on the multiconductor transmission line equations", *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 7, pp. 1277–1291, 2006.

[29] F. Canete, L. Diez, J. Cortes, and J. Entrambasaguas, "Broadband modelling of indoor power-line channels", *IEEE Transactions on Consumer Electronics*, vol. 48, no. 1, pp. 175–183, 2002.

[30] T. Esmailian, F. R. Kschischang, and P. G. Gulak, "In-building power lines as high-speed communication channels: Channel characterization and a test channel ensemble", *International Journal of Communication Systems*, vol. 16, no. 5, pp. 381–400, 2003.

[31] F. Canete, J. Cortes, L. Diez, and J. Entrambasaguas, "A channel model proposal for indoor power line communications", *IEEE Communications Magazine*, vol. 49, no. 12, pp. 166–174, 2011.

[32] A. M. Tonello, F. Versolatto, B. Bejar, and S. Zazo, "A fitting algorithm for random modeling the PLC channel", *IEEE Transactions on Power Delivery*, vol. 27, no. 3, pp. 1477–1484, 2012.

[33] S. Galli, "A simplified model for the indoor power line channel", in *2009 IEEE International Symposium on Power Line Communications and Its Applications*, Institute of Electrical and Electronics Engineers (IEEE), 2009.

[34] ——, "A novel approach to the statistical modeling of wireline channels", *IEEE Transactions on Communications*, vol. 59, no. 5, pp. 1332–1345, 2011.

[35] V. Degardin, M. Lienard, P. Degauque, A. Zeddam, and F. Gauthier, "Impulsive noise on indoor power lines: Characterization and mitigation of its effect on PLC systems", in *2003 IEEE International Symposium on Electromagnetic Compatibility, 2003. EMC 2003.*, Institute of Electrical and Electronics Engineers (IEEE), 2003.

[36] l.C. Papaleonidopoulos, C. Capsalis, C. Karagiannopoulos, and N. Theodorou, "Statistical analysis and simulation of indoor single-phase low voltage power-line communication channels on the basis of multipath propagation", *IEEE Transactions on Consumer Electronics*, vol. 49, no. 1, pp. 89–99, 2003.

[37] Y. Xiaoxian, Z. Tao, Z. Baohui, Y. Fengchun, D. Jiandong, and S. Minghui, "Research of impedance characteristics for medium-voltage power networks", *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 870–878, 2007.

[38] M. Raugi, T. Zheng, M. Tucci, and S. Barmada, "On the time invariance of PLC channels in complex power networks", in *ISPLC2010*, Institute of Electrical and Electronics Engineers (IEEE), 2010.

[39] F. Versolatto, A. M. Tonello, C. Tornelli, and D. D. Giustina, "Statistical analysis of broadband underground medium voltage channels for PLC applications", in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Institute of Electrical and Electronics Engineers (IEEE), 2014.

[40] N. Nasiriani, R. Ramachandran, K. Rahimi, Y. P. Fallah, P. Famouri, S. Bossart, and K. Dodrill, "An embedded communication network simulator for power systems simulations in PSCAD", in *2013 IEEE Power & Energy*

Society General Meeting, Institute of Electrical and Electronics Engineers (IEEE), 2013.

[41] M. Wei and Z. Chen, "Distribution system protection with communication technologies", in *IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society*, Institute of Electrical and Electronics Engineers (IEEE), 2010.

[42] A. Cataliotti, D. D. Cara, R. Fiorelli, and G. Tine, "Power-line communication in medium-voltage system: Simulation model and onfield experimental tests", *IEEE Transactions on Power Delivery*, vol. 27, no. 1, pp. 62–69, 2012.

[43] J. R. Wait, "Theory of wave propagation along a thin wire parallel to an interface", *Radio Science*, vol. 7, no. 6, pp. 675–679, 1972.

[44] M. D'Amore and M. Sarto, "Simulation models of a dissipative transmission line above a lossy ground for a wide-frequency range. i. single conductor configuration", *IEEE Transactions on Electromagnetic Compatibility*, vol. 38, no. 2, pp. 127–138, 1996.

