

A Model for Bursty Traffic and Its Impact on the Study of Cognitive Radio Networks

by
Sofia Cristina Alvarenga Chu

**A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in Partial Fulfilment of the Requirements of the degree of**

Master of Science

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg

© by Sofia Cristina Alvarenga Chu, July 26, 2012

Abstract

In this thesis, we investigate the impact of channels that have a bursty nature in a cognitive radio network scenario. Our goal is to design a general channel usage model that can handle bursty primary user channel usage. The proposed model describes idle periods with a discrete platoon arrival process and describes busy periods with a discrete phase type distribution. The performance of the proposed model is compared with two more traditionally encountered channel usage models in three different secondary user access schemes.

First, we design a reactive access scheme to show the poor performance results an investigator can potentially obtain when ignoring bursty data traffic. We have also analyzed two proactive secondary network access schemes. Numerical results show that the achievable utilization and interference probability of the network are affected when traditional channel models are used in a bursty PU channel.

Acknowledgements

I would first like to express my sincere gratitude to my advisor Dr. Attahiru S. Alfa for his constant guidance, patience, wise advices and encouragement in the entire development of this research. This thesis would not have been possible without Dr. Alfa thorough recommendations and insightful ideas.

I would like to thank my co-advisor Dr. Jun Cai for providing very useful suggestions, ideas and technical support during this thesis process.

Thanks to Dr. Pradeepa Yahampath and Dr. Ramanathan Sri Ranjan for their comments and valuable suggestions.

I would like to thank all my professors at the University of Manitoba, for their patience in teaching and motivation.

I wish to thank my family; my mother, my father and sisters for helping me get through the difficult times, for all the emotional support and caring they provided. A special thanks to the Aveiro family for their lovely events, support and continuous generosity.

I would like to thank my friends and colleagues in the CNER lab for their encouragement, humour and stimulating discussions.

I would like to acknowledge the PTI organization of the Itaipu Binacional in Paraguay for their financial support in this research.

List of Acronyms

ACK	Acknowledgement
BFW	Broadband Fixed Wireless
CMDP	Constrained Markov Decision Process
DNS	Domain Name System
DSA	Dynamic Spectrum Access
DTMC	Discrete Time Markov Chain
EM	Expectation Maximization
FCC	Federal Communications Commission
Geo	Geometric distribution
GPRS	General Packet Radio Service
HMM	Hidden Markov Models
HTTP	Hypertext Transfer Protocol
ISM	Industrial, Scientific and Medical
LAN	Local Area Network
LRD	Long-Range Dependence

MAC	Media Access Control
MAP	Markov Arrival Process
NACK	Negative Acknowledgement
P2P	Peer-to-Peer
PAP	Platoon Arrival Process
PH	Phase Type
POMDP	Partially Observable Markov Decision Process
PU	Primary User
SU	Secondary User
TCP	Transport Control Protocol
UDP	User Datagram Protocol
VoIP	Voice over IP
WLAN	Wireless LAN

Contents

Abstract	ii
Acknowledgements	iii
List of Acronyms	iv
List of Tables	viii
List of Figures	ix
1 Introduction	1
1.1 Previous Work	2
1.2 Outline	3
2 Cognitive Radio Networks	5
2.1 Cognitive Radio Network Architecture	6
2.2 Objectives and Requirements	7
2.3 Cognitive Radio Functions	8
2.3.1 Spectrum Sensing	8
2.4 Reactive and Proactive Schemes	9
2.4.1 Reactive and Proactive Sensing	10
2.4.2 Reactive and Proactive Spectrum Access	11
3 Special Discrete Time Arrival Processes	13
3.1 Phase Type Distribution	13
3.1.1 Examples of Phase Type Distributions	14
3.2 Markov Arrival Process	16
3.3 Platoon Arrival Process	17
3.3.1 Statistical Characteristics	20
4 Channel Usage Models	22
4.1 Traffic Characterization	23
4.1.1 Self-similarity and Long-Range Dependence	24
4.1.2 Self-similarity and LRD in Data Traffic	25
4.2 Spectrum Usage Models	28
4.3 The Geo-Geo traffic model	31
4.4 The PH-PH traffic model	31
4.5 The PAP-PH traffic model	32

5	Statistical Fitting Procedures	34
5.1	Realization Characteristics	35
5.2	Maximum Likelihood Parameter Estimation for the Geo-Geo Traffic Model	35
5.3	EM Algorithm	37
5.3.1	An EM Algorithm for Fitting Phase-Type Distributions	39
5.3.2	An EM Algorithm for Fitting a Platoon Arrival Process	42
5.4	Fitting Process Examples	51
5.4.1	Low Channel Usage Example	52
5.4.2	High Channel Usage Example	58
5.4.3	A Geo-Geo PU channel	61
5.5	General comments of the fitting process	63
6	A Reactive Spectrum Access Scheme	64
6.1	The System Model	65
6.2	Average Transmission Delay	67
6.3	Average Number of Time Slots between Collisions	69
6.4	Numerical and Simulation Results	72
6.4.1	Simulation Results with Low Traffic Example	73
6.4.2	Simulation Results with High Traffic Example	75
6.4.3	Different Platoon sizes	77
7	Proactive Spectrum Access Schemes	79
7.1	Traffic Prediction and Decision Policy Formulation	80
7.1.1	Examples of the traffic prediction formulas	82
7.2	Performance Measures	84
7.3	Fixed Frame System Model	84
7.3.1	Simulation Results of the Low Traffic Example	87
7.3.2	Simulation Results of the High Traffic Example	91
7.3.3	Simulation Results of the Geo-Geo PU Channel Example	93
7.4	Fixed Transmission Length System Model	95
7.4.1	Simulation Results of the Low Traffic Example	97
7.4.2	Simulation Results of the High Traffic Example	99
7.5	Comparison and Comments	100
8	Concluding Remarks	102
8.1	Conclusion and Discussion	102
8.2	Proposal for future work	104
	References	106

List of Tables

5.1	PAP idle samples fitting procedures results for the low traffic model	56
5.2	PAP idle samples fitting procedures results for the high traffic model	60
5.3	Fitting procedures results for the Geo-Geo PU channel	63

List of Figures

2.1	Periodic sensing over a primary user channel	10
3.1	A data traffic channel realization with idle periods described by a PAP.	18
3.2	A PAP realization (idle periods).	18
5.1	Estimated mean idle period duration for the low traffic model	54
5.2	Estimated mean busy period duration for the low traffic model	54
5.3	Log-likelihood of idle samples with estimates for the low traffic model	55
5.4	Cumulative distribution function for the low traffic model	58
5.5	Estimated mean idle period duration for the high traffic model	59
5.6	Log-likelihood of idle samples with estimates for the high traffic model	59
5.7	Cumulative distribution function for the high traffic model	60
5.8	Estimated mean idle period duration for the high traffic model	62
5.9	Cumulative distribution function for the Geo-Geo PU channel	62
6.1	Reactive spectrum access model with $L = 3$	66
6.2	State Diagram to compute the average transmission delay $E[D_L]$	68
6.3	State Diagram to obtain the average number of time slots between collisions $E[T_{iv}]$	71
6.4	Average transmission delay $E[D_L]$ for the low traffic example	74
6.5	Average number of time slots between collisions $E[T_{iv}]$ for the low traffic example	74
6.6	Average transmission delay $E[D_L]$ for the high traffic example	76
6.7	Average number of time slots between collisions $E[T_{iv}]$ for the high traffic example	76
6.8	Average transmission delay $E[D_L]$ versus constant platoon size K	78
6.9	Average number of time slots between collisions $E[T_{iv}]$ versus constant platoon size K	78
7.1	$E_{on}(9, 3)$: channel has remained idle for 9 time slots and PU returns in 3 time slots	80
7.2	Probability that the PU returns in the next time slot given that the channel has been idle for τ time slots for the low traffic example	82
7.3	Probability that the PU returns in the next time slot given that the channel has been idle for τ time slots for the high traffic example	83
7.4	Secondary user access model with fixed frame structure of length M	85

7.5	Optimal transmission length h^* versus τ for $p_{thr} = 0.05$ and $M = 10$	87
7.6	Optimal transmission length h^* versus τ for different p_{thr} values with $M = 10$ in the low traffic example	88
7.7	Fixed frame scheme: p_{iv} as a function of M in the low traffic example with $p_{thr} = 0.05$	89
7.8	Fixed frame scheme: Φ versus M in the low traffic example with $p_{thr} = 0.05$	89
7.9	Fixed frame scheme: p_{iv} versus p_{thr} in the low traffic example with $M = 10$	90
7.10	Fixed frame scheme: Φ versus p_{thr} in the low traffic example with $M = 10$	91
7.11	Fixed frame scheme: p_{iv} as a function of M in the high traffic example with $p_{thr} = 0.05$	92
7.12	Fixed frame scheme: Φ versus M in the high traffic example with $p_{thr} = 0.05$	92
7.13	Fixed frame scheme: Φ versus p_{thr} in the high traffic example with $M = 10$	93
7.14	Fixed frame scheme: p_{iv} as a function of M in the Geo-Geo example with $p_{thr} = 0.05$	94
7.15	Fixed frame scheme: Φ versus M in the Geo-Geo example with $p_{thr} = 0.05$	94
7.16	Secondary user access model with fixed transmission length L	95
7.17	Fixed transmission length scheme: p_{iv} as a function of L in the low traffic example with $p_{thr} = 0.05$	98
7.18	Fixed transmission length scheme: Φ versus L in the low traffic example with $p_{thr} = 0.05$	98
7.19	Fixed transmission length scheme: p_{iv} as a function of L in the high traffic example with $p_{thr} = 0.05$	99
7.20	Fixed transmission length scheme: Φ versus L in the high traffic example with $p_{thr} = 0.05$	100

Chapter 1

Introduction

Many works, such as [1–5], have demonstrated that for some network scenarios, the data traffic pattern is not well modelled by exponential distributions. The network traffic is usually modelled statistically as a Poisson process due to its analytical simplicity. However, the network traffic can exhibit burstiness and correlation between the interarrivals in a wide range time scale. The traffic can also be self-similar. Existing research, e.g. [6, 7], make us believe that when the traffic is bursty, the performance measures will be impacted when using models that involve exponential distributions. Hence, using models that involve exponential interarrivals or geometric interarrivals may not be adequate in these systems because the results may be misleading.

Recent developments in dynamic and opportunistic spectrum allocation [8–10] have caused us to revisit this issue. Given the scarcity of spectrum and the underutilization of licensed bands, dynamic spectrum access (DSA) has brought a lot of attention as a possible solution for a more efficient use of the limited spectrum. The main idea in these DSA networks is to reuse the statically allocated spectrum of primary users (PUs). These PUs or licensed users are those who hold official use of some specific radio spectrum bands. DSA can be done by introducing cognitive radio networks, where secondary users (SUs) opportunistically access the channel when it is unused by primary users. The term cognitive comes from the fact that these secondary users should adapt to the dynamic environment

by using their ability to sense and to learn.

In a cognitive radio network, the interference to PUs should be limited. At the same time, sufficient throughput for a successful communication in the secondary network should be maintained. Hence, a critical task is to identify white spaces/spectrum holes/spectrum opportunities, which strongly depend on the traffic patterns of the primary users. Accurate characterization of the usage of spectrum is needed [11]; thus, we reconsider the issue about bursty arrivals.

Channel usage patterns of PUs can be represented by an alternating process of idle and busy (OFF and ON) periods. The traffic can exhibit burstiness; hence, instead of modelling the channel usage with geometric or exponential distributions as done in many previous works [12–14], we will develop a more general spectrum use model that can also characterize bursty traffic patterns. Besides investigating the benefits in the performance of the cognitive radio network when using the more general model in a bursty PU's channel, we also aim to study the possible impact of the geometric distribution assumption in a bursty PU's channel.

1.1 Previous Work

Many authors [15–17] have considered different designs to model a bursty arrival process. One popular arrival process is the Markov modulated Poisson process [18] and its discrete time counterpart the Markov modulated Bernoulli process. The authors in [15] proposed the use of a Hyper-Erlang distribution to model the interarrival times. A Hyper-Erlang distribution can exhibit fat tail behaviour with adequate parameters. Another option is to have packet arrivals described by a platoon arrival process (PAP). The authors in [7] designed a platoon arrival process to model burstiness encountered in the intersections of roads. We adopt this latter model to describe our bursty channel model.

Commonly encountered channel usage models involve exponentially or geometrically distributed busy and idle periods [12, 19] because of their analytical simplicity. On the

other hand, stemming from the fact that data packets arrivals can be bursty, one of the suggestions in [20] was to use a mixture distribution that involves uniform and generalized Pareto distributions to represent the idle periods. Hidden Markov models (HMM) have also been proposed in [21] and [22] to predict when the channel will be idle. The authors in [23], focused only on capturing the statistical characteristics of TV-broadcast scenarios. Adaptive sensing to maximize the discovery of channel opportunities was explored in [24] using renewal theory.

Similarly as in [20], we want to design a general channel usage model that also takes into account bursty data scenarios while maintaining reasonable complexity. Thus, we have decided to describe the idle periods with a platoon arrival process and the busy periods with a phase type (PH) distribution. This channel usage model is denoted as the PAP-PH model. Another important reason for this design is that there are known algorithms for the statistical fitting procedure of both processes, for a discrete PAP in [25], and a few fitting procedures choices for PH are explained in [26].

Assuming that the idle durations are independent of the busy durations in a PU channel, a simple model is the alternating geometric distributed idle and busy periods model (Geo-Geo). A more general model that includes the Geo-Geo case is the alternating phase type distributed idle and busy periods model (PH-PH). In this thesis, the Geo-Geo and PH-PH channel usage models will be referred to as commonly encountered or traditional channel usage models. Our proposed channel usage model, the PAP-PH model, will be compared with the Geo-Geo and PH-PH models in three spectrum access schemes.

1.2 Outline

The objective in this research work is to develop a general data traffic model that captures bursty traffic scenarios. Under a reactive and proactive spectrum access, our general channel usage model will be compared to those commonly encountered channel usage models. Before describing the general data traffic model, a description of the cognitive radio net-

work with a special focus towards spectrum access schemes will be given in Chapter 2.

The necessary background with respect to discrete time processes will be explained in Chapter. 3. The proposed general channel usage model, i.e. PAP-PH, and other more traditional channel usage models will be presented in Chapter 4, which also outlines the reasons of our PAP-PH model design. The channel usage models can only be implemented in a cognitive radio network scenario if automatic fitting procedures exist to estimate the model's parameters. Thus, Chapter 5 describes the fitting procedures that has been used to estimate the parameters of each channel usage model applied in this research work.

To analyse the potential effects of a bursty PU channel on the performance measures of a cognitive radio network scenario, a reactive spectrum access has been designed and presented in Chapter 6. Given the benefits of a proactive spectrum access, in Chapter 7 we use two different proactive spectrum access schemes to show the increased spectrum utilization and reduced PU interference that can be achieved with an accurate characterization of the spectrum usage. One of these spectrum access schemes was developed in [27], and we have developed an additional scheme which can be considered as a modification of that proposed in [27]. For all these spectrum access schemes, we will show the results of two examples: a low channel utilization example and a high channel utilization usage example.

Finally, the conclusion and proposals for future works are presented in Chapter 8.

Chapter 2

Cognitive Radio Networks

Given the rapid deployment of wireless devices and applications, the demand for the radio spectrum resource has been constantly increasing. Unfortunately, for decades, most of the available radio frequencies have been assigned in a static manner to licensed users by governmental organizations, like the Federal Communications Commission (FCC). These licensed users are also known as *primary users* (PUs). Studies, such as [28], indicated that the current assignment policy underutilizes the radio spectrum. Some bands are heavily used while others are not used or are only partially used. Hence, a new spectrum management approach needs to be developed for a more efficient use of the network resources.

Stemming from the fact that changing the existing network architecture is not viable, a possible solution will be to apply a dynamic spectrum access (DSA) technique. In this technique, *secondary users* (SUs) who do not have spectrum licenses are allowed to opportunistically access the licensed spectrum when it is unused. These unused spectrum bands are also known as white spaces, spectrum opportunities, spectrum holes or simply opportunities.

To accomplish DSA, the FCC has proposed the use of a cognitive radio, for which we provide the following definition [8, 10, 29]: *a cognitive radio is an intelligent wireless communication radio that has the ability to sense and to gather information from its surrounding environment. It has the capacity to learn and to adjust its transmission parameters in*

an autonomous manner.

From the definition of a cognitive radio, the key tasks in a cognitive radio cycle can be divided into four [30]. The first task involves a real time sensing of the radio environment. The second task consists of a rapid analysis of the sensing information to predict the current channel state and channel capacity. Based on the results of the second task, cognitive radio functions should determine the best action or response strategy in the third task. During the fourth task, cognitive radios reconfigure their operating parameters accordingly and perform a strategic action. The transmission parameters that a radio device can adapt according to its dynamic surroundings can consist of [8, 9, 30]: transmission power, carrier-frequency, modulation type, and communication technology. The strategic action can consist of remaining idle, sensing and/or transmission.

A secondary link is formed by a secondary user receiver and a secondary user transmitter. Transmitter and receiver synchronization can be achieved by the use of a feedback channel, where control information can flow between the receiver and transmitter. Depending on the nodes involved in the cognitive radio network, different tasks can be performed in an individual or collective manner.

Each cognitive radio device will have a set of characteristics. The cognitive device can employ a wideband spectrum transmission or narrowband spectrum transmission. In this thesis, we focus on narrowband spectrum transmissions. This access technology is known as overlay spectrum access [9, 11], where SUs only try to transmit when the channel is not being used by PUs. In [31], this access approach is known as the interweave paradigm.

2.1 Cognitive Radio Network Architecture

As explained in [8,9], the CR architecture consists of two components: the primary network and the secondary network. For each of these networks, a centralized or distributed architecture is used. As the name suggests, in a centralized architecture, decisions of spectrum allocation are enforced by a central entity of the network. In a distributed architecture, no

central entity is present, and nodes of the network must organize and communicate between themselves to achieve their goals.

The primary network is composed of primary users and the existing infrastructure surrounding them; such as, a base station for a centralized architecture. On the other hand, the secondary network is formed by users or nodes who have no privilege on the frequency bands. A secondary base station can be present in a centralized architecture to coordinate the access between secondary users in the secondary network. A secondary base station is also the bridge to other cognitive radio networks and can perform other functions in a centralized manner. A third component of the secondary network is the spectrum broker or scheduling server [9] whose function is to administer the usage of the spectrum resources among different cognitive radio networks.

In this work, one of the main assumptions is that a single primary user channel is analyzed. The primary network can consist of one or more primary users that access a single channel. In the same way, the secondary network can consist of one or more secondary users, who opportunistically try to access the primary user channel. We do not focus on the scheduling among secondary users.

2.2 Objectives and Requirements

The main objective in a cognitive radio network is to identify and efficiently use the available white spaces. However, for the success of a cognitive radio network, we must guarantee that the amount of interference to the primary users is limited. Collisions caused by secondary users' transmissions may significantly deteriorate primary users' quality of service. Thus, while avoiding collisions with primary users, secondary users should maintain a stable communication between themselves to fulfill their own performance requirements.

The more information a secondary network possesses, the higher the chances for an intelligent decision. To achieve the aforementioned objectives, there is a need for an accurate spectrum usage characterization. With the adequate channel usage model, more complex

upper layer functions can be successfully deployed.

2.3 Cognitive Radio Functions

To efficiently utilize the dynamic available spectrum, a DSA network has many functions which can be grouped as [9, 30]: spectrum sensing and analysis, spectrum management, spectrum handoff, and spectrum sharing.

- **Spectrum sensing and analysis:** is a critical task from which a secondary user can detect spectrum holes or white spaces [30]. After a white space detection, an SU can proceed to utilize the available spectrum. Similarly, an SU can detect a PU presence by sensing the channel.
- **Spectrum management:** the objective of spectrum management is to select the best available channels to achieve the communication requirements.
- **Spectrum handoff:** according to the time varying frequency channel, the secondary users may need to hop among different PU channels to maintain seeming-less communication. A spectrum hand-off scheme goal is to determine how and to which PU channel, a pair of SU should switch. The switching process should be such that the SUs have high probability of achieving the required quality of service in the new channel.
- **Spectrum sharing:** the secondary network may consist of many secondary nodes, each of which have data packets to transmit and quality of services to be fulfilled. In this case, spectrum sharing is necessary to administer the available frequency bands between the secondary users.

2.3.1 Spectrum Sensing

Secondary users perform sensing to discover spectrum opportunities. In this research, the concept of a white space follows the conventional definition of a spectrum opportunity

which has been described as [10]: “*a band of frequencies that are not being used by the primary user of that band at a particular time in a particular geographic area*”.

The most common spectrum sensing techniques include [10]: energy detection, waveform-based sensing, cyclostationary-based sensing and matched-filtering. A binary hypothesis test can be formulated [9, 13]:

$$\begin{aligned} \mathcal{H}_0 & : \text{ primary user channel is idle,} \\ \mathcal{H}_1 & : \text{ primary user channel is busy.} \end{aligned} \tag{2.1}$$

Let us define that the sensing result is described by random variable θ , with $\theta \in \{\text{busy, idle}\}$. The probability of detection P_d is defined as the probability the channel is sensed busy given that the primary user channel is active, i.e.,

$$P_d = \Pr(\theta = \text{busy} | \mathcal{H}_1). \tag{2.2}$$

The probability of missed detection is $1 - P_d$. The probability of false alarm P_f is defined as the probability of sensing the channel as busy given that the PU channel is idle, i.e.,

$$P_f = \Pr(\theta = \text{busy} | \mathcal{H}_0). \tag{2.3}$$

Given that the outcome of a sensing can be assumed to be idle or busy (OFF or ON), a primary user channel can be modelled as an idle and busy (OFF and ON) random process.

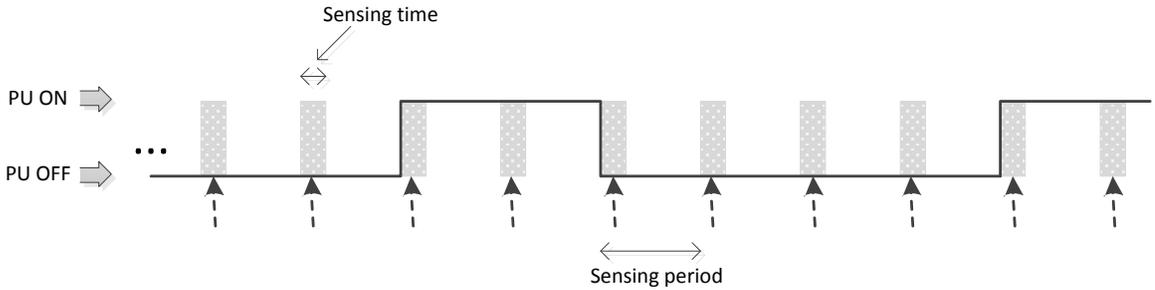
2.4 Reactive and Proactive Schemes

Depending on the amount of sensing information the secondary network uses in the cognitive radio functions, these functions can be classified as a reactive or proactive schemes.

2.4.1 Reactive and Proactive Sensing

In the sensing task, a reactive sensing implies that a secondary user solely monitors the primary user channel when it has packets to transmit or to receive. This means that a secondary user does not exploit the inherent PU channel pattern information. On the other hand, in a proactive sensing, secondary users continuously monitor the channel according to their respective sensing period. This periodic sensing allows the secondary network to gather information about each channel usage behaviour. In addition, channel usage patterns can be estimated using the collected sensing data. Using a statistical channel usage model, the performance of cognitive radio functions can be further improved.

Figure 2.1: Periodic sensing over a primary user channel



In Figure 2.1, we show an example of a primary user channel with a proactive sensing scheme. To obtain a sensing sample, sensing should be performed through a length of time to achieve a given false alarm probability P_f . This length of time will be referred to as sensing time. The length between two consecutive sensing samples is denoted by sensing period. One of the main disadvantages of a proactive sensing is that it can cause more sensing-overhead than a reactive sensing [23, 24]. Even though a secondary user does not have packets to transmit, it needs to periodically sense the channel. Spectrum sensing costs time and is energy demanding. As a result, there is a trade-off between performance and complexity. For adequate performance, the optimal sensing period should be determined for each spectrum band. In [24], the authors proposed a sensing-period adaptation algorithm for a set of PU channels. After collecting a sequence of samples or sensing results from each PU channel, the channel usage behaviour is predicted and the

sensing period is adapted accordingly. A framework for an adaptive spectrum sensing and access has also been proposed in [32]. They minimize the number of sensing assignments in an occupied channel by implementing a back-off time resolution algorithm.

2.4.2 Reactive and Proactive Spectrum Access

In a reactive spectrum access approach, an SU's action of when to transmit is solely based on the latest sensing information whereas no historical sensing data is used, and an SU only reacts after detecting a change in the channel state. Although, a reactive spectrum access is simple, it can cause higher interference to the primary users when compared to a proactive spectrum access scheme [21]. A proactive spectrum access consists of designing a model that captures the primary's usage channel patterns by using historical sensing information. This model can be used to predict the channel availability in the future, and the channel access can be planned accordingly among the SUs. Throughout this thesis, the terms proactive and reactive will denote whether or not the SU network employs a traffic model.

In the literature, reactive and proactive spectrum access schemes have also been used to denote how and when an SU chooses to switch to another PU channel. For instance, in [12], a comparison of a reactive and proactive switching is discussed. In a reactive switching scheme, an SU only switches to a different channel when the current channel is detected busy and it randomly chooses a channel to switch. On the other hand, before a collision with a PU, an SU can switch in advance to a channel that has a larger mean remaining idle period or a higher probability of remaining idle time. These latter schemes are known as proactive switching. When the proactive switching scheme proposed in [12] is used instead of a reactive random search, the authors showed that SU's channel utilization can be increased by 5% and the primary user hit rate can be reduced by 30%.

When a secondary user needs to find an idle channel to continue transmitting, it should determine the optimal order of channels to sense such that the channel-search delay is minimized. The authors in [24], proposed an ordering of channels based on the conditional

probability of finding the channel idle given the previous sensing history. Another option is to use the expected remaining idle time, as in [33]. When no channel ordering is used, the results in [24] indicate that the delay increases as the number of channels to sense increases. However, with channel ordering, the search delay remains almost constant and invariant of the number of channels to sense. Based on prediction, the authors in [19], proposed a channel ordering for an SU with the constraints of maintaining the quality of services requirements and maximizing spectrum utilization.

Markov decision processes have been applied to find optimal or sub-optimal solutions for different opportunistic spectrum access schemes. These schemes usually involve a set of PU channels and a set of requirements and/or constraints. In [14], a constrained Markov decision process (CMDP) is proposed to determine the optimal policy (sensing and transmission) for an SU. Moreover, a partially observable Markov decision process (POMDP) framework for optimal sensing and access strategies has been addressed in [34]. A POMDP framework with energy constraints was designed in [13]. A PU's channel idle and busy periods were characterized by exponential or geometric distributions in [13,14,34]. However, general busy period distributions were considered in the CMDP proposed by [35]. The authors from [35] compared three random access schemes to determine the optimal solution (sensing scheme, access scheme, transmission length) of a problem with throughput and collision constraints.

The performance of a proactive access scheme will heavily depend on how accurately the model can characterize and predict the channel usage. This is the main issue in this thesis and is investigated in *Chapter 7*.

Chapter 3

Special Discrete Time Arrival Processes

In this Chapter, we want to provide the necessary background to understand the channel usage models that will be defined in *Chapter 4*. A general overview of phase type distributions and platoon arrival processes will be explained in the next sections.

3.1 Phase Type Distribution

A discrete phase type (PH) distribution describes the number of time slots until absorption. Consider a discrete time absorbing Markov Chain $\{X_n\}$ with state space $\{0, 1, 2, \dots, m < \infty\}$. State $\{0\}$ is an absorbing state and states $\{1, 2, \dots, m\}$ are transient states. This Markov chain is described by the following transition matrix

$$P = \begin{bmatrix} 1 & \mathbf{0} \\ \boldsymbol{\tau} & T \end{bmatrix}, \quad (3.1)$$

where T is a $m \times m$ substochastic matrix representing the transition probabilities between transient states. Column vector $\boldsymbol{\tau}$ with dimension $m \times 1$ represents the probability of absorption in each transient state. We have that $\boldsymbol{\tau} = \mathbf{1} - T\mathbf{1}$, where the notation $\mathbf{1}$ indicates a column vector of ones. A start probability vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_m]$ represents the probability of starting the process in each phase. The probability of instantaneous

absorption is $\alpha_0 = 1 - \alpha \mathbf{1}$. This phase type distribution has representation $\text{PH}(\alpha, T), m$. The constant m is known as the order of a phase type distribution.

If the random variable Y denotes the time until absorption and is further phase type distributed, $\text{PH}(\alpha, T), m$, then, the probability mass function of Y is

$$\Pr(Y = 0) = \alpha_0, \quad \Pr(Y = i) = \alpha T^{i-1} \tau, \quad i \geq 1. \quad (3.2)$$

The cumulative distribution function of Y is

$$\begin{aligned} F = \Pr(Y \leq i) &= 1 - \Pr(Y > i) \\ &= 1 - \alpha T^i \mathbf{1}. \end{aligned} \quad (3.3)$$

In addition, the i th factorial moment of the time to absorption is

$$\begin{aligned} E[Y^i] &= \sum_{n=1}^{\infty} n^i \Pr(Y = n) \\ &= i! \alpha T^{i-1} (I - T)^{-i} \mathbf{1}. \end{aligned} \quad (3.4)$$

The mean is

$$E[Y] = \alpha (I - T)^{-1} \mathbf{1}. \quad (3.5)$$

A phase type distribution can be used to represent most special statistical distributions encountered in tele-traffic.

3.1.1 Examples of Phase Type Distributions

Suppose we have a sequence of Bernoulli experiments or trials. For each experiment, the probability of a success is p , and the probability of failure is $(1 - p)$.

1. Geometric distribution:

If random variable X represents the number of experiments needed for a success then X has a geometric distribution. We denoted this by $X \sim \text{Geo}(p)$. A geometric

random variable has a probability mass function of $\Pr(X = i) = (1 - p)^{i-1}p$, which can be represented by a PH distribution of order $m = 1$ with $\alpha = 1$, $T = (1 - p)$, and $\tau = p$.

