Exploiting Feedback Information in Cognitive Radio Systems

Author: Sara Abozeid Attila
Supervisor: Dr. Karim Seddik
          Dr. Amr Elsherif
          Dr. Sherif Rabia

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electronics and Communications Engineering

January, 2020
Declaration of Authorship

I, Sara Abozeid Attlla, declare that this thesis titled, “Exploiting Feedback Information in Cognitive Radio Systems” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

____________________________________

Date:

____________________________________
Abstract

School of Sciences and Engineering
Electronics and Communications Engineering Department

Doctor of Philosophy

Exploiting Feedback Information in Cognitive Radio Systems

by Sara Abozeid Attila

In this thesis, we consider a cognitive radio (CR) network where the primary network’s feedback information is utilized to design access schemes for the secondary network, to exploit the underutilized primary spectrum resources. Secondary users (SUs) identify the spectrum opportunities by sensing the spectrum for primary users (PUs) activities and by listening to the PUs feedback. The feedback signals monitored in this research work are the channel quality indicator (CQI) and automatic repeat request (ARQ) available in the PUs network. For detecting the PUs activities, SUs employ hard/soft energy sensing. The secondary access decisions are optimized to maximize the SUs service rate while maintaining the PUs queues’ stability. The proposed systems are modeled as multi-dimensional Markov chains (MCs) that capture the number of packets in the PUs queues and the state of the one/two observed PUs’ feedback signals. The performance of the proposed systems are evaluated by deriving the SUs service rate and the average PUs packet delay. We compare the performance of the proposed systems with other baseline systems utilizing different types of PUs’ feedback signals. Results reveal the improvement in the SUs service rate and the PUs’ delay of the proposed systems compared to the baseline systems. This improvement is mainly due to the fact that in our proposed system SUs have access to extra information, in terms of PUs feedback, as compared to
other systems. Therefore, SUs in our proposed systems can have better inference on the PUs’ activities; thus more collisions between the PUs and the SUs can be avoided, resulting in significant performance gains in terms of SUs’ throughput and PUs’ average delay. Finally, we propose a multi-layer perceptron (MLP) Q-learning based spectrum access scheme for cognitive radio networks. In the proposed scheme, SUs overhear the ARQ feedback available in the PUs’ network and exploit it to learn the PUs behavior. Since the SUs observe only the PUs’ ARQ feedback and have no information about the PUs packet arrival rates or the states of their queues, the system is modeled as a partially observable Markov decision process (POMDP). The proposed MLP Q-learning access scheme is used to solve this POMDP and find the best SUs’ actions (channel access probabilities) based on the observed PUs’ ARQ feedback and past experiences. The performance of the proposed scheme is shown to be on par with that of other feedback-based access schemes, with the added strength of only having partial information about the PUs and the primary network.
I would like to thank Allah for completing this work. Without Allah’s help, this thesis would not have finalized.

I also would like to thank my family for always being supportive during the whole time in my life. I am deeply grateful to my husband who supports me to finish my PhD. degree. This acknowledgement would be incomplete without thanking my baby, Hamza.

I would give a special thanks to my thesis advisors, Dr. Karim Seddik, Dr. Amr Elsherif, and Dr. Sherif Rabia. They gave me technical guidance during my PhD. work.

Contents

Declaration of Authorship iii
Acknowledgements vii
List of Figures xiii
List of Tables xvii

1 Introduction 1
  1.1 Cognitive radio ............................................. 1
  1.2 Cognitive radio Cycle ..................................... 2
  1.3 Spectrum management framework ......................... 3
    1.3.1 Spectrum sensing ...................................... 4
    1.3.2 Spectrum decision .................................... 5
    1.3.3 Spectrum sharing ..................................... 6
    1.3.4 Spectrum mobility .................................... 6
  1.4 Dynamic spectrum access schemes ....................... 6
  1.5 PU feedbacks ............................................... 7
    1.5.1 Automatic repeat request (ARQ) ....................... 7
    1.5.2 Channel quality indicator (CQI) ..................... 8
  1.6 Learning in CR network ................................... 9
  1.7 Organization of the thesis ................................ 10
  1.8 Thesis contributions ..................................... 11

2 CQI feedback based access system 13
  2.1 Introduction .............................................. 13
  2.2 CQI feedback- HD based access system ................. 14
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1</td>
<td>System model</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Channel model</td>
<td>16</td>
</tr>
<tr>
<td>2.2.2</td>
<td>SU spectrum access schemes</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>The baseline systems</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Proposed CQI feedback HD-based access system</td>
<td>18</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Performance analysis</td>
<td>18</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Performance results</td>
<td>31</td>
</tr>
<tr>
<td>2.3</td>
<td>CQI feedback-soft detection (SD) based access system</td>
<td>34</td>
</tr>
<tr>
<td>2.3.1</td>
<td>System model</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Channel model</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Spectrum soft-sensing scheme</td>
<td>35</td>
</tr>
<tr>
<td>2.3.2</td>
<td>SU’s spectrum access schemes</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>The baseline systems</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>The CQI feedback-SD based access system</td>
<td>38</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Performance analysis</td>
<td>38</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Multi-user case</td>
<td>44</td>
</tr>
<tr>
<td>2.3.5</td>
<td>Effect of CQI erasures at the SUs</td>
<td>48</td>
</tr>
<tr>
<td>2.3.6</td>
<td>Performance results</td>
<td>49</td>
</tr>
<tr>
<td>3</td>
<td>Hybrid feedback-based access scheme</td>
<td>59</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>59</td>
</tr>
<tr>
<td>3.2</td>
<td>ARQ-CQI feedback-hard detection (HD) based access scheme</td>
<td>59</td>
</tr>
<tr>
<td>3.2.1</td>
<td>System model</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Channel model</td>
<td>59</td>
</tr>
<tr>
<td>3.2.2</td>
<td>SU spectrum access schemes</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>The No feedback system</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>The ARQ feedback-based system</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>The CQI feedback-based system</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Proposed hybrid feedback HD-based access system</td>
<td>61</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Performance analysis</td>
<td>62</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Performance results</td>
<td>70</td>
</tr>
<tr>
<td>3.3</td>
<td>ARQ-CQI Feedback-Soft detection (SD) based access scheme</td>
<td>73</td>
</tr>
</tbody>
</table>
3.3.1 System model ........................................... 74
Channel model ........................................... 74
3.3.2 SU spectrum access schemes ....................... 75
3.3.3 Performance analysis ................................. 77
3.3.4 Performance results ................................. 84

4 Learning to communicate with multi-agent reinforcement learning 93
4.1 Introduction ............................................ 93
4.2 Background ............................................ 94
  4.2.1 Reinforcement learning (RL) ...................... 94
  4.2.2 Partial observability ............................ 98
4.3 System model ........................................... 98
  4.3.1 Channel Access Model ............................ 98
  4.3.2 The POMDP model ............................. 99
4.4 RL architecture and implementation using MLP .... 102
4.5 Numerical results ................................... 104

5 Conclusions and future work ............................ 109
5.1 Conclusions ........................................... 109
5.2 Future directions ................................... 110

Bibliography ............................................ 111
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Cognitive radio cycle</td>
<td>3</td>
</tr>
<tr>
<td>2.1</td>
<td>The system model</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>The channel model</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>The PU queue MC model for the two-state channel model</td>
<td>19</td>
</tr>
<tr>
<td>2.4</td>
<td>The PU queue MC model for the perfect sensing CQI feedback based-access system (in this figure, the notation $\bar{q}$ is used to denote $1 - q$)</td>
<td>28</td>
</tr>
<tr>
<td>2.5</td>
<td>The SU throughput for different access schemes ($p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.9, p_f = 0.1$)</td>
<td>30</td>
</tr>
<tr>
<td>2.6</td>
<td>The SU throughput for different access schemes ($p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1$)</td>
<td>30</td>
</tr>
<tr>
<td>2.7</td>
<td>The SU optimal access probabilities for different access schemes ($p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.9, p_f = 0.1$)</td>
<td>31</td>
</tr>
<tr>
<td>2.8</td>
<td>The SU optimal access probabilities for different access schemes ($p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1$)</td>
<td>32</td>
</tr>
<tr>
<td>2.9</td>
<td>The PU packet delay for different access schemes ($p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.9, p_f = 0.1$)</td>
<td>33</td>
</tr>
<tr>
<td>2.10</td>
<td>The PU packet delay for different access schemes ($p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1$)</td>
<td>34</td>
</tr>
<tr>
<td>2.11</td>
<td>Soft decision sensing scheme [10]</td>
<td>36</td>
</tr>
<tr>
<td>2.12</td>
<td>The binary erasure channel model</td>
<td>48</td>
</tr>
<tr>
<td>2.13</td>
<td>The SU throughput for different access schemes ($M_p = 1, M_s = 1, p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1$)</td>
<td>50</td>
</tr>
<tr>
<td>2.14</td>
<td>The SU throughput for different access schemes ($M_p = 1, M_s = 0.4, p_G = \zeta_g = 0.6, p_d = 0.9, p_f = 0.1$)</td>
<td>50</td>
</tr>
</tbody>
</table>
2.15 The SU throughput for different access schemes \((M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.7, p_f = 5.8717 \times 10^{-4})\).

2.16 The SU optimal access probabilities for the PU CQI Feedback-SD based access system \((M_p = 1, M_s = 1, p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1)\).

