Rana Al Halaseh

Interference-Aware Spectrum Occupancy Prediction for Cognitive Radio Networks

kassel university press

Rana Al Halaseh

Interference-Aware Spectrum Occupancy Prediction for Cognitive Radio Networks

kassel university press

This work has been accepted by the Faculty of Electrical Engineering / Computer Science of the University of Kassel as a thesis for acquiring the academic degree of Doktor der Ingenieurwissenschaften (Dr.-Ing.).

Supervisor: Univ.-Prof. Dr. sc. techn. Dirk Dahlhaus, University of Kassel
Co-Supervisor: Univ.-Prof. Dr. rer. nat. Rudolf Mathar, RWTH Aachen
Defense day: 19th April 2018

Bibliographic information published by Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.dnb.de.

Zugl.: Kassel, Univ., Diss. 2018
ISBN 978-3-7376-0552-6 (print)
ISBN 978-3-7376-0553-3 (e-book)
DOI: http://dx.medra.org/10.19211/KUP9783737605533
URN: http://nbn-resolving.de/urn:nbn:de:0002-405532

www.upress.uni-kassel.de

Printed in Germany

Acknowledgment

I would like to express my sincere appreciation to Univ.-Prof. Dr. sc. techn. Dirk Dahlhaus for the opportunity to work at Communications Laboratory (ComLab) towards my Ph.D. study. His continuous support and guidance has always been valuable. Furthermore, I would like to thank Univ.-Prof. Dr. rer. nat. Rudolf Mathar for being my second supervisor. I also would like to thank him for the positive atmosphere he has created at my current position in Aachen.

I would like to thank my colleagues from ComLab in the University of Kassel. I appreciate the valuable discussions we had. Specially, I would like to thank Dr.-Ing. Marc Selig for his readiness to help. I also would like to thank my colleagues in Institute für Theoretische Informationstechnik (TI) at Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen for reviewing my dissertation, and for their valuable comments and the help they provided me through the final stages of my Ph.D. A special thank goes to Dr.-Ing. Gholamreza Alirezaei for giving me the last push, and for his priceless advice.

I dedicate this work to my parents. I would have never reached this point without you both by my side. Also, I dedicate this work to my sisters Rania and Rand, and my brother Khaleel for their support, understanding and continuous encouragement.

Finally, I would like to thank my countless friends in Germany. Thank you for the wonderful times we have had together. I cherish your friendship. A special thank goes to Helmut for all the unforgettable memories, laughter, and photos I shared with him.

Zusammenfassung

Da die Zahl der drahtlosen Protokolle, Dienste und Geräte in den letzten Jahrzehnten stetig zugenommen hat, steigen die Anforderungen und Herausforderungen an die verfügbaren Spektrumsressourcen weiter an. Tag für Tag ist der Zuweisungsprozess von festen Teilen des Spektrums zu bestimmten Diensten weniger geeignet für die zukünftigen technischen Herausforderungen. Dies führt zu Ineffizienzen in der Frequenznutzung, da die nicht lizenzierten Frequenzbänder im Gegensatz zu den zeitweise ungenutzten lizenzierten Bändern überfüllt und überlastet sind. Als Lösung dieser aktuellen Probleme und Herausforderungen kann eine intelligenten Funkplattform eingesetzt werden. Diese Plattform muss in der Lage sein, die verfügbaren Ressourcen effizient zu verwalten, indem sie ihre Parameter an die Umgebung anpasst. Dies kann erreicht werden, indem die spektrale Umgebung basierend auf ihrem vorherigen Verhalten präzise erkundet wird.

Um die oben genannte Aspekte in einem Funksysteme zu adressieren, wurde 1999 das Konzept eines kognitiven Radios (CR) eingeführt mit dem Ziel, das verfügbare Spektrum zwischen lizenzierten primären Benutzern (PUs) und nicht lizenzierten sekundären Benutzern (SUs) effizient aufzuteilen. In diesem Szenario sind die SUs in der Lage, dynamisch auf das Spektrum unabhängig von der primären festen Zuweisungsrichtlinie zuzugreifen. Mit anderen Worten, die SUs sind in der Lage, die lizenzierten Frequenzbänder für ihre eigene Übertragung zu erkunden, sofern die Übertragung der PUs dadurch nicht beeinträchtigt wird. Aus diesem Grund benötigt ein CR-Netzwerk (CRN) ein zuverlässiges Spektrumbelegungsvorhersageschema in Kombination mit Spektrumserfassungs- und Spektrumzugriffsmechanismen. Dies ermöglicht den SUs, die Nutzung der spektralen Ressourcen durch die PUs zu erlernen, ihre zukünftige Aktivität vorherzusagen und, falls nötig, ihre Sendeaktivität auf ein anderes freies Band zu verlagern.

Diese Arbeit stellt verschiedene Spektralbelegungsvorhersagen auf Basis von Hidden-Markov-Modellen (HMMs) für Filterbankmehrträger-Übertragungsverfahren (FBMC) vor. Durch den Einsatz einer diskreten Fourier-Transformation-modulierten Filterbank (DFT-FB) kann der SU-Empfänger die Zeit-Frequenz-Eigenschaften des empfangenen Signals, das in einem geeigneten Vorhersage-Schema verwendet werden soll, effizient untersuchen.

Zunächst stellt diese Arbeit einen neuartigen Vorhersageansatz in Form eines kontinuierlichen HMM (CHMM)-basierten Schemas für ein System unabhängiger Teilbänder als Alternative zu dem weit verbreiteten diskreten HMM-(DHMM) basierten Schema vor. Damit entfallen die Nachbearbeitungsmechanismen, die bei den derzeit verfügbaren Prognoseverfahren zum Einsatz kommen. Zweitens wird der vorgeschlagene CHMM-basierte Prädiktor durch die Einführung eines gekoppelten CHMM-basierten TF-Spektrum-Prädiktionsschemas für ein System von abhängigen Teilbändern erweitert, so dass alle Abhängigkeiten zwischen benachbarten Teilbändern in einem geeigneten Belegungsmodell erfasst und in die Prädiktionsentscheidung einbezogen werden können.

Der Vorhersageprozess für die eingeführten Spektrumbelegungsvorhersageschemata besteht aus drei aufeinanderfolgenden Phasen: 1) Eine anfängliche Pilotphase als Teil der Spektrumserfassung, die erforderlich ist, um ein bedingtes Modell des empfangenen Signals basierend auf der Verfügbarkeit des PU-Signals aufzubauen - 2) Eine Lernphase zum Erlernen des PU-Verkehrsmusters, um die Parameter des betrachteten Vorhersagesystems entsprechend anzupassen - 3) Eine Anwesenheitsvorhersagephase.

Die Leistungsfähigkeit der vorgestellten TF-Vorhersagesysteme wird in einer interferenzbasierten Umgebung untersucht. Im Gegensatz zu den meisten Arbeiten in diesem Bereich, welche größtenteils Gaußsches Rauschen betrachten, modellieren wir Interferenzen in dieser Arbeit als nicht-Gaußsches Rauschen. Diese Interferenzterme ergeben sich aus der Überlagerung von Signalen von mehreren Benutzern, die gleichzeitig über das gleiche Teilband senden. Das Rauschen wird mit Hilfe einer räumliche Matérn-Hard-Core-Verteilung modelliert, die die Interferenz in einem Carrier-Sense-Multiple-Access-/Kollisionsvermeidungs-Netzwerk (CSMA/CA) in einem beschränkten Bereich widerspiegelt.

Abstract

As the number of wireless protocols, services, and devices have increased in the last decades, demands and challenges facing the available spectrum resources are continuing to grow. Day by day, the process of assigning fixed spectrum portions to certain services and protocols becomes more challenging. It creates inefficiency in spectrum usage as the unlicensed spectrum bands are crowded and over-utilized, unlike the mostly idle licensed bands. Therefore, an intelligent radio platform must be employed to solve these current issues and challenges. This solution must be able to manage the available resources efficiently by adapting its parameters according to the surrounding environment. This can be achieved by properly learning the spectral environment based on its previous behavior and behave accordingly.

To address the problem outlined above, the concept of cognitive radios (CRs) was introduced in 1999 to efficiently allocate the available spectral resources between licensed primary users (PUs) and unlicensed secondary users (SUs). In this setting, the SUs are able to dynamically access the spectrum irrespective of the available fixed assignment policy. In other words, the SUs are able to explore the licensed spectrum bands for their own transmission, provided they do not disturb the transmission of any PUs. For this reason, a CR network (CRN) requires a reliable spectrum occupancy prediction scheme in combination with spectrum sensing and spectrum access mechanisms. This allows the SUs to learn the PUs' traffic pattern, predict its future activity, and switch its transmission to another free subband when required.

This work presents different hidden Markov model (HMM) based spectrum occupancy prediction schemes for filter bank based multi-carrier (FBMC) transmission applications. By employing a discrete Fourier transform-modulated filter bank

(DFT-FB), the SU receiver can efficiently explore the time-frequency (TF) characteristics of the received signal to be utilized in a suitable prediction scheme.

First, this work presents a novel prediction approach in the form of a continuous HMM (CHMM) based scheme for a system of independent subbands as an alternative approach to the widely implemented discrete HMM (DHMM) based scheme. Hence, the post-processing mechanisms which are employed by the currently available prediction schemes are eliminated. Second, due to its superior performance with respect to the DHMM based scheme, the proposed CHMM based predictor is expanded by introducing a coupled CHMM based TF spectrum occupancy prediction scheme for a system of dependent subbands, such that any interdependencies between neighboring subbands are captured in a proper occupancy model as well as included in the prediction decision.

The prediction process for the introduced spectrum occupancy prediction schemes consists of the three successive phases. 1) An initial pilot phase as part of spectrum sensing required to construct a conditional model of the received signal based on the availability of the PU signal. 2) A learning phase to learn the PU's traffic pattern in order to accordingly adapt the parameters of the prediction scheme under consideration. 3) An occupancy prediction phase.

The performance of the presented TF prediction schemes is investigated in an interference-based environment. Unlike most approaches in this field which mostly consider Gaussian interference, in this work we consider a non-Gaussian interference. This intereference arises from superposition of signals from multiple users simultaneously transmitting over the same subband. It has been modeled as a Matérn hard-core spatial distribution, which reflects the interference in a carrier sense multiple access/collision avoidance (CSMA/CA) network in a finite area.

Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Motivation

New wireless technologies and services have emerged recently and become an important part of our daily lives. For example, the emerging of the Internet of Things (IoT) has provided a revolution of wireless connectivity beyond the traditional fashion, as data can be transmitted over the air between objects and without human interaction. Furthermore, due to the exponential growth of wireless traffic transmitted over the available spectrum bands in the last decades, the unlicensed spectrum bands – also known as the industrial, scientific and medical (ISM) bands – are found to be highly over-utilized. On the other hand, only 38% of the licensed spectrum bands are actually occupied [3, 4], hence a non-uniform and inefficient use of spectrum utilization is introduced. Due to the introduction of new wireless services and technologies as well as an increasing number of devices, wireless networks – especially those operating in the ISM bands – suffer from performance degradation. This can be in the form of Quality of Service (QoS) reduction due to packet loss, the presence of co-channel interference (CCI), and the inability of transmission due to unavailable free spectrum band [5].

Furthermore, based on the traffic forecast update presented in [6] for the time interval 2015-2020, it has been declared that "mobile data traffic has grown 4,000-fold over the past 10 years and almost 400-million-fold over the past 15 years" alone, and "there will be 11.6 billion mobile-connected devices by 2020". Thus, the concept of meeting the current demands in the form of lower latencies, as well as higher

data rate and traffic capacity [7], has been adopted as the 5G telecommunication standards, where cognitive radio (CR) is one of its basic and key enablers [8].

1.2. Concepts of Cognitive Radio

First introduced by Joseph Mitola in 1999 [9], CR technology has come up with a new intelligent radio architecture or platform which integrates different fundamentals such as signal processing, mathematical and computing theories, and machine learning methods [10]. This intelligent wireless platform is capable of utilizing the available spectrum resources efficiently, as it must be aware of its surroundings and able to adapt to any changes in the environment according to the current requirements [11]. This adaptation is in the form of updating the system's parameters, such as transmit-power, carrier-frequency, and modulation strategy [11]. Moreover, the author in [11] has stated that such adjustments must be performed in real-time, taking into account that reliable communications are available on demand within efficiently utilized radio resources.

Licensed spectrum bands are indeed exploited by unlicensed secondary users (SUs) within CR networks (CRNs) under the condition that they must not collide or interfere with the available licensed primary users (PUs), where the latter would maintain a certain predefined QoS [12]. Therefore, CRN must be intelligent enough to perform the following cognitive functionalities: during *spectrum sensing* the SU receiver detects the availability of any PU by sensing the environment and defining those channels which are free from any PU within the sensed time duration. Real-time spectrum sensing results are afterwards utilized to perform a *spectrum access decision*. However, as the information obtained from spectrum sensing would no longer be up-to-date due to several delays introduced from the hardware itself [13], a collision between SU and PU would still occur. Therefore, spectrum sensing is combined with a reliable *spectrum occupancy prediction* scheme to learn the PU's traffic pattern and predict its availability at future time slots in order to minimize the collision probability. Then, the SU would choose the most appropriate spectrum band to utilize for its own transmission upon its requirements, ensuring a smooth

shifting between free subbands as well as resource sharing among multiple coexisting SUs [14].

The CR approach can be carried out within three cognitive paradigms [15]. A SU attempts to transmit its signal over a certain licensed band *only* if PUs are absent. In other words, the SU must not coexist with a PU in the same channel. This approach is similar to an on-off transmission from the SU's point of view, and it is known as an *interweave* paradigm. Due to this on-off performance, synchronization between the SU transceiver has to be repeatedly performed each time it is allowed to access a channel.

A more efficient yet complicated scheme is to access the licensed bands continuously in an *underlay* paradigm, regardless of the presence of any PU. In this approach, both SU and PU transmit simultaneously over a common channel, however the interference level measured at the PU receiver due to the SU's transmission should be less than a predefined value. Hence, the PU still maintains its predefined QoS. For this matter, different adaptive power schemes come in handy [16] to optimize the transmission power level of the SU signal. As a result, the SU must be able to spread its transmission below the noise/interference floor of the PU, similar to the basic idea behind spread spectrum and ultra-wide band communication techniques.

Finally, an *overlay* paradigm is by far the most advanced and sophisticated approach, assuming the SU to have a full knowledge of the PU transmission; including its message and the codebook. Dirty paper coding techniques are utilized for this purpose [17]. As a result, both SU and PU transmit simultaneously over a common channel, where the SU assigns part of its power to deliver its own message, *and* assists the PU by re-transmitting its messages [18]. Such a mechanism can be performed both in licensed and unlicensed bands, as the user's transmission would be interference-free and improved. On the other hand, the SU carries on with its own transmission and assures efficient spectrum utilization. Moreover, hybrid schemes based on multiple paradigms can also be utilized to maximize the spectrum efficiency as explained in [19].

1.3. Spectrum Occupancy Prediction in Cognitive Radio Networks

Irrespective of the considered paradigm, spectrum occupancy prediction plays an important role in CRNs. Not only that it provides future information regarding PU's existence to avoid any possible collision, spectrum occupancy prediction can also be utilized to optimize the CRN. Future spectrum occupancy information can be fed back to spectrum sensing in order to sense only a subset of channels, and skip those which are predicted to be free [13, 20], hence the energy consumption is reduced. In addition, the channels which are predicted to be free can be shared efficiently by multiple SUs based on their requirements and the predicted quality of the channels [21, 22].

Similar to spectrum sensing, spectrum occupancy prediction depends on the signal received (sensed) by the SU receiver in order to model the PU occupancy at each subband, understand its traffic pattern, and predict its future occupancy. In order to obtain a high resolution of the received signal both in time and frequency domains, the SU receiver employs an appropriate time-frequency representation (TFR) method, where the received signal is mapped from the time-domain into the time-frequency (TF) domain. For this purpose, filter bank (FB) transmission has been considered in this work.

FB based multi-carrier (FBMC) transmission has captured the attention of many in signal processing and wireless communication fields. Initially, it was employed in speech compression and image processing applications [23]. Similar to the widely known and conventional orthogonal frequency division multiplexing (OFDM) scheme, FB provides an efficient method to transmit multiple messages simultaneously via parallel subbands over the available bandwidth. Mainly, FBMC consists of two parts: analysis and synthesis filters. In the analysis FB (AFB), the signal is decomposed or broken down into its TF components or sub-signals, whereas the original signal can be reconstructed given its sub-signals in the synthesis FB (SFB) [24]. Hence, FB transmission becomes the fundamental structure of transceivers within multi-carrier systems, where the transmitter and receiver are configured based on the SFB and AFB, respectively.

Different from OFDM, a properly designed FB - as the filters prototype are designed based on a certain optimized window function or pulse - is capable of eliminating both intra-band and cross-band inter-symbol interference (ISI) [25, 26] without the need of a guard period. Among the different implementation methods and approaches, discrete Fourier transform (DFT) FBs have been widely utilized to implement both AFB and SFB in a FBMC system, due to its simplicity and low implementation costs [26]. Thus, the DFT - modulated FB (DFT-FB) has been utilized in this work. Further discussion is presented in Chapter 4.

1.4. Literature Review

The previously explained CRN functionalities have been widely addressed in the literature. However, spectrum occupancy prediction has been the least addressed functionality, in comparison to spectrum sensing for example. This section presents a literature review concerning spectrum occupancy prediction schemes. Subsequently, the contribution of this work is introduced at the end of this chapter.

Authors in [27, 28] have argued that spectrum access decisions based on real-time measured spectrum sensing signals create collisions with PUs, due to delays introduced by the hardware platform. To overcome this matter, a hidden Markov model (HMM) based scheme has been employed to predict the channel state to improve the system performance. Thus, SU would select the best available channel to utilize for its transmission based on the prediction information obtained via HMM. However, the proposed predictor structure contains the fast Fourier transform (FFT) to map the time-domain received signal into the frequency-domain, followed by a quantization process for discretization purposes to fit it into the designed HMM prediction scheme.

Carrier sense multiple access (CSMA) spectrum allocation methods have been discussed in [29] by introducing a HMM based dynamic spectrum access technique. The proposed technique predicts the duration a channel is going to be idle, hence the SU can utilize it for its own transmission. However, the proposed scheme makes sure

that the SU would switch to another free channel before a legitimate PU, which is assumed to follow a Poisson distribution, starts its transmission on that same channel. In a similar method, [30] proposes a frequency-hopping algorithm combined with a HMM based channel occupancy prediction scheme. The proposed scheme allows the SU to switch its transmission between the unoccupied channels by PUs, according to a specific hopping sequence generated by the SU itself and based on the gained future occupancy information. Furthermore, in an underlay CRN, the interference temperature dynamics of a PU has been modeled as a HMM scheme in [31]. The proposed scheme provides the SU with the future interference temperature in order to select the subbands for its underlay transmission.

Apart from HMM based approaches, a neural network (NN) has been utilized in CRN as another example of machine learning schemes. A NN based multilayer perceptron (MLP) predictor has been introduced in [32, 33], which saves energy consumption during spectrum sensing as well as improves the spectrum utilization. Similar to HMM schemes and any machine learning method, MLP requires employing a set of past occupancy information of a certain channel in order to predict its future occupancy. Such NN based approaches come in handy for a CRN, as they require no prior knowledge of the channel statistics, which is unavailable for SUs in practice. Likewise, [34] employs back propagation NN as a spectrum predictor for improving energy consumption and spectrum access performance. Moreover, an online support vector regression (SVR) scheme has been proposed in [35] for a power regression based online learning technique, in which the PU traffic follows a Poisson distribution. The proposed prediction scheme keeps adapting its parameters based on any changes in the considered traffic.

On the other hand, channel quality prediction has been investigated in [21]. First, the authors present a spectrum sensing approach based on a non-stationary HMM (NSHMM) where its parameters are estimated via Bayesian inference approach and Gibbs sampling. Thus, the NSHMM reflects the expected duration of a channel's state as well as spectrum sensing accuracy, which is to be utilized in a channel quality metric. Due to the non-stationarity behavior of the wireless environment, a NSHMM has been chosen such that the transition probabilities are a function of time duration of the PU existence in a current state.

Different spectrum modeling and occupancy prediction approaches can be found in [36]. For example, time series analysis methods have been utilized in [37] to model and predict the occupancy information of the Global System for Mobile communications (GSM) band via autoregressive moving average (ARMA) approach. A similar approach has also been followed in [38] by utilizing autoregressive (AR) models. Furthermore, AR based spectrum hole prediction has been implemented in [39], based on a second-order scheme denoted as [AR-2] along with Kalman filter. A spectrum hole in this content is another term to describe an unoccupied subband by PU. The coefficients of the [AR-2] model are evaluated via Yule-Walker equations. Moreover, being known as a classical statistical approach for time series modeling and prediction, Kalman smoother has been employed in [40] to predict short and long-term behavior of signals, assuming linear Gaussian assumptions and following a cross-validation approach to determine the noise densities.

As spectrum occupancy prediction scheme is a challenging task which consumes energy [13], authors in [22] discuss spectrum prediction schemes from an energy consumption point of view. Thus, an energy efficient prediction strategy based on numerical results is proposed, in which prediction strategies that maximize the energy efficiency based on a parametric optimization problem are implemented. With respect to energy consumption, [20] proposes an exponential moving average (EMA) based spectrum prediction approach to improve spectrum sensing performance. An EMA based predictor provides the future energy level so that the SU can skip executing spectrum sensing on those channels whose future energy is higher than a predefined threshold. Thus, reducing energy consumption is due to redundant spectrum sensing.

1.5. Thesis Contribution

Apart from speech and pattern recognition, computational molecular biology, and stock market prediction [41], HMM schemes have been widely implemented in CRNs, as stated in Section 1.4. It has been considered for spectrum sensing purposes [42], spectrum occupancy modeling based on the PU traffic [43], and spectrum occupancy prediction [44, 29, 45, 32]. So far, the considered schemes for CRN, as presented in

Section 1.4, are basically based on discrete HMM (DHMM) schemes. They rely on discretization processes to map the observed or received signal by SUs into discrete-valued observations, or in other words, employ spectrum sensing output into a spectrum occupancy prediction scheme.

Nevertheless, DHMM based schemes are normally associated with information loss as they require such discretization process to fit the continuous-valued measurements or observed signal into the DHMM scheme. However, due to information loss the performance of the DHMM based scheme is decreased irrespective of its application. For example, during vector quantization of speech signals, degradation in its performance as a speech recognizer occurs [46]. It is known that a DHMM based scheme is best suited for applications where discretization does not introduce any significant loss; such as object [47] and gesture recognition [48]. Furthermore, the available prediction approaches in literature, stated in Section 1.4, are based on utilizing time-domain signals after conventional subband filtering, which prevent those predictors from fully exploiting the received signal [2]. In addition, none of the available spectrum occupancy prediction schemes in literature have considered the interdependencies between neighboring subbands into a suitable model nor in the prediction decision

This work proposes a novel first-order continuous HMM (CHMM) based spectrum occupancy prediction scheme as a more natural and an alternative approach to the conventional DHMM scheme. Within this approach, the SU receiver employs a proper TFR method to efficiently explore the TF characteristics of the received signal. For this purpose, a properly designed DFT-FB has to be considered within the SU transceiver structure, such that the observed signal at the output of the AFB at the SU receiver is utilized directly. Thus, spectrum sensing, spectrum occupancy prediction, and spectrum access are efficiently implemented in one structure [25].

For performance comparison purposes, a first-order DHMM based TF spectrum occupancy prediction scheme is implemented in this work. To fit the signal at the output of the AFB into the DHMM based scheme, a Neyman-Pearson (NP) test is utilized as part of its implementation. For both DHMM and CHMM based prediction schemes, the neighboring subbands are considered mutually independently

occupied by the PU signal. Hence, the prediction schemes have been performed at each subband separately as presented in a previous work in [2].

Subsequently, due to its superior performance, the proposed CHMM based prediction scheme is expanded in the form of a lag-one first-order coupled CHMM based scheme which takes into account the interdependencies between neighboring subbands for modeling and prediction purposes. It is worth mentioning at this point that none of the prediction schemes available in literature have considered the interdependencies between adjacent subbands. In this work, the future occupancy state of a certain subband depends on its current state as well as the current state of its neighboring subbands, which are later referred to as the parent subbands or nodes.

As the CCI plays a significant role in defining the performance of wireless networks [49], its impact cannot be neglected in this work. Therefore, simulations have been conducted in a previous work in [1] to obtain the CCI signal arising from multiple transmitters simultaneously utilizing the same subband. Thus, the performance of the presented prediction schemes has been evaluated in the form of probability of error due to wrong predictions in different environments characterized by the signal-to-noise ratio (SNR) as well as the signal-to-interference ratio (SIR).

1.6. Outline

After stating the motivation behind this work in Chapter 1 along with the significant part played by spectrum occupancy prediction in CRNs, Chapter 2 presents the important concepts and tools for this work. Chapter 3 presents a short literature overview of the standards operating in the ISM band and their mutual interference while coexisting in the same area and interference spatial modeling. Furthermore, the SU transceiver structure is presented in Chapter 4, along with the considered TFR method. Additionally, modeling the CCI signal based on a previous work in [1] is also explained in Chapter 4.

The proposed (uncoupled) CHMM based TF spectrum occupancy prediction scheme for a system of independent subbands is introduced in Chapter 5, as well as the conventional widely used DHMM based scheme. This chapter is based on a previous

work found in [2]. Chapter 6 presents a novel approach for a system of dependent sub-bands in the form of a lag-one first-order coupled CHMM based prediction scheme. This scheme can be considered as a generalization of the CHMM based scheme.

Performance evaluation and simulation consideration of the prediction schemes are presented in Chapter 7, where a comparison between the reliability of those schemes is drawn in the form of the probability of error or wrong prediction in an interference based environment. Furthermore, the main parameters affecting the reliability of the presented schemes are also discussed.

Finally, conclusions, drawbacks, wireless traffic characteristics along with open issues and further considerations are collected in Chapter 8. Useful concepts, derivations, and algorithms can be found in Appendix A and Appendix B.

2 Chapter 2. Mathematical Notations and Markov Models

After presenting the motivation behind this work and its contribution with respect to the available state of the art approaches, this chapter presents the main concepts and tools to prepare the reader for the following chapters in this work. First, it defines the mathematical notations followed in this work. Second, it presents the main concepts behind graphical models and their types and forms reaching to hidden Markov models. Those concepts are to be put to use in the following chapters.

2.1. Notations

Throughout this work, the capital characters $\mathbf{X}$ and X denote random variables, and the small characters $\mathbf{x}$ and x are realizations or deterministic quantities. Boldface characters define matrices of suitable dimensions including vectors and non-boldface characters denote scalars. For an arbitrary random variable X, the following expressions: $\Pr\{X\}$, $\Pr\{X|Y\}$, $\mathbf{E}\{X\}$, and $|X|$ denote the probability of X, its conditional probability given an arbitrary random variable Y, its expectation, and its absolute value, respectively. The notations $\mathbf{X}_t, X_t, \mathbf{x}_t$ and x_t denote the random process and the corresponding realization observed at slot t, where t in this work refers to a time slot, a frequency slots, or a location, when appropriate. In this work, while representing random processes graphically in terms of graph models, the known processes have been graphically presented as shadowed nodes. Further

explanations in this regard are shown in this chapter. Moreover, $\det \mathbf{A}$, $\operatorname{diag}(\mathbf{A})$ denote the determinant of any arbitrary matrix $\mathbf{A}$ and its diagonal as a column vector, respectively. The transpose and the Hermitian transpose of matrix $\mathbf{A}$ are given by $\mathbf{A}^{\mathrm{T}}$, and $\mathbf{A}^{\dagger}$, respectively. $\mathbf{I}_K$ is a $K \times K$ identity matrix, where K is a positive integer.

An example of discrete time signals is given in the form of $f[n]$, where $n \in \mathbb{Z}$ throughout this work denotes a discrete time index. Its z-transform is denoted as $F(z)$ and given by

$$F(z) = \sum_{n \in \mathbb{Z}} f[n] \, z^{-n}.$$

In this work, the complex conjugate of a function $f[n]$ is given by $f^*[n]$. The projection of an arbitrary discrete value function or vector $f_1[n]$ onto another one $f_2[n]$ via the inner product is denoted by $\langle f_1[n], f_2[n] \rangle$.

The Cartesian product of any arbitrary two closed sets Υ_1 and Υ_2 is given by

$$\Upsilon_1 \times \Upsilon_2 = \{(\upsilon_1, \upsilon_2) \mid \upsilon_1 \in \Upsilon_1, \upsilon_2 \in \Upsilon_2\}.$$

Furthermore, $p_{\mathbf{X}}(\mathbf{x}) \sim \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the probability density function (PDF) of a real-valued $L_{\mathbf{x}}$-variate Gaussian random vector $\mathbf{X}$ of length $L_{\mathbf{x}} \in \mathbb{N}$, where $\mathbf{x}$ is a realization of $\mathbf{X}$. The PDF's mean vector is given by $\boldsymbol{\mu}$ and its covariance matrix by $\boldsymbol{\Sigma}$. The expression $p_Y(y) \sim \mathcal{N}(y; \mu, \sigma^2)$ denotes a real-valued univariate Gaussian distribution of the random variable Y taking a value $Y = y$ with an expected value $\mathbf{E}\{Y\} = \mu$ and a variance σ^2. On the other hand, $p_Z(z) \sim \mathcal{CN}(z\,; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents a complex-valued (bivariate) Gaussian mixture model (GMM), where $z = z_{\mathrm{r}} + \jmath z_{\mathrm{i}}$ is a realization of the complex process Y, such that $z_{\mathrm{r}} = \Re\{z\}$ and $z_{\mathrm{i}} = \Im\{z\}$, with $\jmath^2 = -1$.

Finally, an estimate of a quantity a is represented by $\hat{a}$, and $\hat{a}^{(m)}$ represents the m^{th} update of $\hat{a}$ after m iterations.

2.2. Basic Concepts

The signal observed at the SU receiver over time can be considered as an example of *time series data*, such as daily closing stock prices, annual precipitations, and so on. Hence, probability theory serves as a powerful tool for characterization and analysis purposes by means of modeling such sequential data in the form of stochastic processes [50]. Moreover, combining probability theory with graph theory provides the ability to model and represent a dynamic system of multiple processes graphically. Their behavior and interactions are captured graphically in the form of joint and conditional distributions [51] in a so-called *probabilistic graphical network* or simply *graphical model* [52]. Please refer to [52, 53] for further explanations regarding graph theory.

A graphical model consists of nodes representing random variables and links connecting those nodes based on their probabilistic relationship. Those relations are narrowed down in the form of the sum and product rules, which are the elementary rules of probability theory [52].

For any two discrete random variables X and Y, $\Pr\{X\}$ is described via the sum rule as

$$\Pr\{X = x_0\} = \sum_{y_0} \Pr\{X = x_0, Y = y_0\}, \tag{2.1}$$

which is also known as the marginal probability. On the other hand, $\Pr\{X, Y\}$ is obtained via the the product as

$$\Pr\{X = x_0, Y = y_0\} = \Pr\{Y = y_0 | X = x_0\} \Pr\{X = x_0\}. \tag{2.2}$$

Figure 2.1 represents the different types of graphical models based on the type of their links [54, 52]. A graphical network could either be directed or undirected. On the other hand, graphs such as clique tree, factor graph, chain tree, etc., do not belong to any of those classes [51], and they are referred to as "others" in Figure 2.1.

The links in the directed graphical networks (DGNs) are displayed by arrows heading from one node to another, where the former is usually referred to as a *parent*

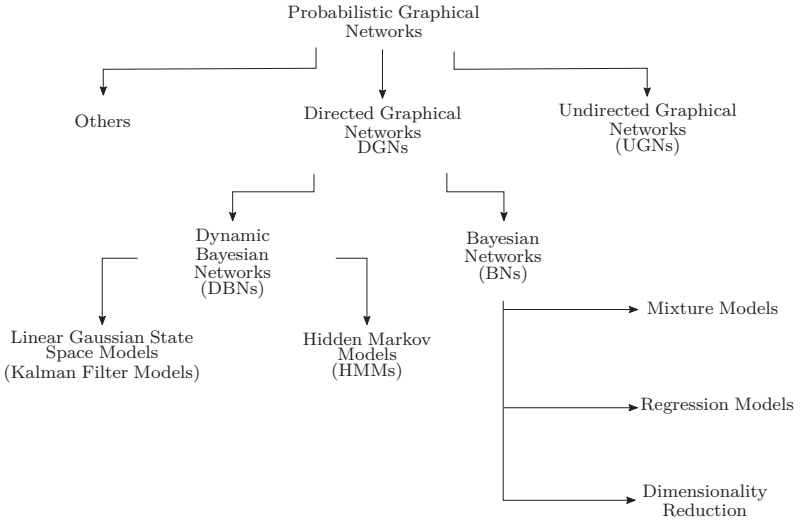

Figure 2.1.: A hierarchical representation of the different derivations of probabilistic graphical models.

node and the latter as a *child node*. Those links represent the conditional relation between the two nodes [52] as to be discussed onwards, as well as the model's causality [55]. Moreover, Bayesian networks (BNs), also known as Belief networks or Bayes networks, are a form of DGNs, yet adding the time factor introduces the so-called dynamic Bayesian networks (DBNs). Therefore, they are suitable to model sequential observations or data [55, 52]. The conditional relation between the nodes in a DBN is usually presented in the form of a conditional probability table [51]. On the other hand, such time stamps and direction orientated links are missing in undirected graphical networks (UGNs). Thus, a Markov random field, which is a well-known example of an UGN, is suitable for image modeling rather than time series analysis.

Figure 2.2 shows an example of both DGN and UGN. The graph to the right represents a DGN consisting of four nodes: E, F, G, and H representing four different discrete random variables. Based on the description mentioned earlier, node E is

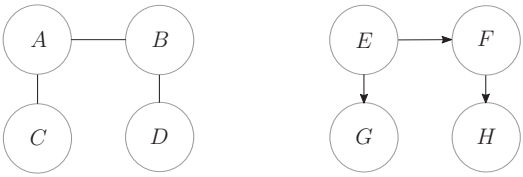

Figure 2.2.: Undirected (left) and directed (right) graphical models.

the parent node for both F and G, as both nodes are conditioned on E, whereas the nodes F and G are child nodes. Correspondingly, the joint probability can be directly deduced from the graph itself and expressed in the form of the product rule as

$$\Pr\{E, F, G, H\} = \Pr\{E\} \Pr\{F|\, E\} \Pr\{H|\, F\} \Pr\{G|\, E\}\,.$$

On the other hand, the graph to the left in Figure 2.2 represents an UGN with the parent-child node relation is absent as the connected nodes reciprocally conditioned on each other. Therefore, the term *blanket* is defined for each node which describes the set of nodes that make a node conditionally independent from the rest of the nodes in a graph [55]. Given the example shown in Figure 2.2, node A is conditionally independent of all the nodes in the graph given its blanket defined by nodes B and C. Further explanation with regard to this is found in [51, 52, 55].