2. Negative Binomial Distribution:

Consider that random variable X denotes the number of trials needed for n successes. Then, X is said to be negative binomial distributed, $X \sim \text{NB}(n, p)$. The probability mass function of X is given by

$$\Pr(X = i) = \binom{i-1}{n-1} p^n (1-p)^{i-n}, \quad i \geq n, \quad (3.6)$$

where $\Pr(X = i)$ is the probability that i trials are needed for n successes. A negative binomial distribution $\text{NB}(n, p)$ can be represented by a $\text{PH}(\alpha, T)$, n with $\alpha = \begin{bmatrix} 1 & 0 & 0 & \dots \end{bmatrix}$ and the following transition matrix

$$T = \begin{bmatrix} (1-p) & p & \dots & 0 \\ 0 & (1-p) & p & \dots & 0 \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & p \\ & & & & (1-p) \end{bmatrix}.$$

3. Constant interarrival time of value n .

One possibility to represent this distribution is with a $\text{PH}(\alpha, T)$, n where $\alpha = \begin{bmatrix} 1 & 0 & 0 & \dots \end{bmatrix}$, and the elements of the transition matrix T are

$$T(i, i+1) = 1, \quad \forall 1 \leq i \leq n-1.$$

For example, assume that $n = 3$, then the probability of start of each phase is $\alpha = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$, and the transition states before absorption are represented by

$$T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

3.2 Markov Arrival Process

Consider that the process starts at $t_0 = 0$ with zero arrivals. We define a discrete time Markov chain with state space

$$\Delta = \{(0) \cup (X_m, J_m); X_m \geq 1, J_m = \{1, 2, \dots, n\}\}, \quad m \geq 0, \quad (3.7)$$

where X_m denotes the number of arrivals that has been observed up to time slot m and J_m denotes the current phase of the arrival process at time slot m .

A *Markov Arrival Process* (MAP) [36] is characterized by two substochastic matrices C_0 and C_1 of dimension n . Matrix C_0 denotes the probabilities of no arrival event, and matrix C_1 denotes the probabilities of an arrival event. A MAP captures correlations between consecutive arrival events. More specifically, say $(C_0)_{i,j}$ and $(C_1)_{i,j}$, with $1 \leq i, j \leq n$, are the elements of matrix C_0 and C_1 , respectively. Element $(C_k)_{i,j}$, with $k = 0, 1$, denotes the transition probability from phase i to phase j with k arrivals. Matrix $C = C_0 + C_1$ is a stochastic and irreducible matrix. The transition matrix of this arrival process whose state space is (3.7) can be written as

$$P = \begin{bmatrix} C_0 & C_1 & 0 & \cdots \\ 0 & C_0 & C_1 & \ddots \\ 0 & 0 & C_0 & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (3.8)$$

3.3 Platoon Arrival Process

A *platoon arrival process* (PAP), which was defined in [7], is a special case of a MAP. A PAP can capture bursty traffic behaviour by using two interarrival distributions: the intraplatoon and interplatoon distributions. A given realization of the platoon arrival process consists of the following repetitive pattern: a consecutive sequence of intraplatoon interarrivals followed by one or more interplatoon interarrivals. A platoon is defined as a sequence of consecutive arrivals separated by intraplatoon interarrivals. The mean intraplatoon interarrival is shorter than the mean interplatoon interarrival. This behaviour gives rise to the bursty nature of the statistical model.

A more general representation of a PAP is given in [25] using a terminating MAP. The PAP model can capture correlation between consecutive arrivals, that is, it can differentiate intervals between arrivals in a platoon, or between the last arrival in a platoon and the first arrival of the following platoon.

In this research, we use a platoon arrival process to characterize idle periods in our proposed traffic model. When we include the platoon concept for idle periods, we can distinguish two types of idle periods denoted as the intraplatoon idle period and the interplatoon idle period. In an arrival process, an idle period represents an interarrival length. To incorporate a PAP without taking into account busy periods, the definition of a platoon is slightly different from the work in [7]. Our definition is that a platoon groups consecutive intraplatoon idle periods into groups, and between platoons there is an interplatoon idle period. Having this definition allows us to preserve the structure of the transition matrix of the PAP process exactly as in [7]. The different definition of a platoon causes the distribution of the number of elements in a platoon to refer to the distribution of the number of consecutive intraplatoon idle periods.

An example of a channel usage realization where idle periods are characterized by a platoon arrival process is given in Figure 3.1. In this figure, random variables Y_i and Z_i represent the duration of the busy period i and the duration of the idle period i , respectively.

Figure 3.1: A data traffic channel realization with idle periods described by a PAP.

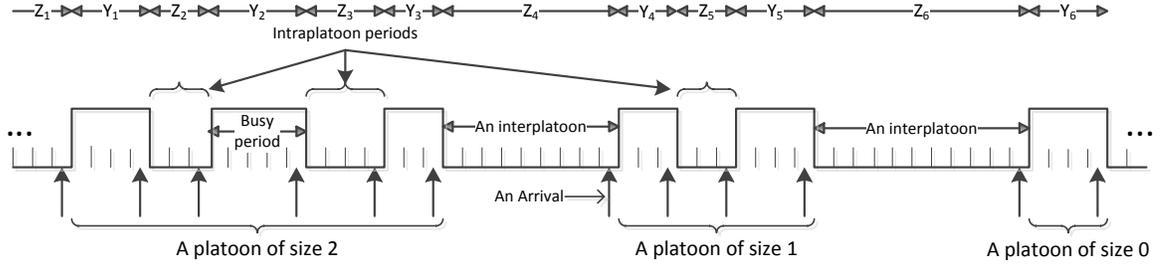
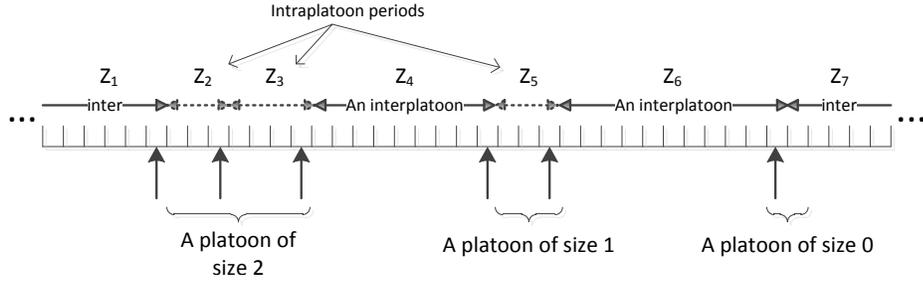


Figure 3.2: A PAP realization (idle periods).



In a cognitive radio network point of view, and as stated in [20], DSA is a promising approach when we can find idle periods that are long enough for an SU packet transmission. If an SU starts transmitting the moment it senses the channel is free, this strategy as mentioned before can result in frequent disruptions to both PUs and SUs. Therefore, we are concerned with differentiating shorter idle period from longer idle periods, and we think that our model can be useful to predict the different types of idle periods.

With the aforementioned definition of a platoon, let us assume that the number of intraplatoon idle periods in a platoon follows a phase type distribution of order K described by $\text{PH}(\boldsymbol{\delta}, \mathbf{D})$, K with $\boldsymbol{\delta}_0 + \boldsymbol{\delta}\mathbf{1} = \mathbf{1}$ and $\mathbf{D}^0 + \mathbf{D}\mathbf{1} = \mathbf{1}$. In this case, $\boldsymbol{\delta}$ is the probability of start for each phase. We allow the possibility of instantaneous absorption with $\delta_0 > 0$, which means that after an interplatoon interval another interplatoon interval can follow. Or in other words, an empty platoon has occurred in the process describing idle periods.

The duration of an idle period Z_i follows one of the following interarrival distributions:

- A $\text{PH}(\boldsymbol{\alpha}_1, T_1)$, m_1 representing the intraplatoon interval distribution with $\boldsymbol{\alpha}_1\mathbf{1} = \mathbf{1}$ and $\boldsymbol{\tau}_1 + T_1\mathbf{1} = \mathbf{1}$.

- A $\text{PH}(\boldsymbol{\alpha}_2, T_2), m_2$ representing the interplatoon interval distribution with $\boldsymbol{\alpha}_2 \mathbf{1} = 1$ and $\boldsymbol{\tau}_2 + T_2 \mathbf{1} = \mathbf{1}$.

The probability of instantaneous absorption, $\alpha_i(0) = 1 - \boldsymbol{\alpha}_i \mathbf{1}$, for $i = 1, 2$, is zero because the duration of an idle period is at least one.

How the two types of idle interarrival distributions are related depends on the $\text{PH}(\boldsymbol{\delta}, D), K$ distribution. This $\text{PH}(\boldsymbol{\delta}, D), K$ distribution describes the probability of how many consecutive intraplatoon idle durations are likely to occur before seeing an interplatoon idle duration.

In discrete time, the underlying Markov chain of a PAP has a transition matrix given by [7]

$$P_{\text{PAP}} = \begin{bmatrix} C_0 & C_1 & 0 & \cdots \\ 0 & C_0 & C_1 & \ddots \\ 0 & 0 & C_0 & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad (3.9)$$

with

$$C_0 = \begin{bmatrix} T_2 & 0 \\ 0 & I_K \otimes T_1 \end{bmatrix}, \quad \text{and} \quad C_1 = \begin{bmatrix} \boldsymbol{\tau}_2 \boldsymbol{\delta}_0 \boldsymbol{\alpha}_2 & \boldsymbol{\tau}_2 \boldsymbol{\delta} \otimes \boldsymbol{\alpha}_1 \\ \mathbf{D}^0 \otimes \boldsymbol{\tau}_1 \boldsymbol{\alpha}_2 & \mathbf{D} \otimes \boldsymbol{\tau}_1 \boldsymbol{\alpha}_1 \end{bmatrix}. \quad (3.10)$$

The sign \otimes represents Kronecker product, the symbol I_K denotes identity matrix of dimension K . Here matrix C_0 represents a non-arrival event, and matrix C_1 represents an arrival event. That is, given the current state, C_0 indicates the probability that an idle period continues in a time slot, and C_1 indicates the probability that an idle period has finished in a time slot. The total number of phases in C_0 and C_1 is $m = m_2 + Km_1$. Matrix $C = C_0 + C_1$ is stochastic, and define the phase stationary vector $\boldsymbol{\pi} = [\boldsymbol{\pi}_{inter} \ \boldsymbol{\pi}_{intra}]$ such that

$$\boldsymbol{\pi} = \boldsymbol{\pi} C. \quad (3.11)$$

The interplatoon interval initial phase is chosen according to vector $\boldsymbol{\alpha}_2$. The matrix T_2

describes phase transitions for the interplatoon interval when the channel is idle, this idle period ends with probability τ_2 in a time slot.

The number of intraplatoon intervals is characterized by the PH type distribution (δ, D) . This probability density initial phase is chosen according to δ . The probability that after an interplatoon interval another interplatoon interval comes is given by δ_0 . The initial phase for an intraplatoon interval is given with probability vector α_1 . In an intraplatoon interval ending, a counter through matrix D indicates that the platoon still continues and with matrix D^0 it shows that the platoon has finished. If the platoon continues, then another intraplatoon interval follows the current intraplatoon period with start probability vector α_1 . The identity matrix, I_K of dimension K , allows that the phase of the counter remain unchanged while the process is in the intraplatoon interval without any arrivals.

A platoon arrival process will be represented by $\text{PAP}(\alpha_1, T_1, \alpha_2, T_2, \delta, D), m$ or equivalently by $\text{PAP}(C_0, C_1), m$.

3.3.1 Statistical Characteristics

Consider that the phase stationary vector probabilities are $\pi = \pi(C_0 + C_1)$ with $\pi\mathbf{1} = 1$. Then, the probability that there is an arrival at an arbitrary time slot in the stationary process is $\lambda = \pi C_1 \mathbf{1}$. This is also known as mean arrival rate [25]. Then, the stationary distribution for the phases immediately after an arrival can be computed as [37]

$$\pi' = \frac{\pi C_1}{\pi C_1 \mathbf{1}}. \quad (3.12)$$

This conditional stationary distribution of the phases π' can be used to calculate the probability mass function and cumulative distribution function of the interarrival times. The probability that the time to absorption is i or the probability that the idle period length is i is given by

$$\Pr(Z = i) = \pi' C_0^{i-1} C_1 \mathbf{1}, \quad i \geq 1, \quad (3.13)$$

and the cumulative distribution function of Z is

$$\begin{aligned}\Pr(Z \leq i) &= 1 - P(Z > i) \\ &= 1 - \boldsymbol{\pi}' C_0^i \mathbf{1}, \quad i \geq 1.\end{aligned}\tag{3.14}$$

We can get the mean duration L_{inter} of the interplatoon period with

$$\begin{aligned}L_{inter} &= \sum_{k=1}^{\infty} k \boldsymbol{\alpha}_2 T_2^{k-1} \boldsymbol{\tau}_2 \\ &= \boldsymbol{\alpha}_2 (I + 2T_2 + 3T_2^2 + \cdots) \boldsymbol{\tau}_2 \\ &= \boldsymbol{\alpha}_2 (I - T_2)^{-2} \boldsymbol{\tau}_2 = \boldsymbol{\alpha}_2 (I - T_2)^{-2} (I - T_2) \mathbf{1} \\ &= \boldsymbol{\alpha}_2 (I - T_2)^{-1} \mathbf{1}.\end{aligned}\tag{3.15}$$

In a similar way, the mean duration L_{intra} of the intraplatoon period is

$$\begin{aligned}L_{intra} &= \sum_{k=1}^{\infty} k \boldsymbol{\alpha}_1 T_1^{k-1} \boldsymbol{\tau}_1 \\ &= \boldsymbol{\alpha}_1 (I - T_1)^{-1} \mathbf{1}.\end{aligned}\tag{3.16}$$

The mean number of consecutive intraplatoon periods in a platoon is

$$\begin{aligned}\bar{E} &= \sum_{k=1}^{\infty} k \boldsymbol{\delta} D^{k-1} \mathbf{D}^0 \\ &= \boldsymbol{\delta} (I - D)^{-1} \mathbf{1}.\end{aligned}\tag{3.17}$$

The mean duration of the idle period is

$$\begin{aligned}I_{ave} &= E[Z] = \sum_{k=1}^{\infty} k \boldsymbol{\pi}' C_0^{k-1} C_1 \\ &= \boldsymbol{\pi}' (I - C_0)^{-1} \mathbf{1}.\end{aligned}\tag{3.18}$$

Chapter 4

Channel Usage Models

In this Chapter, we will specifically define three discrete time channel usage models that will be compared and analyzed in the reactive and proactive schemes in *Chapter 6* and *7*, respectively. Recall that the channel usage model in a proactive scheme will be used to predict the PU's spectrum usage.

Generally, a source traffic model consists of two random processes [1]. These random processes are an arrival process that determines the interarrival time between two events and a process that describes the duration of the event. However, when dealing with channel usage models, as in [12,13,24,33], we also think that the two processes of a traffic model are better described by the distribution of the OFF and ON periods. The terms channel usage model, spectrum usage model, traffic model, and data traffic model in a cognitive radio network context refers to the model used to characterize the usage patterns of a frequency spectrum band.

Channel usage is represented by an alternating process of idle and busy (OFF and ON) periods. Consider that the system is observed at equally spaced points in time; define this interval as a time slot. Let Y_i and Z_i , with $i \geq 1$, be a sequence of two types of discrete random variables representing the duration of the busy period i and the duration of idle period i , respectively. For all the subsequent channel usage models, it is assumed that the duration of busy periods are independent of the duration of idle periods.

When idle and busy periods are described by geometric distributions, we will refer the traffic model as a Geo-Geo channel usage model. A more general traffic model than the Geo-Geo process consists of idle and busy periods that are characterized by phase type distributions, in this case the channel model is named a PH-PH traffic model. Finally, an even more general model capable of including bursty traffic patterns comprises idle periods that follow a platoon arrival process and busy periods that are phase type distributed. This latter model is referred to as a PAP-PH traffic model, and will be the topic in *Section 4.5*.

Before explaining the spectrum usage models, in *Section 4.1*, we intend to provide an overview of the reasons why we have designed a bursty channel model. Our interest comes from the results obtained through the analysis of data network traces. Through traffic characterization, many researches have become concerned about the Poisson process assumption and have questioned its validity when modeling characteristics of the network traffic; such as, packet and connection arrivals, packet and connection sizes. In addition, self-similarity and long-range dependence (LRD), which are briefly explain in *Section 4.1.1*, have become popular terms over the last decade for describing the network traffic.

4.1 Traffic Characterization

When the arrival process is characterized by a Poisson process, we know that the interarrival times are exponentially distributed. To understand why exponential distributions may not be sufficient to accurately model the traffic behaviour, in this Section, we intend to give an overview of different studies that have analyzed the characteristics of different network traffic traces. Given the constant evolution of the internet and network technologies, traffic characterization has been and will continue to be a challenge. Our focus is to provide an introduction of the reasons why exponential models may not be adequate to model some traffic patterns.

It is known that the use of exponential distributions has been widely accepted to describe statistical characteristics of conventional voice applications; such as, the interarrival

time between connections setups and the connection duration [1]. However, next generation mobile communications services are increasingly becoming packet oriented with multiple classes of applications. The use of the Internet has evolved rapidly in terms of the variety of available applications. These not only consist of file transfers, but also consist of music streaming, video streaming, on-line games, voice over IP (VoIP), e-commerce and peer-to-peer (P2P) applications. In addition, the number of users and the variety of technology devices that are connected to the Internet has grown substantially in recent years. New applications, such as those involving variable-bit rate, introduce more complexity to the analysis of the internet traffic. Moreover, the data traffic is highly dependent on the applications and other non-predictable factors. Thus, the internet traffic characterization is complex to understand because of the heterogeneous sources, the different wireless access technologies and the numerous network interconnections.

4.1.1 Self-similarity and Long-Range Dependence

Over the past decades, self-similarity and long range dependence has been an active topic in the network traffic modelling. We will define these two concepts in a discrete time system.

First, we will provide the necessary background of stochastic processes. Consider a discrete time stochastic process $\{X_n\}$ with $n = 0, 1, \dots, N$. The mean of the random process $\{X_n\}$ is defined as [38]

$$m_X(n) = E[X_n]. \quad (4.1)$$

The autocorrelation function is defined by

$$R_X(n, l) = E[X_n X_l], \quad (4.2)$$

and it characterizes the dependence relationship between r.v. X_n and r.v. X_l .

The random process $\{X_n\}$ is wide sense stationary if two conditions are satisfied [38,

39]:

1. $m_X(n) = m_X(n + l) \forall n, l$.
2. $R_X(n, l) = R_X(n - l) = R_X(|n - l|) \forall n, l$.

The first condition means that $m_X(n) = \mu, \forall n$. The secondary condition implies that the autocorrelation function is only dependent of the time difference between two random variables.

A stationary random process $\{X_n\}$ is *long range dependent* if $\sum_{i=1}^{\infty} |R_X(i)| = \infty$ [40]. Unlike the memoryless Poisson arrival process, traffic that exhibits long-range dependence implies that the traffic presents long-term memory. More specifically, as defined in [40], “*LRD means that the behaviour of a time-dependent process shows statistically significant correlations across large time scales*”.

Self-similarity has been defined as “*the phenomenon in which the behaviour of a process is preserved irrespective of scaling in space or time*” [40]. It is important to understand that self-similar processes have different theoretical properties when compared to Poisson processes. As stated in [2], one of the main characteristics of a self-similar and fractal-like traffic is that burstiness is present over a wide range of time scales. Contrary to self-similar processes, aggregated traffic from sources that are described by Poisson processes tend to be less bursty as the number of active sources increases [41].

4.1.2 Self-similarity and LRD in Data Traffic

Popular random processes $\{X_n\}$ in traffic characterization include measuring the packet arrival rate, connection arrival rates, sessions arrival rates in a communication link. The concept of connection and sessions depends on the application involved and assumptions made by the authors to model the traffic. Other random processes measure the distribution of the packet, connection and session sizes.

In the traffic characterization literature, authors commonly use the concept of a flow or connection to discuss the behaviour of the traffic, for which we offer the following

definition [42]: “A flow is defined as a set of packets that share origin and destination addresses, origin and destination ports (if applicable to the transport protocol utilized), transport protocol and are observed within a time-frame.” A flow can be classified based on the size, duration and by burstiness.

Researchers have observed self-similar traffic in local and wide area networks. One of the first works, [41], that identified self-similar traffic involves a local area network (LAN) scenario. Moreover, after analysing traces from a wide-area traffic, the authors in [2] concluded that the commonly used Poisson process for packet and TCP connection arrivals underestimates the burstiness present over a wide range of time scales.

The main objective in [1] was to provide traffic models for data users in wireless networks, particularly general packet radio service (GPRS). As explained in [1], the traffic can present a high degree of bustiness and can be better described by self-similar processes. Not only due to the high variance of the interarrival times between packets, but also from the high variance of the packet length. As indicated by [1], Poisson models in bursty traffic scenarios can lead to an understimation of some performance measures; such as, the average packet delay or maximum queue size.

Authors like [1, 42] claimed that the LRD structure can be caused by: the feedback-oriented nature of TCP control mechanisms, traffic with heavy-tailed flows distributions and heavy-tailed file size distributions. In these cases, the Poisson model is no longer sufficient. Packets of a TCP connection are dependent on each other. Hence, the usage of TCP connections causes a dependence between packets of the same connection that are transmitted over long ranges [3].

In addition, studies as [42] stated that aggregated connections or aggregated network traffic also exhibits LRD correlations and traffic bursts are originated from several active connections. Interestingly, authors in [42] explained that studies indicate that an internet bursty traffic typically arises from a few large dominating flows connections. They further stated that the burstiness is a result of the heterogeneity of the network, and, as in [3], they pointed out that large flows causes LRD.

Over the last few years, there has been a dramatic increase of the P2P traffic in the network. According to [42], the predominant type of traffic involves P2P applications and video sharing applications. Focussing on the size criteria of a flow, these types of traffic cause the “mice and elephant flow” phenomenon in the internet traffic [3, 42]. Namely, internet traffic consists of thousands of mice (small flows) and there is some number of elephants (large flows). These elephants are responsible for a large percentage of the internet traffic volume. Because of this fact, the distribution of the flows is heavy-tailed and the shape of the distribution is very different from that of an exponential distribution.

According to [4], on-line games traffic consists of a large number of small, periodic bursts of UDP packets. The periodicity comes from the predictable and frequent updates of the server to its clients. After analysing on-line game traces in different time-scales, an extremely bursty and a highly periodic pattern was observed in terms of the packet loads at the server.

In [43], upstream internet traffic of a Broadband Fixed Wireless (BFW) access was analysed. Their results indicated that P2P applications are dominant in the user traffic. This is one of the causes for the existence of short flows in the upstream link between the end-user and the head-end. The aggregated upstream traffic consisted mostly of web requests, UDP/DNS packets, TCP acknowledgement and control packets, and other single packet flows. The flow arrival process of a BFW access showed burstiness; which could be caused by the shared bandwidth among users and the underneath MAC protocols.

In [44], the authors analyzed P2P flow-level internet traces on a transpacific link. Their results showed that different characteristics are exhibited by different P2P applications. For the BT and Napster applications, they found that the traffic volume per connection can be modeled by a Weibull and Pareto distribution respectively. Both applications exhibited bursty heavy-tailed phenomenon.

The works explained in this Section are examples that show that the data network traffic sources can be bursty. Given the multi-service internet network, an exact traffic characterization is difficult to develop and depends on many factors. Nonetheless, we think that

general models for predicting channel behaviour can still be useful. Sources of internet traffic are known to be bursty. The merge of different traffic might still be bursty, though we cannot guarantee exactly how the traffic will behave because it will highly depend on the scheduling or network access protocol, routing protocol, running applications among other causes. Nevertheless, we can design a general model that intends to capture the burstiness in the channel usage if such burstiness exists. This model is presented in *Section 4.5*.

4.2 Spectrum Usage Models

We have labelled the Geo-Geo and PH-PH channel model as traditional channel usage models because in the literature they are the most often encountered channel usage models in discrete time. To describe the spectrum use, an alternating exponential spectrum occupancy model, which is the continuous time counterpart of a Geo-Geo spectrum occupancy model, has been used in [12, 14, 19, 24, 32, 33]. A Geo-Geo spectrum usage model has been used in [13, 34, 45] to model each PU channel.

In [35], the idle period distribution is exponential while the busy period distribution is general. Erlang distributed busy and idle periods were considered in [24].

Semi-Markov models have been used in [20, 46] to describe the channel behaviour of WLAN systems. Their empirical data showed heavy-tailed idle periods behaviour. The authors in [46], have defined a semi-Markov model as: “*an extension to a continuous time Markov process with separate statistical specification of the transition behavior and sojourn time within each state*”. Besides the usage of mixture distributions (uniform and Pareto) to model the idle sojourn times, the authors in [20, 46], proposed hyper-Erlang and PH distributions to model heavy-tailed idle periods.

Using real occupancy data measurements, in [47], the authors proposed different time-frequency models for the spectrum use. According to [47], the time domain of a low traffic load 2.4GHz ISM-band can be modelled with a geometric distribution during the day. However, during the night times the geometric assumption is no longer valid. A

key fact of their work is that they used a lower sampling rate than in [20, 46] and their measurement setup cannot capture the details as in [20, 46]. Heavy-tailed behaviour were still found in several traces [47]; such as in an uplink GSM with high load. The authors in [47] proposed a lognormal distribution to match heavy-tailed idle or busy periods.

One of the motivations of our work to represent an idle period with two distributions also comes from the work in [20]. To include bursty transmissions in multi-access communication channels the authors decided to extend the state space, which is usually idle and busy, of the channel usage model. A specific example provided includes a WLAN scenario. After each successful data packet transmission in a WLAN, an acknowledgement (ACK) packet must follow. The white space or idle period between the data packet and the ACK can be shorter than usual. Another characteristic of a WLAN is that after a successful transmission, the stations need to perform the standardized back-off algorithm. Thus, the authors from [20] used two distributions to model the idle periods: one distribution for the contention window and another distribution to model when the stations do not have packets to transmit.

The mixture distribution used to model the idle periods in [20] was a uniform distribution for the contention period and generalized Pareto distribution when the stations have no packets to transmit. They also recommend using PH distributions and hyper-Erlang distributions. From our PAP-PH model perspective; these two types of idle periods can also be modelled with the intraplatoon distribution and interplatoon distributions respectively.

Another interesting fact to notice is that the authors in [20] clarified that the idle period between a data packet and its respective ACK can be too short for secondary users' transmissions. This effect can also be modelled using our PAP-PH traffic model with a mean intraplatoon idle period that is too short for an effective secondary user transmission.

In our model, the concept of a platoon can also represent connection or session. For instance, within a connection, a number of consecutive packets are transmitted. The idle length between two consecutive connections is longer than the idle period between packets in the same connection. Similarly, a session consists of a number of connections. The idle

period between consecutive sessions is expected to be longer than idle periods between connections. For example, a typical HTTP session consists of several TCP connection establishments between the user and the web server [1]. A connection within a session could involve a web request, where the browser retrieves a number of objects or packets from the web server. After retrieving the necessary packets, the user can remain idle while viewing the web.

In [5], the tail heaviness of the transmission delay distribution under a reactive spectrum access scheme is investigated. Besides using light tailed distributions, they compared heavy-tailed busy periods and heavy-tailed idle periods. Although the authors concluded that the tail heaviness of the transmission delay distribution under their reactive spectrum access scheme is not affected by heavy-tailed idle periods, we consider another type of reactive spectrum access scheme and we also design a proactive spectrum access scheme that involves traffic prediction.

The discussed investigations in this Section show that the Geo-Geo traffic model does not describe all possible cases. It should also be noticed that the PAP-PH traffic model can be used to model specific arrival processes. But, in the channel usage pattern context, where there may be many merged application packets, the platoon may not have a straightforward interpretation. In fact, packets grouped in a platoon may come from different applications. As mentioned before, the channel usage pattern will depend on many factors that are difficult to foresee. Hence, the above example of connections and sessions are merely an intuitive physical explanation of the bursty nature of data traffic.

The channel usage models that will be compared in this thesis will be described in the subsequent sections.

4.3 The Geo-Geo traffic model

A geometric idle and busy channel model, $\text{Geo}(q, b)$, has the following transition matrix:

$$P_{\text{Geo}} = \begin{bmatrix} (1-q) & q \\ b & (1-b) \end{bmatrix}, \quad (4.3)$$

where the state of the channel goes from idle to busy in a time slot with probability q and from busy to idle with probability b . The mean duration of the busy period is $E[Y] = 1/b$, and the mean duration of the idle period is $E[Z] = 1/q$.

Assuming that the process is stationary, the stationary probability vector denoted by $\boldsymbol{\pi} = [\pi_I \ \pi_B]$ can be obtained from $\boldsymbol{\pi} = \boldsymbol{\pi} P_{\text{Geo}}$ and $\boldsymbol{\pi} \mathbf{1} = 1$. This results in $\pi_I = b/(q+b)$ and $\pi_B = q/(q+b)$.

4.4 The PH-PH traffic model

A PH-PH traffic model consists of: phase type distributed busy periods with distribution defined by $\text{PH}(\boldsymbol{\gamma}, S), m_b$, and phase type distributed idle periods with parameters $\text{PH}(\boldsymbol{\alpha}, T), m_i$. The transition matrix for the idle and busy states is

$$P_{\text{PH}} = \begin{bmatrix} T & \mathbf{T}^0 \boldsymbol{\gamma} \\ \mathbf{S}^0 \boldsymbol{\alpha} & S \end{bmatrix}, \quad (4.4)$$

with $\mathbf{T}^0 = \mathbf{1} - T\mathbf{1}$, $\mathbf{S}^0 = \mathbf{1} - S\mathbf{1}$.

The mean busy interval is $E[Y] = \boldsymbol{\alpha}(I - T)^{-1}\mathbf{1}$, and the mean idle interval is $E[Z] = \boldsymbol{\gamma}(I - S)^{-1}\mathbf{1}$.

4.5 The PAP-PH traffic model

Our proposed traffic model is the PAP-PH model, where idle periods are described by a platoon arrival process and busy periods are phase-type distributed. To capture the characteristics of a PAP-PH traffic model in a single transition matrix, we can design another Markov chain model to merge the busy and idle period's distributions. Assume that the busy period distribution is $\text{PH}(\gamma, S), m_b$, and the idle periods are characterized by a $\text{PAP}(\alpha_1, T_1, \alpha_2, T_2, \delta, D), m$. Remember that the intraplatoon and interplatoon idle period are $\text{PH}(\alpha_1, T_1), m_1$ and $\text{PH}(\alpha_2, T_2), m_2$ respectively, and the number of consecutive intraplatoon periods have another phase type distribution denoted by $\text{PH}(\delta, D), K$. The total number of phases in a PAP is $m = m_2 + m_1 K$.