[45] P. Amirshahi and M. Kavehrad, "High-frequency characteristics of overhead multiconductor power lines for broadband communications", *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 7, pp. 1292–1303, 2006.

[46] A. G. Lazaropoulos and P. G. Cottis, "Broadband transmission via underground medium-voltage power lines: Part II: Capacity", *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2425–2434, 2010.

[47] P. H. Reynolds and B. L. Dyner, "Safety aspects of high voltage power line fault location techniques", in *1973 EIC 11th Electrical Insulation Conference*, Institute of Electrical and Electronics Engineers (IEEE), 1973.

[48] J. Bridges, "The biological significance of power line and high voltage switchyard environments", in *IEEE 1976 International Symposium on Electromagnetic Compatibility*, Institute of Electrical and Electronics Engineers (IEEE), 1976.

[49] A. Kern, R. Frentzel, and J. Behrens, "Simulation of the transient voltages in the auxiliary power network of a large power plant in case of a direct lightning strike to the high-voltage overhead transmission line", in *2010 30th International Conference on Lightning Protection (ICLP)*, Institute of Electrical and Electronics Engineers (IEEE), 2010.

[50] M. Takanashi, T. Harada, A. Takahashi, H. Tanaka, H. Hayashi, and Y. Hattori, "High-voltage power line communication system for hybrid vehicle", in *2015 IEEE International Symposium on Power Line Communications and Its Applications (ISPLC)*, Institute of Electrical and Electronics Engineers (IEEE), 2015.

[51] J. A. Cortés, L. Diez, F. J. Canete, and J. J. Sánchez-Martínez, "Analysis of the indoor broadband power-line noise scenario", *Electromagnetic Compatibility, IEEE Transactions on*, vol. 52, no. 4, pp. 849–858, 2010.

[52] F. J. C. Corripio, J. A. C. Arrabal, L. D. Del Río, and J. T. E. Muñoz, "Analysis of the cyclic short-term variation of indoor power line channels", *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 7, pp. 1327–1338, 2006.

[53] Y. Hirayama, H. Okada, T. Yamazato, and M. Katayama, "Noise analysis on wide-band plc with high sampling rate and long observation time", 2003.

[54] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications", *Electromagnetic Compatibility, IEEE Transactions on*, vol. 44, no. 1, pp. 249–258, 2002.

[55] V. Degardin, M. Lienard, A. Zeddam, F Gauthie, and P Degauque, "Classification and characterization of impulsive noise on indoor powerline used for data communications", *Consumer Electronics, IEEE Transactions on*, vol. 48, no. 4, pp. 913–918, 2002.

[56] L. Zhuang and M. Zhai, "Modeling the noise in power line communications", in *2010 International Conference on Multimedia Technology*, IEEE, 2010.

[57] D. Middleton, "Statistical-physical models of electromagnetic interference", *Electromagnetic Compatibility, IEEE Transactions on*, no. 3, pp. 106–127, 1977.

[58] D. Middleton, "Canonical and quasi-canonical probability models of class a interference", *Electromagnetic Compatibility, IEEE Transactions on*, vol. EMC-25, no. 2, pp. 76–106, 1983.

[59] M. Katayama, T. Yamazato, and H. Okada, "A mathematical model of noise in narrowband power line communication systems", *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 7, pp. 1267–1276, 2006.

[60] D. Middleton, "Canonical and quasi-canonical probability models of class a interference", *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-25, no. 2, pp. 76–106, 1983.

[61] M. Nassar, K. Gulati, Y. Mortazavi, and B. L. Evans, "Statistical modeling of asynchronous impulsive noise in powerline communication networks", in *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, Institute of Electrical and Electronics Engineers (IEEE), 2011.

[62] M. A. Sonmez and M. Bagriyanik, "Statistical model study for narrowband power line communication noises", in *2013 8th International Conference on Electrical and Electronics Engineering (ELECO)*, Institute of Electrical and Electronics Engineers (IEEE), 2013.