An important quality or rather a restriction of the discussed DGNs is that they must be acyclic. In other words, by moving between the nodes along with the indicated arrows, a closed path does not occur or create [52]. Therefore, they can also be referred to as directed acyclic graphs (DAGs). Figure 2.3 represents a DAG consisting of nodes I, J, O, P and T as discrete random variables. The corresponding joint probability $\Pr\{I, J, O, P, T\}$ can be factorized based on their parent-child relationship as

$$\Pr\{I, J, O, P, T\} = \Pr\{I\} \Pr\{J\} \Pr\{O|\, I\} \Pr\{P|\, J\} \Pr\{T|\, O, P\}\,.$$

Furthermore, there are two important derivations of DBN as shown in Figure 2.1, namely, Kalman filter models as linear Gaussian state space models and HMMs [41]. Generally, a HMM scheme combines two processes: discrete hidden (unknown) states

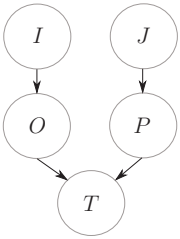

Figure 2.3.: An example of a DAG.

as Markov process *and* shown (known) observations as another process. In this work, known processes have been represented graphically as shadowed nodes. A HMM is actually capable of modeling both discrete observations as in DHMM scheme, also continuous observations as in CHMM scheme. In both cases, the observations are conditioned on the hidden states [41]. Moreover, the observations in DHMM schemes belong to a set of alphabets or symbols modeled as a discrete random process. However, the continuous observations are modeled as conditional PDFs in CHMM schemes. Nevertheless, the main difference between HMM and Kalman filter models is the type of processes representing the underlying states of each model: as a Kalman filter is capable of modeling both discrete and continuous states, HMM can only model discrete states [41]. Further derivations of BN, as mixture models, are also mentioned in Figure 2.1.

As an example of a DBN, a HMM scheme can be expressed via graphs, irrespective of the nature of its observations. Unlike a Markov chain [55], the correlation between the observations and the states is interpreted by the transition between the hidden states over time.

Furthermore, HMM is basically governed by the so-called Markov property (or assumption), also referred to as the (first-order) Markov assumption [55]. It declares the relation between a present hidden state with respect to the model's history, i.e., represents the transition probabilities between states [41]. Figure 2.4 shows a first-order (top) and a second-order (bottom) DHMM scheme, with x and y representing the realizations of the hidden state as a discrete Markov process X and the corresponding observation as the discrete process Y, respectively over time. As $n \in \mathbb{Z}$

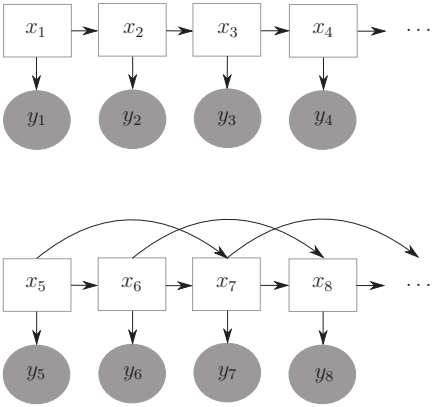

Figure 2.4.: A graphical model representing a first-order (top) and a second-order (bottom) DHMM schemes.

denotes the discrete time index of the observations, the conditional probability of state x_n given the model's history in the first-order scheme is given by

$$\Pr\{x_n | x_{n-1}, x_{n-2}, \dots\} = \Pr\{x_n | x_{n-1}\}, \tag{2.3}$$

which also declares Markov assumption since the state at a certain time depends only on its previous state and irrespective of the system's history. On the other hand, this assumption is slightly relaxed for higher-order HMM schemes, yet the state dependencies in the second-order scheme is given by

$$\Pr\{x_n | x_{n-1}, x_{n-2}, x_{n-3}, x_{n-4}, \dots\} = \Pr\{x_n | x_{n-1}, x_{n-2}\}. \tag{2.4}$$

Similarly, the conditional relations in (2.3) and (2.4) are applicable to the CHMM scheme and can be extended to even higher-order schemes. However, the computational complexity is expected to be increased.

Despite the nature of the modeled observations, there are two types of HMM schemes: *left-to-right* and *ergodic models*. They both differ with respect to the stationarity of the underlying system, which can be represented as extra restrictions to the possible transitions among the hidden states. In the next time step within the left-to-right

HMM scheme, the model may either remain at its same current state or change to a new one, however it does not earn back a past state. In other words, it has a self- and a left-to-right-transition [56]. Such models are best suited to represent transient processes and are widely used to model the beginning and the end of speech signals [56]. Nevertheless, this work assumes that the stationarity assumption is applied to the wireless environment, hence we focus on the most popular ergodic models which have no extra restrictions on their states over time.

Furthermore, representing HMM scheme graphically similar to a DBN helps to

- deduce a probabilistic expression of the whole model based on a simple diagram, and
- understand the conditional dependencies and conditional independencies between the underlying processes,

it also becomes much easier to create different models which are actually derived from a simple HMM scheme [54], irrespective of its order. For example, the memory in an autoregressive HMM (AR-HMM), shown in Figure 2.5, is introduced in the observation-domain as well as the state-domain. In other words, the observed sequence of the discrete process Y is not i.i.d unlike the schemes presented in Figure 2.4, and y_{n-1} helps to predict y_n.

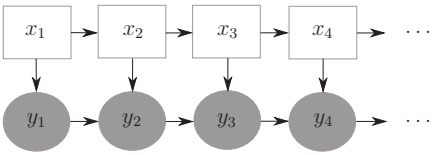

Figure 2.5.: A graph model representing AR-HMM.

Furthermore, an input-output HMM (IO-HMM) in Figure 2.6 represents a probabilistic mapper that maps a certain discrete input process U to another Y. In comparison to (2.3), the transition probability between the system states is hence expressed by $\Pr\{x_n | x_{n-1}, u_n\}$.

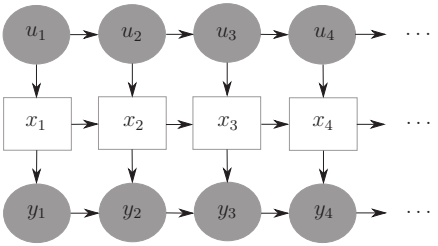

Figure 2.6.: A graph model representing IO-HMM.

A factorial HMM (FHMM) consists of multiple states and an observed sequence which is conditioned on those states. For $D \in \mathbb{N}$ number of Markov chains, the expression

$$\Pr\{x_{n,d} | x_{n-1,d}\} \tag{2.5}$$

is valid, where $d = \{1, \ldots, D\}$. Upon conditioning on the observations those states are coupled [54]. Thus, the discrete observation process Y depends on a vector of states, such that $\Pr\{y_n | x_{n,1}, \ldots, x_{n,D}\}$. Figure 2.7 shows an example of a FHMM scheme for $D = 3$ chains.

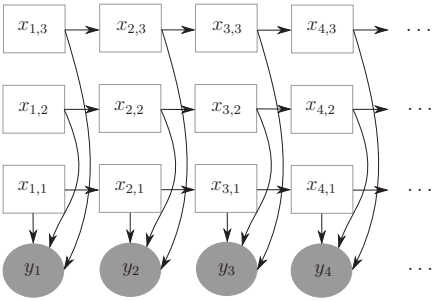

Figure 2.7.: A FHMM with $D = 3$ chains.

Note that in the previously given HMM examples, the states are unknown, and Modeling a single process is usually appropriate via a single HMM scheme, however limitations arise when modeling multiple interacting processes. Neither a single HMM scheme nor its previously mentioned derivations are suitable for this task.

Such limitations are introduced first, due to the Markov assumption governing the HMM scheme, as in (2.3) – or lightly relaxed for higher-order schemes as in (2.4). Second, due to limited system memory and number of states [57].

Thus, given two Markov states processes $X_{n,1}$ and $X_{n,2}$ each along with the corresponding observation process $Y_{n,1}$ and $Y_{n,2}$, respectively, a coupled HMM scheme, shown in Figure 2.8, can be formulated by relaxing the Markov assumption [57] to include neighboring or adjacent chains, such that $\Pr\{X_{n,1}|\, X_{n-1,1}, X_{n-1,2}\}$.

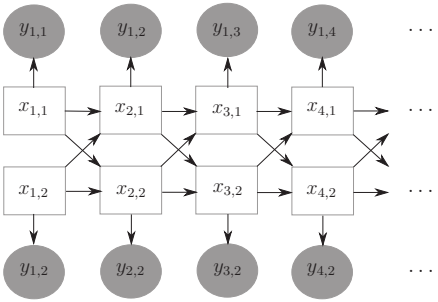

Figure 2.8.: A first-order coupled HMM with two chains.

Different from (2.5), such a coupled HMM scheme is capable of modeling multiple interacting processes and can easily overcome the previously mentioned limitations. Furthermore, being a generalization or a natural expansion of the single HMM scheme, a coupled scheme is a suitable approach to model a dynamic system of multiple interacting processes, and it is capable of modeling both discrete and continuous observations.

Chapter 3.

3 Coexistence in the ISM Band and Interference Spatial Modeling in Literature

The unlicensed spectrum resources, known as the ISM band, have been widely used as a medium for (short-range) communication technologies in industrial applications, controlling remote applications, sensor networks, and medical operations. Although Part 15 rules of the Federal Communications Commission (FCC) [58] govern transmitting over the ISM band with respect to the maximum output power, its users suffer from performance degradation due to the presence of CCI arising from multiple users from different technologies simultaneously transmitting over the same channel. There is no prior planning among users in the ISM band with respect to channel allocation, transmitted power, etc. Therefore, coexistence mechanisms have been subject to investigation over the years to mitigate the interference arising in the ISM band [5].

A short overview of the available technologies operating in the 2.4 GHz ISM band is presented in this chapter based on the available literature, the coexistence mechanism, as well as their mutual interference. Home environment technologies operating over the 2.4 GHz ISM band, including microwave ovens and cordless phones, have not been taken into consideration here due to their negligible effect [59]. Furthermore, different spatial interference models have been presented in this chapter.

3.1. Standards in 2.4 GHz ISM Band

1. **ZigBee**

   ZigBee is a communication protocol suitable for short-range, low power and data throughput applications, as well as applications with long battery life requirements. It is commonly used in sensor networks, remote control applications and medical devices in which the information is transmitted wirelessly over the ISM spectrum band. For further investigation please refer to [59, 60]. The PHY along with the MAC layers of the Open System Interconnect (OSI) stack are both based on the IEEE 802.15.4 standard. It operates in the 2.4 GHz ISM band using the orthogonal quadratic phase shift keying (OQPSK) modulation, with transmission power usually up to 0 dBm over a 100 m range at a rate of 250 Kbps [61, 62]. It transmits over 16 non-overlapping channels each with 2 MHz bandwidth with an offset between central frequencies which reaches up to 5 MHz [61, 62]. Furthermore, to make sure a wireless channel is not occupied by other users, clear channel assessment (CCA) is first performed prior to transmission, which is based on carrier sense multiple access/collision avoidance (CSMA/CA) mechanism [63, 64].

   In general, the ZigBee node is considered to be a weak interferer to other protocols within the 2.4 GHz ISM band, thanks to its very low duty cycle which ensures longer battery life. A device's duty cycle is defined as "the ratio of time it is active to the total time" [59]. A ZigBee's duty cycle is typically less than 1%, in which the device checks the availability of a channel through CCA and transmits its packet only within a few milliseconds [59]. Thus, both CCA and (OQPSK) ensure coexistence with other protocols within the ISM band [63].

2. **Wireless Local Area Network**

   The Wireless Local Area Network (WLAN) is a technology which is based on the family of IEEE 802.11x standards, namely IEEE 802.11b/g/n in the 2.4 GHz ISM band. Its transmission rate varies from 11 Mbps for

IEEE 802.11b reaching up to 54 Mbps for IEEE 802.11g/n [65] over a 100 m range and transmission power up to 30 dBm [59]. It transmits over a total of 14 overlapped channels each with 22 MHz bandwidth and 5 MHz separation. Among the 14 channels, only 3 of them are non-overlapped. Their PHY is designed such that IEEE 802.11b uses the direct sequence spread spectrum (DSSS) technique based on the complementary code keying modulation (CCK), whereas IEEE 802.11g employs an OFDM modulation, as implemented in the 5 GHz ISM band IEEE 802.11a standard [65].

In order to coexist with other protocols, a WLAN device employs CSMA/CA as a channel access technique based on a distributed coordination function (DCF) mechanism [65]. In other words, each WLAN node checks if a specific channel is idle before transmission. If a channel is indeed idle for a certain duration of time, usually called DCF interframe space (DIFS) period, the WLAN node transmits its packets immediately. Otherwise, it is postponed until the end of the current transmission. Plus, it waits for an additional DIFS interval called a backoff timer until it re-attempts to sense the channel. The randomly generated backoff timer is between 0 and a contention window (CW) [66]. The size of the CW increases (doubled) whenever the transmission is not successfully performed, and each time the node's transmission defers [66, 65].

Furthermore, multiple input multiple output-OFDM (MIMO-OFDM) techniques were adapted in 2009 by the Task Group n (TGn), as part of the IEEE 802.11n standard. It enables spatial diversity multiplexing (SDM) to increase throughput up to 100 Mbps and a transmission range up to 250 m [65]. This standard also operates over the ISM band with 40 MHz channels divided into 128 subcarriers. However, the usual 20 MHz channels are used whenever limitations are presented in the available spectrum [67].

3. **Bluetooth**

This wireless standard has been developed by Ericsson along with another four companies: IBM, Intel, Nokia, and Toshiba, forming the Special Interest Group (SIG) in 1998 [65]. It is characterized by its low power con-

sumption and short range operation with no more than 10 m [65]. It transmits up to 1 mW with a rate of 1 Mbps, over a total of 79 radio frequency (RF) channels each with 1 MHz bandwidth. Its signal is modulated using the binary Gaussian frequency shift keying (BGFSK) [68]. Furthermore, a time division multiplexing (TDM) technique is employed by dividing each channel into 625 μs slots. Among the 79 channels, only one is defined or chosen based on a unique pseudo-random frequency hopping sequence determined by a master device – known as piconet – with a maximum of 1600 hops per second [68, 59]. This mechanism helps to reduce the interference with active devices operating in the same band.

4. **WirelessHART**

   Wireless Highway Addressable Remote Transducer (wirelessHART) was the first open wireless communication standard introduced in 2007 as the wireless expansion of the wired Highway Addressable Remote Transducer (HART) 7 protocol [69, 70]. It is a flexible, low cost, robust, and scalable protocol for factory automation industrial and controlling applications. It transmits very low power signals to avoid possible interference scenarios with existing protocols [69].

   Its PHY layer is designed based on the IEEE 802.15.4-2006 standard and operates over the 2.4 GHz ISM band [71, 72]. Similar to ZigBee, it employs DSSS modulation to transmit with a rate of 250 kbps over 15 channels each with 2 MHz bandwidth and 5 MHz central offset frequency [73]. However, different from IEEE 802.15.4-2006, the channel is chosen randomly based on a pseudo-random frequency hopping mechanism in order to minimize the interference within the network [70, 74]. Furthermore, it utilizes Time Division Medium Access (TDMA) as well as CSMA as access mechanisms [71].

   The simulation conducted in [73] has shown that wirelessHART affects the performance of the IEEE 802.11g by increasing the number of dropped packets due to a collision. However, this problem can be eliminated by introducing a slow frequency hopping scheme as well as enabling a blacklist feature. Via its blacklist feature, wirelessHART devices label those channels which are to be

avoided due to coexisting with other networks [70, 74]. Thus, wirelessHART can apparently coexist with other protocols in the 2.4 GHz ISM band without behaving as an interferer to other networks nor reduce its QoS.

3.2. Mutual Interference in the ISM Band

The performance of any wireless network in the presence of interferers is subject to different network topology and PHY-related parameters, such as modulation type, spread spectrum and frequency hopping rate, traffic characteristics, packet size and transmitted power [75]. Among the different standards, the mutual interference between WLAN and ZigBee are the most investigated in the literature [76, 63, 62]. Thus, the possible mutual interference scenarios within the different protocols in the 2.4 GHz ISM band are briefly presented as follows:

- **Mutual Interference Between WLAN and ZigBee**
  Due to the absence of channel allocation planning, protocols such as ZigBee which operate at low power and low transmission rate values are victims which suffer from other existing interferers. On the other hand, as a WLAN node transmits at relatively high power values, it is mostly considered as an interferer. Yet its performance can still suffer from degradation due to the presence of a ZigBee node. Furthermore, [76] has proven that irrespective of the ZigBee's duty cycle, a WLAN's access point is capable of locating a channel to transmit without disturbances or cuts. Following are some remarks pertaining to when WLAN node acts as an interferer to a ZigBee node:

  1. Approximately 90% of ZigBee packets are destroyed due to WLAN transmission when both of them utilize overlapping channels. Specifically, when WLAN transmits over channel 6 (2437 MHz) and ZigBee over channel 16 (2440 MHz) [63, 62].

  2. ZigBee suffers from packet loss when WLAN transmits over variate (mainly higher) transmission rates [75].

3. Based on different experiments conducted in [67], ZigBee nodes suffer the most due to a heavy loaded IEEE 802.11n traffic rather than IEEE 802.11g. Since IEEE802.11n operates over relatively broad channels, it can easily reduce the quality of transmission of a near ZigBee node [67].

The main considerations when ZigBee nodes are causing interference to WLAN nodes are summarized based on the following points:

1. The distance between the interferer's transmitter and the victim's receiver is an important parameter. It is shown in [76] that the performance of a WLAN device is reduced when the distance between the WLAN's receiver and the ZigBee's transmitter is small.

2. Irrespective of the availability of a ZigBee node, it has been noticed in [76] that the performance of a WLAN node is also a function of its packet size. The larger the size of its packet, the worse its performance may be.

3. Both protocols perform CSMA/CA as they sense the environment before transmitting. However, ZigBee is much slower in avoiding collisions given its sensing slot is 320 μs compared to WLAN's 20 μs [62].

4. Experiments conducted in [77] show that WLAN's upper link is affected by the presence of ZigBee more than its downlink.

- **Mutual Interference Between WLAN and Bluetooth**
  The mutual interference between WLAN and Bluetooth is basically a function of time and frequency overlapping. It occurs when Bluetooth's frequency hopping system chooses a channel which occupies 1 MHz of the WLAN's 22 MHz channel [68, 78]. Another important factor to consider is the distance between the two devices [79]. Thus, the degradation in their performance is (10-20)% for both networks. Furthermore, the packet error increases when WLAN transmits large packets in the presence of an active Bluetooth piconet [80].

- **Mutual Interference Between Bluetooth and ZigBee**
  As it has been discussed in [63], no significant packet corruption results due to the coexistence of ZigBee and Bluetooth. The frequency hopping scheme

in Bluetooth is capable of marking a channel as bad due to the presence of interference, a maximum chance of mutual interference between ZigBee and Bluetooth is up to 4% [59].

3.3. Point Processes for Interference Spatial Modeling

Interference has been subject to investigations and research, especially in the crowded ISM band. Aside from thermal noise, interference is an important parameter which has to be taken into account, as it is able to limit a network's performance under heavy interference conditions. As clearly explained in [49], variables such as channel allocation and maximum power transmission are controlled by users within the same network, thus any possible interference arising between homogeneous users can be easily eliminated. However, emerging new technologies and networks have made interference elimination a difficult task. This is due to the fact that each network such as ad hoc, mesh, and sensor networks, has its own specifications including transceiver configurations, channel access schemes, and transmitted power levels.

A Gaussian process has been utilized in signal processing and wireless communication fields to model the noise and/or interference in a certain wireless environment. Since no exact prior knowledge of the noise and/or interference can be collected, the Gaussian assumption is widely considered as it is applicable to the central limit theorem [81, 82]. Furthermore, the Gaussian process is relatively easy to handle; the linear superposition of multiple Gaussian processes are also Gaussian as well as linearly mapping a Gaussian process is also a Gaussian [52, 81]. Hence, many techniques and algorithms have been derived based on this assumption, such as the widely known Kalman filter [83].

Nevertheless, the Gaussian assumption is not entirely an accurate approach to model noise and/or interference, due to the presence of impulsive and heavy-tailed characteristics [84, 85, 86, 87, 82] in the form of sharp spikes and occasional bursts in its observations and measurements [88]. Hence, noise and/or interference are

actually far from being Gaussian in many different environments, such as underwater environments, telephony, man-made noise and indoor wireless communication [88, 89, 90, 91, 92, 93].

The non-Gaussian CCI created between users from different networks is basically subject to uncertainty and randomness due to the following factors [94, 95]:

1. Radio propagation effect on the received signal, such as path loss effect, small-scale and large-scale propagation.

2. Spatial location of interferers and their activities in the network area. The superposition of interference signals affecting a receiver node depends on the geometrical location of the concurrently transmitting interferers within a certain area. For example, the total amount of damage decreases as the distance to the interferer nodes increases [95, 96].

3. Network operation mode as it defines the followed medium access control scheme, such as CSMA and ALOHA.

As the CCI is a function of the network geometry [97], modeling the interference spatially is an important aspect to build an idea behind the performance of the wireless network. Hence, principles of *stochastic geometry and point processes* come in handy. They describe a set of random objects or nodes within finite dimensions (one, two or higher) in a statistical model or approach [98]. As part of their applications, different spatial interference models and outage characterizations have been subject to investigations over decades [96].

Starting with simple models with equally affecting interferers, models with deterministic interferer locations as well as shot noise models, research has led to the well-known Poisson networks where interfering nodes are distributed as a homogeneous Poisson point process (PPP) [99, 100, 101]. Due to its flexibility, it has been by far the most popular interference model [49], especially within CSMA/CA networks [100, 102, 103]. The following describes the PPP along with its drawbacks. Afterwards, an alternative approach which has been considered within this work is subsequently presented.

- **Poisson Point Process**

  Given a network occupying a $\mathbb{R}^2$ plane of area A_{P} with a node intensity ξ_{P} (nodes/area) for a PPP network, the random variable $N_{\mathrm{P}} \in \mathbb{N}$, representing the independent and uniformly (homogeneously) located number of nodes, follows a Poisson distribution as [95, 98, 102]

  $$\Pr\{N_{\mathrm{P}} = n_{\mathrm{P}}\} = \frac{\eta^{n_{\mathrm{P}}}}{n_{\mathrm{P}}!}\mathrm{e}^{-\eta}, \quad \text{for } \; n_{\mathrm{P}} = 0, 1, 2, \ldots$$

  where n_{P} is a realization of N_{P}, and η is the distribution's mean value for a given ξ_{P}, it is given by $\eta = \xi_{\mathrm{P}}(A_{\mathrm{P}})$.

  Let Φ_{P} denote a PPP network characterized as [104, 105]

  $$\Phi_{\mathrm{P}} = \{(\pi_{i_{\mathrm{P}}}, (e_{i_{\mathrm{P}}}, p_{i_{\mathrm{P}}}))\}$$

  where

  - The set $\Phi_{\mathrm{P}} = \{\pi_{i_{\mathrm{P}}}\}$ denotes the location of the independently scattered nodes [98], where $i_{\mathrm{P}} \in \{1, \ldots, N_{\mathrm{P}}\}$ refers to the node's index. In other words, $\pi_{i_{\mathrm{P}}}$ denotes the location of the $i_{\mathrm{P}}^{\mathrm{th}}$ node,

  - $\{e_{i_{\mathrm{P}}}\}_{i_{\mathrm{P}}}$ is the medium access indicator. The $i_{\mathrm{P}}^{\mathrm{th}}$ node is a transmitter if $e_{i_{\mathrm{P}}} = 1$, whereas it is a potential receiver if $e_{i_{\mathrm{P}}} = 0$,

  - and, $p_{i_{\mathrm{P}}}$ is the transmission power of the $i_{\mathrm{P}}^{\mathrm{th}}$ node whose $e_{i_{\mathrm{P}}} = 1$ (i.e., a transmitter node).

  Nevertheless, the assumptions underneath the PPP are actually far from being applicable to characterize the interference in CSMA/CA networks in real-life scenarios. This can be due to the following reasons:

  1. PPP is not a valid model if the number of nodes N_{P} is known. As N_{P} is a Poisson random variable in a PPP network, its realization may reflect either more nodes than reality or even zero nodes [106]. For this reason, PPP does not accurately model a network with a small number of nodes.

2. The node independency assumption is also inaccurate since it fails to capture the case if all the N_{P} nodes are gathered in one location in the network [106].

3. The homogenous assumption, which indicates that the network intensity does not depend on the location of the nodes [96], is inapplicable to model practical networks [95, 107]. As networks, in reality, cover a finite area, nodes located near the center suffer from a higher amount of interference than those located at the boundaries.

4. The nodes in a realistic network are more structured than being entirely random. As an active transmitter in a CSMA/CA network actually first listens to the channel before transmission, a prohibited region is created around the transmitter [95]. Thus, the interferers are actually far from being randomly localized.

- **Matérn Point Process**

  Based on the homogenous and independency weaknesses discussed earlier, a Binomial Point Process (BPP) has been considered as an alternative approach [106] in order to model a more structured network, where nodes are *not* independent [98]. However, there is no common probability density function to characterize this assumption statistically. In order to overcome this issue, the investigation carried out in some literature concerning the BPP is instead performed based on moment generating functions [106, 95].

  In view of the CSMA/CA mode, a so-called Matérn point process II, or Matérn hard-core process. It has been utilized to model nodes spatially in a CSMA/CA network as a more appropriate approach describing users in such networks [108, 105]. It is formulated by applying a thinning (extra condition) on nodes in Φ_{P}, such that nodes are forbidden to lie closer to each other than a certain minimum distance R_{M}, such that $R_{\mathrm{M}} > 0$ [100].

  For this matter, each node $\pi_{i_{\mathrm{P}}} \in \Phi_{\mathrm{P}}$ is independently labeled by a random identifier $m\left(\pi_{i_{\mathrm{P}}}\right)$ drawn uniformly between $(0, 1)$ [95, 98]. The node $\pi_{i_{\mathrm{P}}}$ also belongs to a Matérn hard-core process, denoted by Φ_{M}, if there is no other node from Φ_{P} whose identifier value is smaller than the identifier of node $\pi_{i_{\mathrm{P}}}$

in an circle defined by $B_{\mathrm{M}}(\pi_{i_{\mathrm{P}}}, R_{\mathrm{M}})$, of center $\pi_{i_{\mathrm{P}}}$ and radius R_{M}. This can be expressed by [98]

$$\Phi_{\mathrm{M}} = \left\{ \pi_{i_{\mathrm{P}}} \in \Phi_{\mathrm{P}} : m\left(\pi_{i_{\mathrm{P}}}\right) < m\left(\pi_{i'_{\mathrm{P}}}\right) \forall\, \pi_{i'_{\mathrm{P}}} \in \Phi_{\mathrm{P}} \cap B_{\mathrm{M}}\left(\pi_{i_{\mathrm{P}}}, R_{\mathrm{M}}\right) \setminus \pi_{i_{\mathrm{P}}} \right\}. \quad (3.1)$$

Based on (3.1), any node which belongs to Φ_{M} is separated by at least a minimum distance R_{M} from the next available node. This coincides with the behavior of transceivers in a CSMA/CA network. Unlike BPP, a Matérn hard-core process results in a so-called symmetric α-stable statistical distribution, which represents statistically the interference signal at a certain receiver [95].

Chapter 4.
4 System Description

This chapter presents the system model considered in this work. Similar to spectrum sensing, spectrum occupancy prediction schemes are to be implemented in the SU receiver. On the other hand, different from the prediction schemes presented in Section 1.4, the SU transceiver structure presented in this work utilizes a proper TFR method. Hence, the received signal is efficiently and fully exploited in order to be employed in a proper spectrum occupancy prediction scheme based on the interdependencies between adjacent subbands.

Before dealing with spectrum occupancy schemes, signal representation methods are initially discussed. TFR is an important concept in signal processing, as it is capable of accurately and efficiently analyzing signals with time-varying frequency content [109]. By means of TFR methods, a signal can be viewed in a two-dimensional TF grid in order to examine its features for further processing. Mathematically speaking, this is equivalent to a joint function of two variables: time and frequency [110]. Furthermore, there are two TFR classes or methods in literature: first, linear TFR methods such as the short-time Fourier transform (STFT), wavelet transform and Gabor expansion, which obey the basic superposition principle [110, 111]; second, quadratic TFR methods including distributions of energy and correlation of signals, such as the Wigner distribution and the ambiguity function, respectively. [109, 110].

In this work, the SU transceiver has been considered to utilize a properly designed DFT-FB as a direct implementation of Gabor expansion linear representation method. Hence, the conventional subband filtering mechanisms, which introduce information loss, can be eliminated [2]. Since this work focuses on spectrum occupancy

prediction mechanisms, a short overview of the considered TFR method; i.e., Gabor expansion, is presented briefly in this chapter, along with the SU transceiver structure. Further information with regard to this topic is available in [24, 110].

Furthermore, being an important parameter affecting the network's performance [95, 49, 84], modeling the CCI has also been considered in this chapter. Based on the work in [1], the CCI signal arising from multiple transmitters utilizing the same channel simultaneously in a CSMA/CA network actually follows a Matérn hard-core spatial distribution. The resulting CCI signal in [1] is approximated by a proper statistical and parametric model.

4.1. Gabor Expansion

Gabor expansion, also referred to as Gabor analysis, ideologically represents a (continuous or discrete) time-domain signal into its TF characteristics, known as TF Gabor coefficients. These coefficients are obtained via a family of TF functions created by time-shifting and frequency-modulating a certain window function or pulse, which is referred to as *Gabor atom* [24, 111].

The Gabor coefficients are represented in a joint TF grid which is defined based on two positive integer constants: K and M. Their ratio, known as the TF sampling density [24], defines the size of the TF grid. This aspect is out of the scope of this work, therefore we disregard the different TF sampling cases and directly consider $K = M$ in a critically sampled TF grid Λ defined by

$$\Lambda = \mathbb{Z} \times \Delta,$$

where Δ is give by $\Delta = \{0, \ldots, K-1\}$. Moreover, given the discrete Gabor atom or pulse $g[n] \in L^2(\mathbb{Z})$ of length $L_g \in \mathbb{N}$, with $n \in \mathbb{Z}$, the family of functions $g_{\ell,k}[n]$ is denoted by

$$\{g_{\ell,k}[n] : (\ell, k) \in \Lambda\} \tag{4.1}$$

are derived by $\ell.K$-time translating and k/K-frequency modulating the atom $g[n]$.

The critically sampled Gabor system $g_{\ell,k}[n]$ is defined as

$$g_{\ell,k}[n] = g[n - \ell K] \, \mathrm{e}^{\jmath \frac{2\pi k n}{K}}.$$

The TF coefficients of an arbitrary discrete-time signal $C[n]$, namely $C_{\ell,k}$, are obtained by projecting $C[n]$ into the family of functions in (4.1) as

$$C_{\ell,k} = \langle C[n], g^*_{\ell,k}[n] \rangle = \sum_{n \in \mathbb{Z}} C[n] \; g^*_{\ell,k}[n],$$

with $g^*_{\ell,k}[n]$ representing the complex conjugate of $g_{\ell,k}[n]$. Figure 4.1 demonstrates a critically sampled TF grid.

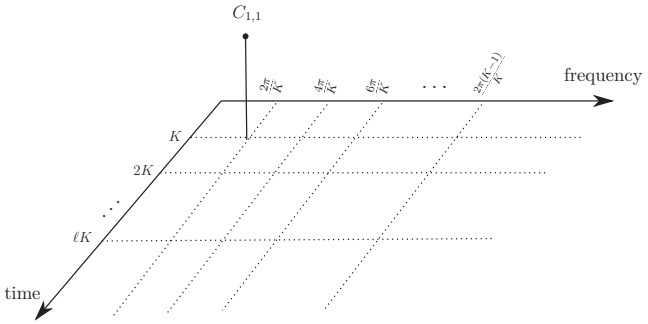

Figure 4.1.: A critically sampled TF grid.

Due to their similarities, a DFT-FB can be utilized to efficiently implement a Gabor system [112, 24, 110, 111]. This can be achieved by properly designing the filter's prototype based on the family of functions defined in (4.1) [112]. Moreover, FBMC based communication, in particular DFT-FB, has been widely utilized in signal processing and particularly in CRN due to the following reasons:

1. FB has shown robust performance by mitigating both CCI and ISI [24, 113, 25].

2. FFT techniques can reduce the complexity of FB's implementation with respect to other implementation approaches [114, 115].

3. A large amount of literature is available regarding FBMC based spectrum sensing for CRNs [116, 117, 118]. Such schemes have shown a superior performance with respect to detecting a PU signal embedded in noise and interference, despite lacking a prior knowledge of the PU signal [1].

4. Different from the widely known OFDM, a properly designed DFT-FB with an orthogonal pulse which does not require additional guard bands [119]. Yet it provides an efficient method of transmitting multiple streams of data within the available bandwidth.

5. FBMC schemes offer higher spectrum flexibility, efficiency, and more robust synchronization than the conventional OFDM scheme [120]. In addition, FBMC is capable of providing a perfect TF localization of a signal's energy, and accessing a fragment of spectrum during spectrum access and spectrum sharing in a simple and straightforward procedure, hence it facilitates the implementation of the 5G telecommunication standards [8]. Moreover, upon considering a DFT-FB based SU transceiver structure, spectrum sensing, spectrum occupancy prediction and spectrum access all share the same hardware structure for a CRN [25, 119].

On the other hand, the choice of having a critically sampled TF grid has been met to provide two advantages: it offers an orthogonal Gabor pulse $g[n]$, and it satisfies the perfect reconstruction property (PRP) [25]. Hence, reconstructing the time-domain signal based on its TF coefficients is possible [24] as

$$C[n] = \sum_{\ell,k\in\Lambda} C_{\ell,k} g_{\ell,k}[n].$$

4.2. Transceiver Structure

Figure 4.2 shows a SU's transceiver [119] as a natural choice for any DFT-FB based SU transceiver for CRNs. The SU receiver is equipped with the AFB, whereas the transmitter is equipped with the SFB. The switch shown provides the SU with the ability to choose between two actions upon required: position A enables the SU to perform spectrum sensing and/or spectrum occupancy prediction. On the

other hand, position B provides spectrum access whenever an appropriate TF slot is available for SU transmission. Hence, the hardware complexity is reduced by implementing multiple CR functionalities in one fundamental design.

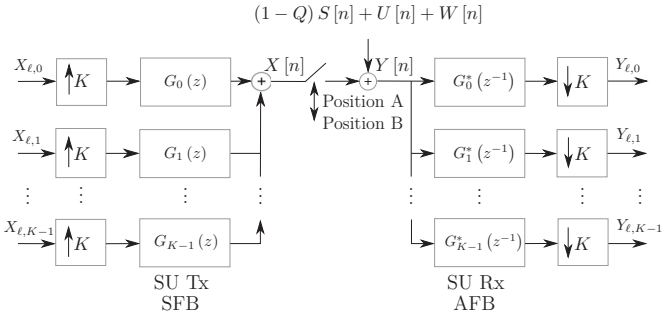

Figure 4.2.: SU transceiver being designed as a critically sampled DFT-FB.

At the receiver side, the prototype of the K filters given by $G_k^*(z^{-1})$ represents the paraconjugate of $G_k(z)$, where $G_k(z) = G\left(zW^k\right)$, $W = \mathrm{e}^{-J\frac{2\pi}{K}}$ and $G(z)$ is given by

$$G(z) = \sum_{n=0}^{L_g-1} g[n]\, z^{-n},$$

for $k = 0, \ldots, K-1$ [2]. The signal at the output of the K filters are subsequently downsampled by a factor of K. Various studies have been carried out to optimize $g[n]$ in order to obtain optimal TF localization properties to ensure minimum signal leakage outside the TF grid [25, 121]. In the scope of this work, $g[n]$ is assumed to be an optimal pulse without any further optimization details.