Consider a discrete time stochastic process $\{X_n\}$ with $n = 0, 1, 2, \dots$, where X_n is a random variable representing the state of the process at time n . The state space is defined as $\mathcal{N} = \{Idle, Busy\}$, with $Idle = \{inter, intra\}$, $inter = \{1, 2, \dots, m_2\}$, $intra = \{1, 2, \dots, Km_1\}$, and $Busy = \{1, 2, \dots, m_b, m_b + 1, m_b + 2, \dots, m_b + Km_b\}$. The cardinality of set \mathcal{N} is $m_2 + Km_1 + m_b + Km_b$. The transition matrix representing the state of the channel (busy or idle) in a given time slot for this Markov Chain is

$$P_{\text{PAP}} = \begin{bmatrix} P_I & P_{IB} \\ P_{BI} & P_B \end{bmatrix}, \quad (4.5)$$

where each element of P_{PAP} is a conditional probability. The substochastic matrix P_I refers to the probability that given that the current state is idle, the next state will also be idle. Hence, matrix P_I describes transitions within the idle period. In a similar way, matrix P_B represents transitions probabilities within the busy period in a given time slot. The matrix P_{BI} denotes transitions from the busy to the idle period, and the matrix P_{IB} denotes transitions from the idle to busy state. Each of these substochastic matrices is

further defined as

$$P_I = \begin{bmatrix} T_2 & \mathbf{0} \\ \mathbf{0} & I_K \otimes T_1 \end{bmatrix}, \quad P_{IB} = \begin{bmatrix} \tau_2 \delta_0 \gamma & \tau_2 \delta \otimes \gamma \\ \mathbf{D}^0 \otimes \tau_1 \gamma & D \otimes \tau_1 \gamma \end{bmatrix},$$

$$P_{BI} = \begin{bmatrix} \mathbf{S}^0 \alpha_2 & \mathbf{0} \\ \mathbf{0} & I_K \otimes \mathbf{S}^0 \alpha_1 \end{bmatrix}, \quad P_B = \begin{bmatrix} S & \mathbf{0} \\ \mathbf{0} & I_K \otimes S \end{bmatrix}.$$

Basic properties for discrete time Markov chain (DTMC) holds, such as the sum of elements in each row of transition matrix P_{PAP} is one, i.e., $P_{\text{PAP}} \mathbf{1} = \mathbf{1}$.

Chapter 5

Statistical Fitting Procedures

An intelligent access scheme can use the previous and present sensing results to build a statistical traffic model which will aid the SUs to make decisions. Hence, the first phase is to estimate the parameters of the statistical model. In this Chapter, the main goal is to explain the statistical fitting procedures that has been used to estimate the parameters of the three channel usage models described in *Chapter 4*. In addition, these fitting procedures will be used in the proactive spectrum access scheme defined in *Chapter 7* to predict future channel behaviour.

To estimate the parameters of the the Geo-Geo traffic model we have used the maximum likelihood parameter estimation described in [24]. This maximum likelihood parameter estimation will be outlined in *Section 5.2*. To estimate the parameters of a PH distribution or a platoon arrival process, we have decided to use the Expectation Maximization (EM) procedure. These two EM algorithms will be discussed in *Section 5.3*.

The parameter estimation results of a low channel usage example is shown in *Section 5.4.1*. In addition, in *Section 5.4.2*, we show the parameters obtained in a high channel usage example. These two examples with their respective estimated parameters will be further used in *Chapter 7*.

5.1 Realization Characteristics

A realization is a sequence of N interarrival times $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$. Assume that this sequence of observations is from a random process $\{X_n\}$. The k th empirical moment can be calculated with [37]

$$E[X_n] = \frac{1}{N} \sum_{i=1}^N (x_i)^k. \quad (5.1)$$

The cumulative distribution function (cdf) of the realization can be computed with [37]

$$F_{X_n}(i) = \Pr(X_n \leq i) = \frac{1}{N} \sum_{i=1}^N \phi(x_i \leq i), \quad (5.2)$$

with

$$\phi(b) = \begin{cases} 1 & : \text{ if } b \text{ is true} \\ 0 & : \text{ otherwise.} \end{cases} \quad (5.3)$$

5.2 Maximum Likelihood Parameter Estimation for the Geo-Geo Traffic Model

Let us assume that we want to describe the data traffic as a Geo-Geo process, with parameters $\text{Geo}(q, b)$, which has been defined in *Section 4.3*. One of the alternatives to estimate the unknown probability density function parameters $\boldsymbol{\theta} = (q, b)$ is to apply the maximum likelihood estimation.

Assume we have historical sensing data of a PU channel. By retrieving the duration of idle and busy periods we obtain a sequence of samples or training sequence. Denote this observed sequence of N samples as \mathbf{x} , *i.e.*, $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$. It can be said that the sequence of samples has been drawn from a joint pmf $p(\mathbf{x}; \boldsymbol{\theta})$ with $\boldsymbol{\theta}$ the unknown parameters. The *likelihood function* is a function of $\boldsymbol{\theta}$ and is defined as [48]

$$L_f(\mathbf{x}; \boldsymbol{\theta}) \equiv p(x_1, x_2, \dots, x_N; \boldsymbol{\theta}). \quad (5.4)$$

Given the first order Markovian property, the likelihood function can be further developed to [24]

$$\begin{aligned} L_f(\mathbf{x}; \boldsymbol{\theta}) &= \Pr(X_1 = x_1; \boldsymbol{\theta}) \prod_{i=2}^N \Pr(X_i = x_i | X_{i-1} = x_{i-1}; \boldsymbol{\theta}), \\ &= \Pr(X_1 = x_1; \boldsymbol{\theta}) (1 - q)^{n_{00}} q^{n_{01}} b^{n_{10}} (1 - b)^{n_{11}}, \end{aligned} \quad (5.5)$$

with n_{01} the total number of transitions from the idle state to the busy state, n_{10} the total number of transitions from the busy state to the idle state. In a similar way, n_{00} and n_{11} are the number of transitions within the idle state and within the busy state respectively.

As stated in [48], the objective of the *maximum likelihood estimation* is to determine the optimal set of parameters $\boldsymbol{\theta}$ that maximizes the likelihood function. Hence, the maximum value corresponds to

$$\hat{\boldsymbol{\theta}}_{\text{ML}} = \arg \max_{\boldsymbol{\theta}} L_f(\mathbf{x}; \boldsymbol{\theta}). \quad (5.6)$$

Because of the monotonicity property of the logarithmic function, we can obtain $\hat{\boldsymbol{\theta}}_{\text{ML}}$ by maximizing the *log-likelihood function* defined as

$$l_f(\mathbf{x}; \boldsymbol{\theta}) = \log L_f(\mathbf{x}; \boldsymbol{\theta}). \quad (5.7)$$

The maximum likelihood $\hat{\boldsymbol{\theta}}_{\text{ML}}$ for the Geo-Geo process can be obtained by calculating the value of $\boldsymbol{\theta}$ where the gradient of the log-likelihood function is zero, that is

$$\frac{\partial l_f(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0,$$

which results in

$$\frac{\partial l_f(\mathbf{x}; \boldsymbol{\theta})}{\partial q} = -\frac{n_{00}}{(1 - q)} + \frac{n_{01}}{q} = 0, \quad (5.8)$$

and

$$\frac{\partial l_f(\mathbf{x}; \boldsymbol{\theta})}{\partial b} = \frac{n_{10}}{b} - \frac{n_{11}}{(1 - b)} = 0. \quad (5.9)$$

Solving (5.8) and (5.9) gives that the maximum likelihood $\hat{\boldsymbol{\theta}}_{\text{ML}} = (\hat{q}, \hat{b})$ are

$$\hat{q} = \frac{n_{01}}{n_{00} + n_{01}}, \quad (5.10)$$

and

$$\hat{b} = \frac{n_{10}}{n_{10} + n_{11}}. \quad (5.11)$$

5.3 EM Algorithm

The parameter estimation of phase type distributions or Markovian arrival processes is an incomplete data problem because the transition phases that lead to an arrival are not observable. Direct application of the maximum likelihood method to the incomplete data might be a very difficult task. It might be easier if we could estimate the true data using the incomplete information and calculate the maximum likelihood based on the estimated true data.

When dealing with missing data problems, the *expectation maximization* (EM) algorithm has been vastly used because of its relative simplicity, as well as its monotonicity property which has been discussed in [49]. Thus, for the estimation of parameters $\boldsymbol{\theta}$, of a phase-type process or a platoon arrival process, we have decided to use the expectation maximization procedure. An EM algorithm for a discrete time PAP has been proposed in [25]. Following the work of [26] and [25], the procedures that have been implemented in this research to estimate the parameters of the distributions will be explained.

As stated in [49], an EM algorithm is an iterated method. In each iteration, the EM algorithm tries to improve the likelihood of the estimates. That is, we want to maximize the conditional expected log-likelihood in each iteration. Let $\boldsymbol{\theta}^{(m)}$ be the estimates obtained from the m -th iteration. First, for $\boldsymbol{\theta}^{(0)}$, we pick a random initial guess. The initial estimates $\boldsymbol{\theta}^{(0)}$ can also be chosen with more refinement if the investigator has more information of the behaviour of the distributions.

With the initial estimates, $\boldsymbol{\theta}^{(0)}$, we execute the following steps of the EM algorithm

until reaching a stopping condition:

1. E-step: given the estimate $\boldsymbol{\theta}^{(m)}$ and the sequence of observations \mathbf{z} , we have to calculate the conditional expected log-likelihood with respect to the complete sequence of observation. Complete sequence of observation means that the state transitions in each time slot is known for every interarrival $z_n \in \mathbf{z}$. Thus, the E-step is given by

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}) &= E_{W|\mathbf{z},\boldsymbol{\theta}^{(m)}}[l_f(W;\boldsymbol{\theta})|Z = \mathbf{z}] \\ &= \sum_{\mathbf{w}} l_f(\mathbf{w};\boldsymbol{\theta})p(\mathbf{w}|\mathbf{z},\boldsymbol{\theta}^{(m)}), \end{aligned} \quad (5.12)$$

where sequence \mathbf{w} denotes the complete state information of the sequence \mathbf{z} .

2. M-step: The $m + 1$ estimation will be obtained from

$$\boldsymbol{\theta}^{(m+1)} = \arg \max_{\boldsymbol{\theta} \in \Omega} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(m)}). \quad (5.13)$$

The stopping condition can be based on the *likelihood ratio* [25], where the EM algorithm is stopped if the following condition holds

$$\frac{L_f(\mathbf{x};\boldsymbol{\theta}^{(m+1)})}{L_f(\mathbf{x};\boldsymbol{\theta}^{(m)})} < 1 + \epsilon. \quad (5.14)$$

with $\epsilon > 0$ a constant set by the investigator according to the wanted accuracy. According to [49], the stopping criteria can also be based on the difference between consecutive estimates, i.e., $|\boldsymbol{\theta}^{(m+1)} - \boldsymbol{\theta}^{(m)}| < \epsilon$ with $\epsilon > 0$. Another option is a set of criteria based on the likelihood function $|l_f^{(m+1)} - l_f^{(m)}| < \epsilon$ for some $\epsilon > 0$.

An EM algorithm for discrete phase type distributions has been proposed in [26]. For completeness, the development of the EM algorithm for PH distributions will be outlined in *Section 5.3.1*. This procedure has been used in the PH-PH traffic model and in the PAP-PH traffic model. Since the busy periods and idle periods distributions are assumed independent, the fitting process for each of these traffic models consists of two independent

processes: a fitting procedure for the idle period distribution and another fitting procedure for the busy period distribution.

In *Section 5.3.2*, the EM algorithm used to estimate the parameters of the platoon arrival process will be explained.

5.3.1 An EM Algorithm for Fitting Phase-Type Distributions

A sequence \mathbf{y} of N samples has been observed, *i.e.*, $\mathbf{y} = \{y_1, y_2, \dots, y_N\}$. Each of these observations is assumed to be phase type distributed with parameters $(\boldsymbol{\gamma}, S)$ with K phases and $\mathbf{S}^0 = \mathbf{1} - S\mathbf{1}$. Using this set of incomplete observations we can use the EM algorithm in the fitting process. All elements of matrix S will be denoted by $s(i, j)$, elements of vector \mathbf{S}^0 are $s^0(i)$ and elements of vector $\boldsymbol{\gamma}$ are $\gamma(i)$ with $1 \leq i, j \leq K$.

The set of parameters that needs to be estimated is $\boldsymbol{\theta} = (\boldsymbol{\gamma}, S, \mathbf{S}^0)$. Given the sequence of independent interarrivals in \mathbf{y} , the incomplete likelihood function is given by [26]

$$L(\mathbf{y}; \boldsymbol{\theta}) = \prod_{n=1}^N \boldsymbol{\gamma} S^{y_n - 1} \mathbf{S}^0. \quad (5.15)$$

For the E-Step we need to calculate the conditional expected log-likelihood of the unknown complete data W given the observed interarrivals \mathbf{y} and the current parameter estimates $\boldsymbol{\theta}^{(m)}$. Assume that the complete information of sequence \mathbf{y} is given in the sequence \mathbf{w} . Then, the complete likelihood of a particular sequence \mathbf{w} with independent phase type distributed samples can be written as

$$L_f(\mathbf{w}; \boldsymbol{\theta}) = \prod_{i=1}^K \gamma(i)^{A(i)} \prod_{i=1}^K \prod_{j=1}^K s(i, j)^{B(i, j)} \prod_{i=1}^K s^0(i)^{C(i)}, \quad (5.16)$$

with $A(i)$ the number of observations that had started on phase i , $B(i, j)$ is the number of transitions from state i to state j , and $C(i)$ is the number of samples that were absorbed in state i . Maximizing the likelihood function $L_f(\mathbf{w}; \boldsymbol{\theta})$ is equivalent to maximizing the

log-likelihood function which is given by

$$l_f(\mathbf{w}; \boldsymbol{\theta}) = \sum_{i=1}^K A(i) \log(\gamma(i)) + \sum_{i=1}^K \sum_{j=1}^K B(i, j) \log(s(i, j)) + \sum_{i=1}^K C(i) \log(s^0(i)). \quad (5.17)$$

We can also write the statistics $A(i)$, $B(i, j)$ and $C(i)$ as

$$A(i) = \sum_{n=1}^N A(i)^n, \quad B(i, j) = \sum_{n=1}^N B(i, j)^n, \quad C(i) = \sum_{n=1}^N C(i)^n,$$

with $A(i)^n \in \{0, 1\}$, $B(i, j)^n, C(i) \in \{0, 1\}$, for $1 \leq i, j \leq K$. For instance, an $A(i)^n = 1$ means that process n started at phase i ; on the other hand, if $A(i)^n = 0$, then the process n did not start at phase i .

Given that the log-likelihood is a linear function of the statistics, in the E-step we need to calculate the expected values of $A(i)$, $B(i, j)$, and $C(i)$ conditioned on \mathbf{y} , which were developed in [26]

$$\begin{aligned} \hat{A}(i) = E_{\boldsymbol{\theta}}[A(i)|Y = \mathbf{y}] &= \sum_{n=1}^N E_{\boldsymbol{\theta}}[A(i)^n | Y_n = y_n] \\ &= \sum_{n=1}^N \frac{e_i^T S^{y_n-1} \mathbf{S}^0 \gamma(i)}{\gamma S^{y_n-1} \mathbf{S}^0}, \end{aligned} \quad (5.18)$$

where e_i denotes a column vector of zeros with a one in position i and superscript T indicates the transpose function. The predictor of $B(i, j)$ is

$$\begin{aligned} \hat{B}(i, j) &= E_{\boldsymbol{\theta}}[B(i, j)|Y = \mathbf{y}] \\ &= \sum_{n=1}^N 1_{\{y_n \geq 2\}} \sum_{k=0}^{y_n-2} \frac{e_j^T S^{(y_n-(k+1)-1)} \mathbf{S}^0 s(i, j) \gamma S^k e_i}{\gamma S^{(y_n-1)} \mathbf{S}^0}. \end{aligned} \quad (5.19)$$

And the predictor of $C(i)$ is

$$\begin{aligned}\hat{C}(i) = E_{\boldsymbol{\theta}}[C(i)|Y = \mathbf{y}] &= \sum_{n=1}^N E_{\boldsymbol{\theta}}[C(i)^n | Y_n = y_n], \\ &= \sum_{n=1}^N \frac{\gamma S^{y_n-1} e_i s^0(i)}{\gamma S^{y_n-1} \mathbf{S}^0}.\end{aligned}\quad (5.20)$$

Using the estimates $\hat{A}(i)$, $\hat{B}(i, j)$, and $\hat{C}(i)$, we want to find the parameter $\boldsymbol{\theta} = (\gamma, S, \mathbf{S}^0)$ such that the expected log-likelihood function defined in Equation (5.17) is maximized. This step is referred to as the M-Step in the EM algorithm. Then, we have the following optimization problem

$$\begin{aligned}\arg \max_{\boldsymbol{\theta}} & E_{W|\mathbf{y}, \boldsymbol{\theta}^{(m)}}(l_f(W; \boldsymbol{\theta}) | Y = \mathbf{y}), \\ \text{subject to} & \sum_{i=1}^K \gamma(i) = 1, \\ & \sum_{j=1}^K s(i, j) + s^0(i) = 1, \quad 1 \leq i \leq K,\end{aligned}$$

which can be converted to an unconstrained optimization problem using the Lagrangian method. Given that the complete log-likelihood is a linear function of the statistics, the expression $E_{W|\mathbf{y}, \boldsymbol{\theta}^{(m)}}(l_f(W; \boldsymbol{\theta}) | Y = \mathbf{y})$ will be written as $l'_f(\mathbf{w}; \boldsymbol{\theta})$ with the statistics replaced by their respective conditional expectations. Then, the Lagrangian is given by

$$LA(\boldsymbol{\theta}) = l'_f(\mathbf{w}; \boldsymbol{\theta}) + \sum_{i=1}^K \left[\lambda_{1i} \left(\sum_{j=1}^K s(i, j) + s^0(i) - 1 \right) \right] + \lambda_2 \left(\sum_{i=1}^K \gamma(i) - 1 \right).$$

We can find the optimal estimates for $\boldsymbol{\theta}$ by maximizing $LA(\boldsymbol{\theta})$ with respect to each parameter in $\boldsymbol{\theta}$. This results in the following estimates developed in [26]. The estimate for $\gamma(i)$ becomes

$$\hat{\gamma}(i) = \frac{\hat{A}(i)}{N}, \quad \forall i \in \{1, 2, \dots, K\}.\quad (5.21)$$

We also get that the estimate for parameter $s^0(i)$ is

$$\hat{s}^0(i) = \frac{\hat{C}(i)}{\sum_{j=1}^K \hat{B}(i, j) + \hat{C}(i)}, \quad \forall i \in \{1, 2, \dots, K\}. \quad (5.22)$$

The estimate for parameter $s(i, j)$ is

$$\hat{s}(i, j) = \frac{\hat{B}(i, j)}{\sum_{s=1}^K \hat{B}(i, s) + \hat{C}(i)}, \quad \forall i, j \in \{1, 2, \dots, K\}. \quad (5.23)$$

The EM algorithm can be summarized as follows:

Step 1: Make an initial estimate for $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\gamma}^{(0)}, S^{(0)}, \boldsymbol{S}^{0(0)})$. Set $m = 0$.

Step 2: Given the sequence of observations \mathbf{y} and the current estimate for $\boldsymbol{\theta}^{(m)}$ calculate the conditional expectation for the statistics using (5.18), (5.19) and (5.20).

Step 3: To get the $m + 1$ estimation of $\boldsymbol{\theta}$ we use the conditional expectations of the statistics from Step 2 and get

$$\gamma(i)^{(m+1)} = \frac{\hat{A}(i)}{N}, \quad (5.24)$$

$$s^0(i)^{(m+1)} = \frac{\hat{C}(i)}{\sum_{j=1}^K \hat{B}(i, j) + \hat{C}(i)}, \quad (5.25)$$

$$s(i, j)^{(m+1)} = \frac{\hat{B}(i, j)}{\sum_{s=1}^K \hat{B}(i, s) + \hat{C}(i)}, \quad (5.26)$$

for all $1 \leq i, j \leq K$.

Step 4: Let $m = m + 1$ and repeat from Step 2 until a stopping criteria is satisfied.

Further details of this algorithm and other maximum likelihood estimators approaches for phase type distributions can be found in [26].

5.3.2 An EM Algorithm for Fitting a Platoon Arrival Process

In this section, the main derivation steps to obtain the EM algorithm proposed by [25] for a discrete time platoon arrival process is shown. The work in [25] does not show the

derivation steps that have been used to obtain their final parameter estimation equations. In addition, the authors focus on deriving the terminating MAP estimates. Hence, we independently will describe the procedures to obtain the parameter estimation equations for the PAP, and we will also highlight some differences from [25] that we have considered.

The procedure to derive the EM algorithm is quite similar to the EM algorithm used for the PH fitting process with some differences. The main difference is that the likelihood function is not a product of independent samples. This causes a more tedious solution to compute the estimated statistics and estimators in an EM algorithm iteration. Towards this goal, let us first assume, that we have an incomplete data sequence of N interarrivals, $\mathbf{z} = \{z_1, z_2, \dots, z_N\}$. This sequence is incomplete because we cannot identify interplatoon and intraplatoon samples and we also cannot distinguish the transition phases to achieve a particular sample. Hence, we are dealing with a hidden data problem.

Define $\boldsymbol{\theta}$ as the set of parameters we want to estimate. For the case where the interplatoon distribution, intraplatoon distribution, and number of interplatoons in a platoon distribution are all phase-types, we have that $\boldsymbol{\theta} = (\delta_0, \boldsymbol{\delta}, D, \mathbf{D}^0, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, T_1, T_2, \boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$. These parameters have been defined in *Section 3.3*. Elements of matrices T_1, T_2, D will be denoted by $t_1(i, j), t_2(i, j)$ and $d(i, j)$ respectively. Elements of vector $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \mathbf{D}^0$ are $\tau_1(i), \tau_2(i)$ and $d_0(i)$ correspondingly. Finally, elements of row vector $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ and $\boldsymbol{\delta}$ are $\alpha_1(i), \alpha_2(i)$ and $\delta(i)$.

Then, the likelihood function of sequence \mathbf{z} given $\boldsymbol{\theta}$ is

$$L(\mathbf{z}; \boldsymbol{\theta}) = \boldsymbol{\pi}' \prod_{n=1}^N (C_0^{z_n-1} C_1) \mathbf{1}_m, \quad (5.27)$$

with matrix C_0 representing a non-arrival event, and matrix C_1 representing an arrival event. Matrices C_0 and C_1 can be obtained from Equation (3.10). We also used the conditional stationary vector $\boldsymbol{\pi}'$ obtained from Equation (3.12). Because we cannot distinguish between an interplatoon and intraplatoon idle duration, we cannot use the product of independent arrivals to compute the likelihood function as in the phase type EM algorithm case.

Say that we have a complete data sequence \mathbf{w} available. That is, the entire phase transitions of the arrivals in sequence \mathbf{w} are known. Then, we can write the complete likelihood function of \mathbf{w} as

$$\begin{aligned}
L_f(\mathbf{w}; \boldsymbol{\theta}) &= \prod_{i=1}^{m_1} \alpha_1(i)^{A_1(i)} \prod_{i=1}^{m_2} \alpha_2(i)^{A_2(i)} \prod_{i=0}^K \delta(i)^{B(i)} \prod_{i=1}^{m_1} \prod_{j=1}^{m_1} t_1(i, j)^{T_1(i, j)} \\
&\times \prod_{i=1}^{m_2} \prod_{j=1}^{m_2} t_2(i, j)^{T_2(i, j)} \prod_{i=1}^K \prod_{j=1}^K d(i, j)^{D(i, j)} \prod_{i=1}^{m_1} \tau_1(i)^{T_1(i)} \\
&\times \prod_{i=1}^{m_2} \tau_2(i)^{T_2(i)} \prod_{i=1}^K d_0(i)^{D(i)}. \tag{5.28}
\end{aligned}$$

where $A_j(i)$ represents the number of observations from sequence \mathbf{w} that start in phase $\alpha_j(i)$, $B(i)$ is the number of observations that are the first intraplatoon period in a platoon and that start at phase $\delta(i)$, $T_k(i, j)$ are the number of transitions from phase i to phase j in matrix T_k that have occurred in sequence \mathbf{w} . In a similar way, $D(i, j)$ is the number of transitions from state i to state j in matrix D , $T_k(i)$ and $D(i)$ denote the number of observations that have been absorbed from state i in their respective distribution.

Each statistic can be calculated by the number of occurrences in each observation.

$$\begin{aligned}
A_j(i) &= \sum_{n=1}^N A_j^n(i), & B(i) &= \sum_{n=1}^N \delta^n(i), & D(i) &= \sum_{n=1}^N D^n(i), \\
T_k(i, j) &= \sum_{n=1}^N T_k^n(i, j), & T_k(i) &= \sum_{n=1}^N T_k^n(i), & D(i, j) &= \sum_{n=1}^N D^n(i, j),
\end{aligned}$$

We want to maximize function $L_f(\mathbf{w}; \boldsymbol{\theta})$, which result will be equivalent to maximizing

the log-likelihood function. The complete log-likelihood function can be written as

$$\begin{aligned}
l_f(\mathbf{w}; \boldsymbol{\theta}) = & \sum_{i=1}^{m_1} A_1(i) \log[\alpha_1(i)] + \sum_{i=1}^{m_2} A_2(i) \log[\alpha_2(i)] + \sum_{i=0}^K B(i) \log[\delta(i)] \\
& + \sum_{i=1}^{m_1} \sum_{j=1}^{m_1} T_1(i, j) \log[t_1(i, j)] + \sum_{i=1}^{m_2} \sum_{j=1}^{m_2} T_2(i, j) \log[t_2(i, j)] \\
& + \sum_{i=1}^K \sum_{j=1}^K D(i, j) \log[d(i, j)] + \sum_{i=1}^{m_1} T_1(i) \log[\tau_1(i)] \\
& + \sum_{i=1}^{m_2} T_2(i) \log[\tau_2(i)] + \sum_{i=1}^K D(i) \log[d_0(i)]. \tag{5.29}
\end{aligned}$$

For the E-step we want to calculate the expected log-likelihood given the current parameter estimates and the sequence of data \mathbf{z} , *i.e.*, $E_{W|z, \theta^{(m)}}[l_f(W; \boldsymbol{\theta}) | Z = \mathbf{z}]$ from Equation (5.12). We can see from Equation (5.29) that the log-likelihood function of w is a linear function of the statistics defined above. Hence, we can compute Equation (5.12) from the conditional expectations of the statistics. The conditional expectation for each statistic S is [26]

$$\hat{S} = E[S | Z = \mathbf{z}] = \sum_{n=1}^N E[S^n | Z = \mathbf{z}] \tag{5.30}$$

with $S \in \{A_j(i), B(i), D^0(i), T_k(i, j), T_k(i), D(i, j)\}$.

In the M-Step we obtain the new estimate $\boldsymbol{\theta}^{(m+1)}$ by maximizing the expected log-likelihood of Equation (5.12). This can be done by applying the method of Lagrange multipliers. Since $E_{W|z, \theta^{(m)}}[l_f(W; \boldsymbol{\theta}) | Z = \mathbf{z}]$ is calculated from the conditional expectations in Equation (5.30), for notation purposes let us write $E_{W|z, \theta^{(m)}}[l_f(W; \boldsymbol{\theta}) | Z = \mathbf{z}]$ as $l'_f(\mathbf{w}; \boldsymbol{\theta})$. Remember that to get the expected log-likelihood, $E_{W|z, \theta^{(m)}}[l_f(W; \boldsymbol{\theta}) | Z = \mathbf{z}]$, we will simply need to replace the statistics by their conditional expectations.

Hence, we have that

$$\begin{aligned}
LA(\boldsymbol{\theta}) &= l'_f(\mathbf{w}; \boldsymbol{\theta}) + \sum_{i=1}^{m_1} \left[\lambda_{1i} \left(1 - \sum_{j=1}^{m_1} t_1(i, j) - \tau_1(i) \right) \right] \\
&+ \sum_{i=1}^{m_2} \left[\lambda_{2i} \left(1 - \sum_{j=1}^{m_2} t_2(i, j) - \tau_2(i) \right) \right] + \lambda_3 \left(1 - \sum_{i=1}^{m_1} \alpha_1(i) \right) \\
&+ \lambda_4 \left(1 - \sum_{i=1}^{m_2} \alpha_2(i) \right) - \lambda_5 \left(1 - \sum_{i=0}^K \delta(i) \right) \\
&+ \sum_{i=1}^K \left[\lambda_{6i} \left(1 - \sum_{j=1}^K d(i, j) - d_0(i) \right) \right]. \tag{5.31}
\end{aligned}$$

Then,

$$\frac{\partial LA}{\partial \alpha_1(i)} = \frac{\hat{A}_1(i)}{\alpha_1(i)} - \lambda_3 = 0 \quad \Rightarrow \quad \lambda_3 = \frac{\hat{A}_1(i)}{\alpha_1(i)}.$$

Using the constraint $\sum_{i=1}^{m_1} \alpha_1(i) = 1$, we have that

$$\sum_{i=1}^{m_1} \alpha_1(i) = 1 \quad \Rightarrow \quad \sum_{i=1}^{m_1} \frac{\hat{A}_1(i)}{\lambda_3} = 1 \quad \Rightarrow \quad \lambda_3 = \sum_{i=1}^{m_1} \hat{A}_1(i).$$

The parameter estimator for $\alpha_1(i)$ is

$$\tilde{\alpha}_1(i) = \frac{\hat{A}_1(i)}{\lambda_3} = \frac{\hat{A}_1(i)}{\sum_{s=1}^{m_1} \hat{A}_1(s)}. \tag{5.32}$$

We repeat the above procedure for all $\alpha_2(i)$ with $1 \leq i \leq m_2$ and for all $\delta(i)$ with $1 \leq i \leq K$.