2.17 The SU optimal access probabilities for the PU CQI Feedback-SD based access system \((M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.9, p_f = 0.1)\).

2.18 The SU optimal access probabilities for the PU CQI Feedback-SD based access system \((M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.7, p_f = 5.8717 \times 10^{-4})\).

2.19 The PU delay for different access schemes \((M_p = 1, M_s = 1, p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1)\).

2.20 The PU delay for different access schemes \((M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.9, p_f = 0.1)\).

2.21 The PU delay for different access schemes \((M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.7, p_f = 5.8717 \times 10^{-4})\).

2.22 The SU throughput for erasure effect \((M_p = 1, M_s = 1, p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1, p_e = 0.1)\).

2.23 The SU throughput for erasure effect \((M_p = 1, M_s = 1, p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1, p_e = 0.3)\).

3.1 The system model.

3.2 The PU queue MC model for ARQ-CQI feedback system.

3.3 The SU throughput for different access schemes \((p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.5, p_f = 0.1)\).

3.4 The SU throughput for different access schemes \((p_B = 0.3, \zeta_B = 0.6364, p_G = 0.6, \zeta_g = 0.6364, p_d = 0.4, p_f = 0.3)\).

3.5 The SU optimal access probabilities for different access schemes \((p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.5, p_f = 0.1)\).

3.6 The SU optimal access probabilities for different access schemes \((p_B = 0.3, \zeta_B = 0.6364, p_G = 0.6, \zeta_g = 0.6364, p_d = 0.4, p_f = 0.3)\).
3.7 The system model. ................................................. 76
3.8 The SU throughput for different access schemes ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.7$, $p_f = 5.7772 \times 10^{-4}$, $r_{PS} = 100$, $M_s = 2$, $M_p = 2$). ................................................. 84
3.9 The SU throughput for different access schemes ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.7$, $p_f = 0.6712$, $r_{PS} = 400$, $M_s = 2$, $M_p = 2$). ................................................. 85
3.10 The SU throughput for different access schemes ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.6$, $p_f = 0.5649$, $r_{PS} = 400$, $M_s = 2$, $M_p = 2$). ................................................. 85
3.11 The SU throughput for different access schemes ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.7$, $p_f = 0.6712$, $r_{PS} = 400$, $M_s = 3$, $M_p = 2$). ................................................. 86
3.12 The SU throughput for different access schemes ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.7$, $p_f = 0.6712$, $r_{PS} = 400$, $M_s = 4$, $M_p = 2$). ................................................. 86
3.13 The SU optimal access probabilities for the Hybrid Feedback-SD based access system ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.7$, $p_f = 5.7772 \times 10^{-4}$, $r_{PS} = 100$, $M_s = 2$, $M_p = 2$). ................................................. 87
3.14 The SU optimal access probabilities for the Hybrid Feedback-SD based access system ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.7$, $p_f = 0.6712$, $r_{PS} = 400$, $M_s = 2$, $M_p = 2$). ................................................. 88
3.15 The PU delay for different access schemes ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.7$, $p_f = 5.7772 \times 10^{-4}$, $r_{PS} = 100$, $M_s = 2$, $M_p = 2$). ................................................. 89
3.16 The PU delay for different access schemes ($p_B = \zeta_B = 0.3$, $p_G = \zeta_G = 0.7$, $p_d = 0.7$, $p_f = 0.6712$, $r_{PS} = 400$, $M_s = 2$, $M_p = 2$). ................................................. 90
4.1 The PU queue MC model [44]. ................................................. 99
4.2 RL architecture using MLP. ................................................. 103
4.3 RL model using MLP summary. ................................................. 104
4.4 SU throughput for different PU feedback history length ($L$) without a spectrum sensing, ($M_s = 1$, $M_p = 1$) ................................................. 105
4.5 SU throughput (per user) for different spectrum access schemes

\( (M_s = 2, M_p = 2, p_d = 0.9, p_f = 0.1, L = 1) \). . . . . . . . . . . . 105

4.6 SU throughput (per user) for different spectrum access schemes

\( (M_s = 2, M_p = 2, p_d = 0.6, p_f = 0.3, L = 1) \). . . . . . . . . . . . 106
List of Tables

2.1 Transition probabilities of the MC of the CQI feedback- (HD) system. ............................................. 19
2.2 Transition probabilities of the MC of the CQI feedback- (SD) system. ............................................. 37
2.3 Transition probabilities of the MC of the multiple SU CQI system. ................................................. 45
3.1 Transition probabilities of the MC for the ARQ-CQI feedback HD system. ........................................... 63
3.2 MC Transition probabilities for the ARQ-CQI feedback SD system. .................................................... 78
4.1 List of network parameters. ........................... 104
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>ARQ</td>
<td>Automatic Repeat Request</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>CQI</td>
<td>Channel Quality Indicator</td>
</tr>
<tr>
<td>CR</td>
<td>Cognitive Radio</td>
</tr>
<tr>
<td>CRC</td>
<td>Cyclic Redundancy Check</td>
</tr>
<tr>
<td>CRSN</td>
<td>Cognitive Radio Sensor Networks</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Indicator</td>
</tr>
<tr>
<td>CSS</td>
<td>Cooperative Spectrum Sensing</td>
</tr>
<tr>
<td>DARPA</td>
<td>Defence Advanced Research Projects Agency</td>
</tr>
<tr>
<td>DSAC</td>
<td>Distributed Spectrum Aware Clustering</td>
</tr>
<tr>
<td>FCC</td>
<td>Federal Communications Commission</td>
</tr>
<tr>
<td>HD</td>
<td>Hard Detection</td>
</tr>
<tr>
<td>iid</td>
<td>Independent Identical Distribution</td>
</tr>
<tr>
<td>MC</td>
<td>Markov Chain</td>
</tr>
<tr>
<td>MDP</td>
<td>Markov Decision Process</td>
</tr>
<tr>
<td>MLP</td>
<td>Multilayer Perceptron</td>
</tr>
<tr>
<td>NTRA</td>
<td>National Telecom Regulatory Authority</td>
</tr>
<tr>
<td>PMI</td>
<td>Pre-coding Matrix Indicator</td>
</tr>
<tr>
<td>POMDP</td>
<td>Partially Observable Markov Decision Process</td>
</tr>
<tr>
<td>PU</td>
<td>Primary User</td>
</tr>
<tr>
<td>QBD</td>
<td>Quasi Birth-and-Death</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RI</td>
<td>Rank Indicator</td>
</tr>
<tr>
<td>RIB</td>
<td>Rate Interference Budget</td>
</tr>
<tr>
<td>RL</td>
<td>Reinforcement Learning</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>SC2</td>
<td>Spectrum Collaboration Challenge</td>
</tr>
<tr>
<td>SD</td>
<td>Soft Detection</td>
</tr>
<tr>
<td>SU</td>
<td>Secondary User</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access</td>
</tr>
<tr>
<td>WRAN</td>
<td>Wireless Regional Area Network</td>
</tr>
</tbody>
</table>
List of Publications

• Journal papers:
  
  

• Conference paper:
  
Chapter 1

Introduction

1.1 Cognitive radio

Radio spectrum, which is a limited resource, is regulated by independent agency in each country such as National Telecom Regulatory Authority (NTRA) in Egypt, and Federal Communications Commission (FCC) in the United States. Radio spectrum is divided to licensed bands and unlicensed bands. The licensed bands are assigned to individual companies after paying a licensing fee to the regulated agency. Licence ensures the licensed bands are used by the licensed users only without interfering from the unlicensed users. Some spectrum bands are over-utilized, while other bands are rarely used. This indicates the inefficient usage of the radio spectrum [1].

In [1], author proposed the Cognitive radio (CR) technology, which used in wireless communication systems to become smart enough to explore the radio spectrum and set its parameters to best use of the available resources. CR enables the unlicensed user or the secondary user (SU) to share the spectrum with the licensed user or the primary user (PU) without harmful effect on the PU network. The SU has the ability to determine the vacant channels in the spectrum and select the best available channel. Therefore, CR solves the spectrum scarcity problem and the inefficient use of the radio spectrum [2]. IEEE 802.22 is the first cognitive radio wireless regional area network standard (WRAN) and proposed in 2009. This standard exploits the white space in the TV frequency spectrum within the rural areas to provide internet connection to home and business [3].
There is a dramatic increase in the wireless communication users. For example, it is expected the number of internet users in the United States increases from 272 million users in 2017 to 283.5 million users in 2022 [4]. Therefore, there is a great interest to the field of radio spectrum management research.

DARPA (Defence Advanced Research Projects Agency) is a research branch of the U.S. Department of Defence. DARPA makes investment to encourage and create breakthrough technologies in many research areas [5]. One of DARPA challenge is Spectrum Collaboration Challenge (SC2). SC2 is a competition aims to solve the spectrum scarcity problem, which is an insisting problem for the world. Competitors will develop strategies to share the radio spectrum between the SUs and PUs while ensuring the best use of the available spectrum. SC2 merging the advantages of artificial intelligence (AI), machine learning, and cognitive radios to achieve the SC2’s goal. SC2 finale was held in Los Angeles on October 23, 2019. The first, second, and third winner team of the competition won $2 million, $1 million, and $750,000, respectively [6].