For spectrum occupancy prediction purposes, the switch in Figure 4.2 must be at Position A, hence the received signal $Y[n]$ can be represented as

$$\mathrm{H}_Q: \quad Y[n] = (1-Q)\,S[n] + U[n] + W[n], \tag{4.2}$$

with the mutually independent processes $S[n]$, $U[n]$ and $W[n]$ representing the PU signal sensed by the SU receiver, the CCI and additive white Gaussian noise (AWGN) signals, respectively. Furthermore, Q is a random process which represents the system state or hypothesis with respect

to channel availability to the SU, such that $Q \in \{0, 1\}$. An available channel for the SU to utilize is indicated by $Q = 1$, whereas an unavailable channel is indicated by $Q = 0$.

Upon projecting the received signal $Y[n]$ into the AFB, it is decomposed into its corresponding TF coefficients as

$$Y_{\ell,k} = \langle Y[n], g^*_{\ell,k}[n] \rangle = \sum_{n \in \mathbb{Z}} Y[n]\ g^*_{\ell,k}[n], \tag{4.3}$$

with $(\ell, k) \in \Lambda$. As explained earlier, the TF coefficients $Y_{\ell,k}$ can be employed for spectrum sensing purposes to detect the presence of the PU [1], as well as for spectrum occupancy prediction. The random quantity Q in (4.2) is also mapped into its equivalent TF random process, namely $Q_{\ell,k}$ with $Q_{\ell,k} \in \{0, 1\}$. It represents the occupancy state of the corresponding TF slot. Similarly, $Q_{\ell,k} = 1$ indicates an available TF slot for the SU to access, whereas $Q_{\ell,k} = 0$ indicates its unavailability due to the presence of a PU signal.

As the CRN allows its users; i.e., SUs, to learn and adapt to the wireless environment, suitable sequences consisting of a sufficient number of symbols of $Y_{\ell,k}$ are observed during the pilot phase and learning phase, as briefly introduced in Section 1.5. Further explanations are presented in the following chapters.

4.3. Interference Modeling

This section deals with modeling the CCI signal observed at the SU receiver discussed in [1]. Characterizing the interference helps to understand the surrounding environment and to evaluate the performance of the proposed spectrum occupancy prediction schemes.

The main purpose of the work conducted in [1] is to perform a TF spectrum sensing scheme based on the generalized likelihood ratio test (GLRT) to detect the presence of a PU signal at the output of the AFB. However, as modeling the CCI signal is required at this point, the observed signal $Y_{\ell,k}$ under hypothesis H_1, i.e., $Q_{\ell,k} = 1$, is

investigated in this section. Note that the hypothesis of PU absence in [1] has been considered to be H_0, different to this work.

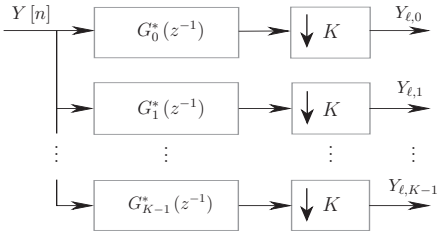

Figure 4.3.: SU receiver structure as a critically sampled AFB.

Based on the structure of the SU receiver shown in (4.2), the time-domain complex baseband process $Y[n]$, which represents the received signal under H_1, is given by

$$\mathrm{H}_1: \qquad Y[n] = U[n] + W[n]. \tag{4.4}$$

At the output of the AFB, the hypothesis test presented in (4.4) can be expressed under H_1 as

$$\mathrm{H}_1: \qquad Y_{\ell,k} = U_{\ell,k} + W_{\ell,k},$$

where the TF coefficients $Y_{\ell,k}$ at the output of the AFB is obtained based on (4.3). The mutually independent processes $U_{\ell,k}$ and $W_{\ell,k}$ represent the TF coefficients of $U[n]$ and $W[n]$, respectively.

Since the CCI shows heavy-tailed and impulsive characteristics, as explained in Chapter 3, the non-Gaussian signal $U[n]$ in a CSMA/CA network follows statistically a complex symmetric α-stable distribution [95, 82], which belongs to the family of stable processes. Both Gaussian and Cauchy are also examples of stable processes [82]. In signal processing, symmetric α-stable distributions have recently gained significant attention, as number of studies have proven that they accurately match and fit real measurements [85, 122, 82].

Given the PDF $p_{U_{\ell,k}}(u) = p_{U_{\ell,k}}(u; \alpha_{\mathrm{int}})$ of the complex random variable $U = U_{\mathrm{r}} + \jmath U_{\mathrm{j}}$ with its realization $u = u_{\mathrm{r}} + \jmath u_{\mathrm{j}}$, the characteristic function $\varphi_U(\omega)$ of the symmetric

α-stable random process U is given by [123]

$$\varphi_U(\omega) = \exp\left(-\int_{A_\mathrm{u}} |\omega_\mathrm{r} u_\mathrm{r} + \omega_\mathrm{i} u_\mathrm{j}|^{\alpha_\mathrm{int}}\, d\Upsilon_{U_\mathrm{r},U_\mathrm{j}}(u_\mathrm{r}, u_\mathrm{j})\right), \tag{4.5}$$

with $\omega = \omega_\mathrm{r} + \jmath\omega_\mathrm{i}$, and $\Upsilon_{U_\mathrm{r},U_\mathrm{j}}$ is the symmetric spectral measure of the random variable U on a unit sphere A_u [1, 123]. Moreover, the characteristic exponent α_int in (4.5), where $\alpha_\mathrm{int} \in (0, 2]$ describes the impulsiveness of the underlying stable process; small values of α_int mean more impulsive characteristics, hence heavier tailed characteristics [88].

The PDF of any random variable can be calculated by applying the Fourier transform on its characteristic function, yet evaluating the integral in (4.5) does not lead to a solution [82]. Therefore, a fixed expression for $p_{U_{\ell,k}}(u; \alpha_\mathrm{int})$ does not actually exist, except for the two special cases with $\alpha_\mathrm{int} = 2$ and 1, representing Gaussian and Cauchy distributions, respectively [88]. Nevertheless, to overcome this obstacle, different approaches can be followed to approximate the required PDF based on a set of observations, such as [82, 124, 88]:

- power series expansion,
- polynomial approximation,
- approximation using finite mixtures; such as Gaussian or Cauchy processes,
- a rational-based partial fraction, or
- power based approximation.

Furthermore, mixture models are examples of BNs as shown in Figure 2.1. They are suitable to model a set of measurements which belong to multiple different subsets, where each subset fits into its appropriate model [52, 55]. The most widely known example of a mixture model is the GMM, which has been commonly used to model non-Gaussian measurements due to its simplicity in the form of a parametric model consisting of the sum of a finite number of weighted Gaussians [82]. Thus, it has been also utilized in [1] to approximate the CCI signal.

Hence, the conditional PDF $p_{Y_{\ell,k}}(y|\,\mathrm{H}_1)$ is approximated by a complex-valued (bivariate) GMM where $y = y_\mathrm{r} + \jmath y_\mathrm{i}$ is a realization of $Y_{\ell,k}$, which can be expressed as

$$p_{Y_{\ell,k}}(\, y|\, \mathrm{H}_1) = \sum_{r=1}^{R} w_r\, \mathcal{CN}\,(\, y\,;\boldsymbol{\mu}_r,\, \boldsymbol{\Sigma}_r)\,, \tag{4.6}$$

consisting of total of R weighted Gaussian components, with $R \in \mathbb{N}$. The weight of the r^{th} Gaussian component, w_r satisfies the conditions

$$\sum_{r=1}^{R} w_r = 1 \quad \text{and} \quad 0 < w_r < 1, \tag{4.7}$$

for all $r = \{1, \ldots, R\}$. Furthermore, $\boldsymbol{\mu}_r$ and $\boldsymbol{\Sigma}_r$ are the mean vector and covariance matrix of the r^{th} component, respectively.

Wireless Environment for Interference Modeling

For simulation purposes, a CSMA/CA wireless network occupying a finite area with a node density of $\xi_{\mathrm{M}} = 5 \cdot 10^{-2}$ nodes/m^2 has been investigated, where $R_{\mathrm{M}} = 10$ m. The nodes are uniformly scattered around the receiver under consideration in an area of 100 m^2. The communication between the nodes is considered over 52 subbands and follows a binary phase-shift keying (BPSK) modulation through a Rayleigh flat fading environment. Based on Monte Carlo simulation, the nodes spatial distribution follows a Matérn hard-core process. The obtained CCI signal results from the superposition of signals from multiple nodes simultaneously transmitting over the same subband. Further simulation considerations are found in [1]. Under hypothesis H_1, i.e., $Y\,[n] = U\,[n] + W\,[n]$, the sequence of i.i.d. symbols denoted by $y_0, \ldots, y_{L_y-1}$ of length L_y, with $L_y \in \mathbb{N}$, is employed to estimate the GMM order R as well as the parameters $w_r, \boldsymbol{\mu}_r$, and $\boldsymbol{\Sigma}_r$ for $r = \{1, \ldots, R\}$. For approximation purposes, the expectation maximization (EM) algorithm along with k-means clustering algorithm [125, 1] are employed to find the best-fit complex (bi-variate) GMM, which best describes the observed sequence.

A heuristic approach based on [126] is employed to estimate the minimum required number of Gaussian components to best approximate the observed sequence, i.e, the model order. For this purpose, the estimated model order $\hat{R}$ has been chosen such that the increase in the corresponding likelihood function $p_{Y_{\ell,k}}(y \mid \mathrm{H}_1)$ given the model order R with respect to the one of order $R-1$ is below a convergence

threshold ϵ. In other words, the minimal model order can be obtained as

$$R = \min \left\{ \hat{R} \in \mathbb{N} : \left| \frac{p_{Y_{\ell,k}}(y)\Big|_{R=\hat{R}} - p_{Y_{\ell,k}}(y)\Big|_{R=\hat{R}-1}}{p_{Y_{\ell,k}}(y)\Big|_{R=\hat{R}-1}} \right| \leq \epsilon \right\}. \tag{4.8}$$

Upon determining $\hat{R}$, the EM algorithm is employed starting with $\hat{w}_r^{(0)}$, $\hat{\boldsymbol{\mu}}_r^{(0)}$ and $\hat{\boldsymbol{\Sigma}}_r^{(0)}$ being the initial values of w_r, $\boldsymbol{\mu}_r$, and $\boldsymbol{\Sigma}_r$, respectively. A certain number of iterations takes place until the relative increase in the corresponding likelihood function [1] between two successive iterations is below the threshold ϵ. Further explanations can be found in Appendix B. Figure 4.4 represents the likelihood function versus an estimate of R, whereas Figure 4.5 represents the relative increase in the likelihood function between two successive model orders.

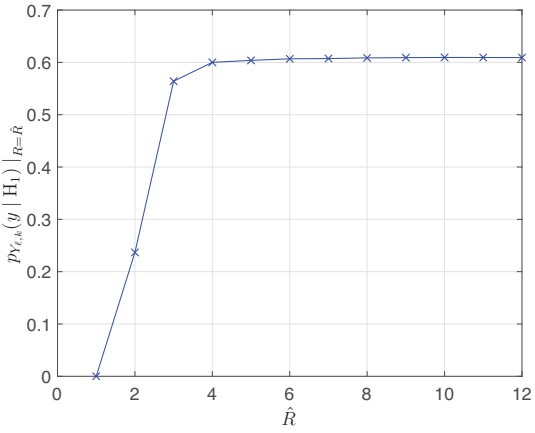

Figure 4.4.: $p_{Y_{\ell,k}}(y)\mid_{R=\hat{R}}$ versus $\hat{R}$ [1].

Based on the simulation carried out for $L_y = 10^4$ [1], the minimal number of Gaussian components has been chosen based on (4.8) to be $\hat{R} = 8$, with $\epsilon = 10^{-4}$. Note that a high number of Gaussian components means more complexity without a significant improvement.

Furthermore, the model and its corresponding approximation for $\hat{R} = 8$ are shown in

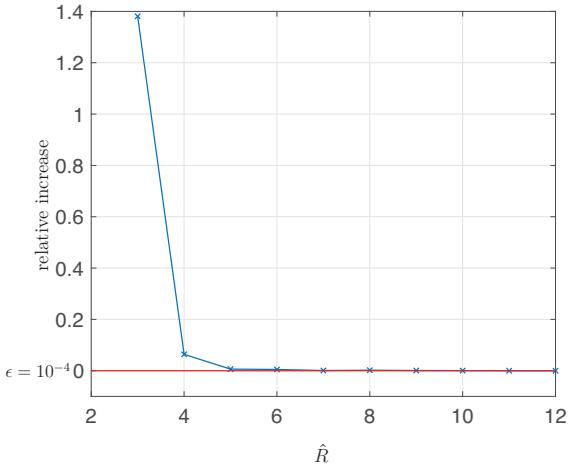

Figure 4.5.: The relative increase in the likelihood function between two successive model orders.

Figure 4.6 and Figure 4.7. The relative frequency of the observed data $y = y_\mathrm{r} + \jmath y_\mathrm{i}$ at the output of the AFB under H_1 is represented in Figure 4.6 for the described Matérn hard-core interference simulation. Figure 4.7 represents the approximated distribution $p_{Y_{\ell,k}}(y|\,\mathrm{H}_1)\,\Delta_\mathrm{area}$ which coincides with the histogram shown in Figure 4.6, based on the conducted GMM estimation. Finally, Δ_area indicates the area of the histogram bin given by $\Delta_\mathrm{area} = \Delta_\mathrm{area,r}\,\Delta_\mathrm{area,i}$, where $\Delta_\mathrm{area,r} = \Delta_\mathrm{area,i} = 0.2$ [1].

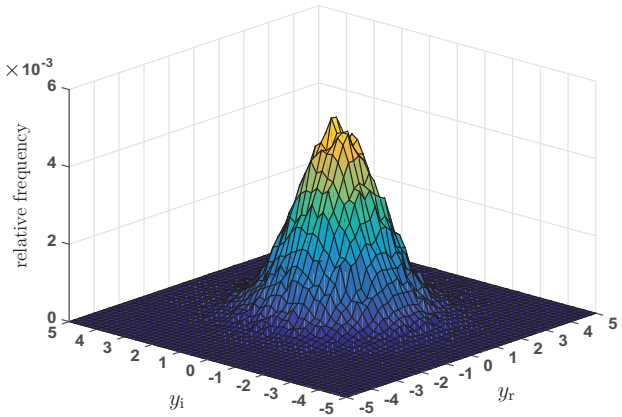

Figure 4.6.: Normalized histogram of observed data $y = [y_\mathrm{r}, y_\mathrm{i}]$ [1].

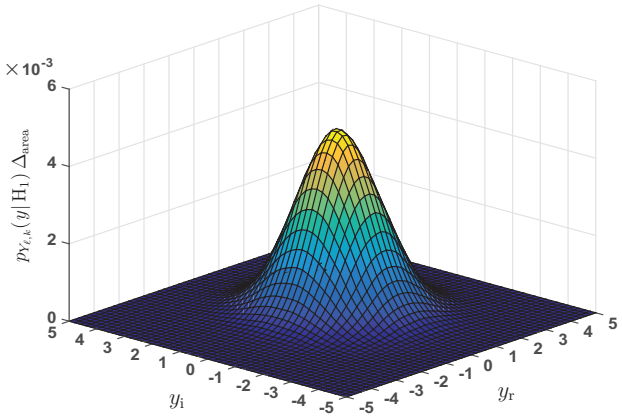

Figure 4.7.: The GMM approximation $p_{Y_{\ell,k}}(y|\,\mathrm{H}_1)\,\Delta_\mathrm{area}$, which corresponds to Figure 4.6 [1].

Chapter 5.

5 Prediction Schemes for a System of Independent Subbands

This chapter proposes TF prediction scheme for a system of K subbands, where the contiguous subbands are considered mutually independently occupied by the PU signal. The SU receiver structure presented in Figure 4.3 has been taken into consideration, similar to spectrum sensing in [1], such that any interdependencies between neighboring subbands are disregarded at this point. In order to be able to explore the occupancy information at TF slot (ℓ, k), the properly designed DFT-FB discussed in Chapter 4 has been considered in this chapter to obtain the TF characteristic of the time-domain received signal. The previously introduced discrete random process $Q_{\ell,k} \in \{0, 1\}$ is used to represent the occupancy information of the TF slot (ℓ, k) based on its availability to the SU. The case with $Q_{\ell,k} = 1$ represents the availability of the (ℓ, k) slot for the SU to access due to the absence of PUs, whereas $Q_{\ell,k} = 0$ indicates its unavailability.

Given those considerations, this chapter discusses the following two TF prediction schemes:

1. The conventional DHMM based scheme, which has been widely implemented in CRNs [42, 127, 43], is first discussed in Section 5.2. Unlike the available DHMM based approaches in literature, the SU utilizes the critically sampled DFT-FB Chapter 4 without the need for conventional subband filtering. Nevertheless,

a proper discretization process is still required to map the continuous-valued signal $Y_{\ell,k}$ into its discrete version which fits into the DHMM scheme, as to be discussed.

2. The alternative novel TF based CHMM prediction scheme is proposed in Section 5.3. This scheme employs $Y_{\ell,k}$ directly to learn the available environment and predict the future spectrum occupancy. Hence, both subband filtering and discretization post-processing methods are eliminated in this scheme.

In this work, the PU signal transmission is performed in frames comprising of a overall pilot sequence of length L_{P}, where $L_{\mathrm{P}} \in \mathbb{N}$, *and* PU spectrum access. This ensures that the SU is able to keep track of the PU transmission and learn its pattern. From the SU's point of view and irrespective of the subband's independencies, the prediction steps consist of three consecutive phases over time: a pilot phase as part of spectrum sensing, learning phase, and spectrum occupancy prediction phase.

During the *pilot phase*, the SU establishes a conditional model describing the received signal based on the PU transmission. In other words, the signal at the output of the AFB is modeled in the form of PDFs conditioned on the vector of subbands state, as can be found in Section 6.2. The conditional PDFs are approximated by GMMs in view of (4.6). However, this general case can be reduced into two hypotheses (occupancy states) whenever the interdependencies between the subbands are neglected. Hence, the parameters of the conditional GMMs are estimated under the two hypotheses in $Q \in \{0, 1\}$ based on a suitable pilot sequence of a certain length. For this matter, we assume a full cooperation from the PU side to transmit the required and sufficient pilot sequences. The successive learning phase allows the prediction scheme to learn and update its underlying parameters based on the current PU transmission upon observing a so-called learning sequence of length L_{L}, where $L_{\mathrm{L}} \in \mathbb{N}$. Finally, estimating the spectrum occupancy states at L_{O}, where $L_{\mathrm{O}} \in \mathbb{N}$ future time slots is carried out during the occupancy prediction phase based on the estimated parameters in the previous two phases. Figure 5.1 illustrates the prediction steps via the PU and SU signals.

As any interdependencies between neighboring subbands are disregarded in this chapter, which could result from the wireless protocol at hand [2], the K subbands

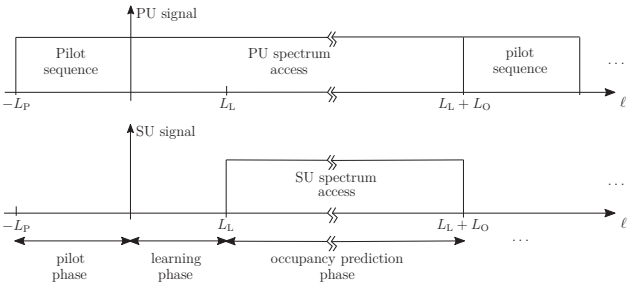

Figure 5.1.: PU and SU signals during the prediction steps.

are modeled independently. Hence, the prediction steps in the form of the three previously described phases are performed for each subband $k \in \Delta$ separately.

The received signal $Y_{\ell,k}$ at the output of the AFB contains the following TF components: $S_{\ell,k}$, $W_{\ell,k}$, and $U_{\ell,k}$ being the PU, AWGN thermal noise, and CCI signals, respectively. For a given subband k', the previously mentioned processes are expressed as $Q_{\ell,k'} = Q_\ell$, $S_{\ell,k'} = S_\ell$, $W_{\ell,k'} = W_\ell$, and $U_{\ell,k'} = U_\ell$, where the dependency on the subband frequency is dropped. Thus, the processes can be simply represented as Q_ℓ, Y_ℓ, S_ℓ, W_ℓ, and U_ℓ. Both S_ℓ and W_ℓ are modeled as circularly symmetric zero-mean complex white Gaussian processes with average powers per subband ν_S and ν_W, respectively. Moreover, the CCI signal U_ℓ, with an average power ν_U per subband, is generated based on the Monte Carlo simulation described in Section 4.3.

In this system of subbands, the state of a certain subband can be expressed based on Y_ℓ in analogy to (4.2) as

$$\mathrm{H}_{Q_\ell} : Y_\ell = (1 - Q_\ell)\, S_\ell + W_\ell + U_\ell, \tag{5.1}$$

with the mutually independent processes S_ℓ, W_ℓ and U_ℓ. The signal Y_ℓ is modeled as a temporally stationary process, where its temporal memory is exclusively introduced in $Q_\ell \in \{0, 1\}$. Assuming a simple spectrum access scheme by the PUs, Q_ℓ is considered to be a wide-sense stationary (WSS) in ℓ. As Q_ℓ is actually *unknown* or *hidden* to the SU receiver, it is modeled as a first-order hidden Markov process

based on the Markov assumption or Markov property [128], such that

$$\begin{aligned} & \Pr\{Q_\ell = q_\ell | \, Q_{\ell-1} = q_{\ell-1}, Q_{\ell-2} = q_{\ell-2}, \ldots\} \\ = \; & \Pr\{Q_\ell = q_\ell | \, Q_{\ell-1} = q_{\ell-1}\}, \end{aligned} \tag{5.2}$$

where $q_\ell \in \{0, 1\}$ is a realization of Q_ℓ. In other words, the current occupancy state of a certain subband depends on its previous state and irrespective of its own occupancy history as well as the state occupancy of its neighboring subbands. Since the phase $\sphericalangle Y_\ell$ of the observed signals at the AFB outputs has no information about the availability of the PUs signal, the outputs Z_ℓ of an energy detector (ED) per subband are considered onwards, where $Z_\ell = |Y_\ell|^2$. Hence, the complexity of the SU receiver, presented in Figure 5.2, is limited [2].

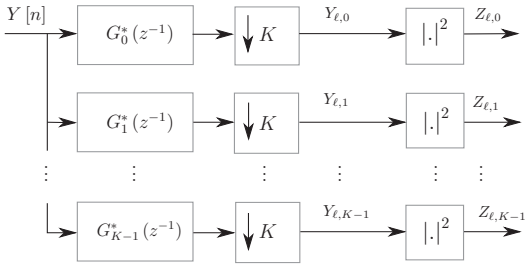

Figure 5.2.: The structure of a SU receiver followed by EDs per subband.

5.1. Pilot Phase

During this initial phase, a sequence of length $\tilde{L}_\mathrm{P} \in \mathbb{N}$ under *each* hypothesis is considered in order to characterize Z_ℓ in the form of conditional PDFs under $Q \in \{0, 1\}$, such that the overall considered pilot sequence $L_\mathrm{P} = 2\tilde{L}_\mathrm{P}$. Notice that the statistics under H_1 have already been estimated in Section 4.3, and the statistics under H_0 is only required at this stage. The conditional PDFs, which are expressed by $p_{Z_\ell}(z | \, Q_\ell = i) = p_Z(z | \, Q = i)$ due to the stationarity assumption with $i \in \{0, 1\}$,

are approximated by GMM as

$$p_Z(z|\,Q=i)=\sum_{r=1}^{R} w_{r,i}\,\mathcal{N}(z;\mu_{r,i},\sigma_{r,i}^2), \tag{5.3}$$

with R number of Gaussian components [2]. The GMM parameters: $w_{r,i}$, $\mu_{r,i}$ and $\sigma_{r,i}^2$ are the Gaussian PDF's weights, expected values and variances, respectively under the state i. The weights $w_{r,i}$ satisfy the conditions in (4.7). As presented in [2], those components are collected in the $(R\times 2)$-dimensional matrices

$$\mathbf{W}=(w_{r,i})\,,\quad \mathbf{M}=(\mu_{r,i})\,,\quad \text{and}\quad \mathbf{\Sigma}=\left(\sigma_{r,i}^2\right),$$

respectively. The EM algorithm along with k-means clustering [2, 125] are employed to estimate $\mathbf{W}$, $\mathbf{M}$, $\mathbf{\Sigma}$ and R, as shown in Section 4.3 [2]. Hence, a best-fit GMM under each hypothesis is obtained. It worth reminding at this point, that the pilot phase is common for both DHMM and CHMM based prediction schemes, which are described in the following sections.

5.2. Discrete Hidden Markov Model Based Prediction Scheme

Given the nature of the observed (received) signal Z_ℓ in (5.3) under each state in Q, a NP test [129] is utilized to map this continuous-valued signal into a discrete-valued observation $Z_{\mathrm{D},\ell}$. This procedure is actually needed to convert the continuous received signal Z_ℓ into a discrete one to fit in the DHMM scheme. The NP test is formulated based on a maximum false-alarm probability P_{FA} as [2]

$$Z_\ell \underset{Z_{\mathrm{D},\ell}=1}{\overset{Z_{\mathrm{D},\ell}=0}{\gtrless}} \tau_{\mathrm{NP}}, \tag{5.4}$$

where $\tau_{\mathrm{NP}}=\tau_{\mathrm{NP}}\left(P_{\mathrm{FA}}\right)$ is a numerically calculated threshold based on the PDFs in (5.3). Consistent with (5.1), H_1 implies the availability of the considered TF slot to the SU to access as $Z_{\mathrm{D},\ell}$ does not contain the PU component S_ℓ. Thus, $Z_{\mathrm{D},\ell}=1$ is determined by $Z_\ell<\tau_{\mathrm{NP}}$ in (5.4), otherwise $Z_{\mathrm{D},\ell}=0$ [2]. In other words, $Z_{\mathrm{D},\ell}\in\{0,1\}$

can be considered as an initial memoryless estimate of Q_ℓ, which however is to be improved via the DHMM based scheme given the memory introduced in (5.2).

In order to properly present the TF based DHMM spectrum occupancy prediction scheme, its underlying parameters are first introduced. Afterwards, both the scheme's learning and prediction phases are subsequently explained. Within the scenario in (5.1), the two hidden states per subband are indicated by $Q_\ell \in \{0, 1\}$. Likewise, there are two possible discrete-valued observations or alphabets indicated in $Z_{\mathrm{D},\ell} \in \{0, 1\}$, and they are conditioned on each state in Q_ℓ. For time slots $\ell = \{0, \ldots, L_\mathrm{L} - 1\}$, with L_L, a first-order DHMM scheme, which is presented graphically in Figure 5.3, is characterized by the following parameters or probabilities [2, 128]:

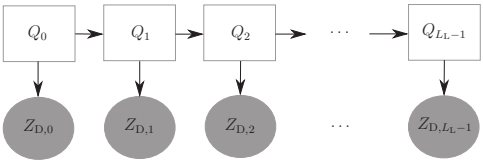

Figure 5.3.: A first-order DHMM for an observed sequence of length L_L [2].

- The transition probabilities $a_{ij}(\ell)$ from state $Q_{\ell-1} = i$ to state $Q_\ell = j$, for $(i, j) \in \{0, 1\} \times \{0, 1\}$, are defined by

$$a_{ij}(\ell) = \Pr\{Q_\ell = j | Q_{\ell-1} = i\}. \tag{5.5}$$

- The conditional emission probabilities, $b_{iz_\mathrm{D}}(\ell)$, based on each state in Q_ℓ, are defined as

$$b_{iz_\mathrm{D}}(\ell) = \Pr\{Z_{\mathrm{D},\ell} = z_{\ell,\mathrm{D}} | Q_\ell = i\}, \tag{5.6}$$

  where $z_{\ell,\mathrm{D}}$ is a realization of $Z_{\mathrm{D},\ell}$, and $(i, z_\mathrm{D}) \in \{0, 1\} \times \{0, 1\}$.

- The initial state probabilities for $k \in \Delta$ are defined as

$$\kappa_{i,k} = \Pr\{Q_{0,k} = i\} = \kappa_i, \tag{5.7}$$

  with $i \in \{0, 1\}$, and κ_i satisfy the condition $\kappa_0 + \kappa_1 = 1$ [128].

Due to the stationarity assumption, (5.5) and (5.6) are subsequently expressed simply by a_{ij} and $b_{iz_{\mathrm{D}}}$, as

$$a_{ij}(\ell) = a_{ij} = \Pr\{Q = j | Q = i\},$$

and

$$b_{iz_{\mathrm{D}}}(\ell) = b_{iz_{\mathrm{D}}} = \Pr\{Z_{\mathrm{D}} = z_{\mathrm{D}} | Q = i\}.$$

All possible transition probabilities are assembled in a $(2 \times 2) -$ dimensional (one-step) transition matrix $\mathbf{A}$, where $\mathbf{A} = (a_{ij})$ and $a_{i0} + a_{i1} = 1$ [130]. Similarly, the (2×2) −dimensional emission matrix $\mathbf{B}$ which contains all probabilities $b_{iz_{\mathrm{D}}}$, such that $\mathbf{B} = (b_{iz_{\mathrm{D}}})$ and $b_{i0} + b_{i1} = 1$ [130].

At this point, it is worth mentioning that the quantities in $\mathbf{A}$ are essential to the presented schemes, whether DHMM or CHMM based prediction approaches. They reflect the PU traffic pattern, as they model the subband's sequential occupancy information in the form of probabilities for all $i, j \in Q$. On the other hand, $\mathbf{B}$ is considered as the probability of correct detection, missed detection and false-alarm in a DHMM based scheme [42].

Furthermore, Figure 5.4 represents a_{ij} and $b_{iz_{\mathrm{D}}}$ for the presented DHMM prediction scheme and the considered CRN scenario. Clearly, one can introduce Γ_{D} as a set

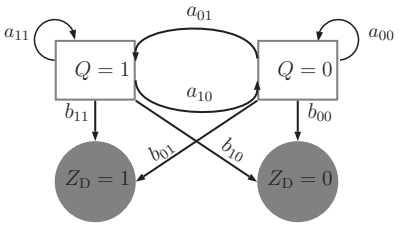

Figure 5.4.: Transition and emission probabilities in a DHMM scheme [2].

which contains the previously explained probabilities as

$$\Gamma_{\mathrm{D}} = (\mathbf{A}, \mathbf{B}), \tag{5.8}$$

with each quantity taking a value in Ω, which is defined as

$$\Omega = \{\, x \in \mathbb{R} |\, 0 \leq x \leq 1 \}\,, \tag{5.9}$$

hence $\Gamma_{\mathrm{D}} \in \Omega_{\mathrm{D}}$, with $\Omega_{\mathrm{D}} = \Omega^8$. Based on Figure 5.3, the joint probability distribution for the observed sequence

$$Z_{\mathrm{D},0} = z_{0,\mathrm{D}}, \ldots, Z_{\mathrm{D},L_{\mathrm{L}}-1} = z_{L_{\mathrm{L}}-1,\mathrm{D}}, \tag{5.10}$$

along with the corresponding hidden states

$$Q_0 = q_0, \ldots, Q_{L_{\mathrm{L}}-1} = q_{L_{\mathrm{L}}-1}, \tag{5.11}$$

based on Γ_{D} is given by [52]

$$\Pr\{\, z_{0,\mathrm{D}}, \ldots, z_{L_{\mathrm{L}}-1,\mathrm{D}}, q_0, \ldots, q_{L_{\mathrm{L}}-1}\,; \Gamma_{\mathrm{D}} \} = \kappa_{q_0} \prod_{\ell=1}^{L_{\mathrm{L}}-1} a_{ij}\,(\ell) \prod_{\ell=0}^{L_{\mathrm{L}}-1} b_{iz_{\mathrm{D}}}\,(\ell)\,. \tag{5.12}$$

5.2.1. Learning Phase

During this phase, the underlying parameters of the DHMM based scheme in (5.8) are estimated based on the previously introduced learning sequence. For this matter, the sequence of length L_{L} defined in (5.10) is considered as the learning sequence, observed directly after the pilot sequence. In other words, a best-fit DHMM based scheme must be estimated in order to best describe the observed PU traffic pattern. Notice that at this point the sequence of states in (5.11) is hidden to the SU receiver, as the spectrum occupancy information is indeed unknown to the SU [13].

Analogous to DBNs with regard to known structure and complete observed data, the maximum-likelihood estimation (MLE) method is a suitable approach for this task [131]. Hence, a maximum-likelihood (ML) estimate of Γ_{D}, namely $\hat{\Gamma}_{\mathrm{D,ML}}$, is expressed by

$$\hat{\Gamma}_{\mathrm{D,ML}} = \arg \max_{\Gamma_{\mathrm{D}} \in \Omega_{\mathrm{D}}} \Pr\{\, z_{0,\mathrm{D}}, \ldots, z_{L_{\mathrm{L}}-1,\mathrm{D}}\,; \Gamma_{\mathrm{D}} \}\,, \tag{5.13}$$

with $\Pr\{\, z_{0,\mathrm{D}}, \ldots, z_{L_{\mathrm{L}}-1,\mathrm{D}}\,; \Gamma_{\mathrm{D}} \}$ is the corresponding likelihood function obtained from marginalizing (5.12) over the states $Q_0, \ldots, Q_{L_{\mathrm{L}}-1}$. As (5.13) can not be solved

analytically [2, 52, 128], dynamic programing methods are required. Known as a generalization of the EM algorithm and as an unsupervised learning method for HMM schemes [52, 132], the forward-backward algorithm – also referred in literature as the Baum-Welch algorithm [130]– is utilized as an iterative procedure based on the ML criterion. Among different learning (training) methods discussed in [43, 45] for this task, this algorithm provides better performance despite its complexity [45]. Via the forward-backward algorithm, the problem in (5.13) is broken down into the following two quantities [128]:

1. The forward probability $\alpha_{\ell,\mathrm{D}}(i)$, defined as the joint probability of observing the sequence $Z_{\mathrm{D},0} = z_{0,\mathrm{D}}, \ldots, Z_{\mathrm{D},\ell} = z_{\ell,\mathrm{D}}$ until time slot ℓ and having the state $Q_\ell = i$ given the model described in Γ_{D}. It is expressed by

$$\alpha_{\ell,\mathrm{D}}(i) = \Pr\left\{ z_{0,\mathrm{D}}, \ldots, z_{\ell,\mathrm{D}}, Q_\ell = i \,; \Gamma_{\mathrm{D}} \right\}. \tag{5.14}$$

2. The backward probability $\beta_{\ell,\mathrm{D}}(i)$, defined as the probability of observing the sequence from $\ell + 1$ until its end, i.e., $Z_{\mathrm{D},\ell+1} = z_{\ell+1,\mathrm{D}}, \ldots, Z_{\mathrm{D},L_{\mathrm{L}}-1} = z_{L_{\mathrm{L}}-1,\mathrm{D}}$, conditioned on the current state $Q_\ell = i$ for the model Γ_{D}. It is expressed by

$$\beta_{\ell,\mathrm{D}}(i) = \Pr\left\{ z_{\ell+1,\mathrm{D}}, \ldots, z_{L_{\mathrm{L}}-1,\mathrm{D}} \,|Q_\ell = i; \Gamma_{\mathrm{D}} \right\}. \tag{5.15}$$

The basic idea behind the forward-backward algorithm is to break down the chain in Figure 5.3 into two parts: the past and the future, conditioned on the current state $Q_\ell = i$ [55] and for the model defined in Γ_{D}. For this purpose we introduce the probability $\gamma_{\ell,\mathrm{D}}(i)$, defined by

$$\gamma_{\ell,\mathrm{D}}(i) \triangleq \Pr\left\{ Q_\ell = i | \, z_{0,\mathrm{D}}, \ldots, z_{L_{\mathrm{L}}-1,\mathrm{D}} \right\}, \tag{5.16}$$

which can be formulated based on the Bayes rule as

$$\gamma_{\ell,\mathrm{D}}(i) = \frac{\Pr\left\{ z_{0,\mathrm{D}}, \ldots, z_{L_{\mathrm{L}}-1,\mathrm{D}} | \, Q_\ell = i \right\} \Pr\left\{ Q_\ell = i \right\}}{\Pr\left\{ z_{0,\mathrm{D}}, \ldots, z_{L_{\mathrm{L}}-1,\mathrm{D}} \right\}}. \tag{5.17}$$

Since the sequence $z_{0,\mathrm{D}}, \ldots, z_{\ell,\mathrm{D}}$ is conditionally independent from the sequence $z_{\ell+1,\mathrm{D}}, \ldots, z_{L_{\mathrm{L}}-1,\mathrm{D}}$ given Q_ℓ and based on the so-called d-separation criterion, the

expression

$$\Pr\{z_{0,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}}|\,Q_\ell=i\}= \Pr\{z_{0,\mathrm{D}},\ldots,z_{\ell,\mathrm{D}}|\,Q_\ell=i\}\Pr\{z_{\ell+1,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}}|\,Q_\ell=i\},$$

is valid [52], and $\gamma_{\ell,\mathrm{D}}(i)$ can be rewritten as

$$\gamma_{\ell,\mathrm{D}}(i)=\frac{\Pr\{z_{0,\mathrm{D}},\ldots,z_{\ell,\mathrm{D}},Q_\ell=i\}\Pr\{z_{\ell+1,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}}|\,Q_\ell=i\}}{\Pr\{z_{0,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}}\}}.$$

Considering the definitions of $\alpha_{\ell,\mathrm{D}}(i)$ and $\beta_{\ell,\mathrm{D}}(i)$ in (5.14) and (5.15), respectively, the following relation

$$\gamma_{\ell,\mathrm{D}}(i)\propto\alpha_{\ell,\mathrm{D}}(i)\,\beta_{\ell,\mathrm{D}}(i)$$

hence is valid. Please refer to Appendix A for further explanations regarding the d-separation criterion.