For the rest of the parameters, we can also derive the respective estimates by calculating the gradients of each pair: $t_1(i, j)$ and $\tau_1(i)$, $t_2(i, j)$ and $\tau_2(i)$, $d(i, j)$ and $d^0(i)$. For instance; for parameters $t_1(i, j)$ and $\tau_1(i)$ we have that

$$\frac{\partial LA}{\partial t_1(i, j)} = \frac{\hat{T}_1(i, j)}{t_1(i, j)} - \lambda_{1i} = 0,$$

$$\frac{\partial LA}{\partial \tau_1(i)} = \frac{\hat{T}_1(i)}{\tau_1(i)} - \lambda_{1i} = 0.$$

Which combined gives

$$\frac{\hat{T}_1(i, j)}{t_1(i, j)} = \frac{\hat{T}_1(i)}{\tau_1(i)} \Rightarrow \sum_{j=1}^{m_1} \hat{T}_1(i, j) \tau_1(i) = \sum_{j=1}^{m_1} \hat{T}_1(i) t_1(i, j).$$

Using the constraint $\sum_{j=1}^{m_1} t_1(i, j) + \tau_1(i) = 1$ results in the following estimates

$$\hat{\tau}_1(i) = \frac{\hat{T}_1(i)}{\sum_{s=1}^{m_1} \hat{T}_1(i, s) + \hat{T}_1(i)},$$

$$\hat{t}_1(i, j) = \frac{\hat{T}_1(i, j)}{\sum_{s=1}^{m_1} \hat{T}_1(i, s) + \hat{T}_1(i)}.$$

In summary, the parameters estimators which we think are equivalent from those specified in [25] are given by

$$\begin{aligned} \hat{\alpha}_1(i) &= \frac{\hat{A}_1(i)}{\sum_{s=1}^{m_1} \hat{A}_1(s)}, & \tilde{\alpha}_2(i) &= \frac{\hat{A}_2(i)}{\sum_{s=1}^{m_2} \hat{A}_2(s)}, & \hat{\delta}(i) &= \frac{\hat{B}(i)}{\sum_{s=0}^K \hat{B}(s)}, \\ \hat{\tau}_1(i) &= \frac{\hat{T}_1(i)}{\sum_{s=1}^{m_1} \hat{T}_1(i, s) + \hat{T}_1(i)}, & \tilde{\tau}_2(i) &= \frac{\hat{T}_2(i)}{\sum_{s=1}^{m_2} \hat{T}_2(i, s) + \hat{T}_2(i)}, \\ \hat{t}_1(i, j) &= \frac{\hat{T}_1(i, j)}{\sum_{s=1}^{m_1} \hat{T}_1(i, s) + \hat{T}_1(i)}, & \tilde{t}_2(i, j) &= \frac{\hat{T}_2(i, j)}{\sum_{s=1}^{m_2} \hat{T}_2(i, s) + \hat{T}_2(i)}, \end{aligned} \quad (5.33)$$

and

$$\hat{d}_0(i) = \frac{\hat{D}(i)}{\sum_{s=1}^{m_2} \hat{D}(i, s) + \hat{D}(i)}, \quad \hat{d}(i, j) = \frac{\hat{D}(i, j)}{\sum_{s=1}^{m_2} \hat{D}(i, s) + \hat{D}(i)}. \quad (5.34)$$

The steps for the EM algorithm for PAP are

Step 1: Obtain an initial estimate for $\theta^{(m)}$ with $m = 0$.

Step 2: Given the sequence of observations \mathbf{z} and the current estimate for $\theta^{(m)}$ calculate the conditional expectation for the statistics using Equation (5.30).

Step 3: Use equations (5.33) and (5.34) to get the $m + 1$ estimation of θ by using the conditional expectations of the statistics from Step 2.

The statistics are calculated based on the conditional expectation given the sequence of observations \mathbf{z} . As in [50], to simplify notations, let us first define the likelihood of observing the sequence of samples up to the $n - 1$ sample of sequence \mathbf{z} as β_{n-1} , which is defined by

$$\beta_{n-1} = \boldsymbol{\pi}' \prod_{s=1}^{n-1} C_0^{z_s-1} C_1. \quad (5.35)$$

In addition, the likelihood that involves the last samples starting from sample $n + 1$ will be denoted as

$$\boldsymbol{\eta}_{n+1} = \prod_{s=n+1}^N C_0^{z_s-1} C_1 \mathbf{1}. \quad (5.36)$$

We can then calculate the statistics estimators with

$$\hat{B}(0) = \sum_{n=1}^N \mathbb{E}(B^n(0)|Z = \mathbf{z}) = \sum_{n=1}^N \frac{\Pr(B^n(0) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.37)$$

and

$$\Pr(B^n(0) = 1 \cap Z = \mathbf{z}) = \beta_{n-1} C_0^{z_n-1} \begin{pmatrix} \boldsymbol{\tau}_2 \\ \mathbf{0} \end{pmatrix} \delta_0(\boldsymbol{\alpha}_2, \mathbf{0}) \boldsymbol{\eta}_{n+1}.$$

Similarly, to estimate $B(i)$ we calculate the following

$$\hat{B}(i) = \sum_{n=1}^N \mathbb{E}(B^n(i)|Z = \mathbf{z}) = \sum_{n=1}^N \frac{\Pr(B^n(i) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.38)$$

with

$$\Pr(B^n(i) = 1 \cap Z = \mathbf{z}) = \beta_{n-1} C_0^{z_n-1} \begin{pmatrix} \boldsymbol{\tau}_2 \\ \mathbf{0} \end{pmatrix} \delta(i) \left(\sum_{h=1}^{m_1} \alpha_1(h) e^{\boldsymbol{\tau}_2^T \mathbf{m}_2 + (j-1)m_1 + h} \right) \boldsymbol{\eta}_{n+1}.$$

We also have that, the statistic estimate $\hat{D}(i)$, which has also been derived in [25], is

$$\hat{D}(i) = \sum_{n=1}^N \mathbb{E}[D^n(i)|Z = \mathbf{z}] = \sum_{n=1}^N \frac{\Pr(D^n(i) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.39)$$

with

$$\Pr(D^n(i) = 1 \cap Z = \mathbf{z}) = \boldsymbol{\beta}_{n-1} C_0^{z_n-1} \left(\sum_{h=1}^{m_1} e_{m_2+(i-1)m_1+h} d_0(i) \tau_1(h) \right) (\boldsymbol{\alpha}_2, \mathbf{0}) \boldsymbol{\eta}_{n+1}.$$

The estimate for $D(i, j)$ is given by

$$\hat{D}(i, j) = \sum_{n=1}^N \mathbb{E}[D^n(i, j) | Z = \mathbf{z}] = \sum_{n=1}^N \frac{\Pr(D^n(i, j) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.40)$$

with

$$\begin{aligned} P(D^n(i, j) = 1 \cap Z = \mathbf{z}) &= \boldsymbol{\pi}' \prod_{s=1}^{n-1} C_0^{z_s-1} C_1 C_0^{z_n-1} \\ &\times \left(\sum_{h=1}^{m_1} \sum_{s=1}^{m_1} e_{m_2+(i-1)m_1+h} d(i, j) \tau_1(h) \alpha_1(s) e_{m_2+(j-1)m_1+s} \right) \prod_{s=n+1}^N C_0^{z_s-1} C_1 \mathbf{1}. \end{aligned}$$

Moreover,

$$\hat{A}_1(i) = \sum_{n=1}^N \mathbb{E}[A_1^n(i) | Z = \mathbf{z}] = \sum_{n=1}^N \frac{\Pr(A_1^n(i) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.41)$$

with

$$\begin{aligned} \Pr(A_1^n(i) = 1 \cap Z = \mathbf{z}) &= \boldsymbol{\beta}_{n-1} C_0^{z_n-1} \left\{ \begin{pmatrix} \boldsymbol{\tau}_2 \\ \mathbf{0} \end{pmatrix} \left(\sum_{h=1}^K \delta(h) \alpha_1(j) e_{m_2+(h-1)m_1+j}^T \right) \right. \\ &\left. + \left(\sum_{r=1}^K \sum_{c=1}^K \sum_{h=1}^{m_1} e_{m_2+(r-1)m_1+h} d(r, c) \tau_1(h) \alpha(i) e_{m_2+(c-1)m_1+i}^T \right) \right\} \boldsymbol{\eta}_{n+1}. \end{aligned}$$

The estimate for statistic $T_1(i)$ is

$$\hat{T}_1(i) = \sum_{n=1}^N \mathbb{E}[T_1^n(i) | Z = \mathbf{z}] = \sum_{n=1}^N \frac{\Pr(T_1^n(i) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.42)$$

with

$$\begin{aligned} \Pr(T_1^n(i) = 1 \cap Z = \mathbf{z}) &= \boldsymbol{\beta}_{n-1} C_0^{z_n-1} \left\{ \left(\sum_{h=1}^K e_{m_2+(h-1)m_1+i} d_0(h) \right) \tau_1(i) (\boldsymbol{\alpha}_2, \mathbf{0}) \right. \\ &\quad \left. + \left(\sum_{h=1}^{m_1} \sum_{r=1}^K \sum_{c=1}^K e_{m_2+(r-1)m_1+i} d(r, c) \tau_1(i) \alpha_1(h) e_{m_2+(c-1)m_1+h}^T \right) \right\} \boldsymbol{\eta}_{n+1}. \end{aligned}$$

Similarly, the estimate for statistic $T_1(i, j)$, which is the same as the work from [25] is

$$\hat{T}_1(i, j) = \sum_{n=1}^N \mathbb{E}[T_1^n(i, j) | Z = \mathbf{z}] = \sum_{n=1}^N \frac{\Pr(T_1^n(i, j) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.43)$$

with

$$\begin{aligned} \Pr(T_1^n(i, j) = 1 \cap Z = \mathbf{z}) &= \boldsymbol{\beta}_{n-2} C_0^{z_n-1-1} \begin{pmatrix} \boldsymbol{\delta} \otimes \boldsymbol{\tau}_2 \boldsymbol{\alpha}_1 \\ D \otimes \boldsymbol{\tau}_1 \boldsymbol{\alpha}_1 \end{pmatrix} \left\{ \sum_{l=1}^{z_n-1} (I \otimes T_1)^{l-1} \right. \\ &\quad \times \left. \left(\sum_{h=1}^K e_{m_2+(h-1)m_1+i} t_1(i, j) e_{m_2+(h-1)m_1+j}^T \right) (I \otimes T_1)^{z_n-l-1} \right\} \\ &\quad \times (D^0 \otimes \boldsymbol{\tau}_1 \boldsymbol{\alpha}_2, D \otimes \boldsymbol{\tau}_1 \boldsymbol{\alpha}_1) \boldsymbol{\eta}_{n+1}. \end{aligned}$$

To obtain $\hat{A}_2(i)$ we have

$$\hat{A}_2(i) = \sum_{n=1}^N \mathbb{E}[A_2^n(i) | Z = \mathbf{z}] = \sum_{n=1}^N \frac{\Pr(A_2^n(i) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.44)$$

with

$$\Pr(A_2^n(i) = 1 \cap Z = \mathbf{z}) = \boldsymbol{\beta}_{n-1} C_0^{z_n-1} \begin{pmatrix} \boldsymbol{\tau}_2 \boldsymbol{\delta}_0 \\ D^0 \otimes \boldsymbol{\tau}_1 \end{pmatrix} \alpha_2(i) e_i^T \boldsymbol{\eta}_{n+1}$$

Analogously, we get $\hat{T}_2(i, j)$ with

$$\hat{T}_2(i, j) = \sum_{n=1}^N \mathbb{E}[T_2^n(i, j) | Z = \mathbf{z}] = \sum_{n=1}^N \frac{\Pr(T_2^n(i, j) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.45)$$

and

$$\begin{aligned} \Pr(T_2^n(i, j) = 1 \cap Z = \mathbf{z}) &= \beta_{n-2} C_0^{z_{n-1}-1} \begin{pmatrix} \tau_2 \delta_0 \boldsymbol{\alpha}_2 \\ \mathbf{D}^0 \otimes \boldsymbol{\tau}_1 \boldsymbol{\alpha}_2 \end{pmatrix} \\ &\times \left(\sum_{l=1}^{z_{n-1}} T_2^{l-1} e_i t_2(i, j) e_j^T T_2^{z_{n-1}-l-1} \right) (\tau_2 \delta_0 \boldsymbol{\alpha}_2, \boldsymbol{\tau}_2 \boldsymbol{\delta} \otimes \boldsymbol{\alpha}_1) \boldsymbol{\eta}_{n+1}. \end{aligned}$$

Finally, the new estimate for $T_2(i)$ is

$$\hat{T}_2(i) = \sum_{n=1}^N \mathbb{E}[T_2^n(i) | Z = \mathbf{z}] = \sum_{n=1}^N \frac{\Pr(T_2^n(i) = 1 \cap Z = \mathbf{z})}{L(\mathbf{z}; \boldsymbol{\theta})}, \quad (5.46)$$

with

$$\Pr(T_2^n(i) = 1 \cap Z = \mathbf{z}) = \beta_{n-1} C_0^{z_{n-1}-1} e_i \tau_2(i) (\delta_0 \boldsymbol{\alpha}_2, \boldsymbol{\delta} \otimes \boldsymbol{\alpha}_1) \boldsymbol{\eta}_{n+1}.$$

5.4 Fitting Process Examples

In this Section, we will show three fitting procedure examples: two PU channels characterized by PAP-PH processes and one PU channel described by a Geo-Geo process. The process describing the PU channel will be referred to as *Original*.

In *Sections 5.4.1* and *5.4.2*, the fitting procedures results will be shown for a low and a high PAP-PH channel usage examples. These two examples will be again used in the reactive and proactive spectrum access in *Chapter 6* and in *Chapter 7*. In *Section 5.4.3*, the fitting procedures will be applied to a PU channel with geometrically distributed idle-busy periods.

In the next examples, we assume that a time slot represents the required length to obtain a sensing sample. The sequence of observations U used for the fitting procedures could have been obtained from historic data stored in a database, or sensing results. We consider that an SU performs perfect sensing, i.e., false alarm and miss-detection probabilities are zero.

Using the original traffic model, a binary sequence was generated, i.e., zeros for idle

and ones for busy. To show the fitting procedure results, each fitting procedure is repeated many times with independent realizations of the PAP-PH process, this is referred to as a *repetition*. A repetition consists of executing the Geo-Geo, PH-PH and PAP-PH fitting procedures once with a sequence of sensing observations U . We chose one repetition as the final parameter estimates.

The first two examples involve a PU channel described by a PAP-PH process where all distributions involved are negative binomial of order 3. This PAP-PH process is characterized by four parameters: $\theta_1, \theta_2, \theta_3, \theta_4$, which correspond to the intraplatoon interval, interplatoon interval, number of intraplatoon in a platoon distribution and busy period respectively. More specifically, the busy period distribution is $\text{PH}(\gamma, S), m_b \sim \text{NB}(3, \theta_4)$, and the parameters of the platoon arrival process for the idle period distribution are $\text{PH}(\alpha_1, T_1), m_1 \sim \text{NB}(3, \theta_1)$, $\text{PH}(\alpha_2, T_2), m_2 \sim \text{NB}(3, \theta_2)$ and $\text{PH}(\delta, D), K \sim \text{NB}(3, \theta_3)$. For these examples, we will use parameters θ_i to describe the original PAP-PH traffic model.

If random variable Y describes the busy duration length and random variable Z characterized the idle duration length, then the *average channel utilization* is defined as [24]

$$\bar{C} = \frac{\text{E}[Y]}{\text{E}[Y] + \text{E}[Z]}. \quad (5.47)$$

5.4.1 Low Channel Usage Example

Low channel usage models are of special interest in opportunistic spectrum access. Many channel-ordering and adaptive sensing frameworks found in the literature give priority to low traffic spectrum bands. In addition, transmissions in longer idle periods are more attractive for secondary users in order to reduce the interference probability with primary users.

Consider a PU channel that behaves as a PAP-PH process with parameters $\theta_1 = 0.2$, $\theta_2 = 0.01$, $\theta_3 = 0.95$ and $\theta_4 = 0.075$. This results in a busy period distribution characterized by parameters $\text{PH}(\gamma, S), m_b \sim \text{NB}(3, 0.075)$, and the parameters of the pla-

toon arrival process for the idle period distribution are $\text{PH}(\boldsymbol{\alpha}_1, T_1), m_1 \sim \text{NB}(3, 0.2)$, $\text{PH}(\boldsymbol{\alpha}_2, T_2), m_2 \sim \text{NB}(3, 0.01)$ and $\text{PH}(\boldsymbol{\delta}, D), K \sim \text{NB}(3, 0.95)$. The statistical characteristics of this original PAP-PH process are: average idle period of $I_{\text{ave}} = 83.54$, average busy period of $B_{\text{ave}} = 40$, $L_{\text{intra}} = 15$, $L_{\text{inter}} = 300$, $\bar{E} = 3.15$ and $\lambda = 0.012$. The average utilization is $\bar{C} = 0.32$.

For each repetition, $U = 13000$ time slots were used in all fitting procedures. The number of idle period interarrivals retrieved from $U = 13000$ time slots was on average $N = 100$. Likewise, the average number of busy period interarrivals obtained from U time slots was also $N = 100$. Both PAP EM algorithm and PH EM algorithm consists of 400 EM iterations.

The initial parameters for all EM algorithms were randomly generated. However, in order to preserve the representation of shorter idle period to intraplatoons and longer idle periods to interplatoons, the reader is advised to select random initial parameters that result in a estimated L_{inter} larger than a estimated L_{intra} . Otherwise, the fitting procedure will most probably converge to a result where the definition of intraplatoon and interplatoon have been exchanged. That is, the interplatoon distribution will characterize intraplatoon idle periods and the intraplatoon distribution will characterize interplatoon idle periods. Although this convergence result is also correct, we use the additional control $L_{\text{inter}} > L_{\text{intra}}$ in the initial parameters to avoid confusion in the previously defined terminology for our platoon arrival process in Section 3.3. If we use this extra control, we are 90% confident that between 95.6% to 100% of the simulations will result in final estimates with a mean interplatoon larger than a mean intraplatoon.

After running for 10 repetitions of $U = 13000$ time slots, the estimated mean idle period and mean busy period for each estimated channel usage model are shown in *Figures 5.1* and *5.2*, respectively. Both of these graphs generally show adequate estimates for the average idle and busy period length as long as the realization used for the fitting procedures accurately describes the underlying idle-busy patterns. Thus, the estimated parameters of each channel usage model heavily depend on the binary realization used in the fitting

Figure 5.1: Estimated mean idle period duration for the low traffic model

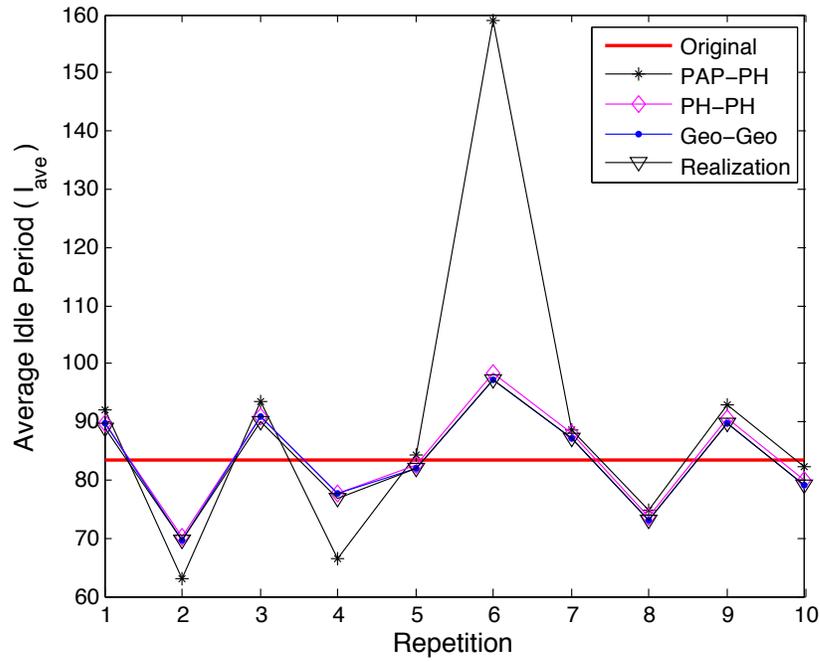
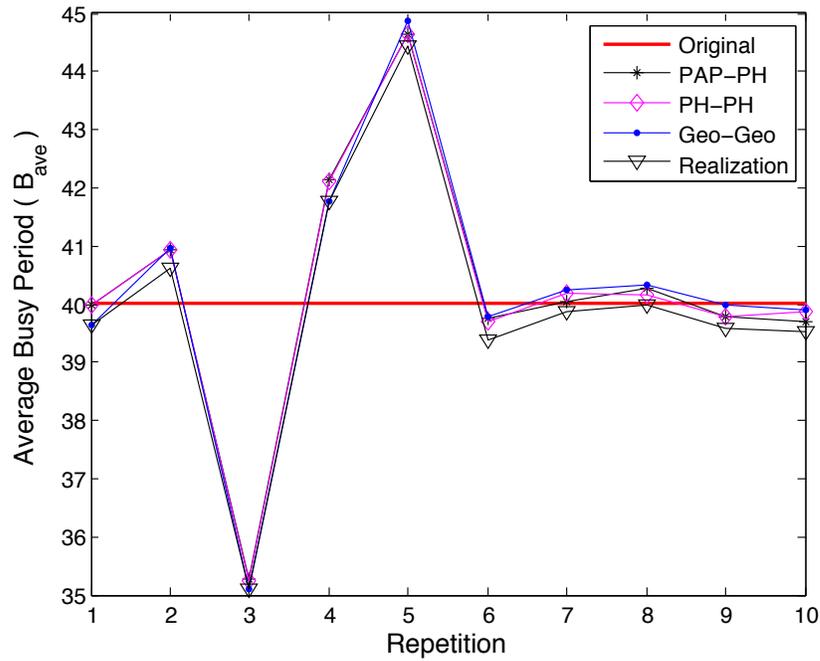


Figure 5.2: Estimated mean busy period duration for the low traffic model



procedures. The EM algorithm for a PAP in repetition 6 converges to an estimate that deviates significantly from the mean idle period as shown in *Figure 5.1*. This can happen because an EM algorithm tries to learn the shape of the distribution and does not focus on

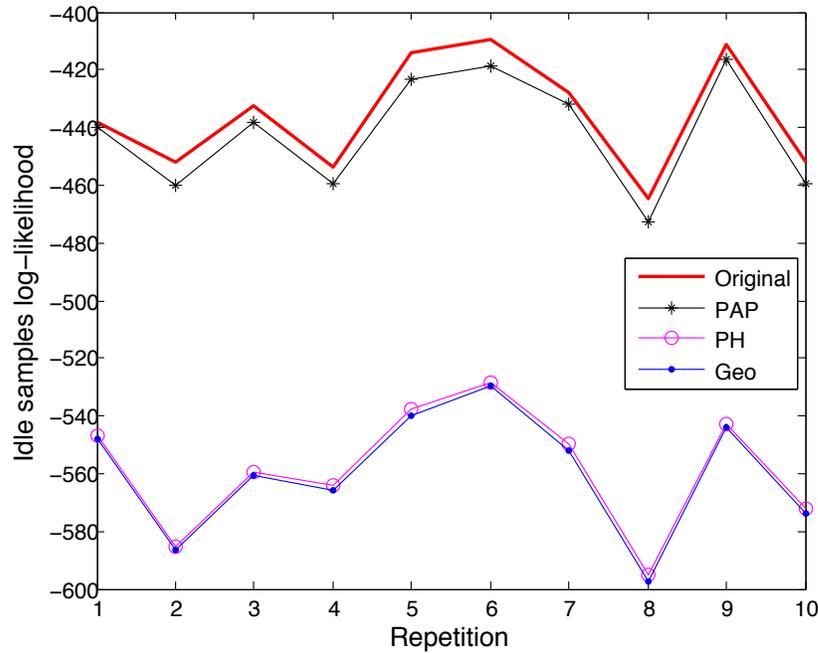
the moment. Hence, it can reach local minimas where the likelihood of the realization with the parameter estimates is high; however, the moment of the estimated model might not closely match the moment of the realization. The best way to avoid this is to restart the EM-algorithm with other initial values.

Consider that sequence \mathbf{z} represents N consecutive idle duration samples or interarrivals. The log-likelihood of the idle distribution parameters estimates, $\log L(\mathbf{z}; \boldsymbol{\theta})$, are compared in *Figure 5.3*. The idle samples likelihood for the Geo-Geo model, $\text{Geo}(q, b)$, is

$$L(\mathbf{z}; \boldsymbol{\theta}) = \prod_{i=1}^N (1 - q)^{z_i - 1} q, \quad (5.48)$$

where each $z_i \in \mathbf{z}$. The idle samples likelihood for the PH-PH and PAP-PH models are given in equations (5.15) and (5.27). As expected, the PAP-PH estimates have a closer likelihood to the original channel usage model when compared to the PH-PH and Geo-Geo estimates.

Figure 5.3: Log-likelihood of idle samples with estimates for the low traffic model



The statistical characteristics of the estimated PAP distribution parameters from repetition 1 to 6 are summarized in *Table 5.1*. What *Table 5.1* shows is that the PAP EM algorithm

is able to capture the bursty nature characteristics of the PU traffic in a reasonable manner. We expect to get better results if more interarrival samples are considered.

Table 5.1: PAP idle samples fitting procedures results for the low traffic model

	Original PAP	Repetition from estimated PAP					
		1	2	3	4	5	6
I_{ave}	83.54	91.93	63.11	93.45	66.54	84.36	158.9
L_{intra}	15	20.67	14.4	16.26	17.6	16.51	15.43
L_{inter}	300	328.2	217.9	331.4	218.22	290.87	615.42
\bar{E}	3.15	3.31	3.17	3.08	3.09	3.04	3.18
λ	0.012	0.011	0.015	0.01	0.015	0.012	0.006

We select the estimates from repetition 5 to be the final parameters estimates of this low traffic PU channel. The repetition selection can be based on the most accurate mean idle and busy repetition. It can also be based on the repetition with higher likelihood. The estimated parameters for the PAP-PH process are:

$$\delta^{400} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad D^{400} = \begin{bmatrix} 0.02 & 0.979 & 0 \\ 0 & 0.022 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$\alpha_1^{400} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad T_1^{400} = \begin{bmatrix} 0.899 & 0 & 0 \\ 0.268 & 0.7318 & 0 \\ 0 & 0.347 & 0.653 \end{bmatrix},$$

$$\alpha_2^{400} = \begin{bmatrix} 0.98 & 0.08 & 0.003 \end{bmatrix}, \quad T_2^{400} = \begin{bmatrix} 0.36 & 0.184 & 0.455 \\ 0.376 & 0.224 & 0.388 \\ 0.345 & 0.309 & 0.344 \end{bmatrix},$$

$$\gamma^{400} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \quad S^{400} = \begin{bmatrix} 0.956 & 0 & 0 \\ 0 & 0.456 & 0.544 \\ 0.081 & 0.348 & 0.57 \end{bmatrix}.$$

The final estimates for the PH-PH channel usage model are:

$$\alpha_{PH}^{(400)} = \begin{bmatrix} 0.8964 & 0 & 0.1036 \end{bmatrix}, \quad T_{PH}^{(400)} = \begin{bmatrix} 0.6475 & 0.0376 & 0.3149 \\ 0.2193 & 0.4803 & 0.2527 \\ 0.0258 & 0.2264 & 0.7472 \end{bmatrix},$$

$$\gamma_{PH}^{(400)} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad S_{PH}^{(400)} = \begin{bmatrix} 0.921 & 0.033 & 0 \\ 0.8829 & 0.1171 & 0 \\ 0 & 0.0477 & 0.9523 \end{bmatrix}.$$

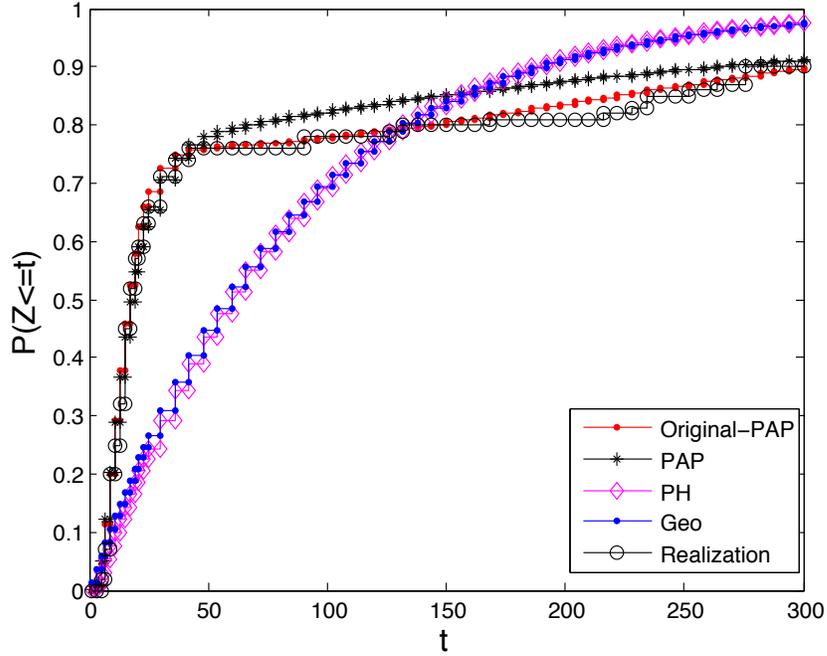
The estimated parameters for the Geo-Geo model are:

$$P_{Geo} = \begin{bmatrix} 0.9878 & 0.0122 \\ 0.0223 & 0.9777 \end{bmatrix}.$$

In *Figure 5.4*, the cumulative distribution function of the estimated idle period distributions from repetition 5 is shown. Besides the three estimated distributions, we have included the original PAP cdf and the empirical realization cdf used in the fitting procedures. The empirical realization cdf has been calculated using equation (5.2). For the cdf plot to be legible we have plotted the cdf versus a subset of the natural numbers. From *Figure 5.4* we can see that the estimated geometric and phase-type distribution functions differ in a noticeable way from the platoon arrival process.

Through different channel usage simulations, we have discovered that sometimes the PH-PH estimates also have the capacity of achieving close likelihoods as the PAP-PH estimates. This means that for some data traffic patterns, the PH-PH model with adequate parameters can approximate the cumulative distribution function of a PAP-PH model in a reasonable way.

Figure 5.4: Cumulative distribution function for the low traffic model



5.4.2 High Channel Usage Example

It is interesting to see the performance of the reactive and proactive spectrum access schemes if the channel utilization is high. Similarly as in the previous Section, consider a PU channel described by a PAP-PH process where all distributions involved are negative binomial of order 3 with parameters $\theta_1 = 0.6$, $\theta_2 = 0.03$, $\theta_3 = 0.2$ and $\theta_4 = 0.075$. The statistical characteristics of the original PAP-PH process are: average idle period of $I_{\text{ave}} = 10.937$, average busy period of $B_{\text{ave}} = 40$, $L_{\text{intra}} = 5$, $L_{\text{inter}} = 100$, $\bar{E} = 15$ and $\lambda = 0.091$. This PAP-PH channel usage model has an average utilization of $\bar{C} = 0.785$.

Similarly as the low traffic model example, 10 repetitions of each fitting procedures was performed. Each repetition used $U = 9000$ time slots which resulted on an average number of $N = 180$ interarrivals. The estimated mean idle period for each estimated channel usage model is shown in *Figure 5.5*. The log-likelihood of the realization with the idle period parameter estimates of each traffic model is plotted in *Figure 5.6*. From these three figures and *Table 5.2*, we can conclude that the estimates of the high traffic example are better than those of the low traffic example in the previous Section. Lower variance in the estimated

mean idle and busy periods is achieved by considering a higher number of interarrivals in the fitting procedures.