1.2 Cognitive radio Cycle

The cognitive radio depends on the cognitive radio cycle, shown in Fig. 1.1 [7]. The cognitive radio must able to observe the environment then adapt its parameters to act in the best way. The cognitive radio should have the ability to learn from its past experiences as cycle repeats. CR system should has the ability of perception, adapting, learning and reasoning [7].

- Cognitive perception: The ability of gathering an information about the transmitted channel such as transmission frequency, bandwidth, power, modulation, etc. Using this information, the SU can identify the best parameters to transmit its data. CR achieves perception by sensing the spectrum so the SU be aware of the surrounding environment.
1.3 Spectrum management framework

Cognitive perception is achieved by the spectrum sensing algorithms. Spectrum sensing techniques enable SU to avoid collision with the PU while using the spectrum. The spectrum management process consists of four major steps:

- **Adaptation**: The SU has the ability to transmit and receive based on the best parameters identified using the cognitive perception.

- **Learning**: It is the ability of the cognitive user to learn from its past experience and current knowledge. The learning process must be efficient to supply the SU with sufficient information which enhances the reasoning ability. Learning process can be achieved through different types of machine learning algorithms.

- **Reasoning**: It enables the SU to select its actions based on the information collected through the learning process.

Over past years, researchers focused only on the perception, and cognitive adapting. The recent researches however gives an interest to learning, and reasoning ability in cognitive radio network [7].
1.3.1 Spectrum sensing

A CR user can allocate only an unused portion of the spectrum. Therefore, a CR user should monitor the available spectrum bands, capture their information, and then detect spectrum holes. This can be accomplished through a real-time wide band sensing capability to detect weak primary signals in a wide spectrum range. The most common spectrum sensing technique is the primary transmitter detection.

— **Primary Transmitter detection:** CR user is able to detect a weak signal from the PU transmitter. Three schemes are used in transmitter detection: matched filter detection, energy detection, and feature detection [8].

- **Matched filter detection:** It makes the decision if the signal is present or not. It is based on the availability of the information of the PU to the SU. The matched filter requires a prior knowledge of the characteristics of the primary user signal [9].

- **Energy detection:** It is the most common method in spectrum sensing as it has low implementation and computational complexity. It also does not need a priori knowledge about the PU. It depends on comparing the transmitted power to a certain threshold which depends on the noise floor. The major drawback of this method is the low performance in low SNR as it is not able to differentiate between the PU signal and the noise. Energy detection is divided to two types:

1. **Hard detection (HD):** The SU makes a one-bit decision by comparing the energy received at the SU from the PU, $||y_{ps}||^2$, to a threshold. If the energy is greater than the threshold, the SU will not access the PU channel, otherwise, the SU will access the PU channel (i.e. simple idle/busy model with one energy detection threshold).

2. **Soft detection (SD):** In this soft-sensing scheme [10], the energy received at the SU from the PU, $||y_{ps}||^2$, is compared to a decision threshold $\eta$. If the energy is above that threshold, the SU defers
from accessing the PU channel as the PU is declared to be active. Below the threshold, the SU accesses the channel with an access probability that depends on the PU received energy. Soft-sensing energy detection will be described in details in Section 2.3.1.

In [11] and [12], it has been shown that soft energy sensing based schemes outperform hard energy sensing based schemes. In soft energy sensing, the soft value for the sensed PU energy is used to control the SU access decisions; this is different from the hard energy sensing scheme, where the SU makes a binary decision of whether the PU is active or idle based on a single threshold. Soft-sensing can result in significant performance gains if compared to hard-sensing, as it better captures the SU belief about the PU activity [10]. In [13], the authors studied the selection of the threshold in the energy detection scheme to enhance the performance of the CR system. In [14], the authors developed a spectrum sensing scheme using a combination of energy detection and Max-Min Eigen value-based detection. This hybrid spectrum sensing scheme has been shown to improve the performance of the CR spectrum sensing.

- Cyclostationary feature detection: This method exploits the cyclostationary feature in the PU signal to identify the presence of the PU. This detection method has a high immunity to the noise as the noise is random and it has not any periodic form. However, it is computationally complex and hence requires significantly longer observation time and also high cost [15].

1.3.2 Spectrum decision

Based on the spectrum availability, CR users can allocate a channel. This allocation not only depends on spectrum availability, but is also determined based on internal (and possibly external) policies. It depends on the channel characteristics and operations of the PUs. Spectrum decision is affected by the activities of other CR users in the network. The spectrum decision has
two steps: 1- The available spectrum bands are characterized. 2- The most appropriate spectrum band can be chosen.

1.3.3 Spectrum sharing

Because there may be multiple CR users trying to access the spectrum, CR network access should be coordinated to prevent multiple users colliding in overlapping portions of the spectrum.

1.3.4 Spectrum mobility

CR users are regarded as visitors to the spectrum. Hence, if the specific portion of the spectrum in use is required by a primary user, the communication must be continued in another vacant portion of the spectrum.

1.4 Dynamic spectrum access schemes

There are three main spectrum access schemes in a cognitive radio network [16] as follows:

- **Underlay paradigm**: In this approach, the SU has information about the level of the interference caused on the PU. The SU and the PU transmit at the same time slot if the interference level on the PU is below some acceptable threshold [17].

- **Overlay paradigm**: In this approach, the SU and the PU transmit simultaneously. The SU has information about the PU’s message and its codebook as well. The SU exploits this information in variety of ways. The SU can cancel the interference caused on the PU by knowing the PU’s message. Moreover, The SU may adjust the signal to noise ratio of the PU signal to cancel the effect caused by the SU’s transmission. The overlay approach enables the SU and the PU to share the band without interfere each other [17].

- **Interweave Paradigm**: Studies made by FCC showed that the radio spectrum is inefficiently used. The licensed band is underutilized and
there are many spectrum holes. Hence, these holes can be used by the cognitive users for their communication. The cognitive user or the secondary user does not transmit at the same time of the primary user. Interweave cognitive radio is very intelligent wireless communication system as it always senses the spectrum holes to minimize the interference with the active users [18].

In this thesis, the interweave approach is considered. In the interweave approach [16] used in cognitive radio systems, the SU senses the PU existence and accesses the PU channel depending on the sensing information. The problem with this model is that the SU has no information about its interference level on the PU network. One solution to this problem is to allow the SU to overhear the feedback sent from the primary receiver to the primary transmitter, and accesses the PU channel based on the overheard feedback. PU feedback provides extra information for the SU to know if the PU is transmitting in the current time slot or not. Therefore, feedback information to the SU reduces the interference between the SU and the PU.

1.5 PU feedbacks

The feedback message between the PU transmitter and receiver has an important role in the cognitive radio network. It gives the PU transmitter an indication about the transmission efficiency at its receiver. Moreover, the SU in the cognitive radio network can use the PU feedback to gain extra information about the PU activity. There are many feedback messages may be available in the PU network as:

1.5.1 Automatic repeat request (ARQ)

The ARQ feedback has two messages: ACK and NACK. The SU refrains from accessing the PU channel upon hearing NACK, assuming that the PU has a good channel, as the PU has to retransmit its undelivered packet to the PU receiver due to unsuccessful transmission in the previous time slot.
Chapter 1. Introduction

The SU tries to access the PU channel if ACK is overheard as the PU packet is successfully received and it may not have a new packet to send.

1.5.2 Channel quality indicator (CQI)

CQI is a common feedback information that is employed in various wireless standards (e.g., LTE) [19]. It informs the PU transmitter of the state of its channel to the PU receiver. Using this feedback information, the PU transmitter adjusts its transmission parameters to achieve the maximum transmission rate or the minimum packet loss rate. In this work, we assume that the CQI feedback have only two states, informing the PU transmitter whether the channel is good, and a successful transmission is expected, or bad, and any transmission is most likely to fail. It is worth noting that practically there would be more than only two CQI feedback levels. For instance, LTE networks have 16 CQI levels. Each CQI level corresponds to the selection of a specific modulation and coding scheme at the eNodeB. At the lowest CQI level, the channel is considered so bad that no transmission can be supported and the transmitter remains silent. Since in our system we are interested only in whether the PU is transmitting or backing off due to bad channel conditions, we only consider two CQI levels. The first level corresponds to the LTE’s lowest CQI level, in which the channel cannot support any transmission attempts and the PU has to refrain from any transmission. The second level corresponds to the remaining 15 levels where a transmission takes place if the PU has packets to send. Therefore, modeling the CQI feedback with only two levels is enough for modeling the PU activity for SU access purposes.

Also, there are Pre-coding matrix indicator (PMI), channel state indicator (CSI), and Rank indicator (RI). However, in this thesis, we only consider ARQ, and CQI [19].

In [20], the SU utilizes the automatic repeat request (ARQ) to maintain a minimal interference level on the PU. Results show that there is an optimal rate-interference budget (RIB). In addition, authors examined how some strategies are far from this RIB by varying the interference budget. In [21],
a distributed power control algorithm among multiple SUs was proposed utilizing the ARQ feedback. It was proven that the feedback information is a good indicator to enable SUs to use the PU channel while achieving PU’s quality of service (QoS). In [22], the SU is assumed to have access to the PU ARQ feedback, which improves the SU’s service rate and the PU’s packet delay. In [10], the SU applies a soft spectrum sensing scheme that combines the PU ARQ feedback information in the SUs access decisions; results show significant performance improvement in the SU throughput and the PU packet delay of the proposed access scheme as compared to systems that utilize hard sensing and/or do not utilize the PU ARQ information. In [23], CR transmission schemes were studied based on the PU ARQ feedback. The derived optimal transmission strategies were designed to maximize the aggregate of primary and secondary service rates.