The forward-backward algorithm, which is the main part of the learning phase, is common for all prediction schemes provided in this work, however with appropriate modifications to fit each prediction scheme. This algorithm is conducted based on the following detailed steps [128, 52, 130]:

1. Starting from initial estimates $\hat{\Gamma}_\mathrm{D}^{(0)}$, given by $\hat{\Gamma}_\mathrm{D}^{(0)}=\left(\hat{\mathbf{A}}^{(0)},\hat{\mathbf{B}}^{(0)}\right)$ with $\hat{\mathbf{A}}^{(0)}$ and $\hat{\mathbf{B}}^{(0)}$ are the initial estimate of $\mathbf{A}^{(0)}$ and $\mathbf{B}^{(0)}$, respectively, the corresponding estimates of $\alpha_{i,\mathrm{D}}(\ell)$ and $\beta_{i,\mathrm{D}}(\ell)$, namely $\hat{\alpha}_{\ell,\mathrm{D}}(i)$ and $\hat{\beta}_{\ell,\mathrm{D}}(i)$, respectively, are calculated. The quantity $\hat{\alpha}_{\ell,\mathrm{D}}(i)$ is calculated iteratively as

   a) initialization
   $$\hat{\alpha}_{0,\mathrm{D}}(i)=\kappa_i\,\hat{b}_{iz_\mathrm{D}}^{(0)}(0),$$

   b) iteration over $\ell=1,\ldots,L_\mathrm{L}-2$
   $$\hat{\alpha}_{\ell,\mathrm{D}}(i)=\hat{b}_{iz_\mathrm{D}}^{(0)}(\ell)\sum_j\hat{\alpha}_{\ell-1,\mathrm{D}}(j)\,\hat{a}_{ji}^{(0)},\tag{5.18}$$

   c) termination
   $$\Pr\{z_{0,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}}\}=\sum_i\hat{\alpha}_{L_\mathrm{L}-1,\mathrm{D}}(i).$$

On the other hand, $\hat{\beta}_{\ell,\mathrm{D}}(i)$ is calculated iteratively as

a) initialization

$$\hat{\beta}_{L_\mathrm{L}-1,\mathrm{D}}(i) = 1,$$

b) iteration over $\ell = L_\mathrm{L} - 2 \ldots, 0$

$$\hat{\beta}_{\ell,\mathrm{D}}(i) = \sum_j \hat{b}^{(0)}_{jz_\mathrm{D}}(\ell+1)\ \hat{a}^{(0)}_{ij}\ \hat{\beta}_{\ell+1,\mathrm{D}}(j). \tag{5.19}$$

The quantities $\hat{b}^{(0)}_{iz_\mathrm{D}}(\ell)$, $\hat{b}^{(0)}_{jz_\mathrm{D}}(\ell)$, $\hat{a}^{(0)}_{ji}$ and $\hat{a}^{(0)}_{ij}$, are the initial estimates of $b_{iz_\mathrm{D}}(\ell), b_{jz_\mathrm{D}}(\ell), a_{ji}(\ell)$ and $a_{ij}(\ell)$, respectively.

Please refer to Appendix A for detailed explanation on deriving the iterative expression for $\alpha_{i,\mathrm{D}}(\ell)$ and $\beta_{i,\mathrm{D}}(\ell)$.

2. For estimation purposes, two intermediate quantities are introduced at this point [128]: *first*, $\gamma_{\ell,\mathrm{D}}(i)$ as defined in (5.16). *Second*, $\zeta_{\ell,\mathrm{D}}(i,j)$, defined as the probability of having the states $Q_{\ell-1} = i$ and $Q_\ell = j$ given the observed learning sequence and the model in Γ_D, namely

$$\zeta_{\ell,\mathrm{D}}(i,j) \triangleq \Pr\{Q_{\ell-1} = i, Q_\ell = j \,|z_{0,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}}; \Gamma_\mathrm{D}\}. \tag{5.20}$$

The right-hand side in (5.20) is rewritten via Bayes rule [52] as

$$\frac{\Pr\{z_{0,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}}| Q_{\ell-1} = i, Q_\ell = j\} \Pr\{Q_{\ell-1} = i, Q_\ell = j\}}{\Pr\{z_{0,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}}; \Gamma_\mathrm{D}\}}. \tag{5.21}$$

As $\Pr\{Q_{\ell-1} = i, Q_\ell = j\}$ in (5.21) is expressed by

$$\Pr\{Q_{\ell-1} = i, Q_\ell = j\} = \Pr\{Q_\ell = j| Q_{\ell-1} = i\} \Pr\{Q_{\ell-1} = i\},$$

and given the parent-child relationship represented in Figure 5.3, the following

relation regarding $\Pr\{z_{0,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}}|\,Q_{\ell-1}=i,Q_\ell=j\}$ in (5.21)

$$\begin{aligned}\Pr\{z_{0,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}}|\,Q_{\ell-1}=i,Q_\ell=j\}=&\\ &\Pr\{z_{0,\mathrm{D}},\ldots,z_{\ell-1,\mathrm{D}}|\,Q_{\ell-1}=i\}\\ &\Pr\{z_{\ell,\mathrm{D}}|\,Q_\ell=j\}\\ &\Pr\{z_{\ell+1,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}}|\,Q_\ell=j\},\end{aligned}$$

is true. Consequently, $\zeta_{\ell,\mathrm{D}}(i,j)$ can be calculated based on the previous relations and in terms of the DHMM scheme's parameters in Γ_D as

$$\zeta_{\ell,\mathrm{D}}(i,j)=\frac{\alpha_{\ell,\mathrm{D}}(i)a_{ij}b_{jz_\mathrm{D}}(\ell+1)\,\beta_{\ell+1,\mathrm{D}}(i)}{\Pr\{z_{0,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}};\Gamma_\mathrm{D}\}}. \tag{5.22}$$

Regarding the estimation problem, estimates of $\zeta_{\ell,\mathrm{D}}(i,j)$ and $\gamma_{\ell,\mathrm{D}}(i)$, namely $\hat{\zeta}_{\ell,\mathrm{D}}(i,j)$ and $\hat{\gamma}_{\ell,\mathrm{D}}(i)$, respectively, can be calculated based on the initial model as

$$\hat{\zeta}_{\ell,\mathrm{D}}(i,j)=\frac{\hat{\alpha}_{\ell,\mathrm{D}}(i)\hat{a}_{ij}^{(0)}\hat{b}_{jz_\mathrm{D}}^{(0)}(\ell+1)\,\hat{\beta}_{\ell+1,\mathrm{D}}(j)}{\Pr\left\{z_{0,\mathrm{D}},\ldots,z_{L_\mathrm{L}-1,\mathrm{D}};\hat{\Gamma}_\mathrm{D}\right\}},$$

and by marginalizing over j, $\hat{\gamma}_{\ell,\mathrm{D}}(i)$ is subsequently obtained as

$$\hat{\gamma}_{\ell,\mathrm{D}}(i)=\sum_{j\in Q}\hat{\zeta}_{\ell,\mathrm{D}}(i,j).$$

3. Hence, a new update of the DHMM scheme's parameters, namely $\hat{\Gamma}_\mathrm{D}$, are calculated as

$$\begin{aligned}\hat{a}_{ij}^{(1)}=&\frac{\text{expected number of transitions from } i \text{ to } j}{\text{expected number of transitions from } i}\\ =&\frac{\sum_{\ell=0}^{L_\mathrm{L}-1}\hat{\zeta}_{\ell,\mathrm{D}}(i,j)}{\sum_{\ell=0}^{L_\mathrm{L}-1}\hat{\gamma}_{\ell,\mathrm{D}}(i)}\end{aligned}$$

and

$$\hat{b}_{iz_{\mathrm{D}}}^{(1)}(\ell) = \frac{\text{expected number of times in state } i \text{ and observing } z_{\mathrm{D}}}{\text{expected number of times in state } i} = \frac{\overset{\text{with } Z_{\mathrm{D},\ell}=z_{\mathrm{D}}}{\overset{L_{\mathrm{L}}-1}{\underset{\ell=0}{\sum}}} \hat{\gamma}_{\ell,\mathrm{D}}(i)}{\overset{L_{\mathrm{L}}-1}{\underset{\ell=0}{\sum}} \hat{\gamma}_{\ell,\mathrm{D}}(i)}$$

4. Similar to the EM algorithm, the forward-backward algorithm takes place with a total $m_{\mathrm{it,D}}$ number of iterations, until the increase in the likelihood function is below the threshold ϵ, according to

$$\left| \frac{\Pr\left\{ z_{0,\mathrm{D}},\ldots,z_{L_{\mathrm{L}}-1,\mathrm{D}};\hat{\Gamma}_{\mathrm{D}}^{\left(m_{\mathrm{it,D}}\right)} \right\} - \Pr\left\{ z_{0,\mathrm{D}},\ldots,z_{L_{\mathrm{L}}-1,\mathrm{D}};\hat{\Gamma}_{\mathrm{D}}^{\left(m_{\mathrm{it,D}}-1\right)} \right\}}{\Pr\left\{ z_{0,\mathrm{D}},\ldots,z_{L_{\mathrm{L}}-1,\mathrm{D}};\hat{\Gamma}_{\mathrm{D}}^{\left(m_{\mathrm{it,D}}-1\right)} \right\}} \right| \leq \epsilon.$$

The resulting model $\hat{\Gamma}_{\mathrm{D}}^{\left(m_{\mathrm{it,D}}\right)}$ of the learning phase, namely $\hat{\Gamma}_{\mathrm{D,ML}} \approx \hat{\Gamma}_{\mathrm{D}}^{\left(m_{\mathrm{it,D}}\right)}$, is to be employed in the final prediction phase in order to estimate the future occupancy for each subband at $\ell \geq L_{\mathrm{L}}$.

Nevertheless, a proper scaling for both $\alpha_{i,\mathrm{D}}(\ell)$ and $\beta_{i,\mathrm{D}}(\ell)$ must be introduced during implementation as their values approach zero quickly with each iteration [52], irrespective of considering DHMM or CHMM schemes. Therefore, $\alpha_{i,\mathrm{D}}(\ell)$ and $\beta_{i,\mathrm{D}}(\ell)$ are normalized at each iterative step by $\sum_i \alpha_{i,\mathrm{D}}(\ell)$, and $\sum_i \beta_{i,\mathrm{D}}(\ell)$, respectively, such that the sum of each quantity at each iteration step equals 1 [133]. The scaled versions of $\alpha_{i,\mathrm{D}}(\ell)$ and $\beta_{i,\mathrm{D}}(\ell)$, denoted by $\tilde{\alpha}_{i,\mathrm{D}}(\ell)$ and $\tilde{\beta}_{i,\mathrm{D}}(\ell)$, respectively, are subsequently defined [52]

$$\tilde{\alpha}_{i,\mathrm{D}}(\ell) = \Pr\left\{ i \middle| z_{0,\mathrm{D}},\ldots,z_{\ell,\mathrm{D}} \right\},$$

and

$$\tilde{\beta}_{i,\mathrm{D}}(\ell) = \frac{\Pr\left\{ z_{0,\mathrm{D}},\ldots,z_{\ell,\mathrm{D}} \middle| Q_\ell = i \right\}}{\Pr\left\{ z_{\ell+1,\mathrm{D}},\ldots,z_{L_{\mathrm{L}}-1,\mathrm{D}} \middle| z_{0,\mathrm{D}},\ldots,z_{\ell,\mathrm{D}} \right\}}.$$

This scaling procedure has been considered as an implementation issue, which does not affect the concept of spectrum occupancy prediction algorithms. Hence, both the forward and backward quantities in the discussed schemes in this work are continued to be defined unnormalized to reduce notations complexity.

5.2.2. Occupancy Prediction Phase

To this end, the sequence of states in (5.11), which has generated the corresponding observed learning sequence in (5.10), is actually hidden to the SU receiver. Therefore, based on $\hat{\Gamma}_{\mathrm{D}}^{\left(m_{\mathrm{it,D}}\right)}$ as the outcome of the learning phase, the hidden states are estimated using the Viterbi algorithm. Considered as an efficient decoding method [2, 134], the Viterbi algorithm estimates the corresponding most likely state sequence [52, 128] starting from 0 until $L_{\mathrm{L}}-1$, namely $\hat{q}_0,\ldots,\hat{q}_{L_{\mathrm{L}}-1}$. Furthermore, the decoding procedure is carried out to estimate $\hat{q}_\ell$ for future time slots $\ell \geq L_{\mathrm{L}}$ [13].

At each time slot $0,1,\ldots,\ell$ the Viterbi algorithm calculates the most likely path until time slot ℓ, such that only a single path survives [52]. For the state j, where $j\in\{0,1\}$, the path until $Q_\ell=j$ is calculated by defining a cumulative best score $\delta_{\ell,\mathrm{D}}\left(j\right)$ which maximizes the joint probability of Q_ℓ and $Z_{\mathrm{D},\ell}$ starting from time slot 0 [13], as [128]

$$\begin{aligned}\delta_{\ell,\mathrm{D}}\left(j\right) &= \max_{q_0,\ldots,q_{\ell-1}} \Pr\left\{ z_{0,\mathrm{D}},\ldots,z_{\ell-1,\mathrm{D}},z_{\ell,\mathrm{D}},q_0,\ldots,q_{\ell-1},Q_\ell=j\,;\hat{\Gamma}_{\mathrm{C}}^{\left(m_{\mathrm{it,D}}\right)}\right\} \\ &= \hat{b}_{jz_{\mathrm{D}}}^{\left(m_{\mathrm{it,D}}\right)}\left(\ell\right)\max_i \delta_{\ell-1,\mathrm{D}}\left(i\right)\hat{a}_{ij}^{\left(m_{\mathrm{it,D}}\right)}.\end{aligned} \tag{5.23}$$

Hence, the estimate $\hat{Q}_\ell$ is obtained based on the following decision

$$\delta_{\ell,\mathrm{D}}\left(1\right) \underset{\hat{Q}_\ell=0}{\overset{\hat{Q}_\ell=1}{\gtrless}} \frac{1}{2}.$$

Since (5.23) produces an optimal sequence of states, it can be noticed that the maximization problem does not take into account the individual probability of occurrence

of Q_ℓ, as in

$$\hat{q}_\ell = \arg\max_i \hat{\gamma}_{\ell,\mathrm{D}}\left(Q_\ell = i\right). \tag{5.24}$$

Thus, any invalid transitions represented by $a_{ij} = 0$ are still taken into consideration in (5.23) unlike in (5.24) [128]. Furthermore, as the Viterbi algorithm utilizes a best-fit model $\hat{\Gamma}_{\mathrm{D}}^{\left(m_{\mathrm{it,D}}\right)}$ being estimated during the learning phase, no backtracking is required in the implemented prediction schemes. In other words, this final step consists of Viterbi decoding, not to be mistaken with Viterbi learning in [128].

5.3. Continuous Hidden Markov Model Based Prediction Scheme

It has been shown that significant performance degradation in speech recognition occurs when utilizing discrete observations to implement the HMM scheme as a speech recognizer in the form of a DHMM based approach. This performance degradation is due to the vector quantization process of speech signals, which introduces information loss [46]. Therefore, an alternative approach is presented in this work by employing a first-order CHMM based scheme which utilizes the signal directly at the output of the EDs and after the AFB, as in Figure 5.2, hence eliminating the need of the threshold in (5.4). During the learning phase associated with the CHMM based prediction scheme, the learning sequence

$$Z_0 = z_0, \ldots, Z_{L_{\mathrm{L}}-1} = z_{L_{\mathrm{L}}-1}, \tag{5.25}$$

is considered and observed directly after the pilot phase as it fits into the CHMM scheme without further processes. For this matter, $\mathbf{B}$ in Γ_{D} is replaced by estimates of the GMM parameters in (5.3). Therefore, only the parameters in $\mathbf{A} = (a_{ij})$ are defined in the set Γ_{C} as

$$\Gamma_{\mathrm{C}} = (\mathbf{A}), \tag{5.26}$$

which represents the alternative first-order CHMM based prediction model. During the learning phase, the parameters in Γ_{C}, namely the transition probabilities a_{ij}, are estimated, where $\Gamma_{\mathrm{C}} \in \Omega_{\mathrm{C}}$, $\Omega_{\mathrm{C}} = \Omega^4$, and Ω is defined in (5.9).

Moreover, Figure 5.5 represents a modified version of Figure 5.4, which describes the fundamental parameters for the CHMM based prediction scheme and the considered scenario as $Q_\ell \in \{0, 1\}$. Similar to the DHMM based prediction approach, a best-fit model is to be estimated during the learning phase as to be utilized thereafter during the prediction phase. Thus, the following sections explain the learning phase and occupancy prediction phase for the CHMM based scheme.

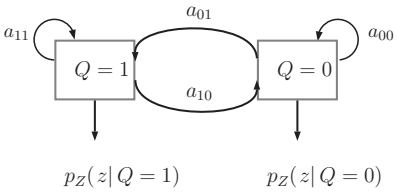

Figure 5.5.: Transition probabilities for the CHMM.

5.3.1. Learning Phase

As declared earlier, during its learning phase the set of parameters characterizing the CHMM scheme defined in Γ_C are estimated given the learning sequence in (5.25). The ML estimate of Γ_C, denoted as $\hat{\Gamma}_{\mathrm{C,ML}}$, is given by

$$\hat{\Gamma}_{\mathrm{C,ML}} = \arg \max_{\Gamma_\mathrm{C} \in \Omega_\mathrm{C}} p\,(\,z_0, \ldots, z_{L_\mathrm{L}-1}\,;\Gamma_\mathrm{C})$$

where $p\,(\,z_0, \ldots, z_{L_\mathrm{L}-1}\,;\Gamma_\mathrm{C})$ represents the corresponding likelihood function for the given learning sequence and the model defined in Γ_C. Similarly, the forward-backward algorithm is utilized. Starting with the model's initial knowledge $\hat{\Gamma}_\mathrm{C}^{(0)}$, its estimate is iteratively calculated by utilizing the forward and backward quantities $\alpha_{\ell,\mathrm{C}}(i)$ and $\beta_{\ell,\mathrm{C}}(i)$, respectively. For the CHMM based scheme and a given Γ_C, the quantities $\alpha_{\ell,\mathrm{C}}(i)$ and $\beta_{\ell,\mathrm{C}}(i)$ are defined as [134, 135]

$$\alpha_{\ell,\mathrm{C}}(i) = p\,(\,z_0, \ldots, z_\ell, Q_\ell = i\,;\Gamma_\mathrm{C})\,,$$

and

$$\beta_{\ell,\mathrm{C}}(i) = p\left(z_{\ell+1}, \ldots, z_{L_\mathrm{L}-1} \,|Q_\ell = i; \Gamma_\mathrm{C}\right),$$

respectively. Similar to the DHMM based scheme, the learning phase in the form of the forward-backward algorithm for the considered CHMM scheme can be outlined in the following steps [46, 128]:

1. Given $\hat{\Gamma}_\mathrm{C}^{(0)} = \left(\hat{\mathbf{A}}^{(0)}\right)$, the quantities $\hat{\alpha}_{\ell,\mathrm{C}}(i)$ and $\hat{\beta}_{\ell,\mathrm{C}}(i)$ are correspondingly obtained as initial estimates of $\alpha_{\ell,\mathrm{C}}(i)$ and $\beta_{\ell,\mathrm{C}}(i)$, respectively. The quantity $\hat{\alpha}_{\ell,\mathrm{C}}(i)$ is iteratively obtained as [136]

   a) initialization
   $$\hat{\alpha}_{0,\mathrm{C}}(i) = \kappa_i \, p_Z(\, z_0 |\, Q_0 = i)\,,$$

   b) iteration over $\ell = 1, \ldots, L_\mathrm{L} - 2$
   $$\hat{\alpha}_{\ell,\mathrm{C}}(i) = p_Z(\, z_\ell |\, Q_\ell = i) \sum_j \hat{\alpha}_{\ell-1,\mathrm{C}}(j)\ \hat{a}_{ji}^{(0)},$$

   c) termination
   $$p\left(z_0, \ldots, z_{L_\mathrm{L}-1}\,; \hat{\Gamma}_\mathrm{C}^{(0)}\right) = \sum_i \hat{\alpha}_{L_\mathrm{L}-1,\mathrm{C}}(i).$$

   Furthermore, $\hat{\beta}_{\ell,\mathrm{C}}(i)$ can be iteratively calculated as

   a) initialization
   $$\hat{\beta}_{L_\mathrm{L}-1,\mathrm{C}}(i) = 1,$$

   b) iteration over $\ell = L_\mathrm{L} - 2 \ldots, 0$
   $$\hat{\beta}_{\ell,\mathrm{C}}(i) = \sum_j p_Z(\, z_\ell |\, Q_\ell = j)\ \hat{a}_{ij}^{(0)}\ \hat{\beta}_{\ell+1,\mathrm{C}}(j).$$

2. In analogy to the quantities $\zeta_{\ell,\mathrm{D}}\,(i,j)$ and $\gamma_{\ell,\mathrm{D}}\,(i)$ in the DHMM based scheme, estimates of $\zeta_{\ell,\mathrm{C}}\,(i,j)$ and $\gamma_{\ell,\mathrm{C}}\,(i,r)$, namely $\hat{\zeta}_{\ell,\mathrm{C}}\,(i,j)$ and $\hat{\gamma}_{\ell,\mathrm{C}}\,(i,r)$, respec-

tively where $r = \{1, \ldots, R\}$, are obtained based on $\hat{\Gamma}_{\mathrm{C}}^{(0)}$ as

$$\begin{aligned}\hat{\zeta}_{\ell,\mathrm{C}}(i,j) &\triangleq p\left(i, j \,|z_0, \ldots, z_{L_{\mathrm{L}}-1}; \hat{\Gamma}_{\mathrm{C}}^{(0)}\right) \\ &= \frac{\hat{\alpha}_{\ell,\mathrm{C}}(i)\hat{a}_{ij}^{(0)} p_Z(z_{\ell+1}|\, Q_{\ell+1} = j)\, \hat{\beta}_{\ell+1,\mathrm{C}}(j)}{p\left(z_0, \ldots, z_{L_{\mathrm{L}}-1}\,; \hat{\Gamma}_{\mathrm{C}}\right)}\end{aligned} \tag{5.27}$$

and

$$\begin{aligned}\hat{\gamma}_{\ell,\mathrm{C}}(i,r) &\triangleq p\left(i, r \,|z_0 \ldots z_{L_{\mathrm{L}}-1}; \hat{\Gamma}_{\mathrm{C}}\right) \\ &= \left[\frac{\hat{\alpha}_{\ell,\mathrm{C}}(i)\hat{\beta}_{\ell,\mathrm{C}}(i)}{\sum\limits_{i} \hat{\alpha}_{\ell,\mathrm{C}}(i)\hat{\beta}_{\ell,\mathrm{C}}(i)}\right]\left[\frac{w_{r,i}\,\mathcal{N}(z; \mu_{r,i}, \sigma_{r,i}^2)}{\sum\limits_{h=1}^{R} w_{h,i}\,\mathcal{N}(z; \mu_{h,i}, \sigma_{h,i}^2)}\right],\end{aligned}$$

where $h = \{1, \ldots, R\}$.

3. A new update, namely $\hat{\Gamma}_{\mathrm{C}}$, is subsequently obtained as

$$\hat{a}_{ij}^{(1)} = \frac{\sum\limits_{\ell=0}^{L_{\mathrm{L}}-1} \hat{\zeta}_{\ell,\mathrm{C}}(i,j)}{\sum\limits_{\ell=0}^{L_{\mathrm{L}}-1} \sum\limits_{r=1}^{R} \hat{\gamma}_{\ell,\mathrm{C}}(i,r)}.$$

4. Likewise, a total number of $m_{\mathrm{it,C}}$ iterations are needed until the increase in the likelihood function is below ϵ as

$$\left|\frac{p\left(z_{0,\mathrm{D}},\ldots,z_{L_{\mathrm{L}}-1,\mathrm{D}}\,;\hat{\Gamma}_{\mathrm{C}}^{\left(m_{\mathrm{it,C}}\right)}\right) - p\left(z_{0,\mathrm{D}},\ldots,z_{L_{\mathrm{L}}-1,\mathrm{D}}\,;\hat{\Gamma}_{\mathrm{C}}^{\left(m_{\mathrm{it,C}}-1\right)}\right)}{p\left(z_{0,\mathrm{D}},\ldots,z_{L_{\mathrm{L}}-1,\mathrm{D}}\,;\hat{\Gamma}_{\mathrm{C}}^{\left(m_{\mathrm{it,C}}-1\right)}\right)}\right| \leq \epsilon,$$

where the resulting model is indicated by $\hat{\Gamma}_{\mathrm{C}}^{\left(m_{\mathrm{it,C}}\right)}$. Hence $\hat{\Gamma}_{\mathrm{C,ML}} \approx \hat{\Gamma}_{\mathrm{C}}^{\left(m_{\mathrm{it,C}}\right)}$.

5.3.2. Occupancy Prediction Phase

In a manner similar to the occupancy prediction phase in the DHMM based prediction scheme, the Viterbi algorithm is employed to decode the hidden states for time slots $\ell = 0, \ldots, L_{\mathrm{L}} - 1$, namely $\hat{q}_0, \ldots, \hat{q}_{L_{\mathrm{L}}-1}$. Afterwards, the procedure continues

to estimate $\hat{q}_\ell$ for future time slots where $\ell \geq L_{\mathrm{L}}$ [2]. Among the available state sequences in the time-states trellis diagram, a single sequence $\hat{q}_0, \ldots, \hat{q}_\ell$ [128, 134] is hence estimated. Analogous to (5.23), the cumulative best score, $\delta_{\ell,\mathrm{C}}(j)$, for the path until the state $Q_\ell = j \in \{0, 1\}$ in the CHMM based scheme defined as

$$\begin{aligned}\delta_{\ell,\mathrm{C}}(j) &= \max_{q_0,\ldots,q_{\ell-1}} p\left(z_0, \ldots, z_{\ell-1}, z_\ell, q_0, \ldots, q_{\ell-1}, Q_\ell = j\,; \hat{\Gamma}_{\mathrm{C}}^{(m_{\mathrm{it,C}})}\right) \\ &= p_Z(z_\ell | Q_\ell = j) \max_i \delta_{\ell-1,\mathrm{C}}(i)\, \hat{a}_{ij}^{(m_{\mathrm{it,C}})},\end{aligned}$$

for $\hat{\Gamma}_{\mathrm{C}}^{(m_{\mathrm{it,C}})}$ being the best-fit model obtained during the learning phase. The estimate $\hat{Q}_\ell$ is subsequently obtained similar to the prediction decision followed in the DHMM based prediction scheme, namely [2]

$$\delta_{\ell,\mathrm{C}}(1) \underset{\hat{Q}_\ell=0}{\overset{\hat{Q}_\ell=1}{\gtrless}} \frac{1}{2}.$$

5.4. Summary

This chapter has introduced the TF spectrum occupancy prediction approach based on the conventional CHMM scheme as an alternative approach to the DHMM based scheme. It applies to cases when adjacent subbands are mutually independently occupied by the PU signal. The CHMM scheme is an appropriate approach due to the nature of the received signal, where no extra post-processing is required. Hence, the information within the received signal is fully explored for prediction purposes.

Chapter 6.

6 Prediction Scheme for a System of Dependent Subbands

After presenting the TF based spectrum occupancy predictor in the form of the first-order CHMM based scheme as an alternative approach to the conventional DHMM scheme, this chapter proposes a more suitable yet complex scheme, which takes the interdependencies between direct adjacent subbands into account. The SU receiver structure in Figure 5.2 has also been considered in this chapter. At this point, it is worth reminding that none of the available spectrum occupancy prediction schemes have yet captured any interdependencies between neighboring subbands in a suitable model, nor within the prediction decision. The available prediction schemes in literature have so far neglected the independencies between signals at neighboring subbands, thus each subband or channel has been modeled and its occupancy has been estimated independently.

Despite its superior performance, the first-order CHMM based prediction scheme presented in Chapter 5 is not sufficient to model and predict the spectrum occupancy when interdependencies between contiguous subbands are taken into consideration. Therefore, a proper coupling mechanism is introduced to the CHMM based scheme in Section 5.3 to present a novel prediction approach in the form of a lag-one first-order coupled CHMM based TF spectrum occupancy prediction scheme. Within this novel approach, the occupancy information of any subband depends on its nearest

directly adjacent TF neighbors in a causal system of multiple subbands. Similar to Chapter 5, the presented scheme also consists of the three successive phases presented in Figure 5.1: an initial pilot phase, a learning phase, and an occupancy prediction phase, as illustrated in Figure 5.1. These phases are similar in concept to those discussed in Section 5.3. However, proper adjustments have been introduced to match the considered coupling mechanism.

Whether modeling discrete or continuous observations, a coupled HMM scheme is known to be a generalization of the single DHMM, or CHMM scheme, respectively [136]. Coupled approaches have been utilized to model dynamic systems and applications with multiple interacting processes [57], either for discrete or continuous observations. They have been employed in audio-visual speech recognition [136, 137], modeling freeway traffic activity [138], as well as various classifications and detection problems [139, 140].

As the previously introduced random process $Q_{\ell,k} \in \{0, 1\}$ describes the occupancy state at the TF slot (ℓ, k), the state vector $\mathbf{Q}_\ell$ given by

$$\mathbf{Q}_\ell = [Q_{\ell,0}, \ldots, Q_{\ell,K-1}]^\mathsf{T} ,$$

characterizes the corresponding occupancy state of the K contiguous subbands. Similar to Chapter 5, $Q_{\ell,k} = 0$ indicates the unavailability of subband $k \in \Delta$ at time slot ℓ for the SU to access due to an existing PU signal $S_{\ell,k}$, whereas $Q_{\ell,k} = 1$ indicates its availability. For this matter, the coefficient vector $\mathbf{Y}_\ell$, defined by

$$\mathbf{Y}_\ell = [Y_{\ell,0}, \ldots, Y_{\ell,K-1}]^\mathsf{T} ,$$

is utilized to obtain the occupancy information. The vector $\mathbf{Y}_\ell$ contains the K TF coefficients at time slot ℓ of the signal $Y[n]$ received by the SU. In other words, the K components of the state vector $\mathbf{Q}_\ell$ at time slot ℓ indicate the availability of the corresponding subbands for the SU's utilization.

Moreover, the vector $\mathbf{S}_\ell = [S_{\ell,0}, \ldots, S_{\ell,K-1}]^\mathsf{T}$, with mutually independent K components in the ℓ-domain, denotes the PU signal vector. The process $\mathbf{S}_\ell$ has been modeled in this work as a zero-mean circularly symmetric complex K-variate Gaus-

sian random vector with a covariance matrix Ξ defined by

$$\Xi = \mathbf{E}\left\{\mathbf{S}_\ell \mathbf{S}_\ell^\dagger\right\}. \tag{6.1}$$

Along with $\mathbf{S}_\ell$, the vector $\mathbf{Y}_\ell$ contains two other TF vector components: the thermal noise

$$\mathbf{W}_\ell = [W_{\ell,0}, \ldots, W_{\ell,K-1}]^\intercal,$$

and the CCI signal

$$\mathbf{U}_\ell = [U_{\ell,0}, \ldots, U_{\ell,K-1}]^\intercal,$$

both having mutually independent components in time and frequency domains. The random vector $\mathbf{W}_\ell$ is modeled as a zero-mean circularly symmetric complex K-variate Gaussian random vector with the covariance matrix

$$\mathbf{E}\left\{\mathbf{W}_\ell \mathbf{W}_\ell^\intercal\right\} = \nu_W \mathbf{I}_K.$$

The non-Gaussian CCI random vector $\mathbf{U}_\ell$ has identical average power ν_U per subband, and its components $U_{\ell,k}$ follows the Matérn hard-core spatial process discussed in Section 4.3. The resulting process $\mathbf{Y}_\ell$ is modeled as a temporally stationary discrete random vector. Its temporal memory is therefore introduced exclusively by the state vector $\mathbf{Q}_\ell$, where the latter contains 2^K possible states or hypotheses correspond to the PU availability in each of the K subbands. In analogy to (4.2), $\mathbf{Y}_\ell$ is represented under each possible hypotheses in $\mathrm{H}_{\mathbf{Q}_\ell}$ as

$$\mathrm{H}_{\mathbf{Q}_\ell} : \mathbf{Y}_\ell = \left(\mathbf{I}_K - \mathrm{diag}\left(\mathbf{Q}_\ell\right)\right)\mathbf{S}_\ell + \mathbf{W}_\ell + \mathbf{U}_\ell, \tag{6.2}$$

with $\mathbf{S}_\ell$, $\mathbf{W}_\ell$, and $\mathbf{U}_\ell$ are mutually independent. Similar to Chapter 5, the components of $\mathbf{Q}_\ell$, namely $Q_{\ell,k}$ are considered to be WSS in the ℓ-domain. To simplify matters, the components of the vector $\mathbf{S}_\ell$ are assumed to be a section of a WSS process in k. Therefore, Ξ has a Toeplitz structure, where the components of its main diagonal represent the average power per subband ν_S.