Figure 5.5: Estimated mean idle period duration for the high traffic model

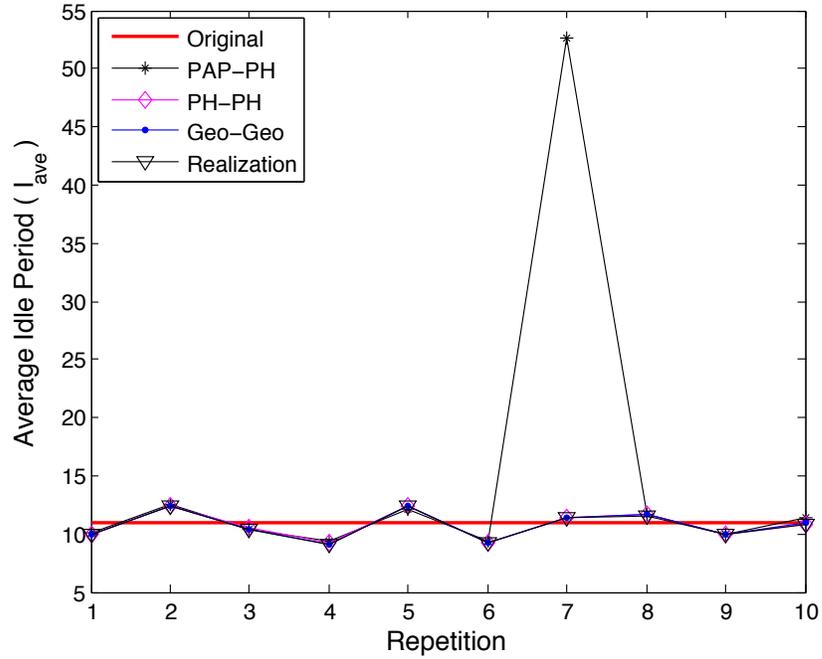


Figure 5.6: Log-likelihood of idle samples with estimates for the high traffic model

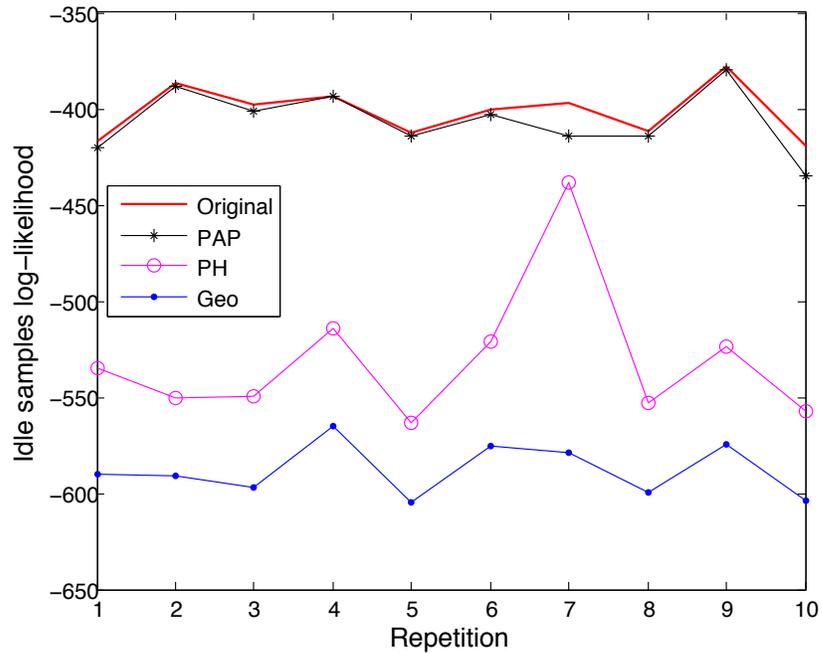
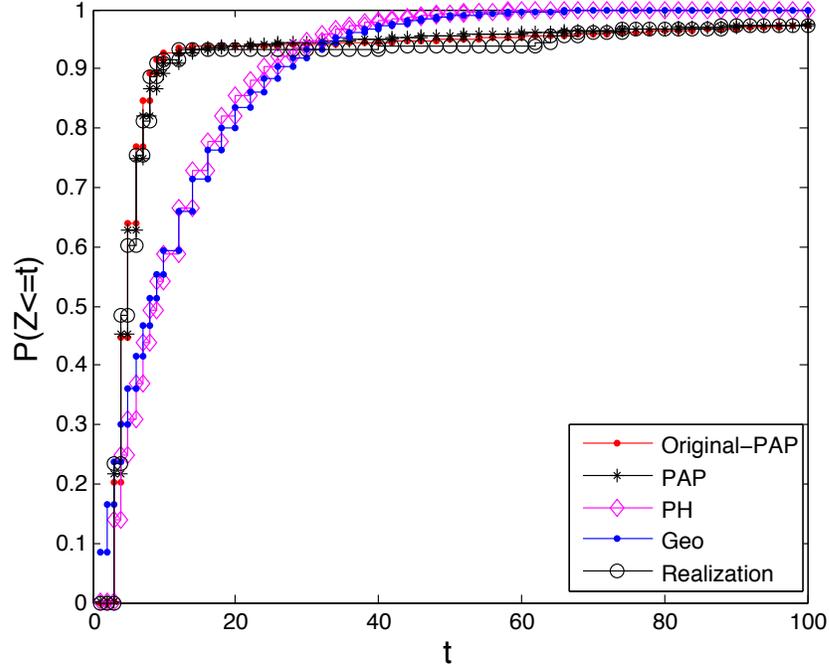


Table 5.2: PAP idle samples fitting procedures results for the high traffic model

	Original PAP	Repetition from estimated PAP					
		3	4	5	6	7	8
I_{ave}	10.937	10.39	9.34	12.16	9.33	52.69	11.69
L_{intra}	5	4.99	5.16	5.16	5.04	5.04	5.08
L_{inter}	100	105	72.5	106.53	71.25	768.02	93.43
\bar{E}	15	17.69	15.11	13.49	14.43	15.01	12.36
λ	0.091	0.09	0.1	0.08	0.1	0.01	0.085

The final parameters estimates for the high traffic PU channel, which will be used in *Chapter 7*, are those obtained from repetition 8. In *Figure 5.7*, the cumulative distribution function of the idle period distributions estimates are shown. The high traffic PAP-PH cdf estimates have a closer resemble than the low traffic PAP-PH cdf estimates in *Figure 5.4*.

Figure 5.7: Cumulative distribution function for the high traffic model



The final estimated parameters for the PAP-PH process are:

$$\delta^{400} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad D^{400} = \begin{bmatrix} 0.779 & 0.2209 & 0 \\ 0 & 0.7315 & 0.2685 \\ 0 & 0 & 0.7569 \end{bmatrix},$$

$$\alpha_1^{400} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad T_1^{400} = \begin{bmatrix} 0.586 & 0 & 0 \\ 0.804 & 0.195 & 0 \\ 0 & 0.7017 & 0.298 \end{bmatrix},$$

$$\alpha_2^{400} = \begin{bmatrix} 0.988 & 0.0001 & 0.0112 \end{bmatrix}, \quad T_2^{400} = \begin{bmatrix} 0.5107 & 0.343 & 0.146 \\ 0.298 & 0.442 & 0.237 \\ 0.219 & 0.463 & 0.309 \end{bmatrix},$$

$$\gamma^{400} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad S^{400} = \begin{bmatrix} 0.831 & 0.168 & 0 \\ 0.0923 & 0.806 & 0.1 \\ 0 & 0 & 0.946 \end{bmatrix}.$$

5.4.3 A Geo-Geo PU channel

From the previous examples, the PAP-PH parameter estimates outperforms the Geo-Geo estimates when the PU channel is characterized by a PAP-PH process. Nevertheless, it is important to analyse the PAP-PH estimates when the PU channel is instead described by a Geo-Geo process. Hence, in this Section, we will show the estimation results of the three traffic models when a PU channel with geometrically distributed busy-idle periods is considered.

The mean idle period is $I_{\text{ave}} = 40$ and the mean busy period of $B_{\text{ave}} = 40$. The estimated mean idle period for the three distributions in 100 repetitions is shown in *Figure 5.8*. In *Figure 5.9*, we show an example of the estimated cumulative distribution functions of a repetition; in this case, repetition 90. As expected, it can be observed that the cdf of the platoon arrival process and phase type distribution closely match the original geometric cdf.

The statistical characteristics of some fitting procedures are outlined in *Table 5.3*. The EM-algorithm, used to estimate the parameters of the PAP, tries to find the maximum likeli-

Figure 5.8: Estimated mean idle period duration for the high traffic model

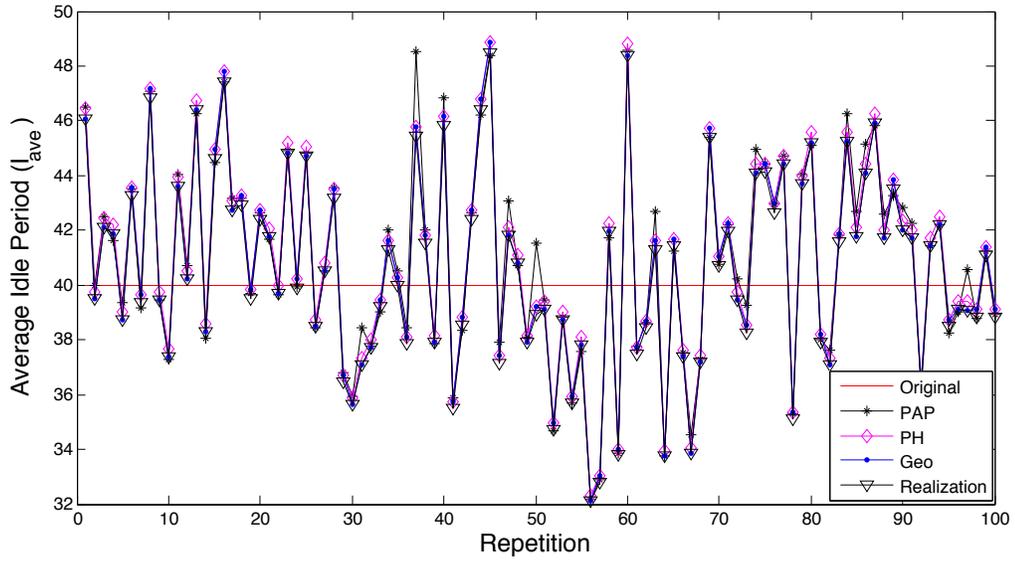
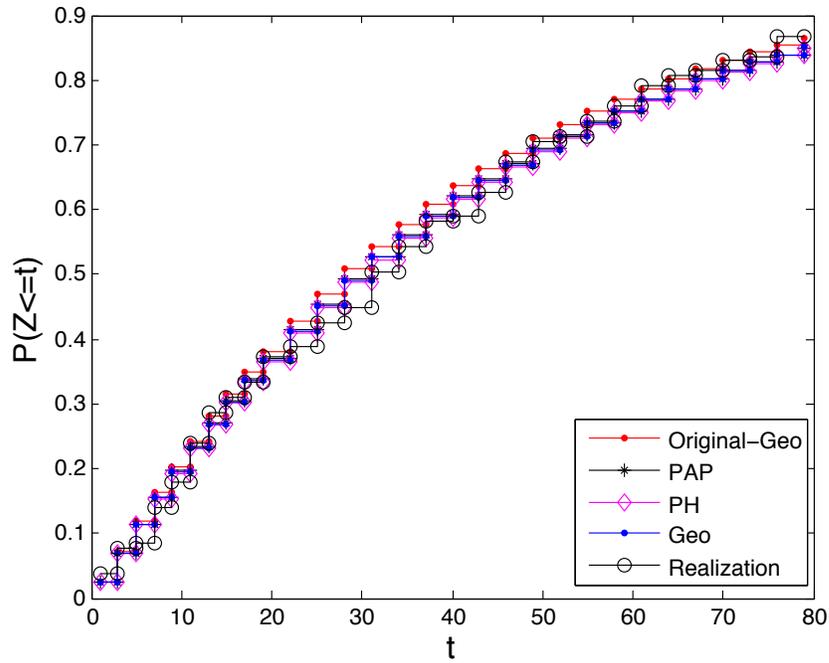


Figure 5.9: Cumulative distribution function for the Geo-Geo PU channel



hood of the idle samples. Hence, the different statistical measures of the PAP estimates are because there are different structures to represent a geometric distribution with a platoon arrival process.

Table 5.3: Fitting procedures results for the Geo-Geo PU channel

	Original Geo-Geo	Repetition N_0 from estimated PAP-PH					
		40	50	60	70	80	90
I_{ave}	40	46.8	41.5	48.6	40.7	45.1	42.8
B_{ave}	40	39.5	39.05	40.34	47.7	421	42.8
L_{intra}		46.06	37.4	36.7	37	35.3	41
L_{inter}		48.25	131.5	76.9	67	69.9	140.5
\bar{E}		1.88	21.82	2.4	7.02	2.55	54.9
λ		0.02	0.02	0.02	0.02	0.02	0.02

5.5 General comments of the fitting process

A computer has limits in the representation of a floating point data. Using the MATLAB software, the range for a double is: $2.22507e-308$ to $1.79769e+308$. This means that the PAP EM algorithm cannot be applied when the likelihood of the realization with the PAP estimates exceeds the MATLAB double representation limit. The likelihood decreases as the number of samples and values of each sample increases. Hence, the major limitation of this algorithm is the number of samples that can be used in the PAP fitting process. An advantage of using a continuous PAP is that the interarrivals can be scaled to smaller values, which will allow more samples to be used in the EM algorithm. For instance; the author in [37] used 10^3 to 10^4 samples in his EM algorithm for a continuous MAP. Each sample had a value between 0 to 1.

Although the EM algorithm for discrete PAP only considers few hundreds of samples, it still yields a reasonable approximation of the original distribution. In addition, the complexity of the EM algorithm increases with the number of samples. In a cognitive radio network, the parameters may need to be adapted frequently. Hence, efficiency in the estimation procedures is necessary.

Chapter 6

A Reactive Spectrum Access Scheme

We have seen in *Chapter 4* that many research works in their investigation assume a Geo-Geo channel usage model for mathematical tractability. However, it is important to analyze the potential effects of this assumption if the channel is behaving in a bursty manner. Hence, in this Chapter we will first design a simple reactive channel access scheme. As explained in *Section 2.4.2*, a reactive spectrum access implies that an SU's action of when to transmit is solely based on the latest sensing information and no historical sensing data is used.

Using the reactive channel access scheme, two performance measures will be analytically derived using Markov chain models. These two measures are the average transmission delay and the average number of time slots between collisions. The analytical results will be compared against simulations to test their accuracies. We investigate the different performance measures an investigator will obtain if a Geo-Geo channel usage model assumption is used over a PU channel with bursty spectrum utilization.

Assuming a PAP-PH primary user channel, our results show that as the channel usage increases, the accuracy of the performance measures derived with the Geo-Geo channel usage model decreases.

6.1 The System Model

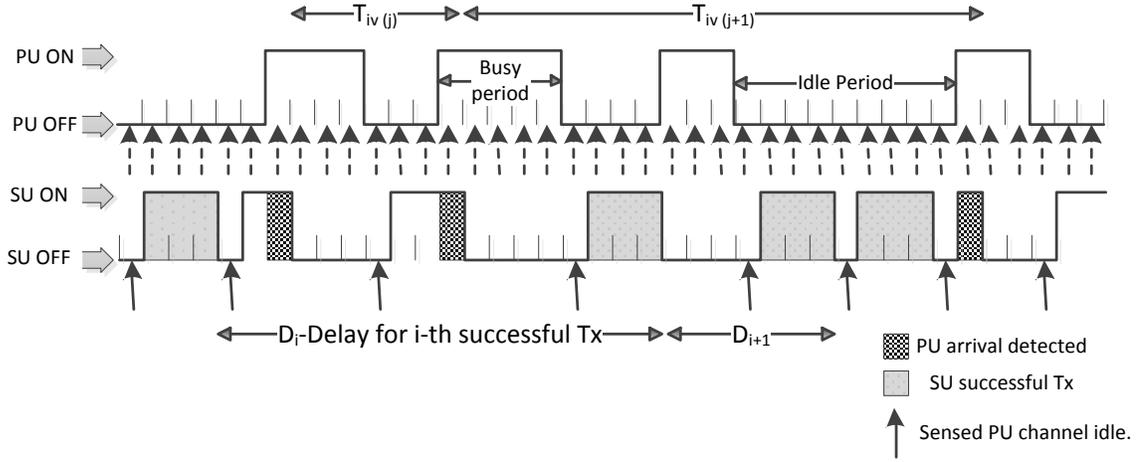
Consider a single PU channel that consists of idle and busy periods (OFF/ON periods) characterized by a bursty PAP-PH process. A discrete time secondary network access scheme is considered, where a time slot is defined as the required length for an SU to perform sensing. It is assumed that in a given time slot; a secondary user is either sensing or transmitting. The listen-before-talk strategy is implemented in this scheme to avoid collision with primary users.

In this access scheme, an SU always needs L consecutive idle time slots for a successful transmission. In other words, a packet length is L . If the channel is sensed idle, the SU starts its transmission in the next time slot. One idle time slot is always used for sensing before a transmission. It is further assumed that the SU always has packets to transmit and continuously stays on the same PU channel. We assume that during transmission, the SU can detect a PU arrival.

The SU always aims to transmit for L time slots; however, it immediately stops if it detects a primary user arrival. If the SU is still transmitting and it detects that the primary user has returned, a collision occurs that lasts for a single time slot. This scheme can be better understood by looking at *Figure 6.1*, where $L = 3$. For a successful transmission the SU needs a total of $L + 1$ idle time slots, one used for sensing and L consecutive time slots for transmission. This SU access model is an optimistic one because an SU may not be able to sense all the time, and it may not be able to detect the PU arrival while transmitting. Not being able to detect a primary user arrival while transmitting will inherently cause more interference to primary users.

In the next sections, we will calculate two performance measures: average transmission delay $E[D_L]$, and the average number of time slots between collisions $E[T_{iv}]$. The *average transmission delay* is defined as the average number of time slots an SU spends for a successful transmission. A given transmission delay includes the time slots used for sensing and time slots that resulted in failed transmissions. As the name suggests, the *average*

Figure 6.1: Reactive spectrum access model with $L = 3$



number of time slots between collisions is the number of time slots that an SU takes on average to collide with a PU given that a collision have just occurred. An example of these two measures can be observed in *Figure. 6.1*, where $T_{iv}(j)$ denotes the number of time slots between collision j and collision $j - 1$, and D_i denotes the delay for the i th successful transmission.

Before defining the Markov chain models used to derive the performance measures, we will first define the standard notation for a channel usage model. The behaviour of the stationary channel can be captured by a transition matrix P of the following form

$$P = \begin{bmatrix} P_I & P_{IB} \\ P_{BI} & P_B \end{bmatrix}. \quad (6.1)$$

Recall that matrices P_I and P_B describe transitions within the idle and busy periods respectively. The matrix P_{BI} denotes transition probabilities from the busy state to idle state, and matrix P_{IB} denotes transitions from the idle state to busy state. Matrix P can be obtained from any discrete time traffic model such as those presented in *Chapter 4*.

6.2 Average Transmission Delay

To calculate the average transmission delay $E[D_L]$ we will design a discrete time absorbing Markov chain to obtain the distribution of the number of time slots required for a successful transmission. Given that the SU has just successfully transmitted a packet, the transmission delay consists of the number of time slots required for the SU to encounter $L + 1$ idle time slots. Therefore, we will design a discrete time absorbing Markov chain that can keep track when successive $L + 1$ idle time slots have been encountered by the SU. In other words, absorption occurs when $L + 1$ idle slots has been encountered. This results in the following state space

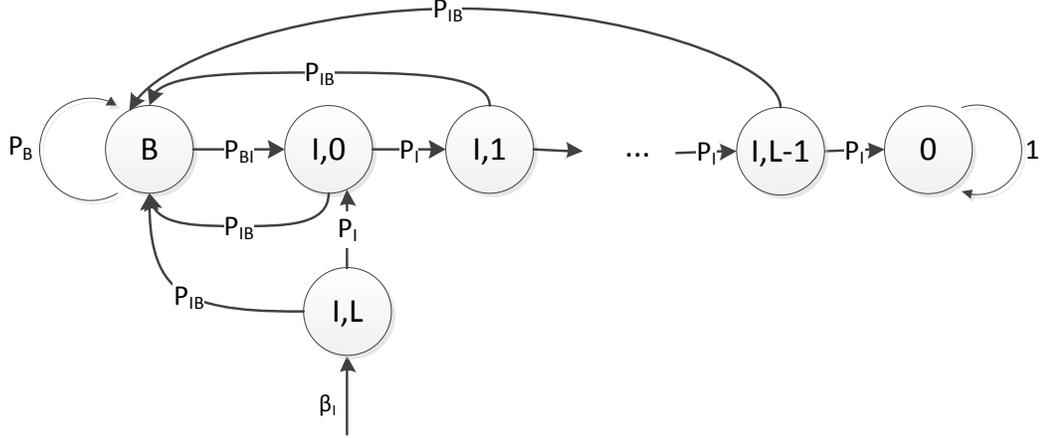
$$\Delta = \{0 \cup B \cup (I, i); 0 \leq i \leq L\},$$

where state $\{0\}$ is the absorbing state, state B represents the channel is busy and can be further decomposed to $B = \{(B, j); 1 \leq j \leq n_B\}$ if the busy state has phases. State (I, i) denotes that the channel is idle. State $(I, 0)$ means that the SU has sensed the channel as idle and the SU will start transmitting in the next time slot. For $1 \leq i \leq L$, state (I, i) signifies that the SU has transmitted for i consecutive idle time slots. Each idle state can further be divided into $(I, i) = \{(I, i, j); 1 \leq j \leq n_I\}$. Constants n_B and n_I are the number of phases transition matrices P_B and P_I have. For the case of a PAP-PH channel-usage model which has been explained in *Section 4.5* we have that $n_B = m_b + Km_b$ and $n_I = m$. For the Geo-Geo traffic model explained in *Section 4.3* we have that $n_B = 1$ and $n_I = 1$.

The state transition diagram is shown in *Figure 6.2*. The process starts at state (I, L) because a successful SU transmission has just occurred. It can be noticed that the distribution of the transmission delay D_L can be represented by a phase-type distribution with parameter $\text{PH}(\beta, V), n_V$ where $n_V = n_B + (L + 1)n_I$. Moreover, the transition matrix of the Markov chain is

$$P_D = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{v} & V \end{bmatrix}, \quad (6.2)$$

Figure 6.2: State Diagram to compute the average transmission delay $E[D_L]$



with

$$V = \begin{bmatrix} P_B & P_{BI} & 0 & 0 & \cdots & 0 & 0 \\ P_{IB} & 0 & P_I & 0 & \cdots & 0 & 0 \\ P_{IB} & 0 & 0 & P_I & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ P_{IB} & 0 & 0 & 0 & \ddots & P_I & 0 \\ P_{IB} & 0 & 0 & 0 & \cdots & 0 & 0 \\ P_{IB} & P_I & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}. \quad (6.3)$$

In this reactive access scheme, an SU continuously monitors the same PU channel regardless of a collision or a successful transmission. Hence, we define the start probability vector $\beta = \begin{bmatrix} 0 & \cdots & 0 & \beta_I \end{bmatrix}$ such that β_I represents the stationary conditional probabilities of finding the channel in the idle phases after a successful transmission. The stationary distribution of finding the PU channel in the idle and busy states in the long run is $\pi = \begin{bmatrix} \pi_I & \pi_B \end{bmatrix}$ which satisfies the following equations

$$\pi = \pi P, \quad \text{and} \quad \pi \mathbf{1} = 1.$$

with P defined in Equation (6.1). Then, the element β_I of the start vector probability for

the transmission delay phase-type distribution is

$$\boldsymbol{\beta}_I = \frac{\tilde{\boldsymbol{\pi}}_I P_I^L}{\tilde{\boldsymbol{\pi}}_I P_I^L \mathbf{1}}, \quad (6.4)$$

with $\tilde{\boldsymbol{\pi}}_I$ the stationary conditional probability of finding the process in each of the idle phases given that the channel is idle, i.e.,

$$\tilde{\boldsymbol{\pi}}_I = \frac{\boldsymbol{\pi}_I}{\boldsymbol{\pi}_I \mathbf{1}}. \quad (6.5)$$

The start probability vector $\boldsymbol{\beta}$ always begins in an idle state because a successful transmission implies that the channel remained idle for $L + 1$ time slots.

The probability that the transmission delay is i is given by

$$\Pr(D_L = i) = \boldsymbol{\beta} V^{i-1} \mathbf{v}, \quad i \geq 1, \quad (6.6)$$

with $\mathbf{v} = \mathbf{1} - V\mathbf{1}$. The average transmission delay for L idle time slots requirement is thus

$$\mathbb{E}[D_L] = \boldsymbol{\beta} (I - V)^{-1} \mathbf{1}. \quad (6.7)$$

6.3 Average Number of Time Slots between Collisions

The average number of time slots between successive collisions will be denoted as $\mathbb{E}[T_{iv}]$. In a similar way as in the preceding section, we can design a discrete time absorbing Markov chain to obtain the distribution of the number of time slots T_{iv} between successive interfered PU time slots. Recall that an interference scenario is caused when the SU starts its transmission given that the channel was found idle and the PU returns at some point of the L consecutive time slots.

The aim in the discrete time absorbing Markov chain design is to keep track of the probabilities that given that an SU starts its transmission, a PU returns before the completion of

the SU transmission. We can define a Markov chain with the following state space

$$\Delta = \{0 \cup B \cup (I, i); 1 \leq i \leq L + 1\},$$

where state $\{0\}$ is the absorbing state, the states $B = \{(B, j); 1 \leq j \leq n_B\}$ represent that the PU channel is busy and $(I, i) = \{(I, i, j); 1 \leq j \leq n_I\}$ represent idle states. More specifically, when the system is in state $(I, 1)$ the SU has sensed that the channel is idle and will start transmitting in the next time slot. For $2 \leq i \leq L + 1$, state (I, i) means that the SU has transmitted for $i - 1$ time slots and the channel has remained idle for i time slots. The number of phases of the busy and idle states are n_B and n_I respectively. As mentioned in the previous section, for the case of a PAP-PH channel-usage model $n_B = m_b + Km_b$ and in the Geo-Geo channel-usage model $n_B = 1$. In a similar way, the variable $n_I = m$ for the PAP-PH channel-usage model, and $n_I = 1$ for the Geo-Geo channel-usage model.

The state transition diagram is shown in *Figure 6.3*. It can be noticed that the distribution of T_{iv} can be represented by a phase-type distribution with parameter $\text{PH}(\boldsymbol{\rho}, U), n_U$ where $n_U = n_B + (L + 1)n_I$. Moreover, the transition matrix of the chain is

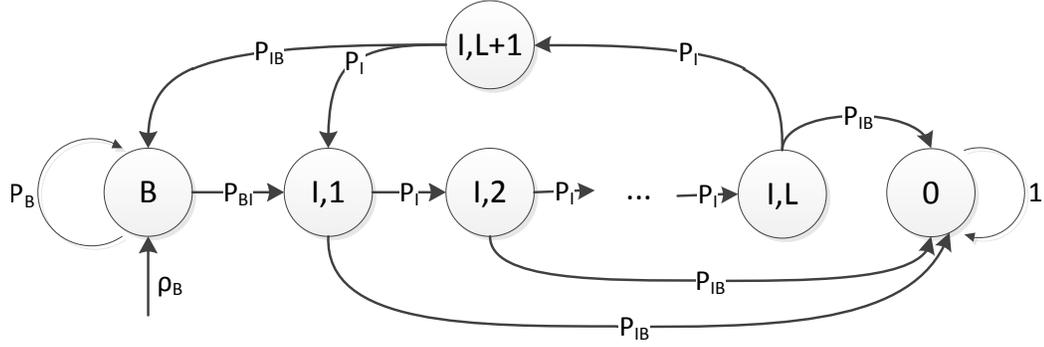
$$P_{IV} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{u} & U \end{bmatrix}, \quad (6.8)$$

with

$$U = \begin{bmatrix} P_B & P_{BI} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & P_I & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & P_I & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & P_I \\ P_{IB} & P_I & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}. \quad (6.9)$$

Using arguments identical to the derivation of the average transmission delay in *Section 6.2*, we can get the average number of time slots required to cause interference to the PU.

Figure 6.3: State Diagram to obtain the average number of time slots between collisions $E[T_{iw}]$



Let us denote the stationary probabilities of finding the PU channel in the idle and busy states by π_I and π_B , respectively. Further define the stationary probability vector of the PU channel as $\pi = \begin{bmatrix} \pi_I & \pi_B \end{bmatrix}$. Assuming that P is the transition matrix defined in Equation (6.1), the stationary vector π can be computed as

$$\pi = \pi P, \quad \text{and} \quad \pi \mathbf{1} = 1.$$

The start probability vector is given by $\rho = \begin{bmatrix} \rho_B & 0 & \dots & 0 \end{bmatrix}$ with ρ_B being the conditional probability of starting in the busy period given that an interference has occurred

$$\rho_B = \frac{\tilde{\pi}_I P_{IB}}{\tilde{\pi}_I P_{IB} \mathbf{1}}, \quad (6.10)$$

and $\tilde{\pi}_I$ is the normalized probability that the PU channel is in the idle state in the long run obtained from Equation 6.5. The phase type distribution, $\text{PH}(\rho, U), n_U$, always starts at a busy phase because random variable T_{iw} denotes the number of time slots between interfered PU time slots.

The probability that the collision with a PU occurs in i time slots is given by

$$\Pr(T_{iw} = i) = \rho U^{i-1} \mathbf{u}, \quad i \geq 1, \quad (6.11)$$

with $\mathbf{u} = \mathbf{1} - U\mathbf{1}$.

The average number of time slots required for a collision for L idle time slots transmission requirement is thus

$$E[T_{iv}] = \boldsymbol{\rho}(I - U)^{-1}\mathbf{1}. \quad (6.12)$$

6.4 Numerical and Simulation Results

Computer simulations of the reactive spectrum access scheme will also be performed to evaluate the analytical formulas (6.7) and (6.12). In this simulation, no PAP-PH fitting process is performed because in a reactive access approach the SU does not use a channel usage model to determine when to transmit.

Say that the PU channel is assumed to behave as a PAP-PH process. If the investigator uses a Geo-Geo model over the bursty PU channel, the estimated Geo-Geo model will have the same average idle period and same average busy period as the PAP-PH model. Hence, in the rest of this Chapter, we will execute the reactive scheme for a PU channel characterized by the estimated Geo-Geo model, and we will also compute the reactive scheme for a PAP-PH PU channel. These two traffic models have the same mean idle period and the same mean busy period. Besides analyzing the performance of the proposed reactive scheme, our objective is to compare the performance measures differences of a Geo-Geo and PAP-PH traffic models. We have generated a long independent realization of each channel model using their respective P matrix from Equation (6.1).