1.6 Learning in CR network

Learning in CR networks aims at collecting and analyzing data over the course of the past experience of the CR user. It then uses the gathered data for the reasoning step. A variety of machine learning models, algorithms and tools [7] are typically used to carry out the learning step in CR networks. Machine learning algorithms can be divided into three major classes, namely, supervised learning, unsupervised learning, and reinforcement learning (RL) [24]. For instance, in [25], a supervised machine learning-based approach is proposed to regulate dynamic handover in cognitive radio networks. The proposed scheme is shown to result in better decisions than the traditional cooperative spectrum sensing (CSS) techniques. In the case of unsupervised learning, the CR user exploits the environment with unlabelled training data. For instance, in [26], the author proposed a Distributed Spectrum-Aware Clustering (DSAC) scheme for Cognitive Radio Sensor Networks (CRSN). The proposed scheme shows preferable performance in terms of its scalability and stability compared to other baseline schemes.
Finally, for RL algorithms, a cognitive radio user takes its actions based on the knowledge received from the environment and the user learns about the environment by interacting with it. The authors in [27] adopt a reinforcement learning approach for the purpose of dynamic channel allocation and transmission power adaptation for spectrum management. The used reinforcement learning approach achieves less interference to the PU, while keeping a high probability of successful transmissions. In [28], the authors employ Q-learning RL for channel allocation in a cognitive radio system. It is observed that Q-learning results in fast convergence to a near-optimal solution in small scale networks [28]. In [29], the authors introduced a deep Q-learning model to maximize the number of successful transmissions in a wireless network.

1.7 Organization of the thesis

This thesis presents some schemes for exploiting the PU feedbacks by the SUs to enhance the cognitive radio network performance. The thesis is organized as follows:

- Chapter (2) presents a proposed scheme that uses the PU CQI feedback. Then we compare the performance of the proposed scheme with other systems, which do not use any PU feedbacks or use other types of PU feedbacks.

- Chapter (3) presents a hybrid access scheme that enables the SU to exploit two PU feedbacks. SUs also can use two sensing scheme, hard sensing scheme or soft sensing scheme. We compare the proposed hybrid scheme with other systems that exploit different amount of feedbacks.

- In Chapter (4), we propose a novel random access scheme for ARQ-based cognitive radio networks using RL.

- In Chapter (5), we conclude our work and present some suggestions as a future work.
1.8 Thesis contributions

In the following, the major contributions of the thesis are listed:

- We propose a secondary access scheme that utilizes the primary ARQ and/or CQI feedback. This allows for better tracking of the PUs activities.

- A PU queue is modeled as a multi-dimensional homogeneous quasi birth-and-death (QBD) MC, which is analyzed to get the steady-state distribution of the queue. This allows us to formulate an optimization problem where we aim at maximizing the secondary network throughput subject to PUs’ queues stability constraints.

- We propose Q-learning algorithm using MLP that enable the SU to exploit its past experience and the PU feedback to enhance the cognitive radio performance in terms of the SU throughput.
Chapter 2

CQI feedback based access system

2.1 Introduction

In this chapter, a spectrum access scheme for SUs based sensing scheme and PU’s CQI feedback has been developed. We consider two energy sensing approaches, namely, hard-sensing and soft-sensing. SUs design their access strategy based on their knowledge of the PUs CQI feedback (which informs a PU transmitter about the channel to its intended receiver) and the PUs behavior for the different states of this feedback. When a SU knows that a PU is refraining from transmission in case of bad channel conditions, it accesses the PU’s channel more aggressively. However, in the case of good channel conditions, SUs rely on sensing scheme to detect the presence or absence of PUs. The effect of exploiting the CQI feedback by the SU from the point of view of queuing theory was studied. The system with a single PU and a single SU has been modeled as a two-dimensional MC. The model is then extended to the multiple PUs and multiple SUs case. Performance evaluation is done in terms of SU throughput and PU delay compared to the systems that do not benefit from CQI feedback.

The work in [30] comes close to our work in the use of PU’s CQI feedback. The authors in [30] developed a spectrum sharing scheme for the SU based on primary CQI feedback. They derived the optimal transmit power and transmission rate for the SU when no or perfect primary CQI
is available at the secondary transmitter. Their objective was to maximize the SU’s average throughput while satisfying a maximum rate loss constraint at the primary system. The main difference between our work and the work in [30] lies in the modeling of PU and SU data arrivals, channel models and access models. We provide a queue-theoretic analysis of the system. Each user is assumed to have a queue to store its data packets with some arrival process. We also consider a collision channel model (i.e., any concurrent transmissions of more than one packet means that all the packets are lost). The work in [30] considers an interference channel model with power adaptation at the secondary users. Our model and the model in [30] consider different primary and secondary access schemes. What is unique about queuing analysis of the system is that it allows for calculating important PU performance metrics like delay, which cannot be calculated using the model in [30]. Based on our model, we can define diverse PU’s QoS constraints, like delay constraint and PU queue stability. Furthermore, our analysis considers the case of multiple PUs and multiple SUs, while the work in [30] considered only the case of single PU and single SU, and the extension to the multiple users’ case is not clear or explained. Finally, we study the case of imperfect CQI feedback at the SUs which is not considered in [30].

The rest of this chapter is organised as follows. In section 2.2, and section 2.3, the analysis of the CQI feedback based access system using hard sensing scheme and soft sensing scheme, respectively, are presented.

### 2.2 CQI feedback- HD based access system

In this section, the effect of the availability of the PU CQI feedback to the SU and the SU’s sensing the PU’s activity on the performance of the cognitive radio system is studied. In this section, we consider the the hard sensing. By overhearing the CQI feedback, the SU can exploit the time slots where the PU channel is bad to access the channel with probability one knowing that the PU is idle for sure. We also compare our proposed system with
two baseline systems. It is assumed in all systems that the SU accesses the channel based on the HD decisions [31]. The first baseline system has no PU CQI feedback. The second baseline system has a PU CQI feedback but the SU cannot access this feedback information. In all systems, the SU access probability is determined by solving an optimization problem that maximizes the SU’s throughput subject to a constraint on the PU’s queue stability. We study and compare the performance of the three systems by finding closed form expressions of the secondary throughput and the PU delay for each system.

2.2.1 System model

We consider a cognitive system consisting of one PU and one SU as shown in Fig. 2.1. The system is time-slotted, and it is assumed that the duration of one time slot equals the time of one packet transmission. It is assumed that the packets arrive at the start of the time slot, which means that a packet can be served in the same time slot it arrives at. The PU accesses the channel at the start of each time slot whenever there is a packet to transmit and the channel is in the good state. It is assumed that the channel state does not change during one time slot. Furthermore, collision channel model is assumed, i.e., if both the PU and SU transmit in the same time slot, then a collision occurs and both packets are lost. The PU and SU have an infinite buffer for storing fixed length packets. The arrival process at the PU queue is a Bernoulli process with mean $0 < \lambda_p < 1$. The SU is assumed to always have packets to transmit in its queue.
In our model, The SU accesses the channel with access probability which are selected such that the SU throughput is maximized and the stability of the PU queue is guaranteed. Stability can be loosely defined as having a certain quantity of interest kept bounded, in our case, the queue size. For more information about stability, see [32] and [33]. If the arrival and service processes of a queuing system are strictly stationary, one can apply Loynes’ theorem to check for stability [34]. This theorem states that if the average arrival rate is less than the average service rate of a queuing system, whose arrival and service processes are strictly stationary, then the queue is stable, otherwise it is unstable.

**Channel model**

The channel between the PU transmitter and receiver is modeled as a two-state MC as shown in Fig. 2.2. The two states are the good state and the bad state. When the channel is in the good state, it is most likely that the PU packet is transmitted successfully. However, the PU packet is not transmitted successfully in the bad state of the channel. The probabilities of the channel staying in the good state and in the bad state are \( p_G \) and \( p_B \), respectively. The probability of the channel moving from a good state to a bad state is \( 1 - p_G \) and moving from a bad state to a good state is \( 1 - p_B \). The steady state probabilities of the channel being in the good state and in the bad state are \( \zeta_G \) and \( \zeta_B \), respectively and can be calculated using the following equations:

\[
\zeta_G = \frac{1 - p_B}{2 - p_B - p_G}, \quad \text{and} \quad \zeta_B = \frac{1 - p_G}{2 - p_B - p_G}.
\]  

A Rayleigh flat fading channel with additive white Gaussian noise (AWGN) is used to model the PU-SU channel. The received signal at the SU transmitter (S) from the PU transmitter (P) in time slot \( t \) is given by [35]:

\[
y_{PS}^t = \sqrt{G_P r_{PS}^t} h_{PS}^t x_P^t + n_S^t,
\]  

2.2. CQI feedback- HD based access system

\[ G_p (1 - p_G) (1 - p_B) \]

**Figure 2.2**: The channel model.

where \( G_p \) is the transmitted power, \( r_{PS} \) is the distance between the PU transmitter and the SU transmitter, and \( \gamma \) is the path loss exponent (which accounts for the large-scale fading effect). The symbol \( x'_p \) denotes the transmitted signal, which is assumed to be drawn from a zero-mean, unit variance constellation. The term \( h'_{PS} \) denotes the channel gain between the PU transmitter and the SU transmitter, modeled as a circularly symmetric complex Gaussian random variable with zero-mean and unit variance. The noise term \( n'_S \) is modeled as a circularly symmetric complex Gaussian random variable with zero-mean and variance \( N_0 \). We assume that the channel gain \( h'_{PS} \) is stationary and independent from slot to slot, thus, the superscript \( t \) will be dropped for convenience.