6.1. Subband Modeling

Similar to Chapter 5, as the occupancy information at each TF slot is unknown to the SU receiver, the components of $\mathbf{Q}_\ell$, namely $Q_{\ell,k}$, are modeled in the coupled scheme as a first-order hidden Markov process [13, 2], such that $Q_{\ell,k}$ depends on its nearest TF neighbors in a causal system. Therefore, the previously discussed Markov assumption in Chapter 2 is relaxed [57, 136], allowing the adjacent subbands to be included in the model. For this matter, the vector $\mathbf{\Psi}_{\ell,k}$ is introduced to define the parent subband state of $Q_{\ell,k}$. The state dependencies in the coupled model is given by

$$\begin{aligned} & \Pr\left\{ Q_{\ell,k} = q_{\ell,k} \middle| \, \mathbf{Q}_{\ell-1} = \mathbf{q}_{\ell-1}, \mathbf{Q}_{\ell-2} = \mathbf{q}_{\ell-2}, \ldots \right\} \\ = \; & \Pr\left\{ Q_{\ell,k} = q_{\ell,k} \middle| \, \mathbf{\Psi}_{\ell,k} = \boldsymbol{\psi}_{\ell,k} \right\}, \end{aligned} \tag{6.3}$$

where $\mathbf{q}_{\ell-1} = [q_{\ell,0}, \ldots, q_{\ell,K-1}]^\intercal$ is a realization of $\mathbf{Q}_{\ell-1}$. The TF parent subbands state $\mathbf{\Psi}_{\ell,k}$ is defined as

$$\mathbf{\Psi}_{\ell,k} = \begin{cases} [Q_{\ell-1,k}, Q_{\ell-1,k+1}]^\intercal, & \text{for } k = 0 \\ [Q_{\ell-1,k-1}, Q_{\ell-1,k}, Q_{\ell-1,k+1}]^\intercal, & \text{for } 1 \leq k \leq K-2 \\ [Q_{\ell-1,k-1}, Q_{\ell-1,k}]^\intercal, & \text{for } k = K-1, \end{cases}$$

with the corresponding realization vectors

$$\boldsymbol{\psi}_{\ell,k} = \begin{cases} [q_{\ell-1,k}, q_{\ell-1,k+1}]^\intercal, & \text{for } k = 0 \\ [q_{\ell-1,k-1}, q_{\ell-1,k}, q_{\ell-1,k+1}]^\intercal, & \text{for } 1 \leq k \leq K-2 \\ [q_{\ell-1,k-1}, q_{\ell-1,k}]^\intercal, & \text{for } k = K-1. \end{cases}$$

Similar to the DBNs introduced in Chapter 2, the relation between $Q_{\ell,k}$ and its parent subband $\mathbf{\Psi}_{\ell,k}$ for $k = 1, \ldots, K-2$ is represented graphically in Figure 6.1. It can be deduced based on Figure 6.1 that the future occupancy information of $k = 1, \ldots, K-2$ subbands depends on its current state as well as the current state of its neighboring subbands.

Furthermore, a biased boundary condition has been introduced assuming that the proposed system of $k \in \Delta$ subbands belongs to a subsystem of an infinite number

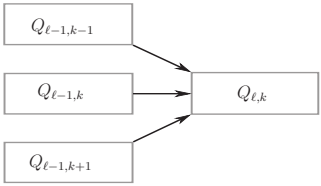

Figure 6.1.: A graphical representation of the considered lag-one first-order coupling scheme.

of subbands, such that $Q_{\ell,k'} = 0$ for $k' \notin \Delta$, as they are actually not available to the SU. Hence, the definition of $\mathbf{\Psi}_{\ell,k}$ is modified as

$$\mathbf{\Psi}_{\ell,k} = \begin{cases} [0, Q_{\ell-1,k}, Q_{\ell-1,k+1}]^\intercal, & \text{for } k = 0 \\ [Q_{\ell-1,k-1}, Q_{\ell-1,k}, Q_{\ell-1,k+1}]^\intercal, & \text{for } 1 \leq k \leq K-2 \\ [Q_{\ell-1,k-1}, Q_{\ell-1,k}, 0]^\intercal, & \text{for } k = K-1. \end{cases}$$

Similar to the reasons stated in Chapter 5, the output signal $\mathbf{Z}_\ell$ denoted by $\mathbf{Z}_\ell = [Z_{\ell,0}, \ldots, Z_{\ell,K-1}]^\intercal$ of the K EDs which follow the AFB is considered onwards, where $Z_{\ell,k} = |Y_{\ell,k}|^2$. Thus, the signal $\mathbf{Z}_\ell$ is employed in the following for spectrum occupancy purposes. Furthermore, as the prediction process is carried out in the previously explained three successive phases, let us first consider the initial pilot phase in terms of complexity. During the pilot phase, modeling $\mathbf{Z}_\ell$ based on the availability of the PU in the set of K subbands is required to be obtained. Basically, the model is presented in the form of PDFs conditioned on the 2^K hypotheses in $\mathbf{Q}_\ell$ based on the ML mechanism. This, however, can be extremely complicated for larger values of K. Therefore, to reduce the complexity, the block of L-component vector $\mathbf{Z}_{\ell,k}$ given by

$$\mathbf{Z}_{\ell,k} = [Z_{\ell,k}, \ldots, Z_{\ell,k+L-1}]^\intercal \tag{6.4}$$

is instead considered in the prediction process for $k = 0, \ldots, K-L$ along with the corresponding state vector $\mathbf{Q}_{\ell,k}$ given by

$$\mathbf{Q}_{\ell,k} = [Q_{\ell,k}, \ldots, Q_{\ell,k+L-1}]^\intercal . \tag{6.5}$$

Further details about the technique followed to predict the occupancy state of each subband is discussed in Section 6.4. The value of the parameter L is important for

implementation purposes in order to obtain a simple yet reliable prediction scheme, where $L = \{1, 2, \dots\}$ and $L < K$. The case where $L = 1$ is equivalent to the scheme presented in Chapter 5. Accordingly, the L-component vector of the PU signal $\mathbf{S}_{\ell,k}$ is by

$$\mathbf{S}_{\ell,k} = [S_{\ell,k}, \dots, S_{\ell,k+L-1}]^{\intercal},$$

with its covariance matrix Ξ_L given by

$$\Xi_L = \nu_S \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{L-1} \\ \rho & 1 & \rho & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{L-1} & \rho^{L-2} & \dots & \dots & 1 \end{bmatrix}, \tag{6.6}$$

which corresponds to a submatrix of (6.1) for matching subbands with a correlation coefficient ρ.

The idea behind considering a block of L subbands for complexity reduction purposes is similar to the concept of a resource block in the LTE standard [141]. For example, the interdependencies between $L = 5$ contiguous subbands is represented graphically in Figure 6.2. Irrespective of the block size L, the occupancy information of a subband depends on its parent subbands which are within the defined block.

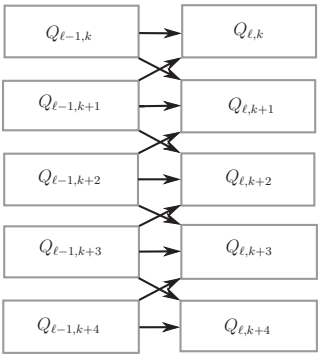

Figure 6.2.: A graphical representation of the considered lag-one first-order coupling scheme within a block of $L = 5$ contiguous subbands.

6.2. Pilot Phase

In order to establish the conditional PDFs of $\mathbf{Z}_{\ell,k}$ given $\mathbf{Q}_{\ell,k}$ within a certain block of length L, $\tilde{L}_{\mathrm{P}}$ instances of $\mathbf{Z}_{\ell,k}$ are observed at past $\ell < 0$ time slots for each of the N possible $\mathbf{Q}_{\ell,k}$ state vectors within the L-block of subbands, where $N = 2^L$. In this work, the block of subbands has been considered to be $L = \{1, 2, 3, 4, 5\}$, and the overall number of observed instances of $\mathbf{Z}_{\ell,k}$ required during the pilot phase, i.e, L_{P}, in order to model the system of K subbands is

$$L_{\mathrm{P}} = \begin{cases} N\tilde{L}_{\mathrm{P}}K/L, & \text{for } L = 1, 4 \\ N\tilde{L}_{\mathrm{P}}(K+2)/L, & \text{for } L = 2, 3 \\ N\tilde{L}_{\mathrm{P}}(K+1)/L, & \text{for } L = 5, \end{cases}$$

with additional boundary subbands for the cases of $L = 2, 3$ and 5. The introduced boundary subbands are included to ensure proper blocks modeling during implementation. Note that the case with $L = 1$ represents the system of independent subbands explained in Chapter 5. Furthermore, the PDFs $p_{\mathbf{Z}_{\ell,k}}(\mathbf{z}_{\ell,k}|\,\mathbf{Q}_{\ell,k} = \mathbf{p})$ are assumed to be independent of k and approximated by L-variate GMM scheme as

$$p_{\mathbf{Z}_{\ell,k}}(\mathbf{z}_{\ell,k}|\,\mathbf{Q}_{\ell,k} = \mathbf{p}) = \sum_{r=1}^{R} w_r \mathcal{N}(\mathbf{z}_{\ell,k}; \boldsymbol{\mu}_{r,\mathbf{p}}, \boldsymbol{\Sigma}_{r,\mathbf{p}}), \tag{6.7}$$

where $\mathbf{p} \in \mathbf{Q}_{\ell,k}$, and the number of Gaussian components R is independent of the states. The PDFs' mean vector $\boldsymbol{\mu}_{r,\mathbf{p}}$, the covariance matrix per state $\boldsymbol{\Sigma}_{r,\mathbf{p}}$, the weights w_r, and R depend neither on ℓ nor k. The weights w_r satisfy the conditions in (4.7). The parameters in (6.7) are estimated via the EM algorithm and k-means clustering in analogy to Section 6.2.

6.3. Learning Phase

In order to model the underlying traffic model, the transition probabilities must be estimated. The previously defined transition probabilities are modified to suit the scheme in Figure 6.1. Therefore, we define $a_{q,\psi}(\ell, k)$ as an extension or generalization of a_{ij} in (5.5), with $i, j \in Q_{\ell,k}$ for the single scheme in Chapter 5, where the

contiguous subbands are mutually independently occupied by the PU signal. The quantities $a_{q,\boldsymbol{\psi}}(\ell,k)$ at the corresponding TF slot are given by

$$a_{q,\boldsymbol{\psi}}(\ell,k) = \Pr\{Q_{\ell,k} = q | \boldsymbol{\Psi}_{\ell,k} = \boldsymbol{\psi}\},$$

which may be expressed due to stationarity as $a_{q,\boldsymbol{\psi}}(k)$, given by

$$a_{q,\boldsymbol{\psi}}(\ell,k) = a_{q,\boldsymbol{\psi}}(k) = \Pr\{Q_k = q | \boldsymbol{\Psi}_k = \boldsymbol{\psi}\}, \tag{6.8}$$

where $a_{q,\boldsymbol{\psi}}(k) \in \Omega$, and for a common $\boldsymbol{\psi}$ the condition $a_{0,\boldsymbol{\psi}}(k) + a_{1,\boldsymbol{\psi}}(k) = 1$ is valid.

In analogy to a DBN, the probabilities $a_{q,\boldsymbol{\psi}}(k)$ are arranged in a conditional probability table to define the conditional dependencies between a subband and its parent subbands. Moreover, $a_{q,\boldsymbol{\psi}}(k)$ is used to formulate the $N \times N$ transition matrix between the N possible states in the state vector $\mathbf{Q}_{\ell,k}$ for a given L, as in single HMM schemes. The transition probabilities $a_{\mathbf{pq}}(\ell)$ between any two state vectors $\mathbf{q}$ and $\mathbf{p}$, such that $\mathbf{p} = [q_{\ell-1,k}, \ldots, q_{\ell-1,k+L-1}]^\intercal$ and $\mathbf{q} = [q_{\ell,k}, \ldots, q_{\ell,k+L-1}]^\intercal$, with $\mathbf{p}, \mathbf{q} \in \mathbf{Q}_{\ell,k}$, are given by

$$\begin{aligned} a_{\mathbf{pq}}(\ell) = a_{\mathbf{pq}} &= \Pr\{\mathbf{Q}_k = \mathbf{q} | \mathbf{Q}_k = \mathbf{p}\} \\ &= a_{q,\psi_{\mathrm{R}}}(k)\, a_{q,\psi_{\mathrm{L}}}\left(\tilde{k}\right) \prod_{c=k+1}^{\tilde{k}-1} a_{q_c,\psi_c}(c), \end{aligned} \tag{6.9}$$

where $\tilde{k} = k + L - 1$. Furthermore, $a_{q,\psi_{\mathrm{R}}}(k)$ and $a_{q,\psi_{\mathrm{L}}}\left(\tilde{k}\right)$ are the transition probabilities of subbands k and $\tilde{k}$ after marginalizing out the left most and the right most parent subband, respectively for a certain values of L and κ_i. The quantities $a_{q,\psi_{\mathrm{R}}}(k), a_{q,\psi_{\mathrm{L}}}\left(\tilde{k}\right) \in \Omega^2$ are defined as

$$a_{q,\psi_{\mathrm{R}}}(k) = \Pr\{Q_{\ell,k} = q | \boldsymbol{\Psi}_{\mathrm{R}} = \boldsymbol{\psi}_{\mathrm{R}}\},$$

and

$$a_{q,\psi_{\mathrm{L}}}\left(\tilde{k}\right) = \Pr\{Q_{\ell,k} = q | \boldsymbol{\Psi}_{\mathrm{L}} = \boldsymbol{\psi}_{\mathrm{L}}\},$$

where $\boldsymbol{\psi}_{\mathrm{R}} = [q_k, q_{k+1}]^\intercal$ and $\boldsymbol{\psi}_{\mathrm{L}} = \left[q_{\tilde{k}-1}, q_{\tilde{k}}\right]^\intercal$ are realizations of $\boldsymbol{\Psi}_{\mathrm{R}}$ and $\boldsymbol{\Psi}_{\mathrm{L}}$, respec-

tively. So that

$$\mathbf{\Psi}_{\mathrm{R}} = [Q_k, Q_{k+1}]^{\intercal},$$

and

$$\mathbf{\Psi}_{\mathrm{L}} = \left[Q_{\tilde{k}-1}, Q_{\tilde{k}}\right]^{\intercal}.$$

The marginalization is performed using the equal priors κ_i [142]. For explanation purposes, let A, B_1, B_2, and B_3 be four arbitrary discrete random variables with a known conditional probability $\Pr\{A|\, B_1, B_2, B_3\}$. The probability $\Pr\{A|\, B_1, B_2\}$ is given by

$$\Pr\{A|\, B_1, B_2\} = \frac{\Pr\{A, B_1, B_2\}}{\Pr\{B_1, B_2\}},$$

where $\Pr\{A, B_1, B_2\}$ can be evaluated from $\Pr\{A|\, B_1, B_2, B_3\}$ as

$$\Pr\{A, B_1, B_2\} = \sum_{B_3} \Pr\{A|\, B_1, B_2, B_3\} \cdot \Pr\{B_1, B_2, B_3\}.$$

Concerning the presented scheme, the set of transition probabilities for the block of L components defined in (6.4) and (6.5) are given by the set Γ, such that

$$\Gamma = \left(a_{q,\psi_{\mathrm{R}}}(k), a_{q,\psi_{\mathrm{L}}}\left(\tilde{k}\right), a_{q,\psi}(c)\right), \tag{6.10}$$

where $c = k+1, \ldots, \tilde{k}-1$, $\Gamma \in \Omega_{\mathrm{CC}}$, where Ω_{CC} is defined by $\Omega_{\mathrm{CC}} = \Omega^L$.

Despite being a direct example of a DBN, it is still recommended to consider the forward-backward algorithm to estimate the transition parameters in (6.10) in order to reduce the complexity of the learning phase [136, 138]. During the learning phase, the learning sequence

$$\mathbf{Z}_{0,k} = \mathbf{z}_{0,k}, \ldots, \mathbf{Z}_{L_{\mathrm{L}}-1,k} = \mathbf{z}_{L_{\mathrm{L}}-1,k}, \tag{6.11}$$

is observed to estimate the so-called best-fit model which best describe the underlying traffic model. In (6.11), the vector $\mathbf{z}_{\ell,k} = \left[z_{\ell,k}, \ldots, z_{\ell,\tilde{k}}\right]^{\intercal}$ is a realization of $\mathbf{Z}_{\ell,k}$. In analogy to the learning phase for the single CHMM scheme in Section 5.3.1, the

ML estimate of Γ, namely $\hat{\Gamma}_{\mathrm{ML}}$, reads

$$\hat{\Gamma}_{\mathrm{ML}} = \arg \max_{\Gamma \in \Omega_{\mathrm{CC}}} p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{L_{\mathrm{L}}-1,k}\,; \Gamma\right), \tag{6.12}$$

where the expression $p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{L_{\mathrm{L}}-1,k}\,; \Gamma\right)$ in (6.12) represents the corresponding likelihood function for a given Γ, L and the learning sequence considered in (6.11). The optimization problem in (6.12) is broken down via the forward-backward algorithm into two parts: the forward and the backward quantities, $\alpha_\ell\left(\mathbf{p}\right)$ and $\beta_\ell\left(\mathbf{p}\right)$, respectively, which are defined in [136] as

$$\alpha_\ell\left(\mathbf{p}\right) = p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{\ell,k}, \mathbf{Q}_{\ell,k} = \mathbf{p}\,; \Gamma\right),$$

and

$$\beta_\ell\left(\mathbf{p}\right) = p\left(\mathbf{z}_{\ell+1,k}, \ldots, \mathbf{z}_{L_{\mathrm{L}}-1,k}\,|\mathbf{Q}_{\ell,k} = \mathbf{p}; \Gamma\right).$$

Similar to Subsection 5.3.1, iterative estimation procedure takes place by evaluating the estimates of $\alpha_\ell\left(\mathbf{p}\right)$ and $\beta_\ell\left(\mathbf{p}\right)$, referred as $\hat{\alpha}_\ell\left(\mathbf{p}\right)$ and $\hat{\beta}_\ell\left(\mathbf{p}\right)$, respectively, which are used consequently to update the estimates of Γ.

For a certain estimate $\hat{\Gamma}$, $\hat{\alpha}_\ell\left(\mathbf{p}\right)$ and $\hat{\beta}_\ell\left(\mathbf{p}\right)$, are evaluated as

$$\begin{aligned}\hat{\alpha}_\ell\left(\mathbf{p}\right) &= p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{\ell,k}, \mathbf{Q}_{\ell,k} = \mathbf{p}\,; \hat{\Gamma}_{\mathrm{C}}\right) \\ &= p_{\mathbf{Z}_{\ell,k}}(\mathbf{z}_{\ell,k}|\,\mathbf{Q}_{\ell,k} = \mathbf{p}) \sum_{\mathbf{q}} \hat{a}_{\mathbf{qp}}\, \hat{\alpha}_{\ell-1}\left(\mathbf{q}\right),\end{aligned} \tag{6.13}$$

and

$$\begin{aligned}\hat{\beta}_\ell\left(\mathbf{p}\right) &= p\left(\mathbf{z}_{\ell+1,k}, \ldots, \mathbf{z}_{L_{\mathrm{L}}-1,k}\,|\mathbf{Q}_{\ell,k} = \mathbf{p}; \hat{\Gamma}_{\mathrm{C}}\right) \\ &= \sum_{\mathbf{q}} p_{\mathbf{Z}_{\ell,k}}(\mathbf{z}_{\ell,k}|\,\mathbf{Q}_{\ell,k} = \mathbf{q})\, \hat{a}_{\mathbf{pq}}\, \hat{\beta}_{\ell+1}\left(\mathbf{q}\right),\end{aligned} \tag{6.14}$$

similar to the CHMM based scheme. The basic idea behind the forward-backward algorithm has been excessively discussed in Chapter 5. In this chapter, the update of the parameters in Γ will be presented. Starting from an initial guess $\hat{\Gamma}^{(0)}$, indicated as

$$\hat{\Gamma}^{(0)} = \left(\hat{a}^{(0)}_{q_k,\psi_{\mathrm{R}}}\left(k\right), \hat{a}^{(0)}_{q_{\tilde{k}},\psi_{\mathrm{L}}}\left(\tilde{k}\right), \hat{a}^{(0)}_{q,\psi}\left(c\right)\right),$$

the estimates $\hat{\alpha}_\ell(\mathbf{p})$ and $\hat{\beta}_\ell(\mathbf{p})$ are evaluated iteratively based on (6.13) and (6.14), respectively. The quantities $\hat{a}^{(0)}_{q_k,\psi_{\mathrm{R}}}(k)$, $\hat{a}^{(0)}_{q_{\tilde{k}},\psi_{\mathrm{L}}}\left(\tilde{k}\right)$, and $\hat{a}^{(0)}_{q,\psi}(c)$ represent the initial estimates of $a_{q,\psi_{\mathrm{R}}}(k), a_{q,\psi_{\mathrm{L}}}\left(\tilde{k}\right)$, and $a_{q,\psi}(c)$, respectively. Correspondingly, a new estimate of $\hat{\Gamma}$ is calculated. In analogy to Subsection 5.3.1, the new estimate of the transition probabilities in $\hat{\Gamma}$, indicating the transition from $\mathbf{p} = \left[q_{\ell-1,k}, \ldots, q_{\ell-1,\tilde{k}}\right]^{\intercal}$ to $\mathbf{q} = \left[q_{\ell,k}, \ldots, q_{\ell,\tilde{k}}\right]^{\intercal}$, are obtained as [143]

$$\hat{a}^{(m+1)}_{q,\psi_q}(k) = \frac{\sum_{\ell=0}^{L_{\mathrm{L}}-1} \sum_{\mathbf{q}} \hat{\alpha}^{(m)}_\ell(\mathbf{p})\ \hat{a}^{(m)}_{\mathbf{pq}}\ p_{\mathbf{Z}_{\ell,k}}(\mathbf{z}_{\ell+1,k} | \mathbf{Q}_{\ell+1,k} = \mathbf{q})\ \hat{\beta}^{(m)}_{\ell+1}(\mathbf{q})}{p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{L_{\mathrm{L}}-1,k}\,; \hat{\Gamma}^{(m)}\right)},$$

where $\boldsymbol{\psi}_q$ defines the parent state within the block L, such that $q \in \mathbf{q}$, $\boldsymbol{\psi}_q \in \mathbf{p}$, and $\hat{a}^{(m)}_{\mathbf{pq}}$ is the estimated value of $a_{\mathbf{pq}}$ at the m^{th} iteration, which is obtained based on (6.9) for the model $\hat{\Gamma}^{(m)}$.

A total of m_{it} number of iterations are needed to update $\hat{\alpha}_\ell(\mathbf{p})$ and $\hat{\beta}_\ell(\mathbf{p})$, and subsequently $\hat{\Gamma}$, until the increase in the corresponding likelihood function is below the threshold ϵ [136], in analogy to (5.27) as discussed in Section 5.3.1, as

$$\left| \frac{p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{L_{\mathrm{L}}-1,k}\,; \hat{\Gamma}^{(m_{\mathrm{it}})}\right) - p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{L_{\mathrm{L}}-1,k}\,; \hat{\Gamma}^{(m_{\mathrm{it}}-1)}\right)}{p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{L_{\mathrm{L}}-1,k}\,; \hat{\Gamma}^{(m_{\mathrm{it}}-1)}\right)} \right| \leq \epsilon.$$

Thus, $\hat{\Gamma}_{\mathrm{ML}} \approx \hat{\Gamma}^{(m_{\mathrm{it}})}$.

6.4. Occupancy Prediction Phase

Finally, the resulting model $\hat{\Gamma}^{(m_{\mathrm{it}})}$ obtained during the learning phase is utilized to estimate $\mathbf{Q}_{\ell,k}$ at $\ell = L_{\mathrm{L}}, L_{\mathrm{L}}+1, \ldots$ future time slots during the prediction phase. While shifting the block – referred tp as a (sliding) window in this part – of L contiguous subbands throughout the k-domain, the occupancy prediction phase is performed such that within each shift a new adjacent subband at the window's right most is decoded at future time slots. At the end, the obtained prediction information is to be utilized in a suitable spectrum access scheme for the SU transmission.

As clarified in Chapter 5, the Viterbi algorithm is an efficient method to estimate the most likely sequence of states [2, 134]. For $\ell = 0, 1, \ldots$, the corresponding most likely sequence of states $\hat{\mathbf{Q}}_{0,k} = \hat{\mathbf{q}}_{0,k}, \ldots, \hat{\mathbf{Q}}_{\ell,k} = \hat{\mathbf{q}}_{\ell,k}$ is decoded for the L contiguous subbands. The path until $\mathbf{Q}_{\ell,k} = \mathbf{q}$ is calculated based on characterizing the cumulative best score $\delta_\ell(\mathbf{q})$ according to

$$\begin{aligned}\delta_\ell(\mathbf{q}) &= \max_{\mathbf{q}_{0,k},\ldots,\mathbf{q}_{\ell-1,k}} p\left(\mathbf{z}_{0,k}, \ldots, \mathbf{z}_{\ell,k}, \mathbf{q}_{0,k}, \ldots, \mathbf{q}_{\ell-1,k}, \mathbf{Q}_{\ell,k} = \mathbf{q}\,; \hat{\Gamma}_{\mathrm{C}}^{(m_{\mathrm{it}})}\right) \\ &= p_{\mathbf{Z}_{\ell,k}}(\mathbf{z}_{\ell,k} | \, \mathbf{Q}_{\ell,k} = \mathbf{q}) \max_{\mathbf{p}} \delta_{\ell-1}(\mathbf{p}) \, \hat{a}_{\mathbf{pq}}^{(m_{\mathrm{it}})},\end{aligned}$$

given the best-fit model $\hat{\Gamma}^{(m_{\mathrm{it}})}$ obtained during the learning phase. Correspondingly, the estimate $\hat{Q}_{\ell,k} = \hat{q}_{\ell,k}$ is obtained based on the prediction decision

$$\hat{\mathbf{q}} = \arg\max_{\mathbf{q}} \delta_\ell(\mathbf{q}),$$

where the state vector $\hat{\mathbf{q}} = [\hat{q}_{\ell,k-L+1}, \ldots, \hat{q}_{\ell,k-1}, \hat{q}_{\ell,k}] \in \mathbf{Q}_{\ell,k}$, contains the estimated states $\hat{q}_{\ell,k-L+1}, \ldots, \hat{q}_{\ell,k-1}$ at a previous window shift. Hence, the occupancy prediction phase is simplified.

6.5. Summary

In cases when the interdependencies between adjacent subbands are to be taken into account, the presented CHMM based prediction scheme is modified in the form of a lag-one coupled CHMM scheme. Hence, in this novel approach, the current occupancy of a certain subband depends on its previous state as well as the previous state of its directly adjacent subbands.

Chapter 7.
7 Performance Analysis

This chapter presents the performance of the previously discussed TF based spectrum occupancy prediction schemes. To measure the reliability of those schemes, the probability of the wrong prediction is considered in terms of different interference based environments. Furthermore, the variables or parameters affecting the schemes' reliability are also investigated. A simple indoor flat Rayleigh fading environment has been considered with the presence of CCI arising from multiple transmitters simultaneously utilizing the same subband [1]. The CCI signal is obtained based on the Monte Carlo simulation discussed in Chapter 3. It is important to note that the predicted occupancy information would be forwarded for utilization within a proper spectrum access scheme, thus any wrong prediction can also be avoided. Moreover, the prediction results can be employed for power consumption optimization purposes, as discussed in Chapter 2.

7.1. Evaluation of the Independent System

In this section, the performance of the TF prediction schemes introduced in Chapter 5 is investigated for the case when the system of subbands are mutually independently occupied by PU signal. In addition, this section discusses the improvements obtained by directly implementing the received signal into the CHMM based scheme instead of the DHMM based scheme, as the discretization process is no longer required. For simulation and performance evaluation purposes, two PU

traffic scenarios: heavy and light traffic patterns, have been considered in implementing both the DHMM and CHMM based schemes. The terms "heavy" and "light" actually reflect the probability a subband is occupied by at least one PU based on its previous occupancy state. Upon this definition, note that a subband is mostly idle while experiencing a light traffic pattern, unlike the heavy one. As the matrix $\mathbf{A}$ in the DHMM and CHMM prediction schemes contains the transition probabilities between *free* and *occupied* states, the heavy and light traffic patterns are hence expressed based on the value the matrix $\mathbf{A}$ takes. The heavy traffic scenario is denoted by $\mathbf{A} = \mathbf{A}_{\text{heavy}}$, where $\mathbf{A}_{\text{heavy}}$ is given by

$$\mathbf{A}_{\text{heavy}} = \begin{bmatrix} 0.7 & 0.3 \\ 0.9 & 0.1 \end{bmatrix}. \tag{7.1}$$

During the learning phase, its initial value, $\hat{\mathbf{A}}^{(0)}_{\text{heavy}}$, has been chosen for estimation purposes as

$$\hat{\mathbf{A}}^{(0)}_{\text{heavy}} = \begin{bmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{bmatrix}.$$

On the other hand, the light traffic scenario is expressed by $\mathbf{A} = \mathbf{A}_{\text{light}}$, where $\mathbf{A}_{\text{light}}$ is given by

$$\mathbf{A}_{\text{light}} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \mathbf{A}_{\text{heavy}} = \begin{bmatrix} 0.3 & 0.7 \\ 0.1 & 0.9 \end{bmatrix}, \tag{7.2}$$

whereas, the initial values in $\hat{\mathbf{A}}^{(0)}_{\text{light}}$ are considered as

$$\hat{\mathbf{A}}^{(0)}_{\text{light}} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \hat{\mathbf{A}}^{(0)}_{\text{heavy}}.$$

Based on (7.1) and (7.2), one can notice that the probability a subband being occupied given that it was previously occupied in a heavy traffic scenario is 0.7, whereas it is 0.3 in a light traffic scenario. The probabilities in $\mathbf{B}$ along with the initial estimates in $\hat{\mathbf{B}}^{(0)}$ for the DHMM based prediction scheme are assumed during simulation to be

$$\mathbf{B} = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix} \quad \text{and,} \quad \hat{\mathbf{B}}^{(0)} = \begin{bmatrix} 0.91 & 0.09 \\ 0.09 & 0.91 \end{bmatrix},$$

for both heavy and light traffic scenarios. Furthermore, simulation has been conducted for $\kappa_i = 0.5$ and a system of $K = 64$ independent subbands. For more clarifications, both the CHMM and DHMM prediction scheme generated based on $\mathbf{A}_{\text{light}}$ and $\mathbf{A}_{\text{heavy}}$, as well as $\mathbf{B}$ are considered as the real-life models, where the estimated models during the learning phase are compared to for performance evaluation purposes.

During the pilot phase, $\tilde{L}_{\text{P}} = 2000$ instances of Z_ℓ under each hypothesis in $Q_\ell \in \{0, 1\}$ are considered. The number of Gaussian components in (5.3) is $R = 8$. During the learning phase, on the other hand, the length of the learning sequence is $L_{\text{L}} = 1500$. The performance for predicting the next $L_{\text{O}} = 5 \cdot 10^5$ future time instant for $\ell \geq L_{\text{L}}$ is characterized based on the probability of error P_{e} for all $k \in \Delta$ subbands, given by

$$P_{\text{e}} = \Pr\left\{\hat{Q}_\ell \neq Q_\ell\right\}, \tag{7.3}$$

which is evaluated under different environment conditions characterized by SNR value θ, as well as SIR value ϑ, where θ and ϑ are respectively defined by

$$\theta = \frac{\nu_S}{\nu_W}, \text{ and } \quad \vartheta = \frac{\nu_S}{\nu_U}. \tag{7.4}$$

Table 7.1 summarizes the parameters for the DHMM as well as the CHMM based prediction schemes and their considered values for both traffic scenarios. Figure 7.1

Table 7.1.: Parameters for the DHMM and CHMM based prediction schemes.

Parameter	Parameter Value
K	64
P_{FA}	0.05
$\tilde{L}_{\text{P}}$	2000
L_{L}	1500
L_{O}	$5 \cdot 10^5$
R	8
ϵ	10^{-4}
κ_i	0.5

and Figure 7.2 present P_{e} as a function of θ and under various ϑ values for the heavy and light traffic scenarios defined in (7.1) and (7.2), respectively.

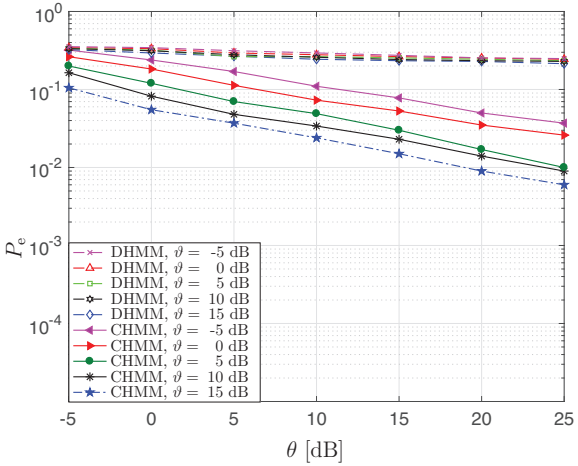

Figure 7.1.: P_e versus θ for a heavy traffic scenario with $\mathbf{A} = \mathbf{A}_{\text{heavy}}$ [2].

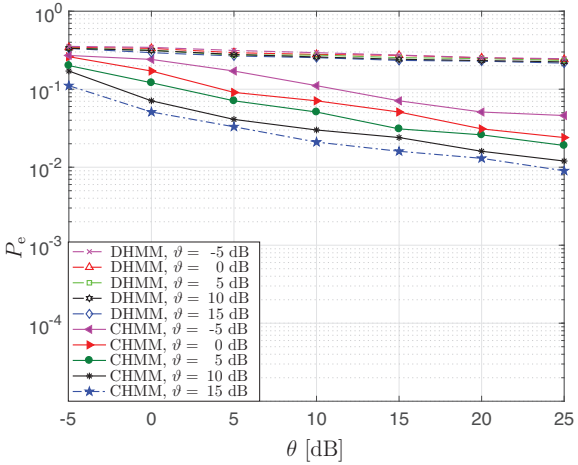

Figure 7.2.: P_e versus θ for a light traffic scenario with $\mathbf{A} = \mathbf{A}_{\text{light}}$.

Under both traffic scenarios, it can be noticed that P_{e} decreases as the effect of interference reduces, i.e., by increasing ϑ for both implemented predictors. Similarly, P_{e} improves by increasing θ up to $\theta = 25$ dB. However, huge improvement can be noticed by implementing the HMM prediction approach via employing the received signal Z_ℓ observed directly at the output of AFB into the CHMM prediction scheme. In Figure 7.1, the probability of error in DHMM based prediction scheme is $P_{\mathrm{e}} \approx 0.214$ for an environment characterized by $\theta = 25$ dB and $\vartheta = 15$ dB. On the other hand, the performance has indeed improved for similar environment conditions by implementing the CHMM scheme as $P_{\mathrm{e}} \approx 6 \cdot 10^{-3}$.

Further remarks can be drawn from Figure 7.1 and Figure 7.2. Concerning the light traffic scenario in Figure 7.2, it is noticed that for an environment characterized by $\theta = 25$ dB and $\vartheta = 15$ dB the probability of error for the DHMM based prediction scheme is $P_{\mathrm{e}} \approx 0.216$, whereas $P_{\mathrm{e}} \approx 9 \cdot 10^{-3}$ for the CHMM based scheme. In comparison with P_{e} for the heavy traffic scenario under similar conditions, one can deduce to this end that the implemented CHMM based prediction scheme is rather robust to different PU traffic patterns, as it is able to adapt its parameters during the learning phase given the traffic pattern under consideration in the form of the learning sequence. Due to its superior performance, CHMM based prediction scheme will be the subject of further analysis and investigation onwards.