Assume that $L + 1$ idle time slots are needed for a single SU transmission. In a simulation run or realization, $D_i(L)$ denotes the delay the SU experiences to transmits the i th packet of length L . The sample mean transmission delay is given by

$$\bar{D}(L) = \frac{1}{N_L} \sum_{i=1}^{N_L} D_i(L), \quad (6.13)$$

with N_L the number of successful SU transmissions in the simulation.

Simulations are also performed to test the accuracy of the absorbing Markov chain to

obtain the measure $E[T_{iv}]$. In this case, if $T_{iv(j)}(L)$ denotes the number of time slots for the j th collision, then the sample T_{iv} average is

$$\bar{T}_{iv}(L) = \frac{1}{C_L} \sum_{j=1}^{C_L} T_{iv(j)}(L), \quad (6.14)$$

with C_L the total number of collisions.

6.4.1 Simulation Results with Low Traffic Example

Consider a PAP-PH and Geo-Geo traffic model with mean idle period $I_{ave} = 83.54$ and mean busy period $B_{ave} = 40$. As defined in *Section 5.4.1*, The PAP-PH model parameters are $\theta_1 = 0.2$, $\theta_2 = 0.01$, $\theta_3 = 0.95$ and $\theta_4 = 0.075$. The corresponding Geo-Geo transition matrix P , which has been defined in Equation (6.1), is

$$P_{Geo} = \begin{bmatrix} 0.988 & 0.012 \\ 0.025 & 0.975 \end{bmatrix}. \quad (6.15)$$

Using these two traffic models, the reactive scheme has been applied to a realization of five million time slots. The analytical and simulation results are compared in *Figures 6.4* and *6.5*. The average transmission delay versus L is shown in *Figure 6.4*. As expected, the average transmission delay increases with the transmission length L . As L gets larger, the SU experiences more difficulty to achieve a successful transmission. For a channel with low traffic the average transmission delay is not seriously affected by a Geo-Geo model assumption. In this channel with low traffic, the simulation results obtained with the variation of L reflects that almost no difference is observed between the Geo-Geo and PAP-PH model.

The average number of time slots between collisions for both of the traffic models are shown in *Figure 6.5*. The analytical formulas to derive $E[T_{iv}]$ have better accuracy than the analytical formulas used to obtain $E[D_L]$. It can be seen from *Figure 6.5* that $E[T_{iv}]$ decreases with longer transmission length L . This means that higher probability of

Figure 6.4: Average transmission delay $E[D_L]$ for the low traffic example

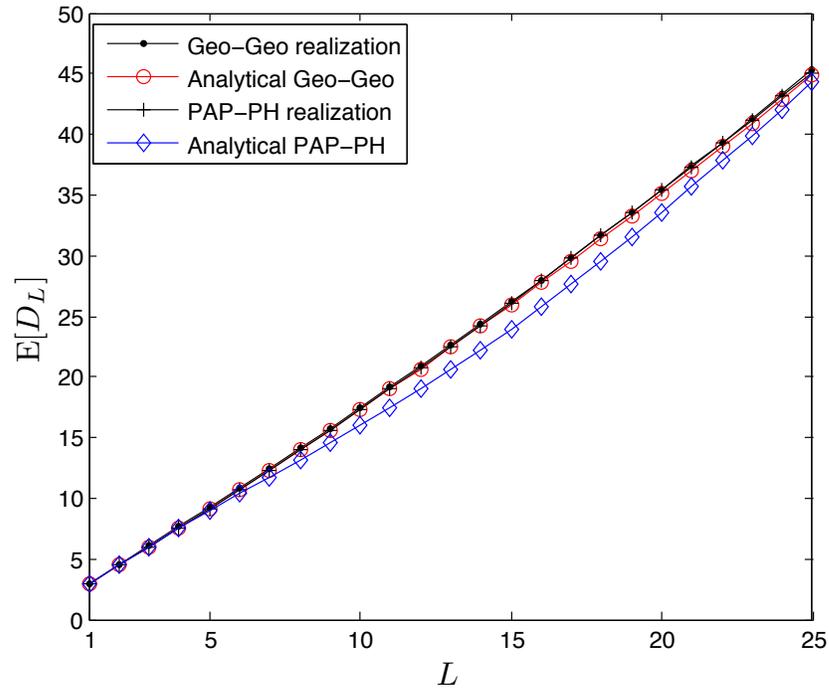
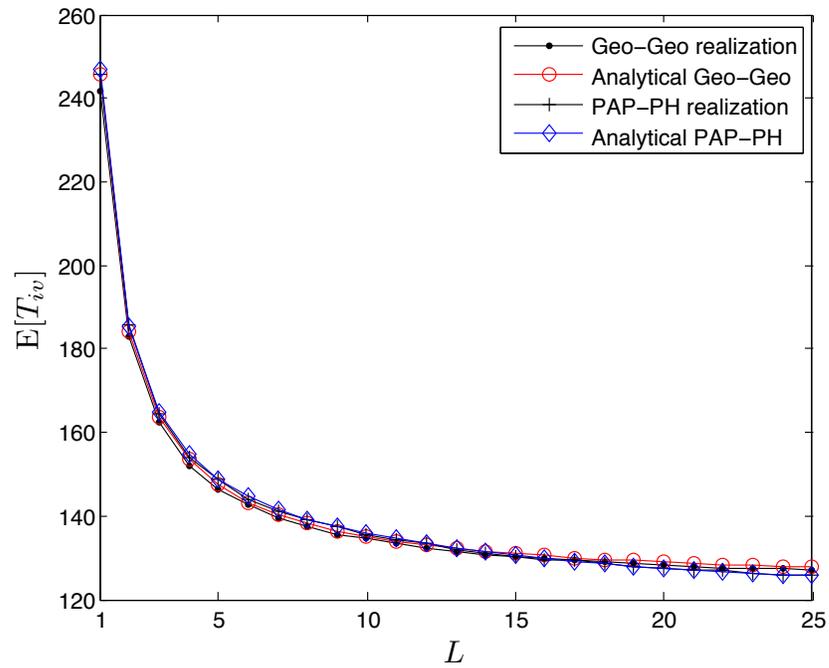


Figure 6.5: Average number of time slots between collisions $E[T_{iv}]$ for the low traffic example



collision occurs with larger L . Hence, an SU should use small values of L to avoid severe interference to primary users.

6.4.2 Simulation Results with High Traffic Example

Consider a PU channel with average utilization of $\bar{C} = 0.785$. The PAP-PH and the Geo-Geo channel model have a mean idle period of $I_{\text{ave}} = 10.937$ and mean busy period of $B_{\text{ave}} = 40$. As defined in *Section 5.4.2*, the PAP-PH model parameters are $\theta_1 = 0.6$, $\theta_2 = 0.03$, $\theta_3 = 0.2$ and $\theta_4 = 0.075$. The resulting Geo-Geo transition matrix P , which has been defined in Equation (6.1), is

$$P_{\text{Geo}} = \begin{bmatrix} 0.9086 & 0.0914 \\ 0.025 & 0.9750 \end{bmatrix}. \quad (6.16)$$

Similarly as the low traffic example, the reactive scheme has been applied to the two traffic models. By comparing *Figures 6.4* and *6.6*, we can see that the average transmission delay is higher as the channel usage increases. In addition, the average transmission delay increases with a longer SU transmission requirement length L . A larger difference in $E[D_L]$ between the Geo-Geo and PAP-PH channel models is observed as the channel utilization increases. However, if the transmission length L is kept small, such as $L < 10$, the Geo-Geo assumption over a bursty PU channel still achieves reasonable accuracy. On the other hand, if $L > 10$, then the Geo-Geo assumption is no longer valid to calculate the average transmission delay over a bursty PU channel.

The analytical approximation underestimates $E[D_L]$ when the value of L is close to the mean intraplatoon idle period. This is especially noticeable when deriving $E[D_L]$ for the PAP-PH traffic model case.

The average number of time slots between collisions versus L is shown in *Figure 6.7*. As L increases, the rate increases until reaching a convergence value. This convergence value is the worst case, where the SU causes interference almost each time the PU returns to the channel. When comparing the results of the Geo-Geo and PAP-PH models in *Figures 6.5* and *6.7*, we notice that a higher traffic causes a higher difference in $E[T_{iv}]$. Recall that for the busy-idle patterns in *Figure 6.7*, the average intraplatoon idle period is $L_{\text{intra}} = 5$.

Figure 6.6: Average transmission delay $E[D_L]$ for the high traffic example

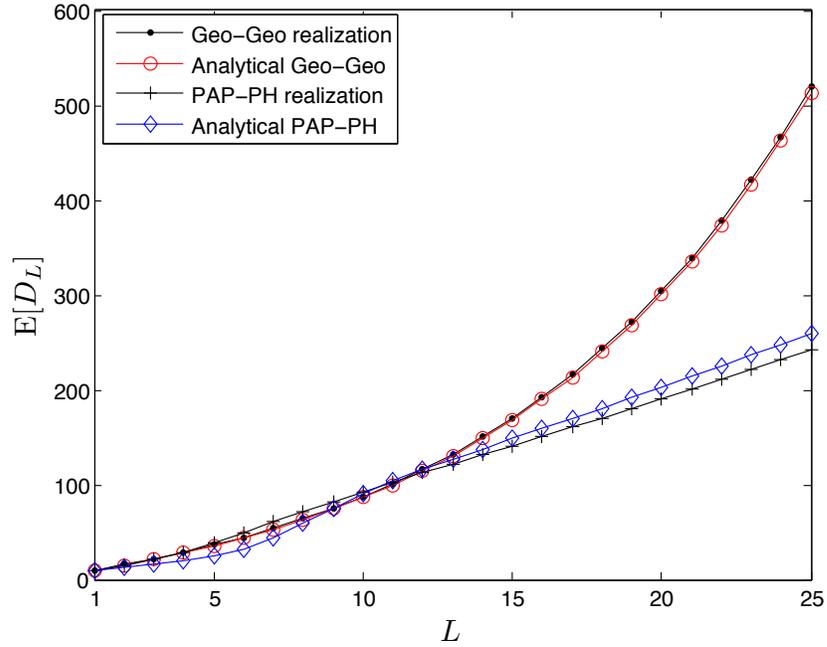
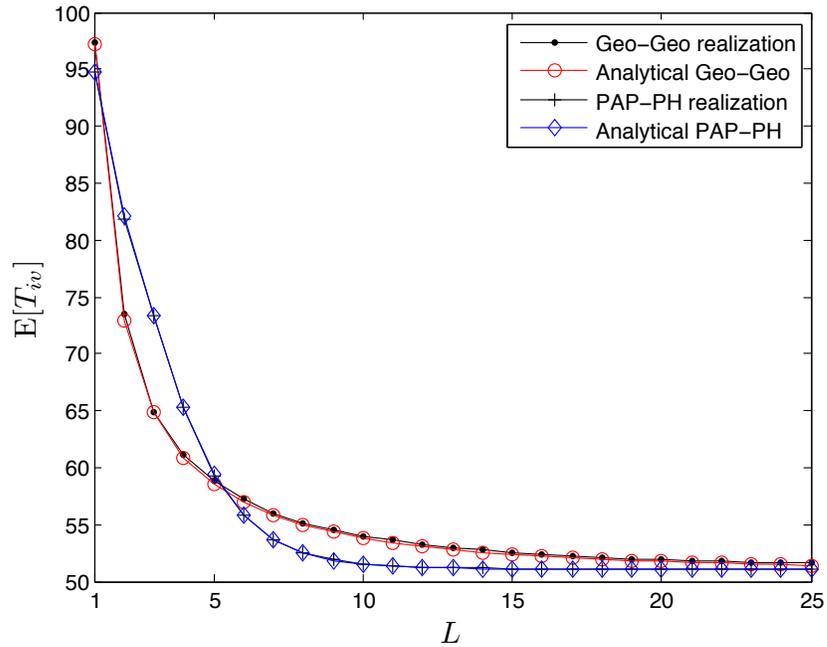


Figure 6.7: Average number of time slots between collisions $E[T_{iv}]$ for the high traffic example



If the idle-busy patterns of the PU channel is characterized by a PAP-PH process, then for $L \leq 5$ the Geo-Geo process underestimates $E[T_{iv}]$, while for $L > 5$ the Geo-Geo process overestimates $E[T_{iv}]$.

The analytical formulas used to get the average number of time slots between colli-

sions has excellent accuracy for both channel usage models. The average transmission delay formulas achieve very good accuracy for the Geo-Geo model. Even though the average transmission delay does present less accuracy for the PAP-PH model, it does provide enough insight of the performance of the secondary network. Hence, in the next section, further analysis of the performance measures will be performed using only the analytical formulas.

6.4.3 Different Platoon sizes

In this section, the reactive scheme will be simulated on PU channels characterized by different platoon sizes. The PAP-PH process is characterized by three negative binomial distributions with parameters $\theta_1 = 0.2$, $\theta_2 = 0.01$, $\theta_4 = 0.075$. The size of a platoon is constant with value equal to K , which is the order of the platoon size phase-type distribution $\text{PH}(\boldsymbol{\delta}, D), K$ defined in *Section 3.3*. The Geo-Geo process is estimated such that it matches the mean idle period and mean busy period of the PAP-PH process.

The average transmission delay versus the constant platoon size K for three different SU transmission length L is plotted in *Figure 6.8*. As the size of the platoon increases the channel utilization also becomes higher. Hence, we observe that $E[D_L]$ increases with K . This increase is more subtle for small values of L , while more dramatic for larger values of L . As noted earlier, we can see that the Geo-Geo model assumption is valid for small values of L , and the average transmission delay of the Geo-Geo diverge from that of the PAP-PH process as L increases.

Figure 6.9 shows the average number of time slots between collisions versus the platoon size K . Lower interference is achieved with smaller SU transmission lengths. For this example, the Geo-Geo model achieves good accuracy. In general, the Geo-Geo channel model achieves good accuracy to estimate $E[T_{iv}]$ and this accuracy is higher on channels with low utilization.

Figure 6.8: Average transmission delay $E[D_L]$ versus constant platoon size K

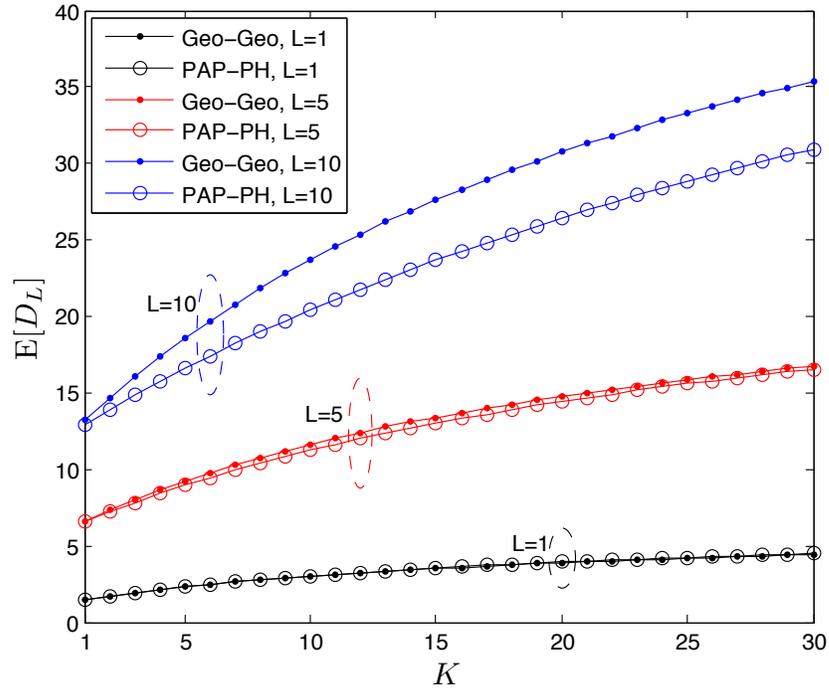
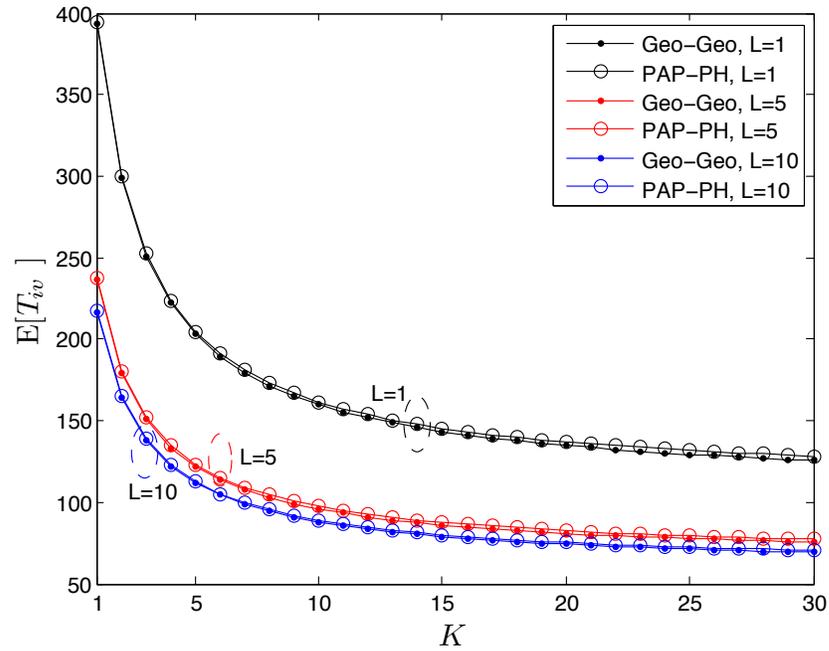


Figure 6.9: Average number of time slots between collisions $E[T_{iv}]$ versus constant platoon size K



Chapter 7

Proactive Spectrum Access Schemes

An intelligent access scheme can use the previous and present sensing results to build a statistical channel usage model which will aid the SU to make decisions, e.g. handoff. Therefore, the first phase to apply a proactive spectrum access scheme is the parameter estimation of the statistical channel usage model. With the estimated channel usage model the performance of different cognitive radio functions can be improved. Although we only focus on one PU channel, the usability of the proactive access schemes can be extended to multiple PU channels.

The authors in [27] have developed a fixed frame structure proactive access model. A frame is a set of consecutive time slots where a mandatory sensing is performed at the first time slot and the rest of the time slots are available for SU transmission. Their discrete time secondary user access model will be explained with more detail in *Section 7.3*. Based on a probability of interference constraint, their fixed frame system model dynamically decides how many time slots are used in each frame for an SU transmission. Using the work from [27], we have developed a proactive scheme where the secondary user system has no frame structure but has a fixed transmission length. This fixed transmission length system will be further described in *Section 7.4*.

The goal in this Chapter is to show the potential advantages of an accurate statistical characterization of a bursty PU channel model. Three traffic models will be compared on

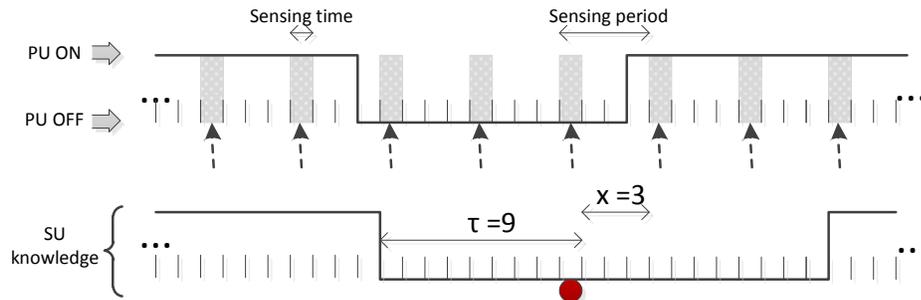
different proactive spectrum access schemes. These three traffic models are: the Geo-Geo model, the PH-PH model and the PAP-PH model. These models are compared in a low traffic scenario and in a high traffic scenario. The simulation results are shown in *Section 7.4.1* and in *Section 7.4.2*, respectively.

7.1 Traffic Prediction and Decision Policy Formulation

In this Section, the fundamental statistical formulas to predict the traffic will be developed. These formulas will then be used in the proactive spectrum access schemes in the subsequent sections to aid the SU make decisions.

First, consider that an SU monitors a single PU channel. The secondary network is a discrete time system where each time slot is the required length to sense the channel. After sensing the channel idle, an SU keeps track of the number of time slots the channel has remained idle. This quantity is denoted as τ . Say that the SU concludes that the channel has been idle for τ time slots and let us denote by $E_{on}(\tau, x)$ the event that the PU returns to the channel in x time slots after τ . That is, given that the channel has remained idle for τ slots, $E_{on}(\tau, x)$ is the event that the channel remains idle only for an additional $x - 1$ time slots. In *Figure 7.1*, the event $E_{on}(9, 3)$ is shown as an example.

Figure 7.1: $E_{on}(9, 3)$: channel has remained idle for 9 time slots and PU returns in 3 time slots



Let the random variable Z denote an idle period duration which has a given distribution function. The probability that the PU will return in x time slots given that the channel has

been idle for τ slots can be computed by [27]

$$\begin{aligned}\Pr[\mathbf{E}_{on}(\tau, x)] &= \Pr(Z = \tau + x - 1 | Z \geq \tau) \\ &= \frac{\Pr(Z = \tau + x - 1)}{\Pr(Z \geq \tau)}.\end{aligned}\tag{7.1}$$

In this Chapter, it is assumed that an SU cannot detect a PU arrival while transmitting. Now, consider that the channel has been idle for τ slots. If the SU transmits for the next h consecutive time slots, the mean number of time slots with interference is [27]

$$\psi_{iv}(\tau, h) = \sum_{x=1}^h \Pr[\mathbf{E}_{on}(\tau, x)](h + 1 - x).\tag{7.2}$$

The expression $\psi_{iv}(\tau, h)$ was computed assuming that the transmission length of an SU will not exceed a busy period duration. It should be noticed that $\psi_{iv}(\tau, h)$ depends on the assumed traffic model. Hence, we will develop $\Pr[\mathbf{E}_{on}(\tau, x)]$ for the three traffic models used in this thesis.

Let the random variable Z be represented by a PAP(C_0, C_1), m . The stationary distribution of the phases is obtained with $\boldsymbol{\pi} = \boldsymbol{\pi}(C_0 + C_1)$ and $\boldsymbol{\pi}\mathbf{1} = 1$. Then, $\Pr[\mathbf{E}_{on}(\tau, x)]$ becomes

$$\Pr[\mathbf{E}_{on}(\tau, x)] = \frac{\boldsymbol{\pi}' C_0^{\tau+x-2} C_1 \mathbf{1}}{\boldsymbol{\pi}' C_0^{\tau-1} \mathbf{1}},\tag{7.3}$$

with $\boldsymbol{\pi}'$ defined in (3.12).

If the channel is thought to be characterized by a PH-PH traffic model, where the idle periods are described by PH($\boldsymbol{\alpha}, T$), m_i we have that

$$\Pr[\mathbf{E}_{on}(\tau, x)] = \frac{\boldsymbol{\alpha} T^{\tau+x-2} \mathbf{T}^0}{\boldsymbol{\alpha} T^{\tau-1} \mathbf{1}},\tag{7.4}$$

with $\mathbf{T}^0 = \mathbf{1} - T\mathbf{1}$.

For a Geo-Geo idle-busy periods with parameters Geo(b, q), $\Pr[\mathbf{E}_{on}(\tau, x)]$ is indepen-

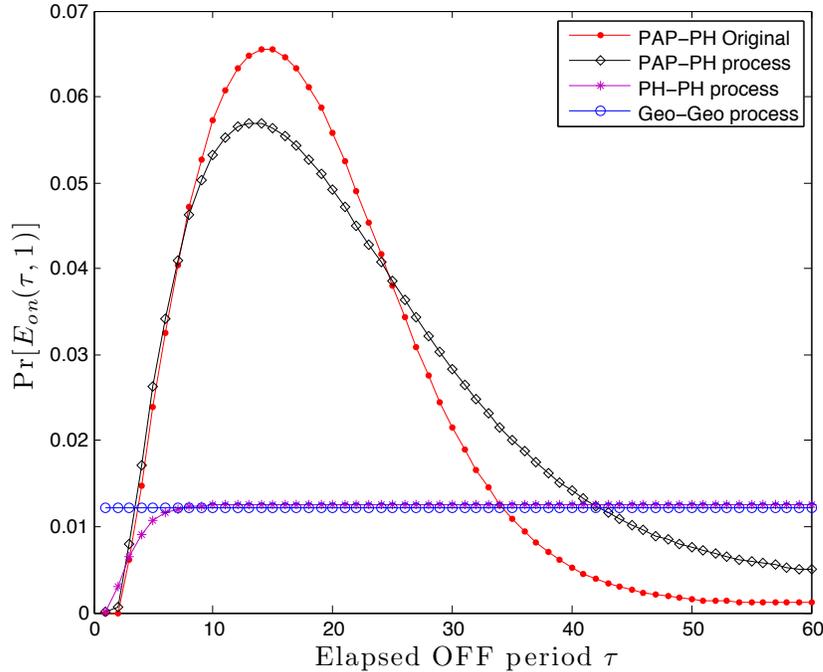
dent of τ and can be calculated as

$$\Pr[E_{on}(\tau, x)] = (1 - p)^{x-1}p. \quad (7.5)$$

7.1.1 Examples of the traffic prediction formulas

If the secondary network has a statistical model for the PU channel patterns, then Equation (7.1) can be used to gain more knowledge of the channel usage. For instance, consider the low traffic example explained in *Section 5.4.1*. The probability that the PU will return in the next time slot given that the channel has been idle for τ time slots is shown in *Figure 7.2*. If we use a geometric distribution to characterize the idle periods, $\Pr[E_{on}(\tau, 1)]$ remains constant and is independent of τ . Because of the memoryless property of geometric random variables, we cannot exploit more information of the idle periods by using variable τ .

Figure 7.2: Probability that the PU returns in the next time slot given that the channel has been idle for τ time slots for the low traffic example

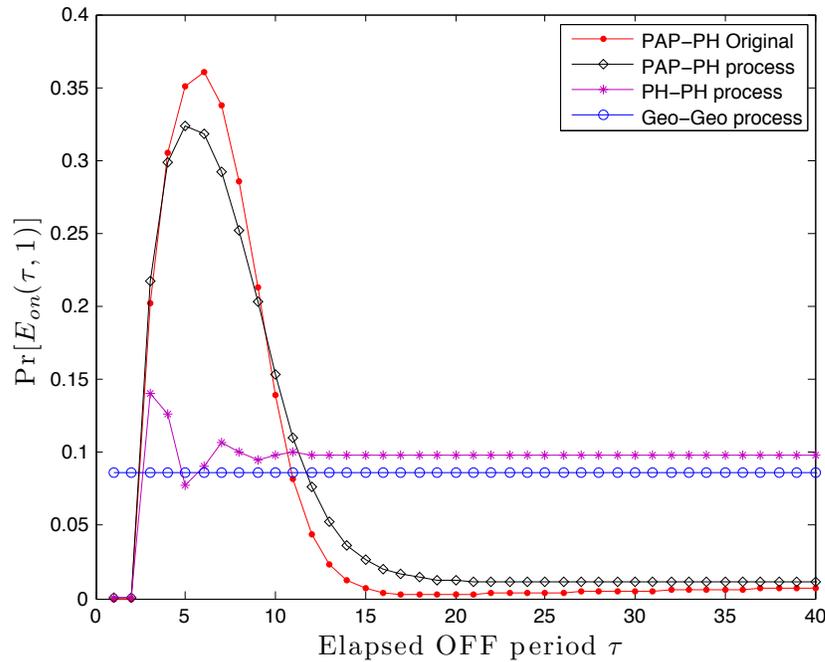


If a phase-type distribution is used to model the idle durations, this provides more information than the geometric case; however, for τ larger than 10, $\Pr[E_{on}(\tau, 1)]$ remains almost constant. This means that the the estimated parameters of the phase type distribution, of

order 3, is able to capture that the PU is less likely to return when the elapsed idle period is less than 10. Remember that for the low-usage channel example the mean idle intraplatoon period is $L_{\text{intra}} = 15$.

After comparing the shape of $\Pr[E_{\text{on}}(\tau, 1)]$ in *Figure 7.2*, we can see that the results obtained from the estimated PAP-PH process resembles the results of the original PAP-PH process. The shape of $\Pr[E_{\text{on}}(\tau, 1)]$, belonging to the PAP-PH process, gives us substantially more information of the behaviour of the PU channel model. By looking at the results obtained from the original PAP-PH process, we can see that there is a higher probability that the PU will return in the next time slot when τ is around 15, which is the mean intraplatoon L_{intra} idle period of this example. As τ gets closer to L_{intra} there is a greater chance that the idle period is an intraplatoon. On the contrary, as τ gets larger than L_{intra} , the probability that the idle period is an interplatoon gets higher.

Figure 7.3: Probability that the PU returns in the next time slot given that the channel has been idle for τ time slots for the high traffic example



Similar results are obtained for the high channel-usage example as shown in *Figure 7.3*. Similar curve shapes for the original and estimated PAP-PH process are observed. According to the original PAP-PH process, as τ gets closer to six, the PU has a higher

probability of returning in the next time slot. Recall that for the high channel usage example specified in *Section 5.4.2*, the mean intraplatoon idle period is $L_{\text{intra}} = 5$.

7.2 Performance Measures

To evaluate the performance of our simulations we have applied the same performance measures defined in [27]. For each simulation we have calculated the spectrum utilization Φ and the interference violation probability p_{iv} . The spectrum utilization Φ has been defined as [27]

$$\Phi = \frac{E[V_{suc}]}{E[Z]}, \quad (7.6)$$

where V_{suc} is the number of time slots per idle period that are the result of a successful transmission, and $E[Z]$ is the mean idle period. The maximum utilization achievable by an SU is defined as [27]

$$\Phi_{\max} = \frac{E[Z] - E[\eta_s]}{E[Z]}, \quad (7.7)$$

where random variable η_s is the number of time slots during an idle period that has been used for a mandatory sensing. The meaning of mandatory sensing will be further explained when each proactive scheme is discussed in the next sections.