[63] L. Tang, P. So, E. Gunawan, S. Chen, T. Lie, and Y. Guan, "Characterization of power distribution lines for high-speed data transmission", in *PowerCon 2000. 2000 International Conference on Power System Technology. Proceedings (Cat. No.00EX409)*, Institute of Electrical and Electronics Engineers (IEEE).

[64] V. Degardin, M. Lienard, A. Zeddam, F. Gauthier, and R. Degauque, "Classification and characterization of impulsive noise on indoor power line used for data communications", *IEEE Transactions on Consumer Electronics*, vol. 48, no. 4, pp. 913–918, 2003.

[65] J. A. Cortes, L. Diez, F. J. Canete, and J. J. Sanchez-Martinez, "Analysis of the indoor broadband power-line noise scenario", *IEEE Transactions on Electromagnetic Compatibility*, vol. 52, no. 4, pp. 849–858, 2010.

[66] M. Bogdanovic and S. Rupcic, "Generalized background noise modeling in power line communication", in *2012 20th Telecommunications Forum (TELFOR)*, Institute of Electrical and Electronics Engineers (IEEE), 2012.

[67] L. D. Bert, P. Caldera, D. Schwingshackl, and A. M. Tonello, "On noise modeling for power line communications", in *2011 IEEE International Symposium on Power Line Communications and Its Applications*, Institute of Electrical and Electronics Engineers (IEEE), 2011.

[68] T. Schaub, "Spread frequency shift keying", *IEEE transactions on communications*, vol. 42, no. 234, pp. 1056–1064, 1994.

[69] M. Doelz, E. Heald, and D. Martin, "Binary data transmission techniques for linear systems", *Proceedings of the IRE*, vol. 45, no. 5, pp. 656–661, 1957.

[70] L. Zhen, W. Weihua, Z. Wenan, and S. Junde, "A modified sub-optimum adaptive bit and power allocation algorithm in wideband ofdm system", in *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on*, IEEE, vol. 3, 2003, pp. 1589–1592.

[71] S. Kasturia, J. Aslanis, and J. Cioffi, "Vector coding for partial response channels", *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 741–762, 1990.

[72] S Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform", *IEEE transactions on Communication Technology*, vol. 19, no. 5, pp. 628–634, 1971.

[73] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[74] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas", *Wireless personal communications*, vol. 6, no. 3, pp. 311–335, 1998.

[75] J. Mietzner, R. Schober, L. Lampe, W. H. Gerstacker, and P. A. Hoeher, "Multiple-antenna techniques for wireless communications-a comprehensive literature survey", *Communications Surveys & Tutorials, IEEE*, vol. 11, no. 2, pp. 87–105, 2009.

[76] L. Stadelmeier, D. Schill, A. Schwager, D. Schneider, and J. Speidel, "Mimo for inhome power line communications", in *Source and Channel Coding (SCC), 2008 7th International ITG Conference on*, VDE, 2008, pp. 1–6.

[77] C. Giovaneli, J. Yazdani, P. Farrell, and B. Honary, "Application of space-time diversity/coding for power line channels", 2002.

[78] R. Hashmat, P. Pagani, A. Zeddam, and T. Chonavel, "Mimo communications for inhome plc networks: Measurements and results up to 100 mhz", in *Power Line Communications and Its Applications (ISPLC), 2010 IEEE International Symposium on*, IEEE, 2010, pp. 120–124.

[79] B. Nikfar and H. Vinck, "Space diversity in mimo power line channels with independent impulsive noise", in *Proc. of the Sixth Workshop on Power Line Communications*, 2012, pp. 1–3.

[80] ——, "Combining techniques performance analysis in spatially correlated mimo-plc systems", in *Power Line Communications and Its Applications (ISPLC), 2013 17th IEEE International Symposium on*, IEEE, 2013, pp. 1–6.

[81] B. Nikfar, T. Akbudak, and H. Vinck, "Mimo capacity of class a impulsive noise channel for different levels of information availability at transmitter", in *Power Line Communications and its Applications (ISPLC), 2014 18th IEEE International Symposium on*, IEEE, 2014, pp. 266–271.

[82] B. Nikfar and H. Vinck, "Analysis of the mtl configuration for mimo plc with mutual coupling", in *Proc. of the Eighth Workshop on Power Line Communications*, 2014.