### 2.2.2 SU spectrum access schemes

#### The baseline systems

- The No CQI feedback- based access system (Baseline system 1 No FB): In this system, the PU has ACK/NACK feedback and has no CQI feedback information. Therefore, the PU transmits its packets regardless of the state of the channel. The SU does not have access to the PU’s ACK/NACK feedback. The SU accesses the channel with an access probability \( a_s \) in every time slot based on the hard decision sensing scheme.

- The PU CQI Feedback- based Access System (Baseline System 2 No FB SU): The PU has both ACK/NACK feedback and the CQI feedback of the channel state \(^1\) in the next time slot. The CQI feedback

\(^1\)For simplicity of presentation, we assume a zero delay CQI feedback channel.
is an indicator of how good/bad the channel between the PU transmitter and receiver is. If a good CQI feedback is observed, the PU transmits whenever it has packets in its queue. Observing a bad PU CQI feedback, the PU backs-off since it knows that the packet will not be received correctly. However, the SU does not monitor the PU feedbacks. The SU accesses the channel with an access probability $a_s$ in every time slot based on the hard decision sensing scheme.

**Proposed CQI feedback HD-based access system:**

The proposed system model is shown in Fig. 2.1, in which the PU has both ACK/NACK and CQI feedback of the channel state in the next time slot, and the SU only listens to this CQI feedback. The SU accesses the channel depending on the hard decision sensing scheme and the primary CQI feedback. If a good CQI feedback is observed and the SU does not detect the PU’s existence, the SU accesses the channel with access probability $a_s$. If a bad CQI feedback is observed, the SU exploits the knowledge that the PU will back-off during the next time slot to transmit with probability 1.

**2.2.3 Performance analysis**

The PU’s queue in the two baseline systems and the proposed system is modelled using the same two-dimensional Markov model shown in Fig. 2.3; the same MC is used for the three systems as the PU queue dynamics in the bad channel states will not be affected by the SU access decisions (as the PU will always fail in the case of a bad channel). Moreover, in the PU good channel states, the SU accesses the channel with an access probability of $a_s$ if the SU does not detect the PU’s existence (where the access probabilities that will maximize the secondary throughput in each of the three systems will be different). Finally, each system will have a different expression for the SU throughput as will be explained later.
2.2. CQI feedback- HD based access system

Figure 2.3: The PU queue MC model for the two-state channel model.

Table 2.1: Transition probabilities of the MC of the CQI feedback- (HD) system.

<table>
<thead>
<tr>
<th>Number</th>
<th>Transition Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>((1 - \lambda_p) + \lambda_p ((1 - a_s)(1 - p_d) + p_d))p_G)</td>
</tr>
<tr>
<td>2</td>
<td>((1 - \lambda_p)(1 - p_B))</td>
</tr>
<tr>
<td>3</td>
<td>((1 - \lambda_p)p_B)</td>
</tr>
<tr>
<td>4</td>
<td>((1 - \lambda_p) + \lambda_p ((1 - a_s)(1 - p_d) + p_d))(1 - p_G))</td>
</tr>
<tr>
<td>5</td>
<td>((1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)p_G)</td>
</tr>
<tr>
<td>6</td>
<td>((1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)(1 - p_G))</td>
</tr>
<tr>
<td>7</td>
<td>(\lambda_p((1 - a_s)(1 - p_d) + p_d) + \lambda_p a_s(1 - p_d))p_G)</td>
</tr>
<tr>
<td>8</td>
<td>(\lambda_p(1 - a_s) + (1 - \lambda_p)a_s)(1 - p_G))</td>
</tr>
<tr>
<td>9</td>
<td>((1 - \lambda_p)(1 - p_B))</td>
</tr>
<tr>
<td>10</td>
<td>((1 - \lambda_p)p_B)</td>
</tr>
<tr>
<td>11</td>
<td>(\lambda_p a_s(1 - p_d)p_G)</td>
</tr>
<tr>
<td>12</td>
<td>(\lambda_p a_s(1 - p_d)(1 - p_G))</td>
</tr>
<tr>
<td>13</td>
<td>(\lambda_p(1 - p_B))</td>
</tr>
<tr>
<td>14</td>
<td>(\lambda_p p_B)</td>
</tr>
</tbody>
</table>
Chapter 2. CQI feedback based access system

The MC has two types of states \((K, G)\) and \((K, B)\), where \(K\) is the number of PU packets in the queue, \(G\) means that the PU’s channel is in the good state and \(B\) means that the PU’s channel is in the bad state. More specifically, we have a MC \(\{X(n), n = 0, 1, 2, ...\}\), whose state space is given by \(S=\{(K, T) : K = 0, 1, 2, ..., T \in \{G, B\}\}\).

The transitions between states are as follows:

- The transition from state \((K, G)\) to \((K + 1, G)\): the transition in this case occurs according to the following equation:

\[
\Pr(X(n+1) = (K + 1, G) \mid X(n) = (K, G)) = \Pr(\text{a new packet arrives at the PU queue}) \cap \text{(SU does not detect the PU presence and decides to access the channel}) \cap \text{(the channel in the next time slot remains in the good state}) = \lambda_p a_s (1 - p_d) p_G, \]

where \(p_d\) is the detection probability of the spectrum sensor.

- From \((K, G)\) to \((K + 1, B)\): it is same as the above transition but \(p_G\) is replaced by \(1 - p_G\). Therefore the transition probability equals to \(\lambda_p a_s (1 - p_d)(1 - p_G)\).

- The rest of the transition probabilities are shown in Fig. 2.3 are listed in Table 2.1 and can be deduced easily.

1. The steady state distribution calculation

We start by calculating the steady state distribution of the MC shown in Fig. 2.3 so that we can get an expression for the SU throughput of the three systems. The steady state distribution vector is given by

\[
v = [\pi^G_0, \pi^B_0, \pi^G_1, \pi^B_1, ....].
\]

Define the vector \(v_k = \begin{pmatrix} \pi^G_k \\ \pi^B_k \end{pmatrix}\). Note that \(v_0 = \begin{pmatrix} \pi^G_0 \\ \pi^B_0 \end{pmatrix}\).

The state transition matrix of the MC shown in Fig. 2.3 can be written
as

\[
\Phi = \begin{pmatrix}
B & A_0 & 0 & 0 & \ldots \\
A_2 & A_1 & A_0 & 0 & \ldots \\
0 & A_2 & A_1 & A_0 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix},
\]

(2.3)

where \(B, A_0, A_1, A_2\) are shown in equation (2.4).

\[
B = \begin{pmatrix}
(1 - \lambda_p) + \lambda_p((1 - a_s)(1 - p_d) + p_d)p_G & (1 - \lambda_p)(1 - p_B) \\
(1 - \lambda_p) + \lambda_p((1 - a_s)(1 - p_d) + p_d)(1 - p_G) & (1 - \lambda_p)p_B
\end{pmatrix},
\]

\[
A_0 = \begin{pmatrix}
(1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)p_G & 0 \\
(1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)(1 - p_G) & 0
\end{pmatrix},
\]

\[
A_1 = \begin{pmatrix}
[\lambda_p((1 - a_s)(1 - p_d) + p_d) + (1 - \lambda_p)a_s(1 - p_d)]p_G & (1 - \lambda_p)(1 - p_B) \\
[\lambda_p((1 - a_s)(1 - p_d) + p_d) + (1 - \lambda_p)a_s(1 - p_d)](1 - p_G) & (1 - \lambda_p)p_B
\end{pmatrix},
\]

\[
A_2 = \begin{pmatrix}
\lambda_p a_s(1 - p_d)p_G & \lambda_p(1 - p_B) \\
\lambda_p a_s(1 - p_d)(1 - p_G) & \lambda_p p_B
\end{pmatrix},
\]

(2.4)

The state transition matrix \(\Phi\) is a block-tridiagonal matrix; therefore the MC shown in Fig. 2.3 is a homogeneous quasi birth-and-death (QBD) MC.

The steady state distribution of the MC shown in Fig. 2.3 satisfies the following equation[36]:

\[
v_k = R^k v_0, \quad k > 0,
\]

(2.5)

where the rate matrix \(R\):

\[
R = \begin{pmatrix}
r_{11} & r_{12} \\
r_{21} & r_{22}
\end{pmatrix},
\]
which is the solution of the following equation

\[ A_2 + (A_1 - I_2)R + A_0 R^2 = 0_{2 \times 2}, \tag{2.6} \]

where \( I_2 \) is the \( 2 \times 2 \) identity matrix. Equation (2.6) can be obtained by substituting equation (2.5) in the next equation

\[ v_k = A_2 v_{k-1} + A_1 v_k + A_0 v_{k+1}, \quad k \geq 1. \tag{2.7} \]

Equation (2.7) can be easily derived using the states balance equations.