The interference violation probability can be calculated with [27]

$$p_{iv} = \frac{E[V_{iv}]}{E[Z] - E[\eta_s]}, \quad (7.8)$$

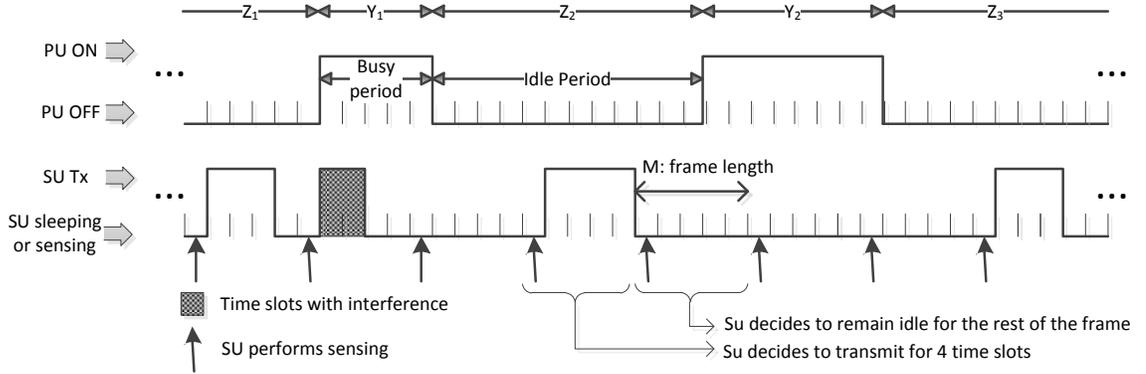
where V_{iv} is the number of time slots per idle period that had interference with the PU.

7.3 Fixed Frame System Model

In this Section, we present the discrete time secondary access model that has been designed in [27]. Consider an SU access model with a fixed frame structure of length M . The first time slot of each frame is used for sensing; this is mandatory. If the channel is found idle,

the SU proceeds to calculate the number of consecutive time slots h of the frame which would be used for transmission. The maximum length of the SU transmission is $M - 1$ which is determined by the length of the frame. A diagram for the secondary user access model is shown in *Figure 7.4*. The SU only starts a transmission in the second time slot of a frame.

Figure 7.4: Secondary user access model with fixed frame structure of length M



Given an interference probability constraint in a time slot, which is denoted as p_{thr} , the number of time slots used for a transmission is determined per frame. The value set to p_{thr} will depend on the quality of service required for the PU network. Given that the SU has knowledge that the PU channel has remained idle for τ time slots and that the interference probability constraint in a time slot is p_{thr} , we have the following linear optimization problem

$$\begin{aligned}
 h^* &= \text{maximize } h \\
 &\text{subject to } \psi_{iv}(\tau, h) \leq (M - 1)p_{thr}, \\
 &0 \leq h \leq M - 1.
 \end{aligned} \tag{7.9}$$

Notice that for each τ , the optimal number of time slots for transmission h^* can be previously calculated and stored in the secondary network as long as the underlying channel pattern remains unchanged. Then, after sensing the channel is idle at the beginning of the frame, the SU can use the optimal number of time slots for transmission. Notice that if the channel is sensed idle after a mandatory sensing, an SU can also decide not to transmit if

no positive h satisfies constraint (7.9) for the corresponding τ .

If the channel is found busy after a mandatory sensing, the SU has two options: either go to sleep or continue to sense for the rest of the frame. We will call these two schemes as:

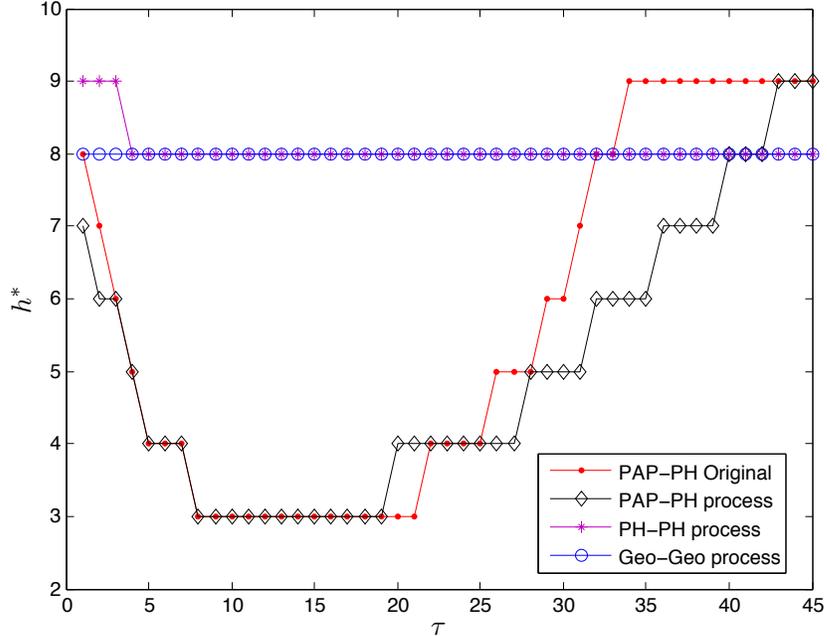
- A) **Fixed frame model of length M with sleep state:** in this scheme an SU has three states: sensing, transmitting and sleeping. A mandatory sensing is performed in the first time slot of each frame. If the PU channel is found busy in a frame, the SU sleeps for the rest of the frame. This scheme can also be thought of having a fixed sensing period of length M . That is, the secondary user only senses the channel every M time slots. The advantage of this scheme is that the SU can conserve its energy.
- B) **Fixed frame length model of length M with no sleep state:** in this scheme an SU has two states: sensing and transmitting. Mandatory sensing is performed in the first time slot of each frame. If the PU channel is sensed busy in the first time slot of a frame, the SU continues to sense for the rest of the frame. The SU however does not start a transmission if at some point of the frame the PU channel is sensed idle. Possible SU transmission only occurs if the PU channel is sensed idle at the beginning of the frame. The advantage of this scheme is that the SU has a better tracking of the elapsed idle period τ ; hence, we expect that the prediction of how many slots to be used for transmission is more accurate than the scheme where the SU sleeps.

For both of these schemes the SU receives a negative acknowledgement (NACK) of those time slots that had suffered a collision with a PU. Using this information, the secondary user can update the state of the PU channel accordingly. By having packet acknowledgements, the elapsed idle period τ is tracked more accurately. For Scheme A, where the SU is allowed to sleep, the historical busy-idle patterns of the PU channel are inherently not exact. On the other hand, assuming perfect sensing, no channel errors and NACK, Scheme B accurately captures the state of the PU channel at all times.

7.3.1 Simulation Results of the Low Traffic Example

Consider that the fixed frame scheme, with frame length $M = 10$, is applied to the low traffic example specified in *Section 5.4.1*. This means that 9 time slots in each frame are available for transmission. For this PU channel that behaves as a PAP-PH process, the optimal transmission length h^* versus τ is plotted for each estimated channel usage model in *Figure 7.5*. If the SU uses a Geo-Geo model to predict, then it will always dedicate $h^* = 8$ time slots in a frame for transmission. In a similar way, the PH-PH model will also use 8 time slots in general; however, when $\tau \leq 3$ the SU is more aggressive and all available time slots, $M - 1 = 9$, are used for transmission.

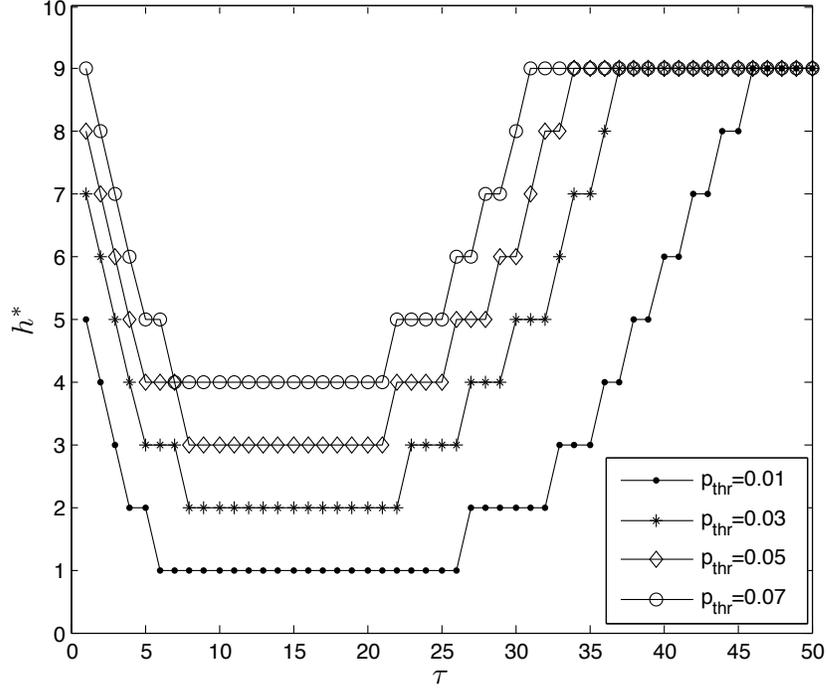
Figure 7.5: Optimal transmission length h^* versus τ for $p_{thr} = 0.05$ and $M = 10$.



An SU using the original or estimated PAP-PH models is more cautious in the transmission length when τ gets closer to the mean intraplatoon idle period. Hence, we can see that the optimal transmission length decreases as τ gets closer to $L_{intra} = 15$. Moreover, the SU becomes more aggressive as $\tau > L_{intra}$ because the SU is more confident that the current idle period is of interplatoon type.

The effect of the interference constraints, p_{thr} , in the optimal transmission length can be observed in *Figure 7.6*. This figure only shows the results for the original PAP-PH model.

Figure 7.6: Optimal transmission length h^* versus τ for different p_{thr} values with $M = 10$ in the low traffic example



As p_{thr} gets larger, the SU is more aggressive, and more time slots are used for transmission in a frame.

To compute the performance measures defined in *Section 7.2*, Monte Carlo simulations with 30 replications of one million time slots have been performed. First, consider the fixed frame scheme with sleep state with a time slot interference probability of $p_{thr} = 0.05$. The average interference in a time slot, p_{iv} , versus M is shown in *Figure 7.7*. For this traffic, the estimated Geo-Geo and estimated PAP-PH satisfy the $p_{thr} = 0.05$ interference threshold; however, the PH-PH channel model does not satisfy p_{thr} for a frame length larger than 10. Although the estimated Geo-Geo model satisfies the interference threshold, we can observe that the estimated PAP-PH process achieves lower interference for $2 \leq M \leq 15$.

The average utilization as a function of M for the access model with sleep state is plotted in *Figure 7.8*. While the achievable mean utilization of the Geo-Geo process is highly affected by larger frames sizes, the estimated PAP-PH process shows more stability and accomplishes a higher utilization when $9 \leq M \leq 16$. Hence, from *Figures 7.7* and *7.8*, we think that the Geo-Geo model can still be used for small frame sizes at a cost of higher

Figure 7.7: Fixed frame scheme: p_{iv} as a function of M in the low traffic example with $p_{thr} = 0.05$

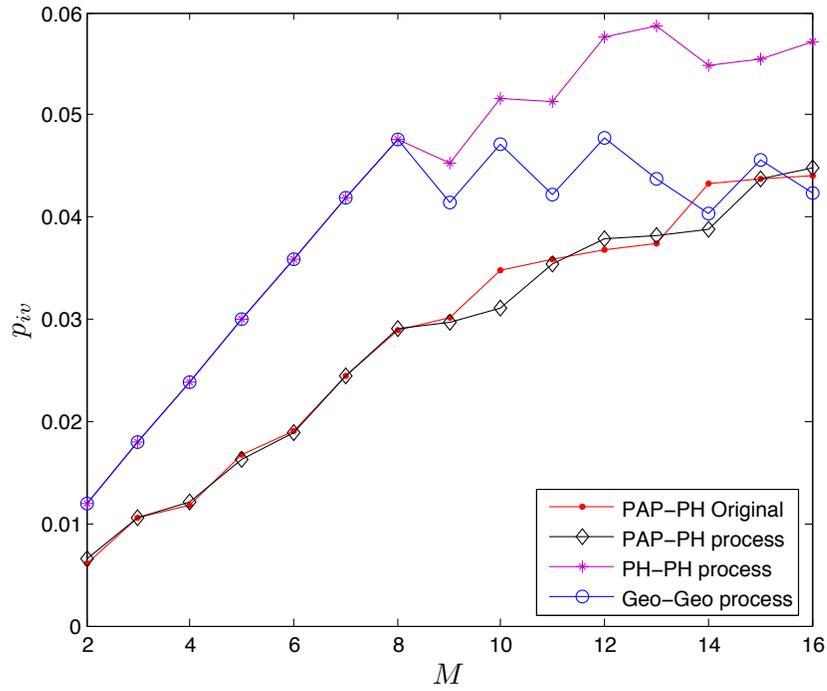
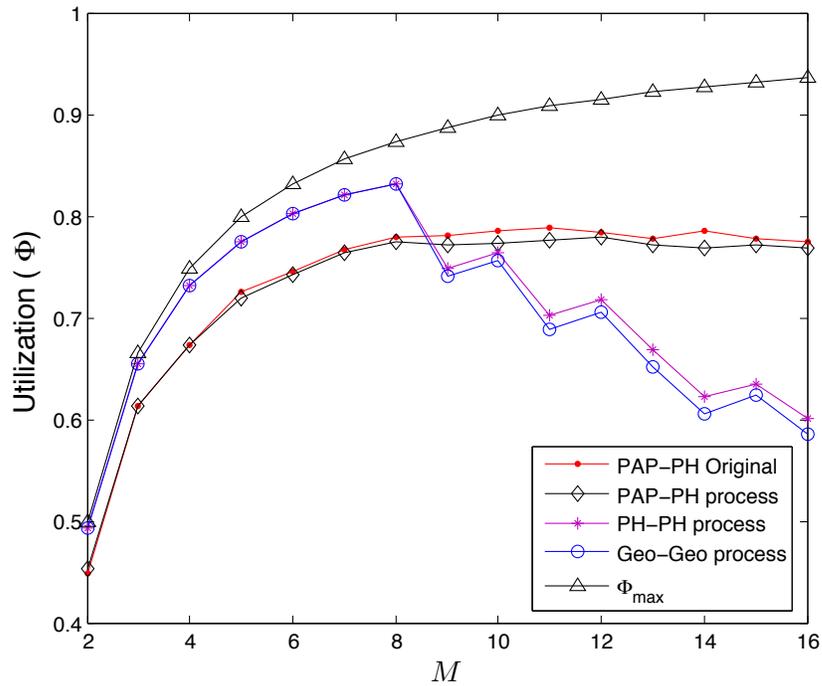
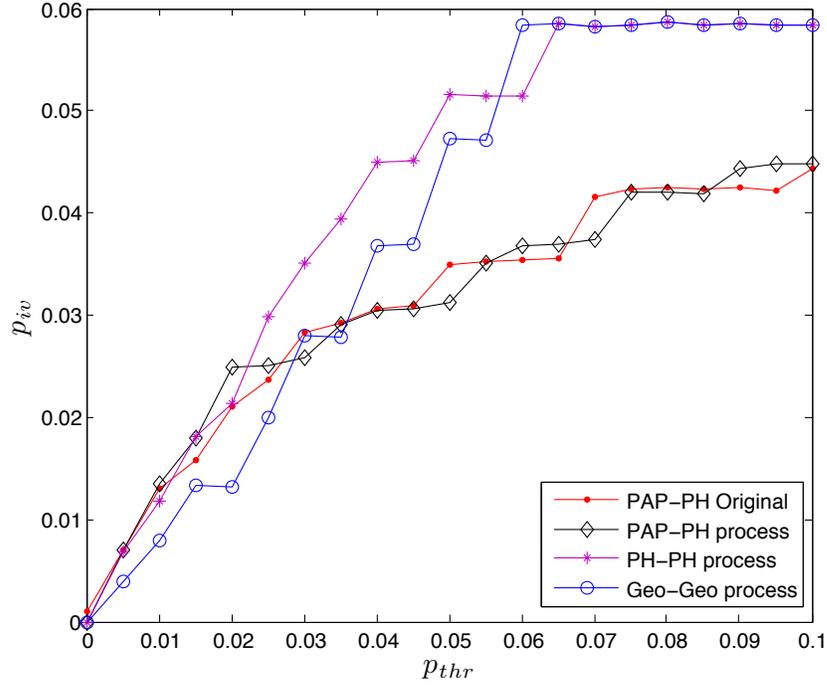


Figure 7.8: Fixed frame scheme: Φ versus M in the low traffic example with $p_{thr} = 0.05$



interference. Nevertheless, for $M > 10$ the spectrum efficiency of the Geo-Geo model decays rapidly. Likewise, the estimated PH-PH model has similar performance behaviour. Thus, the Geo-Geo and PH-PH models are less reliable than the PAP-PH model, especially

Figure 7.9: Fixed frame scheme: p_{iv} versus p_{thr} in the low traffic example with $M = 10$

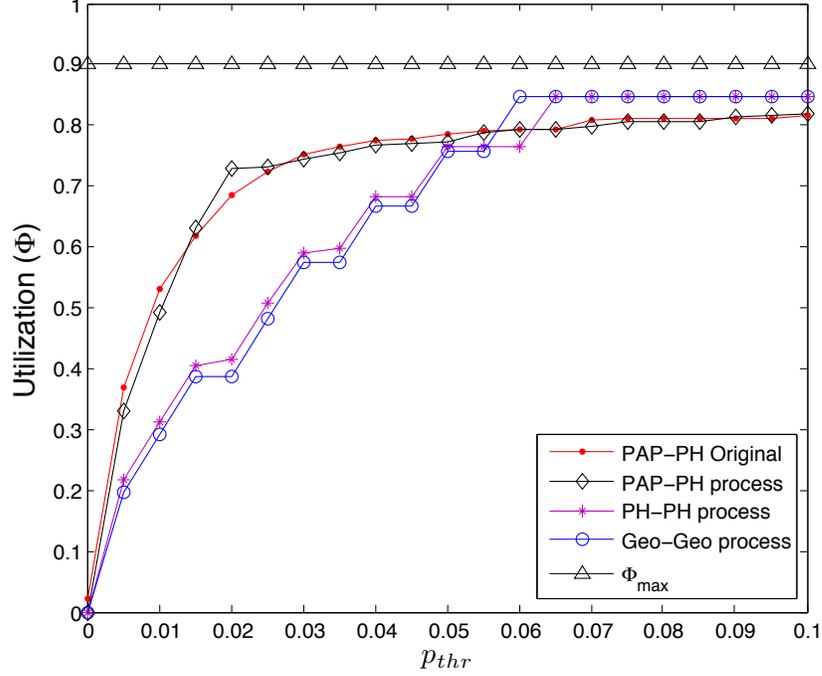


as M gets larger.

In *Figure 7.9*, we show the interference violation probability as a function of the interference threshold when $M = 10$. Even though the Geo-Geo and PH-PH achieves comparable interference with the PAP-PH process when $p_{thr} \leq 0.05$, the estimated PAP-PH process will allow a higher utilization of the PU channel, as shown in *Figure 7.10*. Nevertheless, when the interference constraint probability is somewhat large, such as $p_{thr} > 0.05$, the estimated Geo-Geo model and estimated PH-PH are still suitable for low-usage channels with a small degree of burstiness.

Even for a large interference constraint p_{thr} , in *Figure 7.10*, the gap between the achievable utilization Φ of a scheme and maximum utilization exists because the SU can only transmit at the beginning of a frame.

Figure 7.10: Fixed frame scheme: Φ versus p_{thr} in the low traffic example with $M = 10$



7.3.2 Simulation Results of the High Traffic Example

We have applied the fixed frame scheme with no sleep state to the PU channel with high traffic example defined in *Section 5.4.2*. With an interference threshold of $p_{thr} = 0.05$, the interference probability and channel utilization as a function of M is shown in *Figures 7.11* and *7.12*.

The oscillating behaviour of p_{iv} for the Geo-Geo model, comes from the fact that for a given threshold p_{thr} and frame length M , the optimal transmission length h^* in a frame is constant and is independent of the elapsed idle period. Hence, in this example, $h^* = 1$ for $3 \leq M \leq 6$, $h^* = 2$ for $7 \leq M \leq 11$ and $h^* = 3$ for $12 \leq M \leq 16$.

As shown in *Figure 7.12*, the estimated PAP-PH process outperforms the Geo-Geo and PH-PH models in terms of channel utilization. Moreover, in this simulation, an SU using the Geo-Geo model has an average underutilization of 58.1% with respect to the estimated PAP-PH model. While efficiently predicting spectrum opportunities, the estimated PAP-PH model causes a mean interference below the interference threshold.

In *Figure 7.13*, we show that the estimated PAP-PH model discovers more opportunities

Figure 7.11: Fixed frame scheme: p_{iv} as a function of M in the high traffic example with $p_{thr} = 0.05$

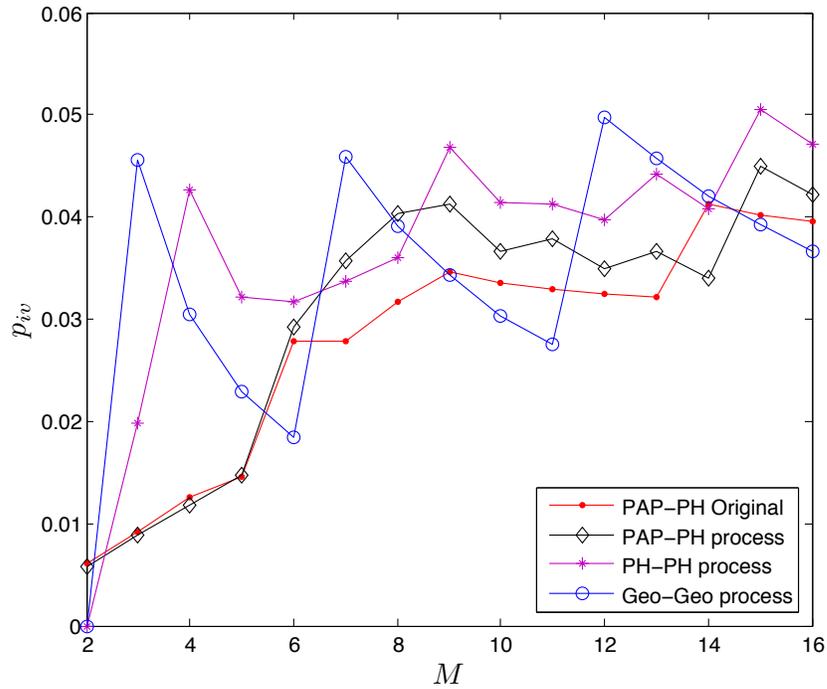
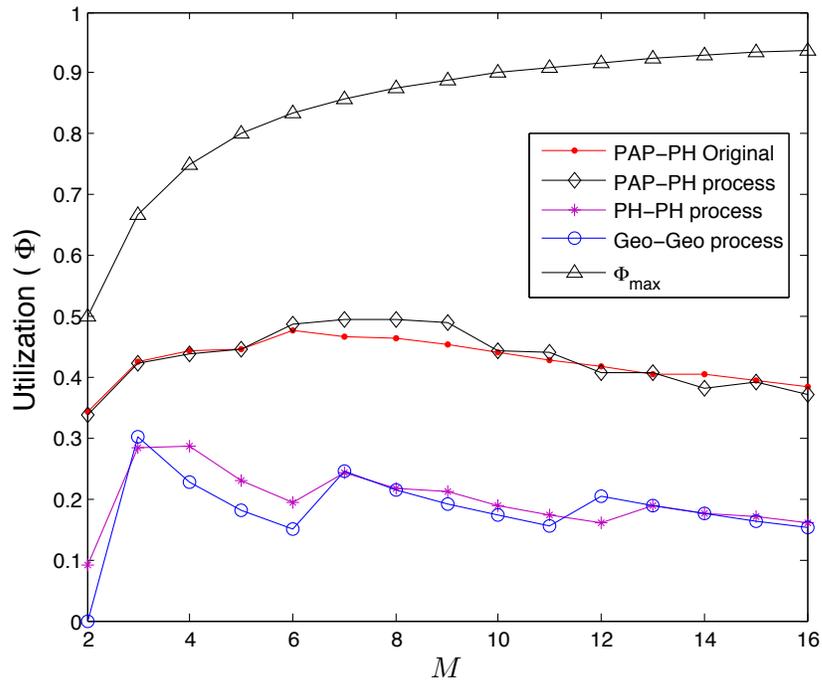


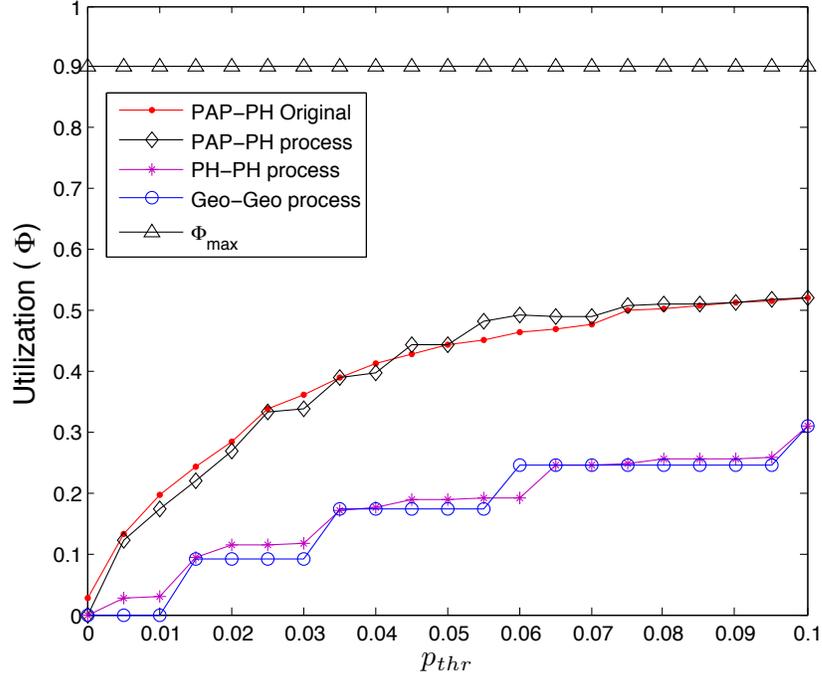
Figure 7.12: Fixed frame scheme: Φ versus M in the high traffic example with $p_{thr} = 0.05$



than the Geo-Geo and PH-PH models for an interference threshold between $0.05 \leq p_{thr} \leq 0.1$. We can see from *Figures 7.13* and *7.10*, that the efficiency of identifying spectrum opportunities decreases as the traffic in the PU channel increases. Moreover, the prediction

of the Geo-Geo model gets worse as the channel usage increases.

Figure 7.13: Fixed frame scheme: Φ versus p_{thr} in the high traffic example with $M = 10$



7.3.3 Simulation Results of the Geo-Geo PU Channel Example

Consider a PU channel with geometrically distributed idle and busy periods, with parameters $\text{Geo}(0.025, 0.025)$. The mean idle period and mean busy period are both 40. We have applied the estimated traffic models, from *Section 5.4.3*, to the fixed frame access scheme with no sleep.

The mean interference and the mean utilization as a function of the frame size is given in *Figures 7.14* and *7.15* respectively. The three estimated traffic models have the same performance. If one achieves lower interference, it also accomplishes lower utilization and viceversa. Besides achieving higher performance with bursty patterns, an SU using the PAP-PH model will also accurately predict the behaviour of a Geo-Geo PU channel. It means that a PAP-PH model can also be used in a Geo-Geo PU channel successfully. Moreover, the same performance in terms of throughput and interference probability as the Geo-Geo model will be obtained.

Figure 7.14: Fixed frame scheme: p_{iv} as a function of M in the Geo-Geo example with $p_{thr} = 0.05$

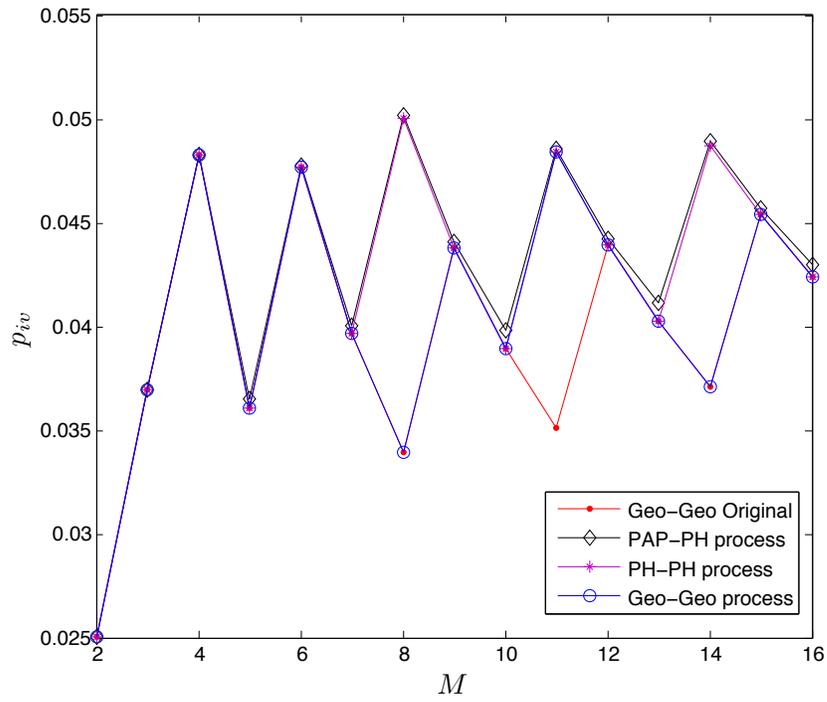
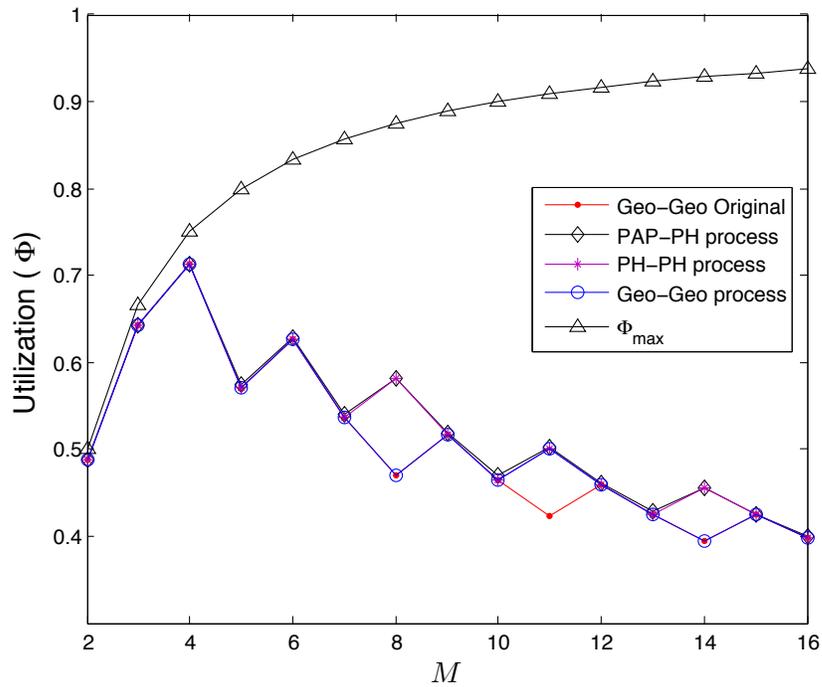


Figure 7.15: Fixed frame scheme: Φ versus M in the Geo-Geo example with $p_{thr} = 0.05$



7.4 Fixed Transmission Length System Model

In this Section, we develop a variation of the proactive scheme presented in *Section 7.3* where the secondary user has a fixed transmission length.