As part of performance analysis, the length of pilot sequence $\tilde{L}_{\mathrm{P}}$ under each hypothesis has been changed to investigate its effect on the predictors under consideration. $\tilde{L}_{\mathrm{P}}$ has been varied within the range $2000 \leq \tilde{L}_{\mathrm{P}} \leq 50000$. Simulation results have remained essentially similar to the results presented in Figure 7.1 and Figure 7.2. The robustness in its performance is due to the fact that the quantization step in (5.4) is not needed in the CHMM scheme, hence, the information in the learning sequence is fully exploited by the prediction scheme to estimate the future occupancy states [2].

Moreover, the length of the observed learning sequence L_{L} has been also considered for performance evaluation. Thus, simulation has been performed over the range $10 \leq L_{\mathrm{L}} \leq 1500$ for the CHMM based prediction scheme. Figure 7.3 shows P_{e} as a function of L_{L} for an environment characterized by $\theta = 25$ dB and $\vartheta = 15$ dB, as the predictors are most reliable under those environment conditions. Unlike

$\tilde{L}_{\mathrm{P}}$, the parameter L_{L} is an important parameter which affects the performance, as in any machine learning method. As expected, the performance improves, the longer the learning sequence, i.e., P_{e} decreases by increasing the L_{L} value. However, this improvement is actually up to a certain value. Starting from $L_{\mathrm{L}} = 1000$, the performance is unchanged and the probability of error is saturated as $P_{\mathrm{e}} \approx 6 \cdot 10^{-3}$. Hence, one would rather employ $L_{\mathrm{L}} = 1000$ rather than 1500 to obtain the same performance results with less computational effort [2].

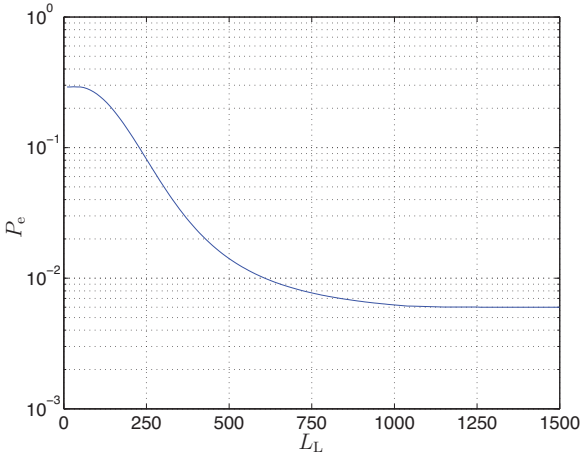

Figure 7.3.: P_{e} versus L_{L} for the CHMM scheme with $\theta = 25$ dB, $\vartheta = 15$ dB, and $\mathbf{A} = \mathbf{A}_{\mathrm{heavy}}$ [2].

Finally, the number of iterations $m_{\mathrm{it,C}}$ required to obtain the best-fit model during the learning phase is also investigated. This parameter has a direct effect on the prediction results, as the resulting model $\hat{\Gamma}_{\mathrm{C}}^{(m_{\mathrm{it,C}})}$ after $m_{\mathrm{it,C}}$ iterations is utilized during the prediction phase. Therefore, simulation has been conducted in [2] to evaluate $m_{\mathrm{it,C}}$ as a function of L_{L} and ϑ for $\theta = 25$ dB as presented in Figure 7.4. It is noticed that the required number of iterations, $m_{\mathrm{it,C}}$, decreases by increasing the values of both L_{L} and ϑ. As $m_{\mathrm{it,C}} \approx 500$ iterations for $L_{\mathrm{L}} = 10$ and $\vartheta = -5$ dB, number of iterations drops down to $m_{\mathrm{it,C}} = 4$ for $L_{\mathrm{L}} = 1500$ and $\vartheta = 25$ dB. Hence, increasing L_{L} for an environment with high interference can be a coping mechanism

for interference mitigation within the CHMM based prediction scheme [2].

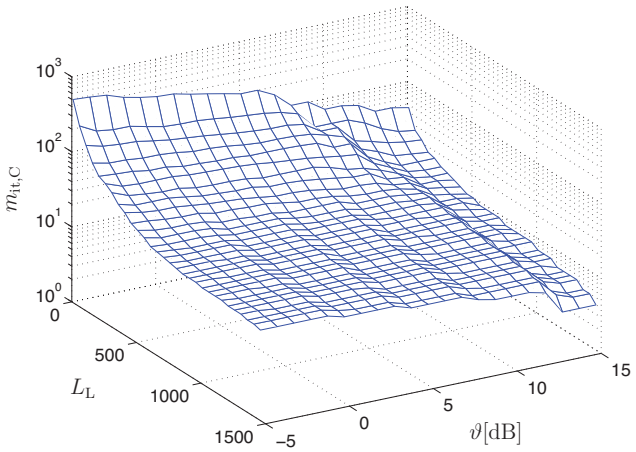

Figure 7.4.: $m_{\mathrm{it,C}}$ versus L_{L} and ϑ for the CHMM scheme, where $\theta = 25$ dB, and $\mathbf{A} = \mathbf{A}_{\mathrm{heavy}}$ [2].

7.2. Evaluation of the Dependent System

A similar wireless environment has been considered in this section in order to evaluate the performance of the prediction scheme for a set of dependent subbands. The SU receiver structure with $K = 64$ critically sampled AFB has also been considered to evaluate the performance of the proposed lag-one first-order coupled CHMM based prediction scheme. Under each of the possible $N = 2^L$ hypotheses in the state vector $\mathbf{Q}_{\ell,k}$, $\tilde{L}_{\mathrm{P}} = 2000$ instances of $\mathbf{Z}_{\ell,k}$ has been considered for $L = \{1, 2, \ldots, 5\}$ during the pilot phase in order to model the PDFs in (6.7). The resulting number of Gaussian components in (6.7) is $R = 8$. In (6.6), the correlation coefficient between directly adjacent subbands has been considered as $\rho = 0.5$.

Furthermore, the learning and occupancy prediction strategies have also been considered similar to the one in Section 7.1. In other words, the concept of heavy and light traffic scenarios are also utilized for simulation and performance evaluation

purposes. The heavy traffic scenario is characterized based on the transition probabilities $a_{q,\psi}(k) = a_{q,\psi,\text{heavy}}(k)$, where $a_{q,\psi,\text{heavy}}(k)$ are arranged in Table 7.2. Those values are also used to build the transition matrix in (6.9) for each block L. The probabilities $a_{q,\psi_\text{R}}(k)$ and $a_{q,\psi_\text{L}}\left(\tilde{k}\right)$ are also obtained based on Table 7.2.

Table 7.2.: The conditional probability table for a heavy traffic scenario.

		Q_k	
		0	1
$\boldsymbol{\Psi}_k$	$[000]^\intercal$	0.8	0.2
	$[001]^\intercal$	0.7	0.3
	$[010]^\intercal$	0.7	0.3
	$[011]^\intercal$	0.8	0.2
	$[100]^\intercal$	0.7	0.3
	$[101]^\intercal$	0.65	0.35
	$[110]^\intercal$	0.8	0.2
	$[111]^\intercal$	0.9	0.1

Based on Table 7.2, within a heavy traffic scenario, the probability a subband k is occupied whereby its parent subbands are also occupied is given by

$$\begin{aligned} a_{0,[000]^\intercal}(k) &= a_{0,[000]^\intercal,\text{heavy}}(k) \\ &= \Pr\left\{Q_k = 0 \middle| \boldsymbol{\Psi}_k = [000]^\intercal\right\} \\ &= 0.8 \end{aligned}$$

As a reasonable approach, the heavy traffic scenario has been designed such that adjacent subbands are also occupied. Furthermore, during the learning phase, the initial guess $\hat{\Gamma}^{(0)}$ required for the forward-backward algorithm has been chosen such that

$$\hat{a}_{0,\psi}^{(0)}(k) = a_{0,\psi,\text{heavy}}(k) - 0.15,$$

and

$$\hat{a}_{1,\psi}^{(0)}(k) = a_{1,\psi,\text{heavy}}(k) + 0.15.$$

In addition, the considered light traffic scenario $a_{q,\psi}(k) = a_{q,\psi,\text{light}}(k)$ is also formu-

lated based on Table 7.2, such that $a_{q,\psi,\text{light}}(k)$ is given by

$$a_{q,\psi}(k) = a_{q,\psi,\text{light}}(k) = 1 - a_{q,\psi,\text{heavy}}(k),$$

with

$$\hat{a}^{(0)}_{0,\psi}(k) = a_{0,\psi,\text{light}}(k) + 0.15,$$

and

$$\hat{a}^{(0)}_{1,\psi}(k) = a_{1,\psi,\text{light}}(k) - 0.15.$$

For both scenarios, $\kappa_i = 0.5$ for all $k \in \Delta$. The learning sequence of length $L_\text{L} = 1500$ has been employed to predict the occupancy state for the next $L_\text{O} = 5 \cdot 10^5$ future time slots with $\ell \geq L_\text{L}$, where $\epsilon = 10^{-4}$.

Simulation has been conducted for $1 \leq L \leq 5$ in order to choose a suitable block length to perform the previously discussed prediction phases to obtain a reliable and simple spectrum occupancy prediction scheme. For this matter, the performance of the proposed scheme in the form of P_e for all $k \in \Delta$, as defined in (7.3), has been evaluated as a function of L for an environment characterized by $\theta = 25$ dB, $\vartheta = 15$ dB, and the parameters of the heavy traffic pattern in Table 7.2. Simulation results are shown in Figure 7.5.

In Figure 7.5, it is observed that the scheme's performance improves by increasing the number of L block components in $\mathbf{Z}_{\ell,k}$ and $\mathbf{Q}_{\ell,k}$ as P_e. For example, $P_\text{e} \approx 6.6 \cdot 10^{-3}$ for $L = 1$, which is actually comparable to the value obtained by the prediction scheme when the K subbands are mutually independently occupied by the PU signal in Section 7.1. Starting from $L \geq 3$, however, the performance is improved by a small rate. The probability of error decreases up to $P_\text{e} \approx 8.15 \cdot 10^{-5}$ for $L = 5$. This is rather a reasonable behavior, as the state of each subband depends only on the state vector of its parent subbands within the block L, irrespective of and independent from the rest of subbands. In the following, the performed simulation and the corresponding results are shown for $L = 5$, where the complexity resulting from schemes with $L > 5$ can be avoided as they provide no distinguishable improvement

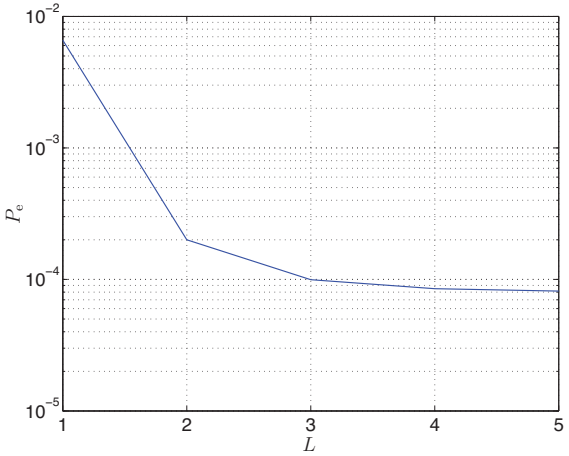

Figure 7.5.: P_e versus L for $\theta = 25$ dB, $\vartheta = 15$ dB and $a_{q,\psi}(k) = a_{q,\psi,\text{heavy}}(k)$.

in the performance. Table 7.3 summarizes the parameters utilized during simulation.

Table 7.3.: Parameters for the coupled CHMM based prediction scheme.

Parameter	Parameter Value
K	64
$\tilde{L}_\text{P}$	2000
L_L	1500
L	5
L_O	$5 \cdot 10^5$
R	8
ϵ	10^{-4}
κ_i	0.5

Furthermore, P_e has been investigated for environments characterized by different θ and ϑ values. Figure 7.6 shows P_e as a function of θ for the heavy traffic scenario, $L = 5$ and different ϑ values. It can be shown that the performance of the implemented scheme improves by increasing θ values up to 25 dB, as well as for higher ϑ

values. On the other hand, the performance of the light traffic scenario is presented in Figure 7.7. Based on Figure 7.6 and Figure 7.7, it can be noticed that the performance of the coupled prediction scheme for both patterns are quite similar. As an example, concerning an environment characterized by $\theta = 25$ dB and $\vartheta = 15$ dB, $P_\mathrm{e} \approx 8.16 \cdot 10^{-5}$ for the heavy traffic scenario, whereas $P_\mathrm{e} \approx 8.48 \cdot 10^{-5}$ for the light one. Hence, it can be deduced that the presented coupled scheme is also robust against the different PU traffic patterns, as in Section 7.1, as it can adapt its parameters during the learning phase based on the traffic pattern under consideration. In view of the performance of the uncoupled CHMM based prediction scheme shown in Figure 7.1 and Figure 7.2, one can deduce that the introduced coupling mechanism in Figure 6.1 presented in the form of the coupled CHMM prediction scheme has indeed improved the performance of the proposed prediction scheme.

Moreover, simulations have been conducted to investigate the parameters affecting the scheme's performance. For this matter, simulation is performed to evaluate P_e as a function of L_L. For an environment characterized by $\theta = 25$ dB and $\vartheta = 15$ dB, Figure 7.8 presents P_e as a function of L_L for the coupled CHMM prediction scheme as well as the uncoupled scheme in Figure 7.3. The corresponding heavy traffic patterns have been considered for both schemes. As probability of error for the uncoupled CHMM prediction scheme saturates at $P_\mathrm{e} \approx 6 \cdot 10^{-3}$ for $L_\mathrm{L} \geq 1000$, it actually saturates at a lower probability of error value and rather shorter learning sequence for the coupled CHMM scheme, such that $P_\mathrm{e} \approx 8.15 \cdot 10^{-5}$ for $L_\mathrm{L} \geq 800$. Hence, one can indeed employ an even shorter learning sequence while implementing the coupled CHMM prediction scheme to obtain better performance results than the one obtained by the uncoupled CHMM prediction scheme.

Finally, simulation has been performed to obtain the required number of iterations to obtain the best-fit model in analogy to Figure 7.4. Thus, Figure 7.9 presents m_it as a function of L_L and ϑ for $\theta = 25$ dB. Similar to Figure 7.4, the number of iterations m_it decreases by increasing L_L and ϑ values. The required number of iterations for $L_\mathrm{L} = 10$ at $\vartheta = -5$ dB is $m_\mathrm{it} \approx 200$, which is even smaller than the number of iterations required for the CHMM based prediction scheme under similar conditions. However, by increasing the ϑ and L_L values, m_it drops down to 4 iterations. Apart from performance improvement, the value of m_it hence decreases by introducing the coupling mechanism in Figure 6.1.

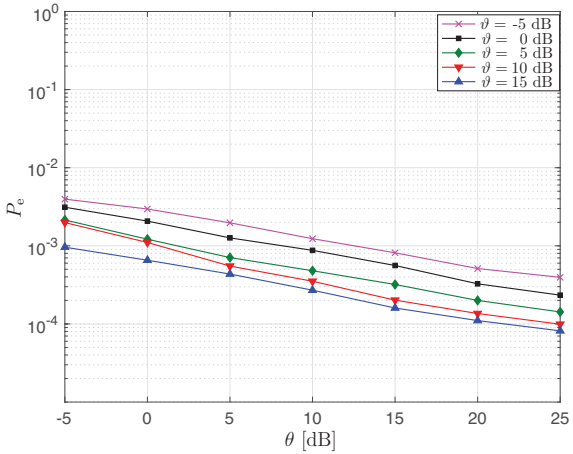

Figure 7.6.: P_e versus θ for $L = 5$ and $a_{q,\boldsymbol{\psi}}(k) = a_{q,\boldsymbol{\psi},\mathrm{heavy}}(k)$.

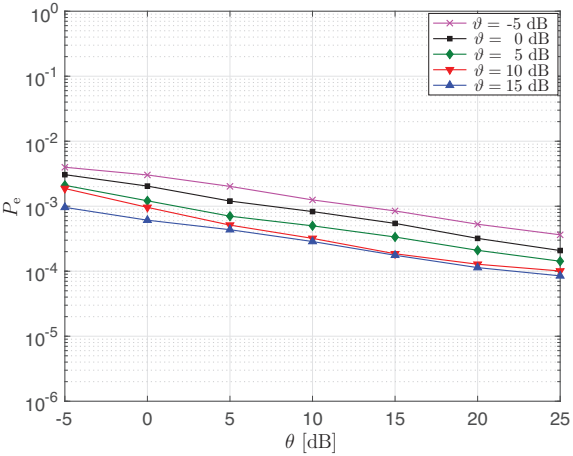

Figure 7.7.: P_e versus θ for $L = 5$ and $a_{q,\boldsymbol{\psi}}(k) = a_{q,\boldsymbol{\psi},\mathrm{light}}(k)$.

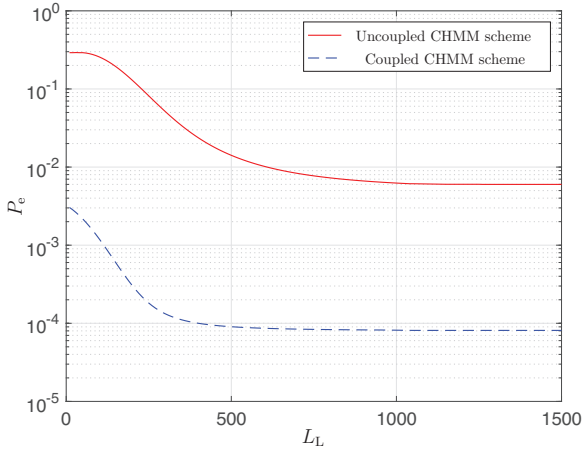

Figure 7.8.: P_e as a function of L_L.

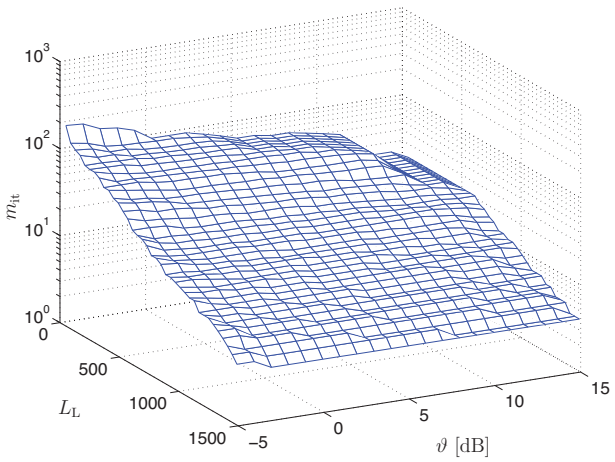

Figure 7.9.: m_{it} versus L_L and ϑ for the coupled CHMM based scheme, and $\theta = 25$ dB.

Chapter 8.
8 Conclusions and Outlook

8.1. Conclusions

The recently discussed CR technology presents a new dynamic and intelligent wireless platform which is able to learn and adapt its parameters to the surrounding environment. Also, it presents promising solutions for the inefficient utilization of the available radio resources; as the unlicensed ISM bands are crowded, over-utilized and suffer from high CCI in comparison to the unexploited licensed bands [3]. Furthermore, CR technology facilitates the emergence of the new 5G telecommunication standards, as it presents solutions for the current challenges resulting from the continuously increasing demands on the available resources due to arising new wireless protocols and increasing number of devices and services. Such a flexible radio platform with extremely high data rate and low latencies is capable of overcoming the challenges facing the current 4G standards [120]. Along with FBMC techniques, massive multiple input multiple output (MIMO) and millimeter wave (mmWave), both CR and machine learning are indeed key enablers to the 5G telecommunication standards [8, 120].

The term *cognitive* indicates awareness, consciousness, perception, adaptation, and judgment [144]. However, this can only be achieved by allowing the radio to *learn* its history to remember past lessons so that it can act in the future accordingly [144]. Therefore, machine learning concepts along with signal processing– and other fundamental theories– are needed to come up with such a CR based dynamic radio platform [10].

CR technology allows SUs to explore the licensed bands whenever a PU is unavailable, or to coexist with an already existing licensed subband in a suitable interference-tolerance mechanism. Thus, the predefined QoS of the PU is maintained. Hence, CRN gets rid of the rigid spectrum assignment policies [14], using [145]:

1. sensing and monitoring the spectrum
2. learning and characterizing the surroundings
3. optimizing a spectrum access decision based on the available constraints in the form of spectrum availability, the followed CR strategy, power loading mechanism, QoS requirements, etc., and
4. maintaining the adaptation process and spectrum sharing schemes.

Nevertheless, spectrum occupancy prediction schemes are essential for CRNs, in order to maintain the PU's QoS. The obtained future spectrum occupancy information can be utilized in a proper spectrum access scheme. For this matter, HMM based predictors have been introduced in this work, as they have shown reliability in speech recognition, pattern classification and prediction [46, 128]. Furthermore, for maximum resolution and utilization of the received signal, the SU employs a properly designed critically sampled DFT-FB as part of its transceiver structure. Hence, spectrum sensing, spectrum occupancy prediction, and spectrum access are all implemented in one transceiver structure [119]. Irrespective of the considered CR paradigm, a spectrum occupancy prediction scheme can be employed for modeling and prediction purposes, as well as energy consumption consideration.

Unlike the available spectrum occupancy prediction schemes, this work has presented two novel TF spectrum occupancy prediction schemes based on the interdependencies between neighboring subbands. For a system of independently occupied subbands by the PU signal, a first-order CHMM based prediction scheme has been introduced to replace the conventional and widely implemented DHMM based prediction approach. Due to its superior performance in comparison with the conventional scheme, the CHMM based approach has been expanded in the form of a lag-one first-order coupled CHMM based prediction scheme, as it is able to model and predict the spectrum occupancy for a system of dependently occupied subbands.

As CCI plays an important role in wireless networks especially in the ISM band, the Matérn hard-core spatial process has been considered to model the non-Gaussian CCI signal spatially in a CSMA/CA network [1], as an accurate approach instead of the PPP. Hence, the performance of the presented prediction schemes has been considered based on different environments characterized by signal-to-noise ratio as well as signal-to-interference ratio. Furthermore, their reliability has also been considered with respect to the length of the learning sequence observed during the learning phase, as an essential parameter in any machine learning based scheme.

On the other hand, in order to identify an appropriate environment or application for the introduced TF based spectrum occupancy prediction schemes, a short overview of the available approaches in literature for modeling the wireless traffic is provided. Statistical traffic characterization and modeling have been investigated for both indoor and outdoor wireless environments as well as for licensed and unlicensed spectrum bands, as it helps to analyze and evaluate the performance of networks [146]. Different studies have shown that wireless traffic has long-range dependence (LRD) properties, as its auto-correlation function decays slowly over time and produces heavy-tailed characteristics [147]. As a result, the wireless traffic displays a rather long-term memory behavior [148]. Furthermore, wireless traffic shows a self-similar nature due to its heavy-tailed behavior [146]. Self-similarity indicates that a segment of wireless traffic over a certain time period would look or behave like a similar segment over a different time period [146]. Therefore, any traditional exponentially and rapidly decaying short-range dependence (SRD) traffic models, such as Poisson and Markov traffic models [147, 149], might rather not be the first choice as appropriate traffic modeling approaches.

Nevertheless, different studies have still employed such SRD approaches, including Markov based schemes, to model LRD traffic due to the following justifications and reasons:

1. It has been clearly declared in [147] that dealing with LRD traffic is “far from straightforward”, and aggregated traffic has indeed a tendency to behave as SRD traffic. Therefore, different studies have indeed utilized Markov approaches to model and predict LRD and self-similar heavy-tailed (internet) traffic [150, 151, 152, 153].

2. Authors in [154] have argued that white spaces, i.e., unoccupied spectrum bands, do exist between bursty WLAN traffic. Hence, a semi-Markov model is introduced in [154] as a suitable channel state classification approach to efficiently utilize those white spaces.

3. Suitable fitting mechanisms can be utilized to model the heavy-tailed characteristics of LRD traffic. For example, traffic with long-tailed distributions has been approximated via a mixture of a finite number of exponentials [155]. Furthermore, before applying Markov based modeling approaches, different approximation fitting techniques have been employed [156, 157] .

4. In resource-limited networks, the wireless traffic is indeed considered as SRD traffic rather than LRD one [158]. Hence, authors in [126] have investigated second-order statistical characteristics of ISM indoor traffic. Based on real-time measurements, it turns out that the traffic has SRD properties. Furthermore, both time and frequency auto-correlation functions have been modeled as exponentially decaying functions.

Thus, the proposed first-order CHMM based schemes are indeed suitable for modeling and predicting the spectrum occupancy for a resource-limited indoor wireless network. A similar approach has been followed in [159] to model and analyze the performance of communication channels in internet protocol (IP) networks. Based on experimental measurements, the authors in [160] and [161] have modeled end-to-end channel behavior, however in terms of packet loss and temporal delays. Furthermore, HMM has been utilized in [150] for modeling internet traffic at packet level based on real measurements.

On the other hand, choosing a rather simple scheme is a reasonable approach to model any dynamic system. The HMM based scheme has been considered a first choice for modeling different systems. It is considered to be analytically simple and tractable scheme where joint dynamics can be easily captured [150]. Moreover, it is rather a reasonable approach to expand the simple first-order HMM scheme into a more complex coupled one. Hence, the interdependencies in TF domain has been captured, instead of investigating the spectrum's history in time in the form of higher-order schemes. The complexity of the proposed schemes in terms of their

memory is considered in a reasonable, objective, and adequate approach for TF concerns.

8.2. Open Issues and Further Considerations

As CR technology provides solutions for the current resource-related issues, new categories of threats have appeared in addition to the traditional ones, as discussed in [3, 145, 162, 163]. Along with eavesdropping and spoofing, several challenges and security concerns such as tampering, PU emulation and denial of service, which are still open issues to this end [145] are introduced.

Any tampering and manipulation in spectrum sensing data create an important security concern and a direct attack to the PUs transmission, as critical decisions regarding spectrum access and management depend on properly learning the available environment based on the signal received by SUs [3]. Furthermore, the concept of a PU emulation attack is mainly performed by an opponent which is able to replicate the PU characteristics and signals [162]. Correspondingly, spectrum sensing results contain false information leading to an improper learning phase as well as denial of service. Hence, the SUs would not be able to access the spectrum for its own transmission.

Apart from that, the SUs actually belong to different networks, such as ad-hoc, sensor networks, etc., where each one of them has its own different parameters defining spectrum quality and its QoS requirements. Hence, applying CR concepts for heterogeneous SUs is a challenging task [162]. On the other hand, a prediction scheme which takes into account sharing the spectrum between multiple SUs has not yet been investigated nor proposed to this end. Such a scheme which predicts the different spectrum requirements in time, frequency and space for heterogeneous users is a nontrivial task. Hence, the outcome of any spectrum occupancy prediction scheme cannot be directly employed for spectrum sharing purposes, an extra mechanism which takes the different requirements of heterogeneous networks into account must be proposed and employed.

On the other hand, due to its complexity, a CR based hardware implementation and development has not yet reached a level as advanced as the theoretical work found in the literature [163]. Furthermore, CR based architecture is extremely sensitive to interference [163]. This alone presents big challenges facing hardware implementation for both spectrum sensing and spectrum occupancy prediction. Moreover, maintaining a properly synchronized system due to ON-OFF transmission in the form of an interweave CRN is an exhausting task; since the re-synchronization mechanism has to be performed each time a subband is labeled to be available.

In view of the presented spectrum occupancy prediction schemes, HMM based approaches have shown and proven their reliability for different recognition, classifications, and prediction applications. Due to their simple structure, DHMM schemes have been widely implemented in CRNs based on discrete observed signals, as discussed earlier. However, for superior performance, CHMM based prediction schemes have been proposed in this work. Whether utilizing overlay, underlay or interweave strategies, CHMM based spectrum occupancy prediction schemes are valuable for CRNs to avoid collision with PUs, maintain the QoS of PUs, and for network optimization purposes in terms of power consumption during the cognitive tasks. As an example, spectrum sensing can be rescheduled based on the prediction information and upon the SU's requirement.

Although it might not coincide with the behavior of a real environment, the learning phases of the previously presented CHMM based schemes assume an environment with long-term stationarity, or at least longer than the duration needed to perform the learning phases [160]. Furthermore, it is basically essential to consider in hardware implementation the time duration required to perform the learning phase, to make sure that the prediction output in terms of future subbands' occupancy is accurate and does not contain old information. Nevertheless, it has to be understood at this point that a sufficient and a proper learning duration is indeed needed to be taken into account so that the SU is able to learn its environment, as learning is an essential part in CRN by definition [11, 144, 10].

Therefore, multiple offline learning phases under different environment conditions are to be considered, where the resulting estimates are recorded in multiple look-up tables. Hence, the computational load performed by the hardware is reduced, and

the resulting occupancy information is up-to-date. Another considered solution is based on a predefined and acceptable probability of collision by the PU. For this reason, an upper bound on the number of iterations required during the forward-backward algorithm as well as the length of the utilized learning sequence can be defined during the learning phase. Ultimately, it is necessary to find a balance between the requirements of the PU, and the capabilities of the CHMM prediction schemes.

The HMM based scheme, unlike other machine learning-based approaches such as support vector machine (SVM), requires a prior knowledge of the PU traffic pattern [35], and it is sensitive to the chosen initial model estimates during its learning phase [128]. In other words, a local maximum might be obtained during the forward-backward algorithm due to an inaccurate initialization of the transition probabilities [128]. To this end, the learning phase can be repeated with several different initializations to obtain a global solution [160]. Furthermore, multiple learning phases for multiple environments facilitate the SU in gaining a proper understanding and comprehension to choose the initial parameters based on the present circumstances, to obtain the MLE of the parameters of the model under consideration, as explained earlier. In this work, k-means clustering has been considered as part of the pilot phase for a proper initialization in order to obtain the conditional PDF of the received signal under the occupancy state of the considered spectrum.

In terms of complexity, it has to be declared that the DHMM based scheme is less complex than the presented CHMM based scheme. Furthermore, the presented coupled CHMM scheme is even more complicated than the single chain schemes. As stated earlier, one can adjust the schemes, in terms of the number of iterations and the length of the learning sequence during the learning phase, for a least computationally complex scheme based on the requirements defined by the PU and the suggestions discussed earlier in this section. Moreover, approaching the interdependencies between subbands in terms of a certain number of subbands defined in a block reduces the complexity to an extent.

Appendix A.

D-separation Concept

The most important concept in probabilistic models is the conditional independency, as it plays an important role in simplifying the structure of the considered model as well as its inference. The advantage of representing a model graphically is to figure out the independence property of the joint distribution directly without the need of performing any complicated analytical manipulations [52]. Given the three nonintersecting nodes X_1, X_2, and X_3 with their corresponding realizations x_1, x_2 and x_3, respectively, representing random variables in a graphical model, the expression

$$\Pr\{X_1 = x_1 | X_2 = x_2, X_3 = x_3\} = \Pr\{X_1 = x_1 | X_3 = x_3\}, \tag{A.1}$$

indicates that X_1 is conditionally independent of X_2 given (conditioned on) X_3. Hence, their joint distribution can be written as

$$\begin{aligned}\Pr\{X_1 = x_1, X_2 = x_2 | X_3 = x_3\} &= \Pr\{X_1 = x_1 | X_2 = x_2, X_3 = x_3\} \Pr\{X_2 = x_2 | X_3 = x_3\} \\ &= \Pr\{X_1 = x_1 | X_3 = x_3\} \Pr\{X_2 = x_2 | X_3 = x_3\},\end{aligned}$$

given (A.1) and the product rule in (2.2). In other words, upon conditioning on X_3, X_1 and X_2 are independent. This property can be represented in notations as

$$X_1 \perp\!\!\!\perp X_2 \mid X_3.$$

Furthermore, the joint probability of X_1, X_2 and X_3 can be expressed as

$$\begin{aligned}&\Pr\{X_1 = x_1, X_2 = x_2, X_3 = x_3\} = \\ &\qquad \Pr\{X_1 = x_1 | X_3 = x_3\} \Pr\{X_2 = x_2 | X_3 = x_3\} \Pr\{X_3 = x_3\}.\end{aligned}$$

A so-called direct separation criterion, or shortly *d-separation*, is important to identify and figure out the conditional independencies between nodes in a graph. Hence, the resulting graph is simply represented, and its joint distribution is easily and directly deduced [52]. Graphically, the conditional dependencies are represented via arrows from node X_3 to node X_1, as well as from X_3 to X_2. Upon conditioning on X_3, the path between X_1 and X_2 is said to be *blocked* by node X_3 [52]. As being such an important property in DAG, the d-separation concept is presented in this part of the thesis.

To properly define the d-separation criteria, let A, B and Cbe three nonintersecting nodes in a DAG. Figure A.1 presents three different examples of paths in a DAG where we condition on C.

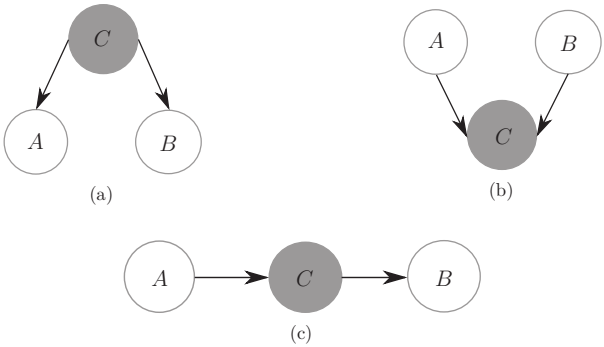

Figure A.1.: Three examples for a path from node A to node B upon conditioning on C. The path with respect to C is said to be: (a) tail-to-tail, (b) head-to-head, and (c) head-to-tail, or equivalently tail-to-head.

The path from from node A to node B is said to be *blocked* by node C [52] if:

- the arrows in the defined path meet either head-to-tail, tail-to-head, or tail-to-tail at node C or any node V in the set of C, or
- the arrows meet head-to-head at node E, where neither E nor its descendent nodes are in the set of C.

In other words, if all possible paths from A to B are blocked upon conditioning on C, then the following remarks are valid:

1. A is d-separated from B by node C.

2. The joint distribution of the nodes in the graph satisfies the relation $A \perp\!\!\!\perp B \mid C$.

The recursive relation for calculating the forward and backward probabilities represented in this work –irrespective of the implemented HMM scheme– are derived and simplified based on the d-separation criteria, along with the Bayes and further probabilistic rules [52]. Due to its notation simplicity, we consider the recursive relations of the forward and backward probabilities for the DHMM scheme presented in Section 5.2. Note that when defining $\alpha_{i,\mathrm{D}}(\ell)$ and $\beta_{i,\mathrm{D}}(\ell)$, the dependencies on the model Γ_{D} still holds, however it is omitted in this part for notation simplicity purposes.