[83] I. E. Commission *et al.*, *Power line communication system or power utility applicationspart 1: Planning of analog and digital power line carrier systems operating over ehv/hv/mv electricity grids*, 2012.

[84] E. T.C.P. T. (PLT), *Powerline telecommunication (plt); basic data relating to lvdn measurements in the 3 mhz to 100 mhz frequency range*, 2004.

[85] M. Ishihara, D. Umehara, and Y. Morihiro, "The correlation between radiated emissions and power line network components on indoor power line communications", in *2006 IEEE International Symposium on Power Line Communications and Its Applications*, IEEE, 2006, pp. 314–318.

[86] A. Schwager, "Powerline communications significant technologies to become ready for integration", PhD thesis, University of Duisburg-Essen, 2010.

[87] T. Banwell and S. Galli, "A novel approach to the modeling of the indoor power line channel part i: Circuit analysis and companion model", *IEEE Transactions on power delivery*, vol. 20, no. 2, pp. 655–663, 2005.

[88] T. ETSI, "101 562-1 v2. 1.1,powerline telecommunications (plt), mimo plt, part 1: Measurements methods of mimo plt,i", Tech. Rep, Tech. Rep., 2012.

[89] C. Giovaneli, B. Honary, and P. Farrell, "Space-frequency coded OFDM system for multi-wire power line communications", in *International Symposium on Power Line Communications and Its Applications, 2005.*, IEEE, 2005.

[90] R. Hashmat, P. Pagani, A. Zeddam, and T. Chonave, "A channel model for multiple input multiple output in-home power line networks", in *2011 IEEE International Symposium on Power Line Communications and Its Applications*, IEEE, 2011.

[91] D. Veronesi, R. Riva, P. Bisaglia, F. Osnato, K. Afkhamie, A. Nayagam, D. Rende, and L. Yonge, "Characterization of in-home MIMO power line channels", in *2011 IEEE International Symposium on Power Line Communications and Its Applications*, IEEE, 2011.

[92] A. Schwager, W. Baschlin, H. Hirsch, P. Pagani, N. Weling, J. L. G. Moreno, and H. Milleret, "European MIMO PLT field measurements: Overview of the ETSI STF410 campaign: EMI analysis", in *2012 IEEE International Symposium on Power Line Communications and Its Applications*, IEEE, 2012.

[93] D. Schneider, A. Schwager, W. Baschlin, and P. Pagani, "European MIMO PLC field measurements: Channel analysis", in *2012 IEEE International Symposium on Power Line Communications and Its Applications*, IEEE, 2012.

[94] F. Versolatto and A. M. Tonello, "An MTL theory approach for the simulation of MIMO power-line communication channels", *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1710–1717, 2011.

[95] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas", *Bell labs technical journal*, vol. 1, no. 2, pp. 41–59, 1996.

[96] S. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[97] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria", in *Proceedings of ICC 1997 - International Conference on Communications*, IEEE.

[98] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs", *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[99] A. Tomasoni, R. Riva, and S. Bellini, "Spatial correlation analysis and model for in-home mimo power line channels", in *Power Line Communications and Its Applications (ISPLC), 2012 16th IEEE International Symposium on*, IEEE, 2012, pp. 286–291.

[100] K. Wiklundh, P. Stenumgaard, and H. Tullberg, "Channel capacity of middleton's class a interference channel", *Electronics letters*, vol. 45, no. 24, pp. 1227–1229, 2009.

[101] E. Telatar, "Capacity of multi-antenna gaussian channels", *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[102] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information", *Information Theory, IEEE Transactions on*, vol. 43, no. 6, pp. 1986–1992, 1997.

[103] D. Palomar and J. Fonollosa, "Practical algorithms for a family of waterfilling solutions", *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 686–695, 2005.

[104] B Hilburn, T. Newman, T Bose, and S. Kadambe, "A survey of basic channel selection techniques for cognitive radios", in *Proc. of the SDR Wireless Innovation Conference*, 2010, pp. 37–41.