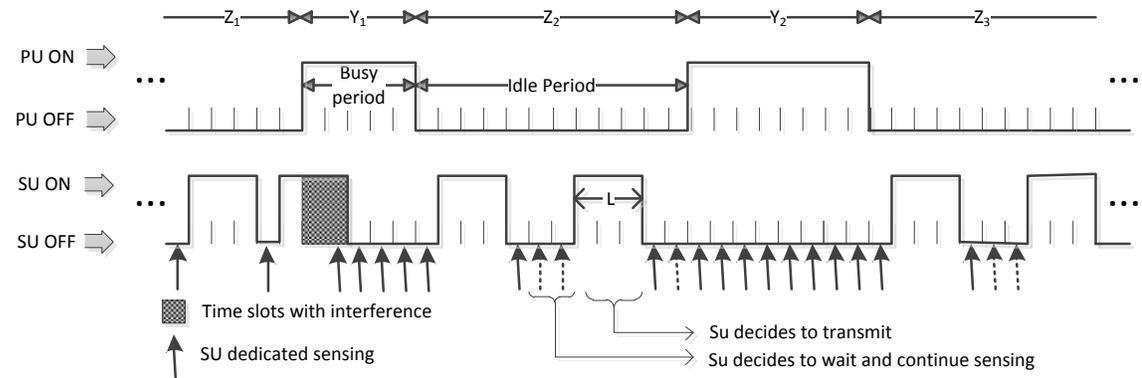
In this scheme, it is assumed that the SU does not sleep and is either sensing or transmitting in a particular time slot. The transmission length of the SU is constant and denoted by L . Specifically, each SU transmission involves L consecutive time slots. When the PU channel is sensed idle, the SU can decide whether or not to start a transmission. If the SU decides not to start a transmission, this action is referred to as waiting. During the waiting state, the SU continues to sense the channel.

As the previous proactive schemes, the decision whether or not to start a transmission is based on a time slot interference constraint p_{thr} . For this proactive scheme, the SU transmits for L consecutive time slots only if the following condition holds

$$\psi_{iv}(\tau, L) \leq \beta \times p_{thr}, \quad (7.10)$$

with p_{thr} the interference violation probability threshold for one time slot, and β is the number of consecutive time slots where the SU remained idle during an OFF period waiting for the condition (7.10) to hold plus L .

Figure 7.16: Secondary user access model with fixed transmission length L



The SU access model with a fixed transmission length is shown in *Figure 7.16*. The variable β is the number of time slots the SU has in its advantage to preserve the proba-

bility of interference p_{thr} constraint of the PU channel. More specifically, after each SU transmission of length L the variable is reset to $\beta = L$. Consider that at time slot $n = 1$ the channel is sensed idle and it has remained idle for τ time slots. If the condition (7.10) does not hold, the SU will restrain from transmission in the next time slot, $n = 2$. In this case, variable β is increased by one, i.e., $\beta = \beta + 1$. If the PU channel still remains idle in the next time slot, $n = 2$, the SU may have larger chances of satisfying condition (7.10). If it does satisfy the condition, then the SU will proceed to transmit during time slots $n = 3$ to $n = 3 + L - 1$. On the other hand, if it does not satisfy condition (7.10), the action of increasing variable β by one continues as long as the SU keeps on waiting during an PU OFF period and condition (7.10) is not yet satisfied. The moment when the channel is sensed busy, the variable β is reset to $\beta = L$, and the SU continues the sensing task until the channel is idle.

The action of a unit increase in β can be seen as a reward to the SU for remaining idle in the next time slot and not causing potential interference to the PU. If the channel is sensed idle in the next time slots the SU can have a possibly advantage for the condition (7.10) to hold. Although, it will depend ultimately on the average number of interference the SU is likely to cause given that the channel has remained idle for τ time slots.

Remember that when the SU is in the transmission state it cannot detect a PU arrival. Hence, a collision only occurs in the following scenario: the secondary user starts transmitting, and the primary user returns before the SU has finished transmitting. If there has been a collision, the SU receives a NACK during these time slots that had suffered interference. As in the fixed frame length model with no sleep state explained in Section 7.3, this fixed transmission length model also has a perfect knowledge of the PU channel historical behaviour. This is because; perfect sensing and no channel errors are assumed.

Given an interference threshold p_{thr} and a fixed transmission length L , an SU will always start a transmission for specific values of τ . That is, with a channel usage model we can calculate when condition (7.10) will hold and the entire behaviour of the transmitter can be known beforehand. In other words, for a traffic model, the SU will always transmit

at particular values of τ . With perfect sensing, this implies that the receiver would not need additional control messages to know when a transmission has started; however, in reality, perfect sensing is not possible.

7.4.1 Simulation Results of the Low Traffic Example

Let us assume that the idle and busy periods of a PU channel behave as a PAP-PH process with parameters defined in *Section 5.4.1*. Monte Carlo simulations have been performed to obtain the performance measures defined in *Section 7.2* for the fixed transmission length proactive scheme.

We first set the interference violation probability to $p_{thr} = 0.05$. The effective interference probability p_{iv} as a function of L , for the three channel-usage models, is shown in *Figure 7.17*. For values of $L \geq 9$, the Geo-Geo traffic model does not satisfy the interference probability threshold p_{thr} . As the transmission length L increases, an SU causes more interference to PUs in general. Nonetheless, the effect in p_{iv} of increasing L , is much less for a PAP-PH model.

In *Figure 7.18*, we show the achievable utilization for each traffic model as a function of transmission length L . The maximum utilization Φ_{max} has been calculated based on the original traffic model. The maximum utilization is an increasing function of L because when L is larger, less dedicated sensing is performed. When $L > 7$, the SU has more difficulties in predicting the behaviour of the channel, and the SU spends more time in the waiting state for the transmission requirement in (7.10) to hold. Recall that in this example, the mean idle intraplatoon is $L_{intra} = 15$. A transmission length that is close to L_{intra} , should cause the SU to avoid transmission when the current idle period has a high chance of being of intraplatoon type. Thus, the mean utilization, of the PAP-PH process, stays almost constant for $8 \leq L \leq 15$ in *Figure 7.18*. On the other hand, prediction with the Geo-Geo process will cause a decreased spectrum efficiency when $L \geq 8$, because a Geo-Geo process cannot capture the types of idle periods. Not only the Geo-Geo process does not efficiently utilize the spectrum, but also causes a higher interference when the

Figure 7.17: Fixed transmission length scheme: p_{iv} as a function of L in the low traffic example with $p_{thr} = 0.05$

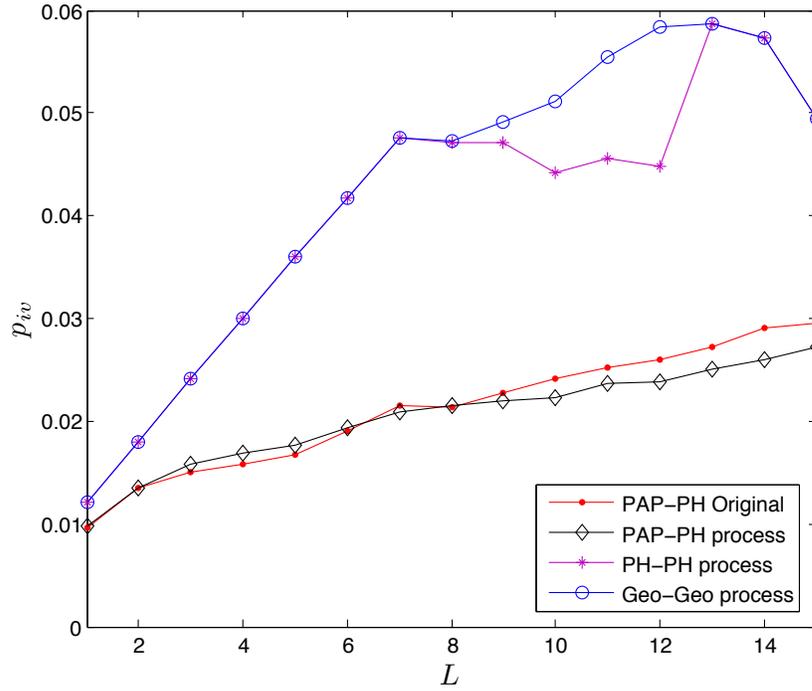
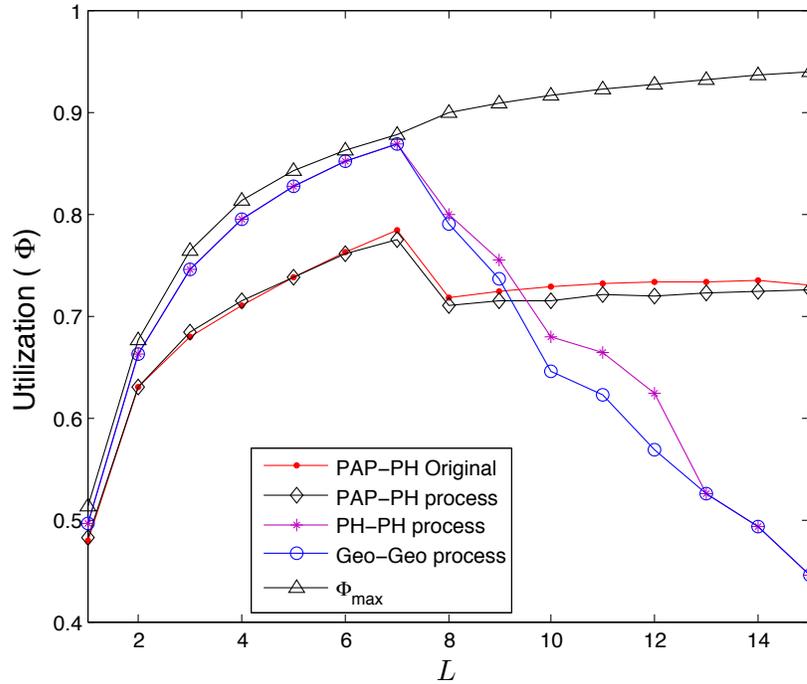


Figure 7.18: Fixed transmission length scheme: Φ versus L in the low traffic example with $p_{thr} = 0.05$



transmission length is close to L_{intra} .

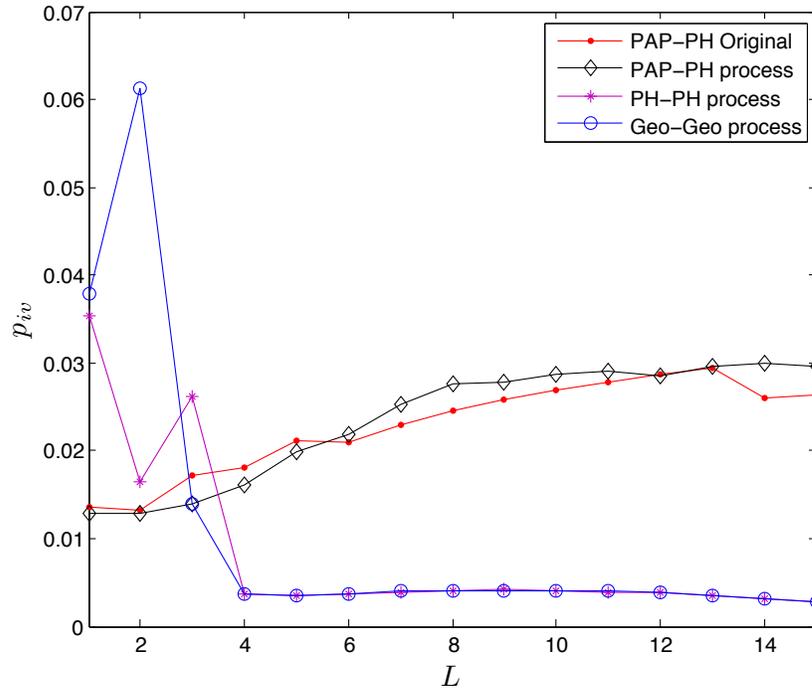
When comparing the results in Figure 7.7 of the fixed frame transmission scheme with

sleep, we can say that the fixed transmission length scheme is more aggressive when $L \leq 7$. In addition, lower interference is achieved in the fixed transmission length scheme because an SU has perfect knowledge of the PU channel when making decisions; while in the fixed frame scheme example, given in *Section 7.3.1*, an SU is allowed to sleep .

7.4.2 Simulation Results of the High Traffic Example

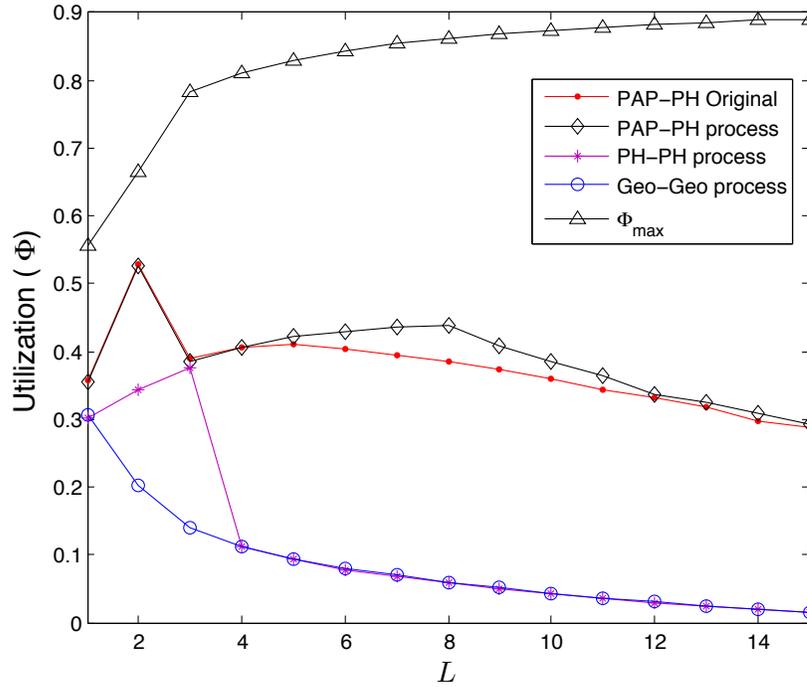
To see the effect of an increased channel utilization, we executed the fixed transmission length scheme in a PU channel with high traffic with parameters defined in *Section 5.4.2*. *Figures 7.19* and *7.20* show the interference probability and utilization as a function of L , respectively.

Figure 7.19: Fixed transmission length scheme: p_{iv} as a function of L in the high traffic example with $p_{thr} = 0.05$



In this scheme, the estimated Geo-Geo model performance is affected by an increased channel usage. Even for small transmission lengths, higher interference and lower spectrum utilization is accomplished with the Geo-Geo model when compared to the estimated PAP-PH model. When $L > 3$ or close to the mean idle intraplatoon period L_{intra} , we can see an

Figure 7.20: Fixed transmission length scheme: Φ versus L in the high traffic example with $p_{thr} = 0.05$



abrupt descent of p_{iv} for the Geo-Geo process and PH-PH process. This is a direct result of the SU remaining in the sensing and waiting state for long periods. This fact can be seen in *Figure 7.20*, where we can identify that for all L , the PAP-PH model discovers more spectrum opportunities than the Geo-Geo and PH-PH models. Therefore, we can also say that a higher utilization and a reasonable interference probability to the PU channel are achieved when the PAP-PH process is used.

7.5 Comparison and Comments

The performance of the Geo-Geo model is better in the fixed frame scheme explained in Section 7.3, rather than the fixed transmission length scheme outlined in Section 7.4. To achieve acceptable prediction results with the Geo-Geo model, an SU should use small frame sizes in the fixed frame scheme. As the channel utilization increases, the advantage of using a PAP-PH process, instead of a Geo-Geo model or a PH-PH model, to predict future channel states becomes more obvious. That is, higher discovery of spectrum opportunities

is accomplished with the PAP-PH traffic model as the channel-usage becomes higher. This is a direct result of the prediction accuracy of the PAP-PH traffic model.

If the fixed transmission length scheme is used over a bursty channel model, the Geo-Geo model becomes unusable when the transmission length is close to the mean intraplatoon period. Hence, we recommend the use of a frame size or transmission length that is smaller than the mean intraplatoon period.

Throughout our simulations, we noticed that the PAP-PH process outperforms the Geo-Geo model in general. If lower interference is caused by the Geo-Geo model, it is very likely that the Geo-Geo model has lower spectrum utilization as well. Although the performance of the PH-PH model has been very similar to the Geo-Geo model, we think that if the order of the PH distributions is larger, the PH-PH model has the capacity of achieving better results.

When comparing the performance of the PAP-PH model in the different proactive schemes, higher spectrum utilization are obtained in the fixed frame scheme in general. However, lower interference probability is achieved in the fixed transmission length scheme. Hence, the achievable performance with the evaluated proactive schemes is quite similar. In terms of interference probability, the fixed frame scheme is more conservative.

Chapter 8

Concluding Remarks

8.1 Conclusion and Discussion

Given that the Geo-Geo model is not adequate for all traffic scenarios, we propose the PAP-PH channel-usage model, which is a more general traffic model that can describe bursty channel behaviour patterns. Our design motivation comes from the fact that for many years, researches have concluded that the data traffic can exhibit bursty characteristics and be described by heavy tailed distributions. Moreover, we think that the PAP-PH model maintains reasonable analytical simplicity and there are fitting procedures available to estimate its parameters. These fitting procedures can be designed such that they work without a manual override. That is, the EM-algorithms only expect as inputs the sequence of interarrival samples and the order of the phase type distributions of the traffic model. This results in a highly adaptable algorithm without continuous user intervention.

We have investigated the potential effect of using traditional statistical models, i.e. the Geo-Geo model and the PH-PH model, over a channel that exhibits burstiness in a cognitive radio network scenario. Different secondary user channel access schemes to a single primary user channel have been designed.

First, a reactive channel access scheme, where the SU uses no historical sensing information, has been designed. In this access scheme, the SU has a fixed transmission length,

and a successful transmission only occurs when the entire transmission length has not suffered a collision. We compare two performance measures that an investigator will likely obtain, if a Geo-Geo model assumption is used in a PU channel characterized by a PAP-PH process. These two performance measures are: the average transmission delay and average number of time slots for a collision with a PU.

According to our results, the Geo-Geo model assumption is appropriate to derive the average number of time slots between successive collisions. The reactive scheme has been designed such that the SU has the ability to sense while transmitting. Hence, the effects of the Geo-Geo model can be quite different if the SU cannot sense during a transmission. Our results show that when the PU channel has a low traffic, the Geo-Geo model can be successfully used to approximate the mean transmission delay of the reactive scheme. However, for a PU channel with high bursty traffic, as the SU transmission length increases, the overestimation in the mean transmission delay of the Geo-Geo model with respect to the PAP-PH model increases. From the high bursty traffic example, it can be concluded that the Geo-Geo model assumption is no longer suitable when the transmission length is close or larger than the mean idle intraplatoon period.

If the secondary network has historical sensing information, applying a proactive access scheme is more suitable. In this thesis, we have analysed two proactive access schemes. We compared the achievable performance of the network in terms of spectrum utilization and interference probability. First, we considered a PU channel characterized by a PAP-PH process. Compared with the estimated PAP-PH process, simulation results show that traditional channel usage models underutilize the available white spaces and cause a larger degree of interference to the primary users. This is because the PAP-PH process describes the channel more accurately and the secondary user can exploit the additional information in its decision policy. As the bursty channel usage increases, accurate channel characterization with the PAP-PH process allows the SU to discover more spectrum opportunities while maintaining reasonable interference.

To show the generality of our proposed channel usage, a Geo-Geo PU channel was also

considered. Our findings show that the PAP-PH process is useful for bursty and non-bursty spectrum use patterns.

Special attention has been given to the Geo-Geo model, because we think that if the order of the idle phase type distribution is larger, a PH-PH model has the capacity of improving its performance.

8.2 Proposal for future work

In this research, we have focused on access schemes that involve a single PU channel. The proactive spectrum schemes can be further investigated for multiple PU channels. As explained in *Section 2.4.2*, many new concerns rise with multiple PU channels. For instance, an SU needs to efficiently choose the channels to continue transmission. We expect that an accurate characterization of the PU channels, in the channel ordering scheme, will improve the identification of the spectrum opportunities. Another example is the implementation of a proactive switching. A proactive switching may allow a higher utilization of the white spaces. Consider bursty channels described by PAP-PH processes. Intuitively, a smart switch will occur when an SU switches to a different channel when the current idle period is likely to be of intraplatoon type. These and other cognitive radio functions can be used with a PAP-PH channel-usage model assumption.

Another interesting research direction is to incorporate imperfect sensing in the secondary network. Further analyses are needed to measure the effects of imperfect sensing in the proactive spectrum access schemes. The effects can be similar to those encountered with the fixed frame model with sleep state explained in *Section 7.3*.

Since the busy duration of the traffic can exhibit heavy-tailed behaviour, these heavy-tailed busy periods can be incorporated to the traffic model by describing the busy periods with a platoon arrival process. Hence, a PH-PAP model is a potential candidate to characterize the heavy-tailed busy periods in discrete time. By alternating the state representation of the PAP-PH model, the heavy-tailed busy periods can be analysed in the simulations.

Even a more complex model can involve the usage of platoon arrival processes for the idle and busy periods of a PU channel. Further investigation is needed to study their effects in deriving the performance measures in the spectrum access schemes.

Finally, the busy period distributions can be used for a smart sensing scheme, where an SU remains idle for longer periods when the channel is likely to remain busy. Energy aware schemes can be further designed.

Bibliography

- [1] P. Tran-Gia, D. Staehle, and K. Leibnitz, "Source traffic modeling of wireless applications," *AEU - International Journal of Electronics and Communications*, vol. 55, no. 1, pp. 27 – 36, 2001.
- [2] V. Paxson and S. Floyd, "Wide area traffic: the failure of Poisson modeling," *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, pp. 226–244, Jun. 1995.
- [3] P. Owezarski and N. Larrieu, "Internet traffic characterization - an analysis of traffic oscillations," *7th IEEE International Conference on High Speed Networks and Multimedia Communications, Toulouse, France*, pp. 96–107, 2004.
- [4] W. chang Feng, F. Chang, W. chi Feng, and J. Walpole, "A traffic characterization of popular on-line games," *Networking, IEEE/ACM Transactions on*, vol. 13, no. 3, pp. 488 – 500, june 2005.
- [5] P. Wang and I. Akyildiz, "On the origins of heavy-tailed delay in dynamic spectrum access networks," *Mobile Computing, IEEE Transactions on*, vol. 11, no. 2, pp. 204 –217, feb. 2012.
- [6] P. Nain, "Impact of Bursty Traffic on Queues," *Statistical Inference for Stochastic Processes*, vol. 5, pp. 307–320, 2002.
- [7] A. S. Alfa and M. Neuts, "Modelling Vehicular Traffic using the Discrete Time Markovian Arrival Processes," *Transportation Science*, vol. 29, no. 2, pp. 109–117, 1995.
- [8] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [9] I. Akyildiz, W. Lee, M. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.
- [10] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.
- [11] Q. Zhao and B. M. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine*, pp. 79–89, May 2007.
- [12] L. Yang, L. Cao, and H. Zheng, "Proactive channel access in dynamic spectrum networks," *Physical Communication*, vol. 1, no. 2, pp. 103–111, Jun. 2008.
- [13] A. T. Hoang, Y.-C. Liang, D. T. C. Wong, Y. Zeng, and R. Zhang, "Opportunistic spectrum access for energy-constrained cognitive radios," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1206–1211, Mar. 2009.

- [14] Q. Zhao, S. Geirhofer, L. Tong, and B. Sadler, "Optimal dynamic spectrum access via periodic channel sensing," *Wireless Communications and Networking Conference*, pp. 33–37, march 2007.
- [15] Y. Fang, "Hyper-Erlang Distribution Model and its Application in Wireless Mobile Networks," *Wireless Networks*, vol. 7, no. 3, pp. 211–219, May 2001.
- [16] M. Neuts and S. Chakravarthy, "A single server queue with Platooned Arrivals and phase type services," *European Journal of Operational Research*, vol. 8, pp. 379–389, 1981.
- [17] A. S. Alfa and H. A. Ghazaleh, "Goodput Analysis Using Terminating MAP for a Class of Discrete-Time Queueing Models," *Stochastic Models*, vol. 27, no. 4, pp. 615–628, Oct. 2011.
- [18] H. Heffes and D. Lucantoni, "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance," *IEEE Journal on Selected Areas in Communications*, vol. 4, no. 6, pp. 856–868, Sep. 1986.
- [19] G. Yuan, R. C. Grammenos, Y. Yang, and W. Wang, "Performance Analysis of Selective Opportunistic Spectrum Access With Traffic Prediction," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 4, pp. 1949–1959, 2010.
- [20] S. Geirhofer, L. Tong, and B. Sadler, "Cognitive radios for dynamic spectrum access—dynamic spectrum access in the time domain: Modeling and exploiting white space," *Communications Magazine, IEEE*, vol. 45, no. 5, pp. 66–72, May 2007.
- [21] T. C. Clancy and B. D. Walker, "Predictive Dynamic Spectrum Access," *SDR Forum Conference*, 2006.
- [22] Z. Chen and R. C. Qiu, "Prediction of channel state for cognitive radio using higher-order hidden Markov model," *Proceedings of the IEEE SoutheastCon*, pp. 276–282, Mar. 2010.
- [23] P. A. Kumar Acharya, S. Singh, and H. Zheng, "Reliable open spectrum communications through proactive spectrum access," *Proceedings of the first international workshop on Technology and policy for accessing spectrum - TAPAS '06*, 2006.
- [24] H. Kim and K. Shin, "Efficient discovery of spectrum opportunities with mac-layer sensing in cognitive radio networks," *Mobile Computing, IEEE Transactions on*, vol. 7, no. 5, pp. 533–545, may 2008.
- [25] L. Breuer and A. Alfa, "An EM Algorithm for Platoon Arrival Processes in Discrete Time," *Operations Research Letters*, vol. 33, no. 5, pp. 535–543, Sep. 2005.
- [26] L. J. Rodriguez Esparza, "Maximum likelihood estimation of phase-type distributions," Ph.D. dissertation, 2010.
- [27] K. W. Sung, S.-L. Kim, and J. Zander, "Temporal spectrum sharing based on primary user activity prediction," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 12, pp. 3848–3855, december 2010.
- [28] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency working group," ET Docket No. 02-135, Nov. 2002. [Online]. Available: <http://www.fcc.gov/sptf/reports.html>
- [29] FCC, "Notice of proposed rule making and order," ET Docket N0 03-222, Dec. 2003.

- [30] B. Wang and K. Liu, "Advances in cognitive radio networks: A survey," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 1, pp. 5–23, feb. 2011.
- [31] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, may 2009.
- [32] D. Datla, R. Rajbanshi, A. Wyglinski, and G. Minden, "Parametric adaptive spectrum sensing framework for dynamic spectrum access networks," *New Frontiers in Dynamic Spectrum Access Networks, 2nd IEEE International Symposium on*, pp. 482–485, april 2007.
- [33] M. Hoyhtya, S. Pollin, and A. Mammela, "Performance improvement with predictive channel selection for cognitive radios," *First International Workshop on Cognitive Radio and Advanced Spectrum Management*, pp. 1–5, Feb. 2008.
- [34] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589–600, Apr. 2007.
- [35] S. Huang, X. Liu, and Z. Ding, "Opportunistic spectrum access in cognitive radio networks," *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pp. 1427–1435, april 2008.
- [36] A. S. Alfa, *Queueing Theory for Telecommunications: discrete time modelling of a single node system*. Springer US, 2010.
- [37] P. Buchholz, "An EM-Algorithm for MAP Fitting from Real Traffic Data," *Computer Performance Evaluation. Modelling Techniques and Tools*, vol. 2794, pp. 218–236, 2003.
- [38] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, 1st ed. Kluwer Academic Publishers, 1992.
- [39] J. G. Proakis and S. Masoud, *Fundamentals of Communication Systems*, 1st ed. Prentice Hall, 2004.
- [40] T. Karagiannis, M. Molle, and M. Faloutsos, "Long-range dependence ten years of Internet traffic modeling," *IEEE Internet Computing*, vol. 8, no. 5, pp. 57–64, Sep. 2004.
- [41] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, 1994.
- [42] A. Callado, C. Kamienski, G. Szabo, B. Gero, J. Kelner, S. Fernandes, and D. Sadok, "A survey on internet traffic identification," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 3, pp. 37–52, third quarter 2009.
- [43] A. Sinha, K. Mitchell, and D. Medhi, "Flow-level upstream traffic behavior in broadband access networks: Dsl versus broadband fixed wireless," *IP Operations and Management, 3rd IEEE Workshop on*, pp. 135–141, oct. 2003.
- [44] G. He, J. Hou, W.-P. Chen, and T. Hamada, "One Size Does Not Fit All: A Detailed Analysis and Modeling of P2P Traffic," *IEEE Global Telecommunications Conference*, pp. 393–398, Nov. 2007.

- [45] M. Wellens, A. de Baynast, and P. Mahonen, "PHY 28-2 - Exploiting Historical Spectrum Occupancy Information for Adaptive Spectrum Sensing," *IEEE Wireless Communications and Networking Conference*, pp. 717–722, Mar. 2008.
- [46] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic Spectrum Access in WLAN Channels: Empirical Model and Its Stochastic Analysis," *Proceedings of International Workshop on Technology and Policy for Accessing Spectrum*, 2006.
- [47] M. Wellens, J. Riihijärvi, and P. Mähönen, "Empirical time and frequency domain models of spectrum use," *Physical Communication*, vol. 2, no. 1-2, pp. 10–32, Mar. 2009.
- [48] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. Elsevier/Academic Press, 2009.
- [49] M. R. Gupta, "Theory and Use of the EM Algorithm," *Foundations and Trends in Signal Processing*, vol. 4, no. 3, pp. 223–296, 2010.
- [50] L. Breuer and A. Kume, "An EM Algorithm for Markovian Arrival Processes Observed at Discrete Times," *Lecture Notes in Computer Science. Measurement, Modelling, and Evaluation of Computing Systems and Dependability and Fault Tolerance*, pp. 242–258, 2010.