By solving equation (2.6), the matrix \( R \) is obtained as follows:

\[ r_{11} = \frac{a_s \lambda_p (1 - p_d)}{(1 - \lambda_p)(a_s p_d - a_s + 1)}, \]

\[ r_{12} = \frac{\lambda_p}{(1 - \lambda_p)(a_s p_d - a_s + 1)}, \]

\[ r_{21} = \frac{a_s \lambda_p (1 - p_d) (1 - p_G)}{(1 - \lambda_p)(a_s p_d - a_s + 1)(\lambda_p p_B - p b - \lambda_p + \lambda_p p_g + 1)} \]

\[ r_{22} = \frac{\lambda_p}{(1 - \lambda_p)(a_s p_d - a_s + 1)(\lambda_p p_B - p b - \lambda_p + \lambda_p p_g + 1)}, \tag{2.8} \]

where \( B_1 \) is shown in the following equation:

\[ B_1 = \lambda_p (a_s + \lambda_p + p_B - a_s \lambda_p - a_s p_B - a_s p_g - \lambda_p p_B - \lambda_p p_g + a_s \lambda_B p_d + a_s \lambda_B p_B + a_s \lambda_B p_g + a_s p_B p_B + a_s p_d p_g - a_s \lambda_B p_d p_B - a_s \lambda_B p_d p_g), \tag{2.9} \]
To get the steady state distribution of the MC, the following normalization requirement is applied:

\[ \sum_{k=0}^{\infty} (\pi_k^G + \pi_k^B) = 1, \]

and using equation (2.5), we have

\[ \bar{I} \left( \sum_{k=0}^{\infty} R^k \right) v_0 = 1, \]

where \( \bar{I} = [1 \quad 1] \).

So, \( \bar{I} \left( \sum_{k=0}^{\infty} R^k \right) v_0 = \bar{I} (I - R)^{-1} \begin{pmatrix} \pi_0^G \\ \pi_0^B \end{pmatrix} = 1. \)

The relationship between \( \pi_0^G \) and \( \pi_0^B \) has to obtained so the previous equation will be function in one variable only. To get the relationship between \( \pi_0^G \) and \( \pi_0^B \), the balance equations around \((0, G)\) and \((0, B)\) are solved. The balance equation around state \((0, G)\) is given by:

\[
[a_s \lambda_p p_G - p_G - a_s \lambda_p p_G + 1] \pi_0^G \\
= (1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)p_G \pi_1^G \\
+ (1 - \lambda_p)(1 - p_B) \pi_0^B.  
\]  \hspace{1cm} (2.10)

The balance equation around state \((0, B)\) is given by:

\[
[\lambda_p + (1 - \lambda_p)(1 - p_B)] \pi_0^B = \\
(1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)(1 - p_G) \pi_1^G \\
+ [(1 - \lambda_p)(1 - p_G) + \lambda_p((1 - a_s)(1 - p_d) + p_d)(1 - p_G)] \pi_0^G.  
\]  \hspace{1cm} (2.11)

Eliminating \( \pi_1^G \) from equation (2.10) and (2.11), we obtain the relation between \( \pi_0^G \) and \( \pi_0^B \) as

\[
\pi_0^B = \frac{(1 - p_G)\pi_0^G}{\lambda_p p_B - p_B - \lambda_p + \lambda_p p_G + 1}.  
\]  \hspace{1cm} (2.12)
24

Chapter 2. CQI feedback based access system

\[ \pi_0^G \text{ is obtained as shown in the following equation:} \]

\[ \pi_0^G = \frac{a_s p_d - 2 \lambda_p - p_B - a_s + a_s p_B + \lambda_p p_B + \lambda_p p_G - a_s p_d p_B + 1}{(1 - \lambda_p)(p_B + p_G - 2)(a_s p_d - a_s + 1)}. \] (2.13)

2. Secondary throughput analysis:
The closed-form expressions of the SU throughput of the baseline systems and the proposed system are derived as follows:

- **The No CQI feedback- HD based access system (Baseline system 1 No FB):**
  This is the first baseline system, in which the PU has no CQI feedback. Therefore, the PU always accesses the channel when it has packets in its queue even if the channel is in the bad state. Moreover, the SU accesses the channel with access probability \(a_s\) in every time slot if it does not detect the PU’s presence. Therefore, the SU transmits its packets with no PU collisions only in the PU empty states \((0, G)\), when the PU does not receive new packet in this time slot and \((0, B)\). Hence, the SU throughput in this system, \(\mu_{s1}\), is given by:

\[ \mu_{s1} = a_s (1 - p_f) \left(1 - \lambda_p\right) \left[\pi_0^G + \pi_0^B\right] \]

\[ = \frac{B_2}{(p_B + p_G - 2)(a_s p_d - a_s + 1)(\lambda_p p_B - p_B - \lambda_p + \lambda_p p_G + 1)}, \] (2.14)

where \(p_f\) is the false alarm probability of the spectrum sensor and \(B_2\) is shown in the following equation:

\[ B_2 = a_s (1 - p_f) \left(\lambda_p + p_B + p_G - \lambda_p p_B - \lambda_p p_G - 2\right) \]

\[ (a_s p_d - 2 \lambda_p - p_B - a_s + a_s p_B + \lambda_p p_B + \lambda_p p_G - a_s p_d p_B + 1). \] (2.15)

- **The PU CQI Feedback- HD based access system (Baseline system 2 No FB SU):**
  This is the second baseline system. In this system, only the PU has access to the CQI feedback, so the SU accesses the channel in the good and bad CQI states with an access probability \(a_s\) if the
SU does not detect the PU’s existence. The PU does not transmit packets in the bad channel states. Thus, the SU succeeds in transmitting a packet in the \((K, B)\) states with probability \(a_s\), as the PU will be backing-off. Also, the SU succeeds in transmitting a packet with probability \(a_s\) if the PU channel is in the empty good state, i.e., \((0, G)\) and the PU does not receive new packet in this time slot. Moreover, the SU has to detect the PU’s absence. Hence, the SU throughput in this system, \(\mu_{s2}\), is given by:

\[
\mu_{s2} = a_s(1 - p_f)[(1 - \lambda_p)\pi_0^G + \sum_{k=0}^{\infty} \pi_k^B]
\]

\[
= a_s(1 - p_f)[(1 - \lambda_p)\pi_0^G + [0 1](I_2 - R)^{-1} \begin{pmatrix} \pi_0^G \\ \pi_0^B \end{pmatrix}]
\]

\[
= a_s(1 - p_f)(1 - a_s - \lambda_p + a_s p_d) \frac{a_s p_d - a_s + 1}{a_s p_d - a_s + 1}.
\]

\[
\tag{2.16}
\]

- The proposed system CQI feedback- HD based access system:

In this system, the SU has access to the PU CQI feedback; therefore, the SU accesses the channel with probability 1 under bad PU CQI feedback, where the PU is backing off. However, under good PU CQI feedback the SU accesses the channel with probability \(a_s\) if the SU decides that the PU is absent. Therefore, the SU transmits its packets collision-free in the bad states \((K, B)\) with probability 1 and in the empty good state, \((0, G)\), with probability \(a_s\). Hence, the SU throughput in this system, \(\mu_{s3}\), is given by:

\[
\mu_{s3} = a_s(1 - p_f)(1 - \lambda_p)\pi_0^G + \sum_{k=0}^{\infty} \pi_k^B
\]

\[
= a_s(1 - p_f)(1 - \lambda_p)\pi_0^G + [0 1](I_2 - R)^{-1} \begin{pmatrix} \pi_0^G \\ \pi_0^B \end{pmatrix}
\]

\[
= \frac{B_3}{(1 - \lambda_p)(2 - p_B - p_G)(a_s p_d - a_s + 1)},
\]

where \(B_3 = (1 - p_G)(1 - a_s - \lambda_p + a_s p_d) + a_s(1 - \lambda_p)(1 - p_f)(a_s p_d -
\]
Chapter 2. CQI feedback based access system

\[ 2\lambda_p - p_B - a_s + a_s p_B + \lambda_p p_B + \lambda_p p_G - a_s p_d p_B + 1) + a_s \lambda_p (1 - p_d) (1 - p_G). \]

- **The Perfect sensing CQI feedback based-access system**

In this system, the PU has a CQI feedback. The SU accesses the PU channel in the bad channel states with probability 1. When the PU channel is in the good state and the PU’s queue is empty, the SU accesses the channel with probability 1 as well (because of the perfect sensing). The MC of the CQI feedback perfect sensing system is also a two-dimensional MC, which is shown in Fig. 2.4. The analysis of this MC is done using the same steps used in the above MC, shown in 2.3. Therefore, we can get the steady state distribution as follows:

\[ \pi^G_0 = \frac{\lambda_p p_B - p_B - 2\lambda_p + \lambda_p p_g + 1}{(\lambda_p - 1)(p_B + p_g - 2)}. \] (2.18)

\[ \pi^B_0 = \frac{(1 - p_g) \pi^G_0}{\lambda_p p_B - p_B - \lambda_p + \lambda_p p_g + 1}. \] (2.19)

\[ R = \begin{pmatrix} 0 & \lambda_p/(1 - \lambda_p) \\ \lambda_p/(1 - \lambda_p) & 1/(1 - \lambda_p) \end{pmatrix}. \] (2.20)

The SU throughput for this system \( \mu_{sp} \), can be calculated using the following expression:

\[ \mu_{sp} = (1 - \lambda_p) \pi^G_0 + \sum_{k=0}^{\infty} \pi^B_k = (1 - \lambda_p) \pi^G_0 + [0 \ 1] (I_2 - R)^{-1} \begin{pmatrix} \pi^G_0 \\ \pi^B_0 \end{pmatrix}. \] (2.21)

Substituting by equations (2.18), (2.19), and (2.20) in equation (2.21), the closed form expression of the SU throughput can be
shown in the following equation:

\[
\mu_{sp} = \begin{cases} 
1 - \lambda_p, & \text{if } \lambda_p < 1 - \zeta_B, \\
\zeta_B, & \text{otherwise}.
\end{cases}
\] (2.22)

When the PU’s arrival rate is less than the probability with which the PU’s channel is in the good state, then the PU will be able to empty its queue regularly. Therefore, the SU throughput is equal to the channel’s idle probability (probability of the PU’s queue being empty which is equal to \(1 - \lambda_p\)). When the PU’s arrival rate is greater than the probability with which the PU’s channel is in the good state, the PU will not be able to empty its queue and the SU will only access the channel when it is in the bad state (which has a probability \(\zeta_B\)).