[105] K. Tsukamoto, Y. Omori, O. Altintas, M. Tsuru, and Y. Oie, "On spatially-aware channel selection in dynamic spectrum access multi-hop inter-vehicle communications", in *Vehicular Technology Conference Fall (VTC 2009-Fall), 2009 IEEE 70th*, IEEE, 2009, pp. 1–7.

[106] F. Hou and J. Huang, "Dynamic channel selection in cognitive radio network with channel heterogeneity", in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, IEEE, 2010, pp. 1–6.

[107] A. Ghrayeb, "A survey on antenna selection for mimo communication systems", in *Information and Communication Technologies, 2006. ICTTA'06. 2nd*, IEEE, vol. 2, 2006, pp. 2104–2109.

[108] C. Jiang and L. J. Cimini, "Antenna selection for energy-efficient mimo transmission", *Wireless Communications Letters, IEEE*, vol. 1, no. 6, pp. 577–580, 2012.

[109] T. A. W. Le Chung Tran, A. Mertins, and J. Seberry, "Transmitter diversity antenna selection techniques for wireless channels utilizing differential space-time block codes", *Journal of Telecommunications and Information Technology*,

[110] B. Nikfar, S. Maghsudi, and H. Vinck, "Multi-armed bandit channel selection for power line communication", in *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, IEEE, 2015, pp. 19–24.

[111] K.-L. Du and M. N. Swamy, *Wireless communication systems: from RF subsystems to 4G enabling technologies*. Cambridge University Press, 2010.

[112] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: understanding broadband wireless networking*. Pearson Education, 2007.

[113] G. Bumiller, "Single frequency network technology for medium access and network management", in *6th International Symposium on Power-Line Communications and its applications*, 2002.

[114] G. Bumiller, L. Lampe, and H. Hrasnica, "Power line communication networks for large-scale control and automation systems", *Communications Magazine, IEEE*, vol. 48, no. 4, pp. 106–113, 2010.

[115] L. Lampe, R. Schober, and S. Yiu, "Distributed space-time coding for multihop transmission in power line communication networks", *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 7, pp. 1389–1400, 2006.

[116] A. M. Tonello, F. Versolatto, and S. D'Alessandro, "Opportunistic relaying in in-home plc networks", in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, IEEE, 2010, pp. 1–5.

[117] V. B. Balakirsky and A. H. Vinck, "Potential performance of plc systems composed of several communication links", in *Power Line Communications and Its Applications, 2005 International Symposium on*, IEEE, 2005, pp. 12–16.

[118] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks", *Communications Magazine, IEEE*, vol. 42, no. 10, pp. 74–80, 2004.

[119] J. N. Laneman, D. N. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior", *Information Theory, IEEE Transactions on*, vol. 50, no. 12, pp. 3062–3080, 2004.

[120] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, "Multihop diversity in wireless relaying channels", *Communications, IEEE Transactions on*, vol. 52, no. 10, pp. 1820–1830, 2004.

[121] O. Oyman, J. N. Laneman, and S. Sandhu, "Multihop relaying for broadband wireless mesh networks: From theory to practice", *IEEE Communications Magazine*, vol. 45, no. 11, p. 116, 2007.

[122] IEEE, "Ieee standard for low-frequency (less than 500 khz) narrowband power line communications for smart grid applications", *Oct*, vol. 31, p. 269, 2013.

[123] ITU-T, "Narrowband orthogonal frequency division multiplexing power line communication transceivers for g3-plc networks", 2013.

[124] E. Z. Tragos, S. Zeadally, A. G. Fragkiadakis, and V. A. Siris, "Spectrum assignment in cognitive radio networks: A comprehensive survey", *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 3, pp. 1108–1135, 2013.

[125] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access", *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, 2007.

[126] R. M. Eletreby, H. M. Elsayed, and M. M. Khairy, "Optimal spectrum assignment for cognitive radio sensor networks under coverage constraint", *Communications, IET*, vol. 8, no. 18, pp. 3318–3325, 2014.