Recursive Definition of the Forward Probability

The recursive relation for the forward probability $\alpha_{i,\mathrm{D}}(\ell)$, as presented in (5.18), is derived based on the steps explained as follows. Starting with its definition

$$\begin{aligned}\alpha_{i,\mathrm{D}}(\ell) &= \Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell,\mathrm{D}}, Q_\ell = i\} \\ &= \Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell,\mathrm{D}} | Q_\ell = i\} \Pr\{Q_\ell = i\}.\end{aligned} \tag{A.2}$$

As the sequence $z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}}$ is actually d-separated from $z_{\ell,\mathrm{D}}$ by Q_ℓ, i.e., $z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}} \perp\!\!\!\perp z_{\ell,\mathrm{D}} \mid Q_\ell$, the expression

$$\Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}} | z_{\ell,\mathrm{D}}, Q_\ell = i\} = \Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}} | Q_\ell = i\}$$

is valid, hence the right-hand side in (A.2) is rewritten as

$$\alpha_{i,\mathrm{D}}(\ell) = \Pr\{z_{\ell,\mathrm{D}} | Q_\ell = i\} \Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}} | Q_\ell = i\} \Pr\{Q_\ell = i\}$$

Upon considering both the Bayes and sum rules, $\alpha_{i,\mathrm{D}}(\ell)$ is written as

$$\begin{aligned}\alpha_{i,\mathrm{D}}(\ell) &= \Pr\{z_{\ell,\mathrm{D}}|\, Q_\ell = i\} \Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}}, Q_\ell = i\} \\ &= \Pr\{z_{\ell,\mathrm{D}}|\, Q_\ell = i\} \sum_j \Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}}, Q_{\ell-1} = j, Q_\ell = i\} \\ &= \Pr\{z_{\ell,\mathrm{D}}|\, Q_\ell = i\} \sum_j \Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}}, Q_\ell = i|\, Q_{\ell-1} = j\} \Pr\{Q_{\ell-1} = j\}.\end{aligned}$$

However, as $z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}} \perp\!\!\!\perp Q_\ell \mid Q_{\ell-1}$, $\alpha_{i,\mathrm{D}}(\ell)$ is rewritten as

$$\begin{aligned}\alpha_{i,\mathrm{D}}(\ell) =& \Pr\{z_{\ell,\mathrm{D}}|\, Q_\ell = i\} \\ & \sum_j \Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}}|\, Q_{\ell-1} = j\} \Pr\{Q_\ell = i|\, Q_{\ell-1} = j\} \Pr\{Q_{\ell-1} = j\},\end{aligned}$$

and via the Bayes rule

$$\alpha_{i,\mathrm{D}}(\ell) = \underbrace{\Pr\{z_{\ell,\mathrm{D}}|\, Q_\ell = i\}}_{b_{iz_\mathrm{D}}} \sum_j \underbrace{\Pr\{z_{0,\mathrm{D}}, \ldots, z_{\ell-1,\mathrm{D}}, Q_{\ell-1} = j\}}_{\alpha_{j,\mathrm{D}}(\ell-1)} \underbrace{\Pr\{Q_\ell = i|\, Q_{\ell-1} = j\}}_{a_{ji}},$$

which is nothing but the expression in (5.18) for a certain model Γ_D.

Recursive Definition of the Backward Probability

Furthermore, deriving the recursive relation of $\beta_{i,\mathrm{D}}(\ell)$ goes in similar approach. Based on the sum rule and the Bayes rules, its main definition can be expressed as

$$\begin{aligned}\beta_{i,\mathrm{D}}(\ell) &= \Pr\{z_{\ell+1,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}}|\, Q_\ell = i\} \\ &= \sum_j \Pr\{z_{\ell+1,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}}, Q_{\ell+1} = j|\, Q_\ell = i\} \\ &= \sum_j \Pr\{z_{\ell+1,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}}|\, Q_\ell = i, Q_{\ell+1} = j\} \Pr\{Q_{\ell+1} = j|\, Q_\ell = i\}.\end{aligned}$$

Since $z_{\ell+1,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}} \perp\!\!\!\perp Q_\ell \mid Q_{\ell+1}$, the expression

$$\Pr\{z_{\ell+1,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}}|\, Q_\ell = i, Q_{\ell+1} = j\} = \Pr\{z_{\ell+1,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}}|\, Q_{\ell+1} = j\},$$

is valid, and $\beta_{i,\mathrm{D}}(\ell)$ can be written as

$$\beta_{i,\mathrm{D}}(\ell) = \sum_j \Pr\{z_{\ell+1,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}} | Q_{\ell+1} = j\} \Pr\{Q_{\ell+1} = j | Q_\ell = j\}.$$

Moreover, $z_{\ell+2,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}} \perp\!\!\!\perp Q_\ell \mid Q_{\ell+1}$, indicating that

$$\Pr\{z_{\ell+2,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}} | Q_{\ell+1} = j, z_{\ell+1,\mathrm{D}}\} = \Pr\{z_{\ell+2,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}} | Q_{\ell+1} = j\},$$

hence, $\beta_{i,\mathrm{D}}(\ell)$ is rewritten as

$$\beta_{i,\mathrm{D}}(\ell) = \sum_j \underbrace{\Pr\{z_{\ell+2,\mathrm{D}}, \ldots, z_{L_\mathrm{L}-1,\mathrm{D}} | Q_{\ell+1} = j\}}_{\beta_{j,\mathrm{D}}(\ell+1)} \underbrace{\Pr\{z_{\ell+1,\mathrm{D}} | Q_{\ell+1} = j\}}_{b_{jz_\mathrm{D}}} \underbrace{\Pr\{Q_{\ell+1} = j | Q_\ell = i\}}_{a_{ij}},$$

which is given in (5.19).

Appendix B.
B Gaussian Mixture Models

For notation simplicity, and in order to focus on estimating the GMM parameters, the state dependencies in the model presented here have been dropped, unlike in (5.3), (4.6) and (6.7). A rather parametric model of an arbitrary D- dimensional random vector $\mathbf{X} \in \mathbb{R}^D$ is discussed here as a general case. Analogous to those mentioned equations, a GMM with $R \in \mathbb{N}$ number of components, i.e, a model of order R, is given by

$$p_{\mathbf{X}}\left(\mathbf{x}\,;\Theta\right) = \sum_{r=1}^{R} w_r\, \mathcal{N}(\mathbf{x};\boldsymbol{\mu}_r,\boldsymbol{\Sigma}_r), \tag{B.1}$$

where $\mathbf{x}$ is a realization of $\mathbf{X}$. In (B.1), the model's parameters are presented in the set $\Theta = \left(\theta^{(1)},\ldots,\theta^{(R)}\right)$, where $\theta^{(r)} = (w_r, \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r)$ are the parameters defining the r^{th} Gaussian component. $w_r, \boldsymbol{\mu}_r$ and $\boldsymbol{\Sigma}_r$ are the weight, mean vector and covariance matrix of the r^{th} component, respectively. The weight w_r in (B.1) satisfies the conditions

$$0 \leq w_r \leq 1 \quad \text{and} \quad \sum_{r}^{R} w_r = 1.$$

Furthermore, the PDF of the r^{th} Gaussian component is given by

$$f\left(\mathbf{x}\,;\boldsymbol{\mu}_r,\boldsymbol{\Sigma}_r\right) = \frac{1}{(2\pi)^{D/2}\left(\det\,\boldsymbol{\Sigma}_r\right)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_r)^{\intercal}\,\boldsymbol{\Sigma}_r^{-1}(\mathbf{x}-\boldsymbol{\mu}_r)\right). \tag{B.2}$$

B.1. Estimating the GMM's Parameters of a Certain Order

Based on a sequence of $L_{\mathbf{x}} \in \mathbb{N}$ mutually independent observations

$$\mathbf{x}_1, \ldots, \mathbf{x}_{L_{\mathbf{x}}}, \tag{B.3}$$

a best-fit GMM is obtained for a predefined R components by estimating the parameters in Θ. For this matter, the EM algorithm is employed iteratively based on the MLE criterion. It requires proper parameter initialization, therefore the k-mean clustering algorithm, known as the hard EM algorithm [55], is chosen for this purpose. The main steps in k-means clustering as well as EM algorithms for GMM are summarized as follows [125, 55].

k-means initialization

1. R number of observations are randomly chosen from the observed sequence in (B.3) to be the initial values of the mean vectors $\boldsymbol{\mu}_r^{(0)}, \forall r = \{1, \ldots, R\}$.

2. The closest cluster center ψ_i for each $\mathbf{x}_i$ with $i = \{1, \ldots, L_{\mathbf{x}}\}$ is found, such that
$$\left|\mathbf{x}_i - \boldsymbol{\mu}_{\psi_i}\right|^2 = \min_{r=1,\ldots,R} \left|\mathbf{x}_i - \boldsymbol{\mu}_r\right|^2.$$

3. Update each cluster center ψ_i iteratively by recalculating the r^{th} mean vector. At the m^{th} iteration the mean vector is given by
$$\boldsymbol{\mu}_r^{(m)} = \frac{1}{\sum_{i=1}^{L_{\mathbf{x}}} I\left(\psi_i = r\right)} \sum_{i=1}^{L_{\mathbf{x}}} I\left(\psi_i = r\right) \mathbf{x}_i,$$
where $I\left(\psi_i = r\right)$ is an indicator function, where $I\left(\psi_i = r\right) = 1$ if the equality $\psi_i = r$ holds, and zero otherwise.

4. Upon exceeding the maximum number of iterations m_{k}, the mean vector is selected such that $\boldsymbol{\mu}_r = \boldsymbol{\mu}_r^{(m_{\mathrm{k}})}$ [125]. Furthermore, w_r and $\boldsymbol{\Sigma}_r$ are correspondingly calculated based on $\boldsymbol{\mu}_r^{(m_{\mathrm{k}})}$ as

$$w_r = \frac{1}{L_{\mathbf{x}}} \sum_{i=1}^{L_{\mathbf{x}}} I\left(\psi_i = r\right),$$

and

$$\boldsymbol{\Sigma}_r = \frac{1}{\sum_{i=1}^{L_{\mathbf{x}}} I\left(\psi_i = r\right)} \sum_{i=1}^{L_{\mathbf{x}}} I\left(\psi_i = r\right) \left(\mathbf{x}_i - \boldsymbol{\mu}_r^{(m_{\mathrm{k}})}\right) \left(\mathbf{x}_i - \boldsymbol{\mu}_r^{(m_{\mathrm{k}})}\right)^{\mathsf{T}}.$$

Those values are then utilized as the best chosen initial estimate to be employed during the EM algorithm, which to be described as follows.

The EM algorithm

Starting from the initial values $\Theta^{(0)} = \left(w_1^{(0)}, \boldsymbol{\mu}_1^{(0)}, \boldsymbol{\Sigma}_1^{(0)}, \ldots, w_R^{(0)}, \boldsymbol{\mu}_R^{(0)}, \boldsymbol{\Sigma}_R^{(0)}\right)$, which are obtained via k-means clustering algorithm, and given the observed sequence in (B.3), iterative update takes place. A new update at the $(m+1)^{\text{th}}$ iteration, i.e., $\Theta^{(m+1)} = \left(\hat{w}_r^{(m+1)}, \hat{\boldsymbol{\mu}}_r^{(m+1)}, \hat{\boldsymbol{\Sigma}}_r^{(m+1)}\right), \forall r = \{1, \ldots, R\}$, is calculated by as

1. E-step: calculating the posteriori probability for the $r = 1, \ldots, R$ components at the m^{th} iteration, and for $i = 1, \ldots, L_{\mathbf{x}}$ as

$$\Pr\left\{r \mid \mathbf{x}_i\,; \theta_r^{(m)}\right\} = \frac{w_r^{(m)} \mathcal{N}(\mathbf{x}_i; \boldsymbol{\mu}_r^{(m)}, \boldsymbol{\Sigma}_r^{(m)})}{\sum_{h=1}^{R} w_h^{(m)} \mathcal{N}(\mathbf{x}_i; \boldsymbol{\mu}_h^{(m)}, \boldsymbol{\Sigma}_h^{(m)})}.$$

2. M-step: updating the parameters as follows:

$$\hat{w}_r^{(m+1)} = \frac{1}{L_{\mathbf{x}}} \sum_{i}^{L_{\mathbf{x}}} \Pr\left\{r \mid \mathbf{x}_i\,; \theta_r^{(m)}\right\},$$

$$\hat{\boldsymbol{\mu}}_r^{(m+1)} = \frac{\sum_{i=1}^{L_{\mathbf{x}}} \Pr\left\{ r \mid \mathbf{x}_i ; \theta_r^{(m)} \right\} \mathbf{x}_i}{\sum_{i=1}^{L_{\mathbf{x}}} \Pr\left\{ r \mid \mathbf{x}_i ; \theta_r^{(m)} \right\}} = \frac{1}{L_{\mathbf{x}}\, \hat{w}_r^{(m+1)}} \sum_{i=1}^{L_{\mathbf{x}}} \Pr\left\{ r \mid \mathbf{x}_i ; \theta_r^{(m)} \right\} \mathbf{x}_i,$$

and

$$\hat{\boldsymbol{\Sigma}}_r^{(m+1)} = \frac{1}{L_{\mathbf{x}}\, \hat{w}_r^{(m+1)}} \sum_{i=1}^{L_{\mathbf{x}}} \Pr\left\{ r \mid \mathbf{x}_i ; \theta_r^{(m)} \right\} (\mathbf{x}_i - \hat{\boldsymbol{\mu}}_r^{(m+1)}) (\mathbf{x}_i - \hat{\boldsymbol{\mu}}_r^{(m+1)})^{\mathsf{T}}.$$

3. This updating procedure continues until the increase in the likelihood function is within a certain threshold ϵ, as

$$\left| \frac{p\left(\mathbf{x}_1, \ldots, \mathbf{x}_{L_{\mathbf{x}}} ; \Theta^{(m+1)}\right) - p\left(\mathbf{x}_1, \ldots, \mathbf{x}_{L_{\mathbf{x}}} ; \Theta^{(m)}\right)}{p\left(\mathbf{x}_1, \ldots, \mathbf{x}_{L_{\mathbf{x}}} ; \Theta^{(m)}\right)} \right| \leq \epsilon.$$

B.2. Estimating the Model Order

Before estimating the parameters in Θ, the model order R has to be predefined. Different criteria, such as Akaike information criterion (AIC) and Bayesian information criterion (BIC), can be employed to find the minimum number of Gaussian components which are needed to best describe the observed sequence in (B.3) [125]. However, those criteria might over-estimate the model's order and provide inaccurate information regarding the accuracy of the estimated model [1]. Therefore, a heuristic approach [126] is performed to find the minimum number of components R required such that the increase in the likelihood function based on a certain R with respect to the one at $R-1$ is below a certain threshold ϵ as

$$R = \min \left\{ \hat{R} \in \mathbb{N} : \left| \frac{p\left(\mathbf{x}_1, \ldots, \mathbf{x}_{L_{\mathbf{x}}}\right) |_{R=\hat{R}} - p\left(\mathbf{x}_1, \ldots, \mathbf{x}_{L_{\mathbf{x}}}\right) |_{R=\hat{R}-1}}{p\left(\mathbf{x}_1, \ldots, \mathbf{x}_{L_{\mathbf{x}}}\right) |_{R=\hat{R}-1}} \right| \leq \epsilon \right\}. \tag{B.4}$$

List of Abbreviations

IoT	Internet of Things
ISM	industrial, scientific and medical
QoS	Quality of Service
CCI	co-channel interference
CR	cognitive radio
SU	secondary user
CRN	CR network
PU	primary user
TFR	time-frequency representation
TF	time-frequency
FB	filter bank
FBMC	FB based multi-carrier
OFDM	orthogonal frequency division multiplexing
AFB	analysis FB
SFB	synthesis FB
ISI	inter-symbol interference
DFT	discrete Fourier transform
DFT-FB	DFT - modulated FB
HMM	hidden Markov model
FFT	fast Fourier transform
CSMA	carrier sense multiple access
NN	neural network
MLP	multilayer perceptron
SVR	support vector regression
NSHMM	non-stationary HMM
ARMA	autoregressive moving average

AR	autoregressive
EMA	exponential moving average
DHMM	discrete HMM
CHMM	continuous HMM
NP	Neyman-Pearson
SNR	signal-to-noise ratio
SIR	signal-to-interference ratio
PDF	probability density function
GMM	Gaussian mixture model
DGN	directed graphical network
BN	Bayesian network
DBN	dynamic Bayesian network
UGN	undirected graphical network
DAG	directed acyclic graph
AR-HMM	autoregressive HMM
IO-HMM	input-output HMM
FHMM	factorial HMM
CCA	clear channel assessment
CSMA/CA	carrier sense multiple access/collision avoidance
WLAN	Wireless Local Area Network
DSSS	direct sequence spread spectrum
DCF	distributed coordination function
DIFS	DCF interframe space
CW	contention window
MIMO-OFDM	multiple input multiple output-OFDM
SDM	spatial diversity multiplexing
TDM	time division multiplexing
wirelessHART	Wireless Highway Addressable Remote Transducer
HART	Highway Addressable Remote Transducer
TDMA	Time Division Medium Access
PPP	Poisson point process
BPP	Binomial Point Process
STFT	short-time Fourier transform
AWGN	additive white Gaussian noise

GLRT	generalized likelihood ratio test
EM	expectation maximization
WSS	wide-sense stationary
ED	energy detector
MLE	maximum-likelihood estimation
ML	maximum-likelihood
MIMO	multiple input multiple output
LRD	long-range dependence
SRD	short-range dependence
SVM	support vector machine
AIC	Akaike information criterion
BIC	Bayesian information criterion

List of Symbols

$\Pr\{X\}$	Probability of an arbitrary random variable X 11
$\Pr\{X \mid Y\}$	Conditional probability of an arbitrary random variable X given another Y 11
$\mathbf{E}\{X\}$	Expectation of X 11
$\lvert X \rvert$	Absolute value of X 11
$\det \mathbf{A}$	The determinant of an arbitrary matrix $\mathbf{A}$ 11
$\operatorname{diag}(\mathbf{A})$	The diagonal of a matrix $\mathbf{A}$ 11
$\mathbf{A}^{\mathrm{T}}$	The transpose of a matrix $\mathbf{A}$ 11
$\mathbf{A}^{\dagger}$	The Hermitian transpose of a matrix $\mathbf{A}$ 11
$\mathbf{I}_K$	An identity matrix of size $K \times K$ 11
A_{P}	Area of a network 28
ξ_{P}	Node intensity in a PPP network 28
N_{P}	Number of nodes in a PPP network 28
η	The mean value of a PPP 29
Φ_{P}	A random process representing a set of nodes in a PPP network 29
$\pi_{i_{\mathrm{P}}}$	The location of the $i_{\mathrm{P}}^{\mathrm{th}}$ node in a network 29
R_{M}	The minimum distance separate a valid interferer node from a user under consideration 30
Φ_{M}	A random process representing a set of nodes in a Matérn hard-core network 30
$B_{\mathrm{M}}(\pi_{i_{\mathrm{P}}}, R_{\mathrm{M}})$	A circle centered at $\pi_{i_{\mathrm{P}}}$ and has a radius of R_{M} 30
K	Number of subbands 34
M	Downsampling factor 34
Λ	The TF index set 34
Δ	A set of subbands defined as $\Delta = \{0, \ldots, K-1\}$ 34

$g[n]$	An optimized Gabor window or pulse 34
L_g	The length of a pulse $g[n]$ 34
n	The index of discrete-time signals 34
$g_{\ell,k}[n]$	A shifted version of the Gabor window $g[n]$ at the TF slot (ℓ, k) 34
ℓ	The index of time slot 34
k	The index of frequency slot 34
$g^*_{\ell,k}[n]$	The complex conjugate of $g_{\ell,k}[n]$ 35
$G^*_k(z^{-1})$	The prototype of the k^{th} AFB being the complex conjugate of $G_k(z)$ 37
$G_k(z)$	The prototype of the k^{th} SFB 37
$G(z)$	The z-transform of $g[n]$ 37
$Y[n]$	A random process representing the time-domain complex baseband received signal at the SU receiver 37
$S[n]$	A random process representing the time-domain PU signal observed at the SU receiver 37
$U[n]$	A random process representing the time-domain CCI signal observed at the SU receiver 37
$W[n]$	A random process representing the time-domain AWGN signal observed at the SU receiver 37
Q	The occupancy state or hypothesis with respect to the PU signal availability 37
$Y_{\ell,k}$	The TF coefficient of $Y[n]$ at the TF slot (ℓ, k) 38
$Q_{\ell,k}$	The occupancy state of the TF slot (ℓ, k), such that $Q_{\ell,k} \in \{0, 1\}$ 38
$U_{\ell,k}$	The TF coefficient of $U[n]$ at slot (ℓ, k) 39
$W_{\ell,k}$	The TF coefficient of $W[n]$ at slot (ℓ, k) 39
$\varphi_U(\omega)$	The characteristic function of a symmetric α-stable random process 39
$\Upsilon_{U_\text{r},U_\text{j}}$	The symmetric spectral measure of the complex symmetric α-stable random variable U, where $U = U_\text{r} + \jmath U_\text{j}$ 40
A_u	A unit shpere 40
α_int	The characteristic exponent of a symmetric α-stable distribution 40

R	The number of components in a GMM scheme, also known as the GMM order 41
w_r	The weight of the r^{th} Gaussian component in a GMM scheme approximating the interference signal, where $r = \{1, \ldots, R\}$ 41
$\boldsymbol{\mu}_r$	The mean vector of the r^{th} Gaussian component in a GMM scheme approximating the interference signal, where $r = \{1, \ldots, R\}$ 41
$\boldsymbol{\Sigma}_r$	The covariance matrix of the r^{th} Gaussian component in a GMM scheme approximating the interference signal, where $r = \{1, \ldots, R\}$ 41
L_y	Length of the observed sequence utilized to estimate the parameters of the GMM scheme at hypothesis H_1 41
$\hat{R}$	The estimated order of the GMM scheme 41
ϵ	The convergence threshold 41
$\hat{w}_r^{(0)}$	The initial estimate of w_r 42
$\hat{\boldsymbol{\mu}}_r^{(0)}$	The initial estimate of $\boldsymbol{\mu}_r$ 42
$\hat{\boldsymbol{\Sigma}}_r^{(0)}$	The initial estimate of $\boldsymbol{\Sigma}_r$ 42
Δ_{area}	Area of histogram bin 42
L_{P}	Length of the overall pilot sequence utilized during the pilot phase, and under all possible states 46
L_{L}	Length of the learning sequence utilized during the learning phase to estimate the underlying parameters of the prediction scheme 46
L_{O}	Length of the predicted instances, i.e., the prediction horizon 46
$S_{\ell,k}$	The TF coefficient of $S[n]$ at slot (ℓ, k) 47
Q_ℓ	The occupancy state of a certain subband at time slot ℓ in a system of independent subbands 47
Y_ℓ	The TF coefficient of $Y[n]$ at time slot ℓ and a certain subband for a system of independent subbands 47
S_ℓ	TF coefficient of $S[n]$ at time slot ℓ and a certain subband for a system of independent subbands 47

W_ℓ	The TF coefficient of $W[n]$ at time slot ℓ and a certain subband for a system of independent subbands 47
U_ℓ	The TF coefficient of $U[n]$ at time slot ℓ and a certain subband for a system of independent subbands 47
ν_S	The average power of the PU signal per subband 47
ν_W	The average power of the noise signal per subband 47
ν_U	The average power of the CCI signal per subband 47
Z_ℓ	The TF coefficient for a certain subband at time slot ℓ observed at the output of an ED 48
$w_{r,i}$	The weight of a univariate r^{th} Gaussian component in a GMM scheme conditioned on a subband's occupancy state $i \in Q_\ell$ in a system of independent subbands 49
$\mu_{r,i}$	The mean value of a univariate r^{th} Gaussian component in a GMM scheme conditioned on a subband's occupancy state $i \in Q_\ell$ in a system of independent subbands 49
$\sigma^2_{r,i}$	The variance of a univariate r^{th} Gaussian component in a GMM scheme conditioned on the occupancy state $i \in Q_\ell$ in a system of independent subbands 49
i	The occupancy state at TF slot (ℓ, k), where $i \in Q_{\ell,k}$ 49
$\mathbf{W}$	A $(R \times 2)$ matrix contains the weights $w_{r,i}$ 49
$\mathbf{M}$	A $(R \times 2)$ matrix contains the mean values $\mu_{r,i}$ 49
$\mathbf{\Sigma}$	A $(R \times 2)$ matrix contains the variance values $\sigma^2_{r,i}$ 49
$Z_{\text{D},\ell}$	A discretized version of the signal Z_ℓ 49
τ_{NP}	NP threshold 49
$a_{ij}(\ell)$	The transition probability from state $Q_{\ell-1} = i$ to state $Q_\ell = j$ for both uncoupled DHMM and CHMM based schemes 50
$b_{iz_{\text{D}}}(\ell)$	The conditional emission probability of observing $Z_{\text{D},\ell} = z_{\text{D}}$ at state $Q_\ell = i$ in a DHMM based scheme 50
κ_i	The initial probability $\kappa_i = Q_{0,k}$ 50
a_{ij}	The transition probability from state i to state j for both uncoupled DHMM and CHMM based schemes 50

$b_{iz_{\mathrm{D}}}$	The conditional emission probability of observing z_{D} at state i in a DHMM based scheme 50
$\mathbf{A}$	A (2×2) transition matrix the transition probabilities a_{ij} for both uncoupled DHMM and CHMM based schemes 51
$\mathbf{B}$	A (2×2) emission matrix contains the emission probabilities $b_{iz_{\mathrm{D}}}$ for the DHMM based scheme 51
Γ_{D}	The set of parameters representing the DHMM based scheme 51
Ω	A set defined by $\{x \in \mathbb{R} \vert\, 0 \leq x \leq 1\}$ 51
Ω_{D}	A set which Γ_{D} belongs to, such that $\Omega_{\mathrm{D}} = \Omega^8$ 52
$\hat{\Gamma}_{\mathrm{D,ML}}$	A ML estimate of Γ_{D} 52
$\alpha_{\ell,\mathrm{D}}(i)$	The forward quantity for an uncoupled DHMM based scheme 53
$\beta_{\ell,\mathrm{D}}(i)$	The backward quantity for an uncoupled DHMM based scheme 53
$\gamma_{\ell,\mathrm{D}}(i)$	The probability of having the state $Q_\ell = i$ given the observed learning sequence until time slot ℓ and Γ_{D} in a DHMM based scheme 53
$\hat{\Gamma}_{\mathrm{D}}^{(0)}$	An initial estimate of Γ_{D} 54
$\hat{\mathbf{A}}^{(0)}$	An initial estimate of $\mathbf{A}$ 54
$\hat{\mathbf{B}}^{(0)}$	An initial estimate of $\mathbf{B}$ 54, 78
$\hat{\alpha}_{\ell,\mathrm{D}}(i)$	An estimate of $\alpha_{\ell,\mathrm{D}}(i)$ for a certain estimate of Γ_{D} 54
$\hat{\beta}_{\ell,\mathrm{D}}(i)$	An estimate of $\beta_{\ell,\mathrm{D}}(i)$ for a certain estimate of Γ_{D} 54
$\hat{b}_{iz_{\mathrm{D}}}^{(0)}(\ell)$	An initial estimate of $b_{iz_{\mathrm{D}}}(\ell)$ 55
$\hat{b}_{jz_{\mathrm{D}}}^{(0)}(\ell)$	An initial estimate of $b_{jz_{\mathrm{D}}}(\ell)$ 55
$\hat{a}_{ji}^{(0)}$	An initial estimate of a_{ji} 55
$\hat{a}_{ij}^{(0)}$	An initial estimate of a_{ij} 55
$\zeta_{\ell,\mathrm{D}}(i,j)$	The probability of having the states $Q_{\ell-1} = i$ and $Q_\ell = j$ given the observed learning sequence and Γ_{D} in a DHMM based scheme 55
$\hat{\zeta}_{\ell,\mathrm{D}}(i,j)$	An estimate of $\zeta_{\ell,\mathrm{D}}(i,j)$ 56
$\hat{\gamma}_{\ell,\mathrm{D}}(i)$	An estimate of $\gamma_{\ell,\mathrm{D}}(i)$ 56
$\hat{\Gamma}_{\mathrm{D}}$	An estimate of Γ_{D} 56

$m_{\text{it,D}}$	Number of iterations required during the learning phase to obtain the best-fit model $\hat{\Gamma}_{\text{D}}^{(m_{\text{it,D}})}$ 57
$\hat{\Gamma}_{\text{D}}^{(m_{\text{it,D}})}$	A best-fit uncoupled DHMM scheme as an output of the learning phase after $m_{\text{it,D}}$ number of iterations 57
$\tilde{\alpha}_{i,\text{D}}(\ell)$	A normalized version of $\alpha_{i,\text{D}}(\ell)$ 57
$\tilde{\beta}_{i,\text{D}}(\ell)$	A normalized version of $\beta_{i,\text{D}}(\ell)$ 57
j	The occupancy state at TF slot (ℓ, k), where $j \in Q_{\ell,k}$ 58
$\delta_{\ell,\text{D}}(j)$	A cumulative best score which maximizes the joint probability of $Z_{\text{D},\ell}$ and Q_ℓ until time slot ℓ for the uncoupled DHMM based scheme 58
$\hat{Q}_\ell$	An estimate of Q_ℓ 58
Γ_{C}	The set of parameters representing the uncoupled CHMM based scheme 59
Ω_{C}	A set which Γ_{C} belongs to, such that $\Omega_{\text{C}} = \Omega^4$ 59
$\hat{\Gamma}_{\text{C}}^{(0)}$	An initial estimate of Γ_{C} 60
$\alpha_{\ell,\text{C}}(i)$	The forward quantity for an uncoupled CHMM based scheme 60
$\beta_{\ell,\text{C}}(i)$	The backward quantity for an uncoupled CHMM based scheme 60
$\hat{\alpha}_{\ell,\text{C}}(i)$	An estimate of $\alpha_{\ell,\text{C}}(i)$ for a certain estimate of Γ_{C} 61
$\hat{\beta}_{\ell,\text{C}}(i)$	An estimate of $\beta_{\ell,\text{C}}(i)$ for a certain estimate of Γ_{C} 61
$\zeta_{\ell,\text{C}}(i,j)$	The analogous quantity of $\zeta_{\ell,\text{D}}(i,j)$ for the uncoupled CHMM given the observed learning sequence and Γ_{C} 61
$\gamma_{\ell,\text{C}}(i,r)$	The analogous quantity of $\gamma_{\ell,\text{D}}(i)$ for the uncoupled CHMM scheme given the learning sequence, the r^{th} Gaussian component and Γ_{C} 61
$\hat{\zeta}_{\ell,\text{C}}(i,j)$	An estimate of $\zeta_{\ell,\text{C}}(i,j)$ 61
$\hat{\gamma}_{\ell,\text{C}}(i,r)$	An estimate of $\gamma_{\ell,\text{C}}(i,r)$ 61
$\hat{\Gamma}_{\text{C}}$	An estimate of Γ_{C} 62
$m_{\text{it,C}}$	Number of iterations required during the learning phase to obtain the best-fit model $\hat{\Gamma}_{\text{C}}^{(m_{\text{it,C}})}$ 62
$\hat{\Gamma}_{\text{C}}^{(m_{\text{it,C}})}$	A best-fit uncoupled CHMM scheme as an output of the learning phase after $m_{\text{it,C}}$ number of iterations 62

$\delta_{\ell,\mathrm{C}}(j)$	A cumulative best score in analogy to $\delta_{\ell,\mathrm{D}}(j)$ for the uncoupled CHMM based scheme 62
$\mathbf{Q}_\ell$	A random vector representing the occupancy state of K subbands at time slot ℓ 66, 67
$\mathbf{Y}_\ell$	The TF coefficient of $Y[n]$ at time slot ℓ and K dependent subbands 66
$\mathbf{S}_\ell$	The TF coefficient of $S[n]$ at time slot ℓ and K dependent subbands 66
Ξ	The covariance matrix of a PU signal given by $\Xi = \mathbf{E}\left\{\mathbf{S}_\ell \mathbf{S}_\ell^\dagger\right\}$ in a system of K dependent subbands 66
$\mathbf{W}_\ell$	The TF coefficients of $W[n]$ at time slot ℓ and K subbands 67
$\mathbf{U}_\ell$	The TF coefficients of $U[n]$ at time slot ℓ and K subbands 67
$\mathbf{\Psi}_{\ell,k}$	The occupancy state vector of the parent subbands for subband k at time slot ℓ. 67
$\mathbf{Z}_\ell$	The TF coefficients at the output of an ED, at time slot ℓ and K subbands 69
L	Number of contiguous subbands considered as a block in the coupled CHMM prediction scheme 69
$\mathbf{Z}_{\ell,k}$	L-components of the random vector $\mathbf{Z}_\ell$ observed at $k,\ldots,k+L-1$ contiguous subbands and time slot ℓ 69
$\mathbf{Q}_{\ell,k}$	The occupancy states for $k,\ldots,k+L-1$ subbands at time slot ℓ 69
$\mathbf{S}_{\ell,k}$	The TF coefficients representing the L-components of $\mathbf{S}_\ell$ at time slot ℓ 69
Ξ_L	A submatrix of Ξ defining the covariance matrix of $\mathbf{S}_{\ell,k}$ 70
ρ	The correlation coefficients between observations of contiguous subbands in a system of dependent subbands 70
$\tilde{L}_\mathrm{P}$	The number of symbols in the pilot sequence observed under each hypothesis. 70

N	Number of possible occupancy state within a block of L contiguous and dependent subbands 70
$\boldsymbol{\mu}_{r,\mathbf{p}}$	The mean vector of the r^{th} Gaussian component in a GMM scheme conditioned on the vector occupancy state $\mathbf{p}$ 71
$\boldsymbol{\Sigma}_{r,\mathbf{p}}$	The covariance matrix of the r^{th} Gaussian component in a GMM scheme conditioned on the vector occupancy state $\mathbf{p}$ 71
w_r	The weight of the r^{th} Gaussian component in a GMM scheme 71
$a_{q,\boldsymbol{\psi}}(\ell, k)$	The transition probability from the parent occupancy state vector $\boldsymbol{\Psi}_{\ell-1,k} = \boldsymbol{\psi}$ to $Q_{\ell,k} = q$ 71
$a_{q,\boldsymbol{\psi}}(k)$	The transition probability from the parent occupancy state vector $\boldsymbol{\Psi}_k = \boldsymbol{\psi}$ to q 72
$a_{\mathbf{pq}}(\ell)$	The transition probability from the state vector $\mathbf{Q}_{\ell-1,k} = \mathbf{p}$ to $\mathbf{Q}_{\ell,k} = \mathbf{q}$ 72
$\mathbf{q}$	An occupancy state vector, such that $\mathbf{q} \in \mathbf{Q}_{\ell,k}$ 72
$\mathbf{p}$	An occupancy state vector, such that $\mathbf{p} \in \mathbf{Q}_{\ell,k}$ 72
$a_{q,\psi_{\text{R}}}(k)$	The transition probability of the first subband in a block of L contiguous and dependent subbands 72
$a_{q,\psi_{\text{L}}}\left(\tilde{k}\right)$	The transition probability of the last subband in a block of L contiguous and dependent subbands 72
$\boldsymbol{\Psi}_{\text{R}}$	The parent vector occupancy state of the first subband in the block of L contiguous and dependent subbands, which are within the defined block 72
$\boldsymbol{\Psi}_{\text{L}}$	The parent vector occupancy state of the last subband in the block of L contiguous and dependent subbands, which are within the defined block 72
Γ	The set of parameters representing the coupled CHMM based scheme within a block of L contiguous subbands 73
Ω_{CC}	A set which Γ belongs to, such that $\Gamma = \Omega^L$ 73
$\hat{\Gamma}_{\text{ML}}$	A ML estimate of Γ 73

$\alpha_\ell(\mathbf{p})$	The forward quantity in a coupled CHMM based scheme for a block of L contiguous and dependent subbands 74
$\beta_\ell(\mathbf{p})$	The backward quantity in a coupled CHMM based scheme for a block of L contiguous and dependent subbands 74
$\hat{\alpha}_\ell(\mathbf{p})$	An estimate of $\alpha_\ell(\mathbf{p})$ for a certain estimate of Γ 74
$\hat{\beta}_\ell(\mathbf{p})$	An estimate of $\beta_\ell(\mathbf{p})$ for a certain estimate of Γ 74
$\hat{\Gamma}$	An estimate of Γ 74
$\hat{\Gamma}^{(0)}$	An initial estimate of Γ 74
$\hat{a}^{(0)}_{q_k,\psi_{\mathrm{R}}}(k)$	The initial estimate of $a_{q,\psi_{\mathrm{R}}}(k)$ 74
$\hat{a}^{(0)}_{q_{\tilde{k}},\psi_{\mathrm{L}}}\left(\tilde{k}\right)$	The initial estimate of $a_{q,\psi_{\mathrm{L}}}\left(\tilde{k}\right)$ 74
$\hat{a}^{(0)}_{q,\psi}(c)$	The initial estimate of $a_{q,\psi}(c)$, where $c = \{k+1,\ldots,k+L-2\}$ 74
$\hat{a}^{(m)}_{\mathbf{pq}}$	The m^{th} estimate of $a_{\mathbf{pq}}$ 75
$a_{\mathbf{pq}}$	The transition probability from the state vector $\mathbf{p}$ to $\mathbf{q}$, such that $\mathbf{p},\mathbf{q} \in \mathbf{Q}_{\ell,k}$ 75
$\hat{\Gamma}^{(m)}$	An estimate of Γ at the m^{th} iteration 75
m_{it}	Number of iterations required during the learning phase to obtain the best-fit model $\hat{\Gamma}^{(m_{\text{it}})}$ 75
$\hat{\Gamma}^{(m_{\text{it}})}$	A best-fit coupled CHMM scheme as an output of the learning phase after m_{it} number of iterations 75
$\delta_\ell(\mathbf{q})$	A cumulative best score in analogy to $\delta_{\ell,\mathrm{C}}(j)$ for the coupled CHMM based scheme and L contiguous and dependent subbands 75
$\hat{Q}_{\ell,k}$	An estimate of $Q_{\ell,k}$ 76
$\mathbf{A}_{\text{heavy}}$	The value of the transition matrix $\mathbf{A}$ representing a heavy pattern scenario for both uncoupled DHMM and CHMM based schemes 77
$\hat{\mathbf{A}}^{(0)}_{\text{heavy}}$	The initial estimate of $\mathbf{A}_{\text{heavy}}$ 78
$\mathbf{A}_{\text{light}}$	The value of the transition matrix $\mathbf{A}$ representing a light pattern scenario for both uncoupled DHMM and CHMM based schemes 78
$\hat{\mathbf{A}}^{(0)}_{\text{light}}$	The initial estimate of $\mathbf{A}_{\text{light}}$ 78

P_{e}	Probability of error 79
θ	Signal-to-noise-ratio (SNR) 79
ϑ	Signal-to-interference-ratio (SIR) 79
$a_{q,\psi,\mathrm{heavy}}(k)$	The transition probabilities $a_{q,\psi}(k)$ representing a heavy pattern for a system of dependent subbands 83
$a_{q,\psi,\mathrm{light}}(k)$	The transition probabilities $a_{q,\psi}(k)$ representing a light pattern for a system of dependent subbands 84

Bibliography

[1] M. Sharma, D. Dahlhaus, R. Al Halaseh, and N. Mansour, "Interference-aware time-frequency based spectrum sensing for cognitive radio networks," in *Proc. Advances in Wireless and Optical Communications (RTUWO)*, October 2015.