3. \textbf{Primary delay analysis}

We derive an expression for the average PU packet delay for the baseline systems and the proposed system using Little’s law as follows:

\[
D_p = \frac{E(Q_p)}{\lambda_p} = \frac{1}{\lambda_p} \sum_{k=0}^{\infty} k(\pi_k^G + \pi_k^B)
\]

\[
= \frac{1}{\lambda_p} [1 - R(I_2 - R)^{-2} \left( \begin{array}{c} \pi_0^G \\ \pi_0^B \end{array} \right)],
\]

where \(D_p\) is the average PU packet delay, and \(E(Q_p)\) is the average number of packets in the PU queue. The closed-form expressions of the average PU packet delay of the two baseline systems and the proposed system are as follows:

\[
D_p = \frac{B_4}{(2 - p_B - p_G)(a_s p_d - 2\lambda_p - p_B - a_s + a_s p_B + \lambda_p p_B + \lambda_p p_G - a_s p_d p_B + 1)}.
\] (2.23)
where \( B_4 \) is shown in the following equation:

\[
B_4 = \lambda_P(3a_s - p_G - 3a_s p_d - 4a_s p_B - 3a_s p_G + a_s p_B^2 + a_s p_G^2 - a_s p_d p_B^2 - a_s p_d p_G^2 - s p_d p_B + 3a_s p_d p_G + 2a_s p_B p_G - 2a_s p_d p_B p_G + 1).
\]

(2.24)

The expression for the PU delay in equation (2.23) is valid for all the systems discussed above. Since the optimal value for the SU channel access probability, \( a_s \), will differ from system to system as shown later, the average PU delay will also be different for each system.

4. Access probabilities calculation

The access probability \( a_s \) has to be selected to maximize the secondary throughput, \( \mu_{si} \), \( i = 1, 2, 3 \), while keeping the PU queue stable. The problem can be formulated as follows.

\[
\max_{a_s} \mu_{si}
\]

subject to

\[
\pi_0^G > 0 \text{ and } \pi_0^B > 0.
\]

By differentiating the expression of \( \mu_{si} \) with respect to \( a_s \) and equating the derivative to zero, the optimal access probability \( a_s^* \) can be

\footnote{Stability of any queue can be guaranteed if the queue has a non-zero probability of being empty [37].}
obtained.

For all systems, the differentiation of $\mu_{si}$ with respect to $a_s$ results in a second degree polynomial in $a_s$. Therefore, there are two solutions of this maximization problem. The solution in the range from 0 to 1 is selected as the value of $a_s$ that results in the maximum secondary user throughput will always guarantee the stability of the PU queue; since if this $a_s$ causes the PU queue to be unstable, this will reduce the SU throughput since the SU will never transmit any packets in the good channel states, as the PU queue will always be backlogged. The maximum secondary throughput is obtained by substitution of $a_s^*$ in the equation of $\mu_{si}$ to get the maximum secondary throughput, $\mu_{s}^{\text{max}}$.

The closed-form expressions of the access probabilities to maximize the secondary throughput of the baseline systems and the proposed system are as follows,

- **Baseline system 1**: $a_s^* = \frac{p_B + \sqrt{\lambda p_B(p_B-1)(p_B+p_G-2) - 1}}{p_d + p_B - p_d p_B - 1}$.
- **Baseline system 2**: $a_s^* = \frac{\sqrt{\lambda_B - 1}}{p_d - 1}$.
- **The proposed system**: $a_s^* = \frac{p_B + \sqrt{\lambda_B(p_B-1)(p_B+p_G-2) - 1}}{p_d + p_B - p_d p_B - 1}$.

It can be noticed that the first baseline system and the proposed system have the same expression for the access probabilities to maximize the secondary throughput of each of them. Note that the difference between these two systems lies in the fact that the SU accesses the channel with probability 1 under bad PU channel state in the CQI FB-based system and with probability $a_s$ in the no-FB system. Under good PU channel state, both access the channel with some $a_s$. So both systems will have the same effect on the PU (since under PU bad state, whatever action that will be taken by the SU will not affect the PU, since it will be either in the back-off state or it will transmit a packet with failure). Therefore, it is expected that these two systems will have the same access probability.
Figure 2.5: The SU throughput for different access schemes
\( (p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.9, p_f = 0.1). \)

Figure 2.6: The SU throughput for different access schemes
\( (p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1). \)
2.2. CQI feedback - HD based access system

2.2.4 Performance results

In this section, a comparative study in terms of the SU throughput of the proposed scheme, the baseline systems, and the perfect sensing CQI feedback based access system, which is an upper bound system, is provided. We also compare the PU delay for the proposed scheme, the two baseline systems and the perfect sensing scheme.

In Fig. 2.5 and Fig. 2.6, the SU throughput is plotted against the PU arrival rate for the different access schemes. In Fig. 2.5 and Fig. 2.6, the steady state probability of the channel being in the bad state equals 0.4 and 0.125, respectively, which can be obtained using equation (2.1). It is clear that the proposed scheme has the highest performance below the perfect sensing CQI feedback based access system since the SU exploits the PU CQI feedback to efficiently access the primary network. Regarding the two baseline systems, it is expected that the performance of the second system is better than the performance of the first one. The PU in the second system has additional information about the channel state, so the PU does not transmit its
Chapter 2. CQI feedback based access system

Figure 2.8: The SU optimal access probabilities for different access schemes 
($p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1$).

packets if the channel is in the bad state, which gives the SU more opportunities to access the channel. However, PU in the first system has no CQI feedback so it transmits its packet independent of channel state.

Moreover, in Fig. 2.5 and Fig. 2.6, it is noticed that for the proposed scheme the secondary throughput does not tend to zero as the PU arrival rate goes to 1, unlike the first baseline systems. The minimum value of the SU throughput in the proposed scheme equals the steady state probability of the PU channel being in the bad state, since this minimum level of SU service is always guaranteed (as in the bad states, the PU will be backing-off and the SU will access the channel, collision-free, with probability 1). It can be seen that the SU throughput remains constant after a certain PU’s arrival rate. After this PU’s arrival rate, the PU’s queue is unstable so it always has packets to transmit, however, this does not affect the SU service rate since in this case the SU only accesses the channel when the PU’s channel in the bad state.

In Fig. 2.7 and Fig. 2.8, the SU optimal access probabilities are plotted against the PU arrival rate for different access schemes, for a steady state
probability of the channel being in the bad state of 0.4 in Fig. 2.7 and 0.125 in Fig. 2.8. It can be noticed that the second baseline system has the highest optimal access probabilities, which are greater than or equal to one for all values of the PU arrival rate so the maximum access probability equals to one for all values of PU arrival rate. It can be easily shown that the SU service rate is concave in $a_s$, so setting the values of access probability to one does not affect the optimality if the optimum value of $a_s$ is greater than one. Moreover, the optimal access probabilities in the first baseline system and the optimal access probabilities of the proposed scheme are equal. It is clear that the SU will be less aggressive in accessing the channel under the good channel state in our proposed system as compared to the second baseline system; since, the SU is guaranteed a service rate of 1 under PU bad channel state. For all systems, there will be no access probabilities after a certain value of the PU arrival rate as the PU’s queue will be unstable and the SU will be backing-off.

In Fig. 2.9 and Fig. 2.10, the PU packet delay are plotted against the PU arrival rate for different access schemes. The PU packet delay in the proposed scheme and the first baseline system are coincident and lower...
### Chapter 2. CQI feedback based access system

#### 2.3 CQI feedback- soft detection (SD) based access system

In this system, an access scheme for CR networks is proposed in which SUs make use of the PUs’ channel quality feedback information. SUs decide their transmission strategy based on the PUs’ CQI feedback available in the PU network and the SUs’ sensing of PUs’ activity. We consider soft-sensing as energy sensing scheme. A SU accesses the primary channel with access probabilities which are selected to maximize the SU service rate while ensuring PUs’ QoS requirements defined as primary queues’ stability. The proposed system is modeled using a two-dimensional MC. Closed-form expressions for SUs’ throughput and the average PUs’ packet delays are

**Figure 2.10:** The PU packet delay for different access schemes

\[ p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1 \]

than the second baseline system. It is expected that the proposed scheme improves the PU packet delay due to the CQI awareness at the SU. The SU exploits this additional information to reduce the collisions with the PU and so the PU delay decreases.
2.3. CQI feedback- soft detection (SD) based access system

derived. Moreover, SUs’ access probabilities selection is formulated as a
constrained optimization problem. Performance of the proposed scheme,
in terms of the secondary user service rate and the primary user delay, is
shown to be superior to other systems in which SUs do not leverage PUs’
CQI feedback information.