[127] K. Kastner, S. McClellan, and W. Stapleton, "Dynamic spectrum allocation in low-bandwidth power line communications",

[128] S. Maghsudi and S. Staczak, "Channel selection for network-assisted d2d communication via no-regret bandit learning with calibrated forecasting", *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1309–1322, 2015.

[129] ——, "Transmission mode selection for network-assisted device to device communication: A levy-bandit approach", in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 7009–7013.

[130] ——, "Joint channel selection and power control in infrastructureless wireless networks: A multiplayer multiarmed bandit framework", *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4565–4578, 2015.

[131] S. Maghsudi and S. Stnczak, "Relay selection with no side information: An adversarial bandit approach", in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, 2013, pp. 715–720.

[132] S. Maghsudi and E. Hossain, "Multi-armed bandits with application to 5g small cells", *IEEE Wireless Communications*, vol. 23, no. 3, pp. 64–73, 2016.

[133] D. A. Berry and B. Fristedt, *Bandit problems: sequential allocation of experiments (Monographs on statistics and applied probability)*. Springer, 1985.

[134] J. Gittins, K. Glazebrook, and R. Weber, *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.

[135] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "Gambling in a rigged casino: The adversarial multi-armed bandit problem", in *Foundations of Computer Science, 1995. Proceedings., 36th Annual Symposium on*, IEEE, 1995, pp. 322 –331.

[136] P. Zhou, Y. Chang, and J. A. Copeland, "Reinforcement learning for repeated power control game in cognitive radio networks", *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 54–69, 2012.

[137] I. Macaluso, D. Finn, B. Ozgul, and L. A. DaSilva, "Complexity of spectrum activity and benefits of reinforcement learning for dynamic channel selection", *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 11, pp. 2237–2248, 2013.

[138] M. Bennis, S. M. Perlaza, P. Blasco, Z. Han, and H. V. Poor, "Self-organization in small cell networks: A reinforcement learning approach", *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3202–3212, 2013.

[139] F. Shams, G. Bacci, and M. Luise, "Energy-efficient power control for multiple-relay cooperative networks using q -learning", *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1567–1580, 2015.

[140] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules", *Advances in applied mathematics*, vol. 6, no. 1, pp. 4–22, 1985.

[141] R. Agrawal, "Sample mean based index policies with o (log n) regret for the multi-armed bandit problem", *Advances in Applied Probability*, pp. 1054–1078, 1995.

[142] P. Auer, "Using confidence bounds for exploitation-exploration trade-offs", *Journal of Machine Learning Research*, vol. 3, no. Nov, pp. 397–422, 2002.

[143] ——, "Using upper confidence bounds for online learning", in *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, IEEE, 2000, pp. 270–279.

[144] A. Garivier and E. Moulines, "On upper-confidence bound policies for non stationary bandit problems", *arXiv preprint arXiv:0805.3415*, 2008.

[145] L. Kocsis and C. Szepesvári, "Discounted ucb", in *2nd PASCAL Challenges Workshop*, 2006, pp. 784–791.

[146] D. E. Koulouriotis and A Xanthopoulos, "Reinforcement learning and evolutionary algorithms for non-stationary multi-armed bandit problems", *Applied Mathematics and Computation*, vol. 196, no. 2, pp. 913–922, 2008.

[147] T. Le, C. Szepesvri, and R. Zheng, "Sequential learning for multi-channel wireless network monitoring with channel switching costs", *IEEE Transactions on Signal Processing*, vol. 62, no. 22, pp. 5919–5929, 2014.

[148] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for gaussian process optimization in the bandit setting", *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3250–3265, 2012.

[149] S. Padakandla, P. K. J., and S. Bhatnagar, "Energy sharing for multiple sensor nodes with finite buffers", *IEEE Transactions on Communications*, vol. 63, no. 5, pp. 1811–1823, 2015.

[150] D. Thierens, "An adaptive pursuit strategy for allocating operator probabilities", in *Proceedings of the 7th annual conference on Genetic and evolutionary computation*, ACM, 2005, pp. 1539–1546.

[151] D. E. Goldberg, "Probability matching, the magnitude of reinforcement, and classifier system bidding", *Machine Learning*, vol. 5, no. 4, pp. 407–425, 1990.