[2] R. Al Halaseh and D. Dahlhaus, "Continuous hidden Markov model based interference-aware cognitive radio spectrum occupancy prediction," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC): Fundamentals and PHY (IEEE PIMRC2016 Fundamentals & PHY)*, pp. 447–452, September 2016.

[3] H. Idoudi, K. Daimi, and M. Saed, "Security challenges in cognitive radio networks," in *Proc. World Congress on Engineering (WCE)*, 2014.

[4] M. Pan, C. Zhang, P. Li, and Y. Fang, "Joint routing and link scheduling for cognitive radio networks under uncertain spectrum supply," in *Proc. IEEE INFOCOM*, pp. 2237–2245, April 2011.

[5] C. F. Chiasserini and R. R. Rao, "Coexistence mechanisms for interference mitigation in the 2.4-GHz ISM band," *IEEE Transactions on Wireless Communications*, vol. 2, pp. 964–975, Sept 2003.

[6] C. S. Inc, "Cisco visual networking index: global mobile data traffic forecast update 2015–2020," 2016.

[7] R. D. Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks: how location information can improve scalability, latency, and robustness of 5G," *IEEE Signal Processing Magazine*, vol. 31, pp. 102–112, Nov 2014.

[8] J. F. Monserrat, G. Mange, V. Braun, H. Tullberg, G. Zimmermann, and Ö. Bulakci, "METIS research advances towards the 5G mobile and wireless

system definition," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–16, 2015.

[9] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, pp. 13–18, Aug 1999.

[10] J. Mitola, *Cognitive radio: an integrated agent architecture for software defined radio*. PhD thesis, Royal Institute of Technology, Sweden, 2000.

[11] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201–220, Feb 2005.

[12] G. Yuan, R. Grammenos, Y. Yang, and W. Wang, "Performance analysis of selective opportunistic spectrum access with traffic prediction," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 1949–1959, May 2010.

[13] X. Xing, T. Jing, W. Cheng, Y. Huo, and X. Cheng, "Spectrum prediction in cognitive radio networks," *IEEE Wireless Communications Magazine*, vol. 20, pp. 90–96, April 2013.

[14] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks Journal (ELSEVIER*, vol. 50, pp. 2127–2159, 2006.

[15] G. Chung, S. Sridharan, S. Vishwanath, and C. S. Hwang, "On the capacity of overlay cognitive radios with partial cognition," *IEEE Transactions on Information Theory*, vol. 58, pp. 2935–2949, May 2012.

[16] G. Bansal, M. Hossain, and V. Bhargava, "Adaptive power loading for OFDM-based cognitive radio systems with statistical interference constraint," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 2786–2791, September 2011.

[17] Y. Sun, W. Xu, and J. Lin, "Trellis shaping based dirty paper coding scheme for the overlay cognitive radio channel," in *Proc. IEEE International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, pp. 1773–1777, Sept 2014.

[18] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: an information theoretic perspective," *Proceedings of the IEEE*, vol. 97, pp. 894–914, May 2009.

[19] J. Oh and W. Choi, "A hybrid cognitive radio system: a combination of under-

lay and overlay approaches," in *Proc. IEEE Vehicular Technology Conference Fall*, pp. 1–5, Sept 2010.

[20] Z. Lin, X. Jiang, L. Huang, and Y. Yao, "A energy prediction based spectrum sensing approach for cognitive radio networks," in *Proc. International Conference on Wireless Communications, Networking and Mobile Computing (WiCom)*, pp. 1–4, Sept 2009.

[21] X. Xing, T. Jing, Y. Huo, H. Li, and X. Cheng, "Channel quality prediction based on Bayesian inference in cognitive radio networks," in *Proc. IEEE INFOCOM*, pp. 1465–1473, April 2013.

[22] J. Yang, H. s. Zhao, X. Chen, J. y. Xu, and J. z. Zhang, "Energy-efficient design of spectrum prediction in cognitive radio networks: prediction strategy and communication environment," in *Proc. International Conference on Signal Processing (ICSP)*, pp. 154–158, Oct 2014.

[23] T. D. Lookabaugh and M. G. Perkins, "Application of the Princen-Bradley filter bank to speech and image compression," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1914–1926, Nov 1990.

[24] H. Feichtinger and T. Strohmer, *Gabor analysis and algorithms: theory and applications*. Austria: Springer Science + Business Media, LLC., 1998.

[25] N. Mansour and D. Dahlhaus, "Interference in DFT modulated filter bank transceivers for cognitive radio," in *Proc. European Wireless Conference*, pp. 1–7, May 2014.

[26] S. M. Phoong, Y. Chang, and C.-Y. Chen, "DFT-modulated filterbank transceivers for multipath fading channels," *IEEE Transactions on Signal Processing*, vol. 53, pp. 182–192, Jan 2005.

[27] Z. Chen, N. Guo, Z. Hu, and R. C. Qiu, "Channel state prediction in cognitive radio, part I: response delays in practical hardware platforms," in *Proc. IEEE Southeastcon*, pp. 45–49, March 2011.

[28] Z. Chen, N. Guo, Z. Hu, and R. C. Qiu, "Channel state prediction in cognitive radio, part II: single-user prediction," in *Proc. IEEE Southeastcon*, pp. 50–54, March 2011.

[29] I. Akbar and W. Tranter, "Dynamic spectrum allocation in cognitive radio using hidden Markov models: Poisson distributed case," in *Proc. IEEE SoutheastCon*, pp. 196–201, March 2007.

[30] S. H. Sohn, S. J. Jang, and J. M. Kim, "HMM-based adaptive frequency-hopping cognitive radio system to reduce interference time and to improve throughput," *KSII Transactions on Internet and Information Systems*, vol. 4, no. 4, pp. 475–490, 2010.

[31] M. Sharma, A. Sahoo, and K. Nayak, "Channel modeling based on interference temperature in underlay cognitive wireless networks," in *Proc. IEEE International Symposium on Wireless Communication Systems (ISWCS '08)*, pp. 224–228, Oct 2008.

[32] V. K. Tumuluru, P. Wang, and D. Niyato, "Channel status prediction for cognitive radio networks," *Wireless Communications and Mobile Computing*, vol. 12, pp. 862–874, July 2012.

[33] V. Tumuluru, P. Wang, and D. Niyato, "A neural network based spectrum prediction scheme for cognitive radio," in *Proc. IEEE International Conference on Communications (ICC)*, pp. 1–5, May 2010.

[34] S. Bai, X. Zhou, and F. Xu, "Spectrum prediction based on improved-back-propagation neural networks," in *Proc. International Conference on Natural Computation (ICNC)*, pp. 1006–1011, Aug 2015.

[35] Z. Zhang, K. Zhang, F. Gao, and S. Zhang, "Spectrum prediction and channel selection for sensing-based spectrum sharing scheme using online learning techniques," in *Proc. IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC'15)*, pp. 355–359, Aug 2015.

[36] Y. Chen and H. S. Oh, "A survey of measurement-based spectrum occupancy modeling for cognitive radios," *IEEE Communications Surveys Tutorials*, vol. 18, pp. 848–859, Firstquarter 2016.

[37] Z. Wang and S. Salous, "Spectrum occupancy statistics and time series models for cognitive radio," *Journal of Signal Processing Systems*, vol. 62, no. 2, pp. 145–155, 2011.

[38] S. Kaneko, S. Nomoto, T. Ueda, S. Nomura, and K. Takeuchi, "Predicting radio resource availability in cognitive radio - an experimental examination," in *Proc. International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, pp. 1–6, May 2008.

[39] Z. Wen, T. Luo, W. Xiang, S. Majhi, and Y. Ma, "Autoregressive spectrum hole prediction model for cognitive radio systems," in *Proc. IEEE International Conference on Communications Workshops*, pp. 154–157, May 2008.

[40] S. Sarkka, A. Vehtari, and J. Lampinen, "Time series prediction by Kalman smoother with cross-validated noise density," in *Proc. IEEE International Joint Conference on Neural Networks*, vol. 2, pp. 1615–1619 vol.2, July 2004.

[41] Z. Ghahramani, "An introduction to hidden Markov models and Bayesian networks," *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)*, vol. 15, no. 1, pp. 9–42, 2001.

[42] C. Ghosh, C. Cordeiro, D. Agrawal, and M. Rao, "Markov chain existence and hidden Markov models in spectrum sensing," in *Proc. IEEE International Conference on Pervasive Computing and Communications*, pp. 1–6, March 2009.

[43] S. D. Barnes and B. T. Maharaj, "Performance of a hidden Markov channel occupancy model for cognitive radio," in *Proc. IEEE AFRICON, 2011*, pp. 1–6, Sept 2011.

[44] T. Black, B. Kerans, and A. Kerans, "Implementation of hidden Markov model spectrum prediction algorithm," in *Proc. International Symposium on Communications and Information Technologies (ISCIT)*, pp. 280–283, Oct 2012.

[45] S. Barnes and B. Maharaj, "Prediction based channel allocation performance for cognitive radio," *International Journal of Electronics and Communications (AEU)*, vol. 68, no. 4, pp. 336–345, 2014.

[46] S. Theodoridis and K. Koutroumbas, *Pattern recognition*. Elsevier Inc., 4 ed., 2009.

[47] Y. K. Ham and R.-H. Park, "3D object recognition in range images using hidden Markov models and neural networks," *Pattern Recognition*, vol. 32, no. 5, pp. 729 – 742, 1999.

[48] C.-L. Huang, M.-S. Wu, and S.-H. Jeng, "Gesture recognition using the multi-PDM method and hidden Markov model," *Image and Vision Computing*, vol. 18, no. 11, pp. 865 – 879, 2000.

[49] M. Haenggi and R. K. Ganti, "Interference in large wireless networks," *Foundations and Trends in Networking*, vol. 3, pp. 127–248, Feb. 2009.

[50] J. D. Cryer and K.-S. Chan, *Time series analysis with applications in R*. Springer texts in statistics, USA: Springer Science + Business Media, LLC, 2 ed., 2008.

[51] D. Barber, *Bayesian reasoning and machine learning*. New York, NY, USA: Cambridge University Press, 2012.

[52] C. M. Bishop, *Pattern recognition and machine learning*. Springer, New York, 2006.

[53] J. L.Gross and J. Yellen, *Handbook of graph theory*. Discrete mathematics and applications, CRC Press, 2004.

[54] K. Murphy, *Dynamic Bayesian networks: representation, inference and learning*. PhD thesis, University of California, Berkeley, 2002.

[55] K. P. Murphy, *Machine learning: a probabilistic perspective*. Adaptive computation and machine learning series, England: Massachusetts Institute of Technology (MIT) Press, 2012.

[56] M. Kijima, *Markov processes for stochastic modeling*. Springer, 1 ed., 1997.

[57] M. Brand, N. Oliver, and A. Pentland, "Coupled hidden Markov models for complex action recognition," in *Proc. Computer Vision and Pattern Recognition (CVPR)*, CVPR '97, (Washington, DC, USA), pp. 994–999, IEEE Computer Society, 1997.

[58] Federal Communications Commission (FCC). `https://www.fcc.gov/general/rules-regulations-title-47`, Last accessed: 15 Nov. 2015.

[59] S. Farahani, *ZigBee wireless networks and transceivers*. Adaptive computation and machine learning series, England: Elsevier Ltd., 2008.

[60] J. Neburka, Z. Tlamsa, V. Benes, L. Polak, O. Kaller, L. Klozar, L. Bolecek, O. Zach, J. Kufa, J. Sebesta, and T. Kratochvil, "Study of the coexistence between ZigBee and Wi-Fi IEEE 802.11b/g networks in the ISM band," in *Proc. International Conference Radioelektronika*, pp. 106–109, April 2015.

[61] I. Howitt and J. A. Gutierrez, "IEEE 802.15.4 low rate - wireless personal area network coexistence issues," in *Proc. IEEE Wireless Communications and Networking (WCNC)*, vol. 3, pp. 1481–1486 vol.3, March 2003.

[62] S. Pollin, I. Tan, B. Hodge, C. Chun, and A. Bahai, "Harmful coexistence between 802.15.4 and 802.11: a measurement-based study," in *Proc. International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, pp. 1–6, May 2008.

[63] A. Sikora and V. F. Groza, "Coexistence of IEEE802.15.4 with other systems in the 2.4 GHz-ISM-Band," in *Proc. IEEE Instrumentation and Measurement Technology*, vol. 3, pp. 1786–1791, May 2005.

[64] . W. Group, "Coexistence analysis of IEEE Std 802.15.4 with other IEEE standards and proposed standards," Tech. Rep. IEEE P802.15-10-0808-00, IEEE P802.15 Working Group for Wireless Personal Area Networks (WPANs), September 2010.

[65] A. S. Tanenbaum and D. J. Wetherall, *Computer networks*. Pearson Education, Inc., 5 ed., 2011.

[66] A. F. Molisch, *Wireless communications*. England: John Wiley & Sons, Ltd, 2005.

[67] M. Petrova, L. Wu, P. Mahonen, and J. Riihijarvi, "Interference measurements on performance degradation between colocated IEEE 802.11g/n and IEEE 802.15.4 networks," in *Proc. International Conference on Networking (ICN)*, pp. 93–93, April 2007.

[68] N. Golmie and F. Mouveaux, "Interference in the 2.4 GHz ISM band: impact on the Bluetooth access control performance," in *Proc. IEEE International Conference on Communications (ICC)*, vol. 8, pp. 2540–2545 vol.8, 2001.

[69] FieldComm Group. `http://en.hartcomm.org/index.html`, Last accessed: 07 May 2016.

[70] A. Limited, "HART communications protocol- what's new in WirelessHART?," 2009.

[71] T. Lennvall, S. Svensson, and F. Hekland, "A comparison of WirelessHART and ZigBee for industrial applications," in *Proc. IEEE International Workshop on Factory Communication Systems (WFCS)*, pp. 85–88, May 2008.

[72] I. C. Society, "IEEE Draft Amendment to IEEE Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANS): Amendment to Add Alternate PHY," *IEEE Std P802.15.4a/D6, Dec 2006*, 2006.

[73] C. M. D. Dominicis, P. Ferrari, A. Flammini, E. Sisinni, M. Bertocco, G. Giorgi, C. Narduzzi, and F. Tramarin, "Investigating WirelessHART coexistence issues through a specifically designed simulator," in *Proc. IEEE Instrumentation and Measurement Technology Conference, 2009 (I2MTC '09)*, pp. 1085–1090, May 2009.

[74] T. Machado, I. Muller, J. Winter, V. Dickow, C. E. Pereira, and J. C. Netto,

"WirelessHART network analyzer with coexistence detection," in *Proc. IEEE International Conference on Industrial Informatics (INDIN)*, pp. 696–701, July 2014.

[75] D. Yang, Y. Xu, and M. Gidlund, "Wireless coexistence between IEEE 802.11- and IEEE 802.15.4- based networks: a survey," *International Journal of Distributed Sensor Networks*, vol. 2011, 2010.

[76] L. Angrisani, M. Bertocco, D. Fortin, and A. Sona, "Experimental study of coexistence issues between IEEE 802.11b and IEEE 802.15.4 wireless networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, pp. 1514–1523, Aug 2008.

[77] K. Shuaib, M. Boulmalf, F. Sallabi, and A. Lakas, "Co-existence of Zigbee and WLAN, a performance study," in *Proc. Wireless Telecommunications Symposium*, pp. 1–6, April 2006.

[78] N. Golmie, N. Chevrollier, and O. Rebala, "Bluetooth and WLAN coexistence: challenges and solutions," *IEEE Wireless Communications*, vol. 10, pp. 22–29, Dec 2003.

[79] M. B. Shoemake, "Wi-Fi (IEEE 802.11b) and Bluetooth coexistence issues and solutions for the 2.4 GHz ISM band," 2001.

[80] C. F. Chiasserini and R. R. Rao, "Performance of IEEE 802.11 WLANs in a Bluetooth environment," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 1, pp. 94–99 vol.1, 2000.

[81] W. Feller, *An introduction to probability theory and its applications, V II*. John Wiley & Sons, 2 ed., 1996.

[82] E. E. Kuruoglu, *Signal processing in α-stable noise environments: a least l_p-norm approach*. PhD thesis, University of Cambrigde, UK, 1998.

[83] J. Durbin and S. J. Koopman, *Time series analysis by state space methods*. Oxford statistical science series, 38, Oxford University Press, 2 ed., 2012.

[84] R. K. Ganti and M. Haenggi, "Interference in ad hoc networks with general motion-invariant node distributions," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, pp. 1–5, July 2008.

[85] J. Ilow and D. Hatzinakos, "Analytic alpha-stable noise modeling in a Poisson field of interferers or scatterers," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1601–1611, Jun 1998.

[86] M. Z. Win, P. C. Pinto, and L. A. Shepp, "A mathematical theory of network interference and its applications," *Proceedings of the IEEE*, vol. 97, pp. 205–230, Feb 2009.

[87] S. A. Kassam, *Signal detection in non-Gaussian noise*. Springer texts in electrical engineering, Springer-Verlag, 1988.

[88] C. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*. Adaptive and learning systems for signal processing, communications and control series, Austria: Wiley-Interscience, 1995.

[89] M. Shao and C. L. Nikias, "Signal detection in impulsive noise based on stable distributions," in *Proc. Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pp. 218–222 vol.1, Nov 1993.

[90] M. Bouvet and S. C. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 621–626, May 1989.

[91] P. Mertz, "Model of impulsive noise for data transmission," *IRE Transactions on Communications Systems*, vol. 9, pp. 130–137, June 1961.

[92] B. W. Stuck and B. Kleiner, "A statistical analysis of telephone noise," *The Bell System Technical Journal*, vol. 53, pp. 1263–1320, Sept 1974.

[93] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 11, pp. 991–1001, Sep 1993.

[94] M. Haenggi, "Outage and throughput bounds for stochastic wireless networks," in *Proc. International Symposium on Information Theory, 2005*, pp. 2070–2074, Sept 2005.

[95] P. Cardieri, "Modeling interference in wireless ad hoc networks," *IEEE Communications Surveys Tutorials*, vol. 12, pp. 551–572, Fourth 2010.

[96] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 3 ed., 2013.

[97] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, pp. 1029–1046, September 2009.

[98] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic geometry and its applications*. Wiley series in probability and statistics, John Wily & Sons. Ltd, 3 ed., 2013.

[99] J. D. Gibson, *Mobile communications handbook*. Taylor & Francis Group, 2013.

[100] A. Busson and G. Chelius, "Point processes for interference modeling in CSMA/CA ad-hoc networks," in *Pro. ACM Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks*, PE-WASUN '09, (New York, NY, USA), pp. 33–40, ACM, 2009.

[101] J. Venkataraman, M. Haenggi, and O. Collins, "Shot noise models for outage and throughput analyses in wireless ad hoc networks," in *IEEE Military Communications Conference (MILCOM)*, pp. 1–7, Oct 2006.

[102] J. Andrews, R. Ganti, M. Haenggi, N. Jindal, and S. Weber, "A primer on spatial modeling and analysis in wireless networks," *IEEE Communications Magazine*, vol. 48, pp. 156–163, November 2010.

[103] X. Yang and A. P. Petropulu, "Joint statistics of interference in a wireless communications link resulted from a poisson field of interferers," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, pp. II–1749–II–1752, May 2002.

[104] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An Aloha protocol for multihop mobile wireless networks," *IEEE Transactions on Information Theory*, vol. 52, pp. 421–436, Feb 2006.

[105] F. Baccelli, B. Blaszczyszyn, and P. Mühlethaler, "A spatial reuse aloha MAC protocol for multihop wireless mobile networks," Research Report RR-4955, INRIA, 2003.

[106] S. Srinivasa and M. Haenggi, "Distance distributions in finite uniformly random networks: theory and applications," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 940–949, Feb 2010.

[107] R. K. Ganti and M. Haenggi, "Regularity, interference, and capacity of large ad hoc networks," in *Proc. Asilomar Conference on Signals, Systems and Computers*, pp. 3–7, Oct 2006.

[108] A. Busson, G. Chelius, and J.-M. Gorce, "Interference modeling in CSMA multi-hop wireless networks," Tech. Rep. 6624, Institut d'électronique fonda-

mentale - IEF , ARES - CITI Insa Lyon / INRIA Grenoble Rhône-Alpes, 2009.

[109] B. Boashash, *Time- frequency signal analysis and processing: a comprehensive reference*. UK: ELSEVIER, 2003.

[110] F. Hlawatsch and G. Boudreaux-bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, April 1992.

[111] Z. Ju, *A filter bank based reconfigurable receiver architecture for universal wireless communications*. PhD thesis, University of Kassel, 2010.

[112] H. Bölcskei, F. Hlawatsch, and H. G. Feichtinger, "Equivalence of DFT filter banks and Gabor expansions," in *Proc. Wavelet Applications in Signal and Image Processing III (SPIE)*, vol. 2569, pp. 128–139, 1995.

[113] Y. Zhang, K. Yang, and M. G. Amin, "Adaptive array processing for multipath fading mitigation via exploitation of filter banks," *IEEE Transactions on Antennas and Propagation*, vol. 49, pp. 505–516, Apr 2001.

[114] S. M. Phoong, Y. Chang, and C.-Y. Chen, "DFT-modulated filterbank transceivers for multipath fading channels," *IEEE Transactions on Signal Processing*, vol. 53, pp. 182–192, Jan 2005.

[115] T. Hunziker, U. Rehman, and D. Dahlhaus, "Spectrum sensing in cognitive radios: Design of DFT filter banks achieving maximal time-frequency resolution," in *Proc. International Conference on Information, Communications and Signal Processing (ICICS)*, pp. 1–5, Dec 2011.

[116] B. Farhang-Boroujeny, "Filter bank spectrum sensing for cognitive radios," *IEEE Transactions on Signal Processing*, vol. 56, pp. 1801–1811, May 2008.

[117] D. Ariananda, M. K. Lakshmanan, and H. Nikookar, "A survey on spectrum sensing techniques for cognitive radio," in *Proc. International Workshop on Cognitive Radio and Advanced Spectrum Management (CogART)*, pp. 74–79, May 2009.

[118] P. Amini, R. Kempter, and B. Farhang-Boroujeny, "A comparison of alternative filterbank multicarrier methods for cognitive radio systems," in *Proc. SDR Technical Conference and Product Exposition*, 2006.

[119] A. Saad, N. Mansour, A. Friedrich, Z. Youssef, D. Dahlhaus, M. Sharma, R. A. Halaseh, E. Majeed, K. D. Kohrt, G. Bruck, R. Knorr, and P. Jung, "Cognitive

radio prototype for industrial applications," in *Proc. 22th European Wireless Conference*, pp. 1–8, May 2016.

[120] L. B. Le, V. Lau, E. Jorswieck, N.-D. Dao, A. Haghighat, D. I. Kim, and T. Le-Ngoc, "Enabling 5G mobile wireless technologies," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–14, 2015.

[121] N. Mansour and D. Dahlhaus, "Post detection integration in DFT modulated filter bank transceivers for cognitive radio," in *Proc. IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), 2015*, pp. 14–18, May 2015.

[122] by Vladimir M. Zolotarev and V. V. Uchaikin, *Chance and stability, stable distributions and their applications*. Modern Probability and Statistics, De Gruyter, 2 ed., 2011.

[123] P. Tsakalides, *Array signal processing with alpha-stable distributions*. PhD thesis, University of Southern California, USA, Dec. 1995.

[124] Y. Chen and J. Chen, "Novel Sα_{int}S PDF approximations and their applications in wireless signal detection," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 1080–1091, Feb 2015.

[125] A. Webb and K. D. Copsey, *Statistical pattern recognition*. Pattern perception - statistical methods, John Wiley & Sons, Ltd, 3 ed., 2011.

[126] K. Ehsan and D. Dahlhaus, "Statistical modeling of ISM data traffic in indoor environments for cognitive radio systems," in *Proc. International Conference on Digital Information, Networking, and Wireless Communications (DINWC)*, pp. 1–6, Feb. 2015.

[127] S. Geirhofer, L. Tong, and B. M. Sadler, "A measurement-based model for dynamic spectrum access in wlan channels," in *Proc. IEEE Military Communications Conference, 2006. MILCOM 2006*, pp. 1–7, Oct 2006.

[128] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–286, Feb 1989.

[129] H. V. Poor, *An introduction to signal detection and estimation*. New York, NY, USA: Springer-Verlag New York, Inc., 2 ed., 1994.

[130] R. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. Wiley & Sons, Inc., 2 ed., 2001.

[131] N. Ghanmy, M. A. Mahjoub, and N. E. B. Amara, "Characterization of dy-

namic Bayesian network," *International Journal of Advanced Computer Science and Applications, (IJACSA)*, vol. 2, no. 7, 2011.

[132] J. Adibi and W.-M. Shen, "Self-similar layered hidden Markov models," in *Proc. European Conference on Principles of Data Mining and Knowledge Discovery*, PKDD '01, (London, UK), pp. 1–15, Springer-Verlag, 2001.

[133] T. Nguyeny, B. L. Mark, and Y. Ephraim, "Hidden Markov process based dynamic spectrum access for cognitive radio," in *Proc. Conference on Information Sciences and Systems (CISS)*, pp. 1–6, March 2011.

[134] S. V. Vaseghi, *Advanced digital signal processing and noise reduction.* John Wiley & Sons, Ltd, 4 ed., 2008.

[135] B. H. F. Juang, "On the hidden Markov model and dynamic time warping for speech recognition– a unified view," *AT & T Bell Laboratories Technical Journal*, vol. 63, pp. 1213–1243, Sept 1984.

[136] A. V. Nefian, L. Liang, X. Pi, X. Liu, and K. Murphy, "Dynamic Bayesian networks for audio-visual speech recognition," *EURASIP Journal on Applied Signal Processing*, vol. 2002, pp. 1274–1288, January 2002.

[137] A. Nefian, L. Liang, X. Pi, L. Xiaoxiang, C. Mao, and K. Murphy, "A coupled HMM for audio-visual speech recognition," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, pp. II 2013–II 2016, May 2002.

[138] K. M. J. Kwon, "Modeling freeway traffic with coupled HMMs," Tech. Rep. Technical Report 405, University of California Berkeley, May 2000.

[139] S. Darmanjian, S.-P. Kim, M. C. Nechyba, J. Principe, J. Wessberg, and M. A. L. Nicolelis, "Independently coupled HMM switching classifier for a bimodel brain-machine interface," in *Proc. IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing*, pp. 379–384, 2006.

[140] N. M. Ghahjaverestan, S. Masoudi, M. B. Shamsollahi, A. Beuchée, P. Pladys, D. Ge, and A. I. Hernández, "Coupled hidden Markov model-based method for apnea bradycardia detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, pp. 527–538, March 2016.

[141] A. Larmo, M. Lindström, M. Meyer, G. Pelletier, J. Torsner, and H. Wiemann, "The LTE link-layer design," *IEEE Communications Magazine*, vol. 47, pp. 52–59, April 2009.

[142] I. Rezek and S. J. Roberts, "Estimation of coupled hidden Markov models with application to biosignal interaction modelling," in *Proc. IEEE International Conference on Neural Network for Signal Processing, 2002*, pp. 804–813, 2002.

[143] L. Xie and Z.-Q. Liu, "A coupled HMM approach to video-realistic speech animation," *Pattern Recognition*, vol. 40, no. 8, pp. 2325–2340, 2007.

[144] C. Clancy, J. Hecker, E. Stuntebeck, and T. O'Shea, "Applications of machine learning to cognitive radio networks," *IEEE Wireless Communications*, vol. 14, pp. 47–52, August 2007.

[145] A. Araujo, J. Blesa, E. Romero, and D. Villanueva, "Security in cognitive wireless sensor networks. Challenges and open problems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–8, 2012.

[146] A. Erramilli, M. Roughan, D. Veitch, and W. Willinger, "Self-similar traffic and network dynamics," *Proceedings of the IEEE*, vol. 90, pp. 800–819, May 2002.

[147] T. Karagiannis, M. Molle, and M. Faloutsos, "Long-range dependence ten years of Internet traffic modeling," *IEEE Internet Computing*, vol. 8, pp. 57–64, Sept 2004.

[148] D. Yamkhin and Y. Won, "Modeling and analysis of wireless LAN traffic," *Journal of Information Science and Engineering*, no. 25, pp. 1783–1801, 2009.

[149] W. Willinger, V. Paxson, and M. S. Taqqu, "Self-similarity and heavy tails: structural modeling of network traffic," in *A practical guide to heavy tails: statistical, techniques and applications* (R. J. Adler, R. E. Feldman, and M. S. Taqqu, eds.), pp. 27–53, Cambridge, MA, USA: Birkhauser Boston Inc., 1998.

[150] A. Dainotti, A. Pescapé, P. S. Rossi, F. Palmieri, and G. Ventre, "Internet traffic modeling by means of hidden Markov models," *Computer Networks*, vol. 52, no. 14, pp. 2645–2662, 2008.

[151] M. Gani, H. Sarwar, and C. M. Rahman, "Prediction of state of wireless network using Markov and hidden Markov model," *Journal of Networks*, vol. 4, no. 10, 2009.

[152] T. Kurasaugi, K. Kobayashi, and Y. Takahashi, "A Markovian model of coded video traffic which exhibits long-range dependence in statistical analysis," *Journal of the Operations Research Society of Japan*, vol. 42, no. 1, 1998.

[153] J. Adibi and W.-M. Shen, "Self-similar layered hidden Markov models," in *Proc. European Conference on Principles of Data Mining and Knowledge Discovery*, (London, UK, UK), pp. 1–15, Springer-Verlag, 2001.

[154] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in the time domain: modeling and exploiting white space," *IEEE Communications Magazine*, 2007.

[155] A. Feldmann and W. Whitt, "Fitting mixtures of exponentials to long-tail distributions to analyze network performance models," *Performance Evaluation*, vol. 31, no. 3-4, pp. 245–279, 1998.

[156] S. Robert and J.-Y. L. Boudec, "On a Markov modulated chain exhibiting self-similarities over finite timescale," *Performance Evaluation*, vol. 27, pp. 159–173, 1996.

[157] A. T. Andersen and B. F. Nielsen, "A Markovian approach for modeling packet traffic with long-range dependence," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 5, pp. 719–732, 2006.

[158] I.-H. Li, "Modeling long-range dependent VBR traffic using synthetic Markov-Gaussian TES models," in *Proc. International Conference on Next Generation Teletraffic and Wired/Wireless Advanced Networking*, Springer, 2008.

[159] K. Salamatian and S. Vaton, "Hidden Markov modeling for network communication channels," *ACM SIGMETRICS Performance Evaluation Review*, vol. 29, no. 1, pp. 92–101, 2001.

[160] P. S. Rossi, G. Romano, F. Palmieri, and G. Iannello, "Joint end-to-end loss-delay hidden Markov model for periodic UDP traffic over the Internet," *IEEE Transactions on Signal Processing*, vol. 54, pp. 530–541, Feb 2006.

[161] G. Iannello, F. Palmieri, A. Pescape, and P. S. Rossi, "End-to-end packet-channel Bayesian model applied to heterogeneous wireless networks," in *Proc. IEEE Global Telecommunications Conference, (GLOBECOM)*, vol. 1, pp. 484–489, Dec 2005.

[162] M. T. Masonta, M. Mzyece, and N. Ntlatlapa, "Spectrum decision in cognitive radio networks: a survey," *IEEE Communications Surveys Tutorials*, vol. 15, pp. 1088–1107, Third 2013.

[163] V. T. Nguyen, F. Villain, and Y. L. Guillou, "Cognitive radio RF: overview and challenges," *VLSI Design*, vol. 2012, 2012.

ISBN 978-3-7376-0552-6