2.3.1 System model

This system has the same system model that was described in Section 2.2.1.

Channel model

The system has the same channel model was shown in Section 2.2.1.

Spectrum soft-sensing scheme

Spectrum sensing can be performed using different techniques; in this sub-
section, we will rely on energy detection because of its simplicity and math-
ematical tractability. It is worth to note that any spectrum sensing technique
can be used in our system model for better detection accuracy. Generally,
the detector will be base on comparison of some test statistic to a threshold.
Specifically, we use soft energy sensing [10], where the SU senses the PU’s
transmitted energy and compares this sensed energy, $||y_{ps}||^2$, to a certain
decision threshold $\eta$ as shown in Fig. 2.11. If the sensed energy is larger
than the threshold $\eta$, the SU decides not to access the PU’s channel as the
PU is active. When the sensed energy level falls below the threshold, the SU
accesses the channel with an access probability that depends on the sensed
energy level. The soft-sensing based access scheme we adopt can be sum-
marized as follows:

- The interval $[0, \eta]$ is divided into $n$ subintervals of equal width as
  shown in Fig. 2.11. Let the set of all subintervals be $I = \{1, 2, ..., n\}$.
- A SU access probability $a_i$ is assigned for each subinterval $i \in I$
- If the sensed energy level $||y_{ps}||^2$ falls in the $i$-th subinterval, the SU
  accesses the PU channel with probability $a_i$. 
Chapter 2. CQI feedback based access system

If $\|y_{ps}\|^2$ is greater than $\eta$, the SU refrain from accessing the PU channel.

The SU selects its access probabilities in such a way to limit collisions with the PU. Therefore, when the sensed energy level falls in subintervals close to the threshold $\eta$, the SU accesses the PU’s channel with relatively low access probabilities, since in these subintervals the SU is less certain about the state of the PU being idle. On the other hand, when the sensed energy level falls in subintervals close to 0, the SU accesses the channel more aggressively, since in these subintervals it is more certain that the PU is idle. The optimum SU’s access probabilities, that maximize the SU’s throughput, are obtained by solving a constrained optimization problem that will be presented in Section 2.3.3.

Given the soft-sensing and spectrum access scheme discussed above, we can evaluate the probabilities that the SU accesses the PU’s channel when the PU is active or inactive as follows:

- When the PU is inactive, the SU will access the channel with probability $\sum_{i \in I} p^0_i a_i$, where $p^0_i$ is the probability that the sensed energy level $\|y_{ps}\|^2$ falls in the $i$-th subinterval and $a_i$ is the SU’s access probability when $\|y_{ps}\|^2$ falls in the $i$-th subinterval in the case of idle PU.

From the channel model discussed above, it can be shown that the probability $p^0_i$ is given by the following equation [10]:

$$p^0_i = \exp \left( \frac{-(i - 1)\eta}{2n\sigma_0^2} \right) - \exp \left( \frac{-i\eta}{2n\sigma_0^2} \right),$$

(2.25)

where $\sigma_0^2$ is the variance of the energy detector’s output when the PU is absent.

- When the PU is active, the SU will access the channel with probability $\sum_{i \in I} p^1_i a_i$, where $p^1_i$ is the probability that the sensed energy level
### 2.3. CQI feedback - soft detection (SD) based access system

#### Table 2.2: Transition probabilities of the MC of the CQI feedback - (SD) system.

<table>
<thead>
<tr>
<th>Number</th>
<th>Transition Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>((1 - \lambda_p) + \lambda_p(1 - \sum_{i \in I} p^1_i a_i) p_G)</td>
</tr>
<tr>
<td>2</td>
<td>((1 - \lambda_p)(1 - p_B))</td>
</tr>
<tr>
<td>3</td>
<td>((1 - \lambda_p)p_B)</td>
</tr>
<tr>
<td>4</td>
<td>((1 - \lambda_p)(1 - \sum_{i \in I} p^1_i a_i)(1 - p_G))</td>
</tr>
<tr>
<td>5</td>
<td>((1 - \lambda_p)(1 - \sum_{i \in I} p^1_i a_i) p_G)</td>
</tr>
<tr>
<td>6</td>
<td>((1 - \lambda_p)(1 - \sum_{i \in I} p^1_i a_i)(1 - p_G))</td>
</tr>
<tr>
<td>7</td>
<td>[\lambda_p(1 - \sum_{i \in I} p^1_i a_i) + (1 - \lambda_p)\sum_{i \in I} p^1_i a_i p_G]</td>
</tr>
<tr>
<td>8</td>
<td>[\lambda_p(1 - \sum_{i \in I} p^1_i a_i) + (1 - \lambda_p)\sum_{i \in I} p^1_i a_i(1 - p_G)]</td>
</tr>
<tr>
<td>9</td>
<td>(p_B)</td>
</tr>
<tr>
<td>10</td>
<td>(p_B)</td>
</tr>
<tr>
<td>11</td>
<td>[\lambda_p(1 - \sum_{i \in I} p^1_i a_i) p_G]</td>
</tr>
<tr>
<td>12</td>
<td>[\lambda_p(1 - p_B)]</td>
</tr>
<tr>
<td>13</td>
<td>[\lambda_p p_B]</td>
</tr>
</tbody>
</table>

\(\|y_{ps}\|^2\) falls in the \(i\)-th subinterval in the case of active PU. It can be shown that the probability \(p^1_i\) is given by the following equation [10]:

\[
p^1_i = \exp\left(-\frac{(i-1)\eta}{2n_2\sigma^2_1}\right) - \exp\left(-\frac{-\eta}{2n_2\sigma^2_1}\right),
\]

where \(\sigma^2_2\) is the variance of the energy detector’s output when the PU is present.

#### 2.3.2 SU’s spectrum access schemes

In the following subsections, different SU access schemes are presented. First, baseline systems with different types of available feedback information at the PU and the SU are presented. Then, our proposed CQI-based access schemes soft-sensing is presented.

**The baseline systems**

- The No CQI feedback- SD based access system (Baseline system 1 No FB): In this baseline system, the PU has ACK/NACK feedback and has no CQI feedback information. Therefore, the PU transmits its packets regardless of the state of the channel. The SU does not have
access to the PU’s ACK/NACK feedback and accesses the channel in every time slot based on its soft-sensing decision.

- The PU CQI Feedback- SD Based Access System (Baseline System 2 No FB SU): In this system, the PU has both ACK/NACK feedback and CQI feedback of the channel state. If a good CQI feedback is observed, the PU transmits whenever it has packets in its queue. Upon observing a bad PU CQI feedback, the PU backs-off since it knows that the packet will not be received correctly. In this system as well, the SU does not monitor the PU feedback and accesses the PU channel based on a soft-sensing energy detector.

The CQI feedback- SD based access system

In this system, the PU has both ACK/NACK and CQI feedback. The SU only has access to the CQI feedback. The SU accesses the PU channel based on the PU CQI feedback and the soft-decision sensing scheme. If the PU channel is in the good state, and the SU does not detect PU’s existence, the SU accesses the channel with probability $a_i$ if the energy detector soft output lies in the $i$-th subinterval of its decision regions (cf. Fig. 2.11). If the PU channel is in the bad state, the SU accesses the PU channel with probability 1 as the SU is sure that the PU will be backing off even if it has packets to transmit.

2.3.3 Performance analysis

The PU’s queue in the proposed system and the baseline systems can be modeled using the two-dimensional Markov model shown in Fig. 2.3 with transition probabilities that are listed in Table 2.2. The transition probabilities between the different states can be calculated as follows:

- From $(K, G)$ to $(K + 1, G)$: the transition in this case occurs according to the following equation:

$$\Pr\{X(n + 1) = (K + 1, G) \mid X(n) = (K, G)\} = \Pr(\text{a new packet arrives at the PU queue}) \cap (\text{SU does not detect the PU presence and...}$$
decides to access the channel) \cap (the channel in the next time slot remains in the good state) = \lambda_p(\sum_{i \in I} p_i^1 a_i)p_G.

- From \((K, G)\) to \((K + 1, B)\): the events leading to this transition are the same as the transition described above, except that the channel state in the next time slot changes from good to bad. Therefore, this transition probability is given by \(\lambda_p(\sum_{i \in I} p_i^1 a_i)(1 - p_G)\).

- The complete set of transition probabilities of the MC in Fig. 2.3 are listed in Table 2.2.

1. **The steady-state distribution calculation**

In this subsection, we derive the steady-state distribution of the MC shown in Fig. 2.3 with transition probabilities given in Table 2.2; this, in turn, enables us to derive expressions for the SU throughput of the proposed system as well as the baseline systems.

For this system, the matrices \(B, A_0, A_1, A_2\) of the transition matrix, shown in equation (2.3), are listed as follows:

\[
B = \begin{pmatrix}
((1 - \lambda_p) + \lambda_p(1 - \sum_{i \in I} p_i^1 a_i))p_G & (1 - \lambda_p)(1 - p_B) \\
((1 - \lambda_p) + \lambda_p(1 - \sum_{i \in I} p_i^1 a_i))(1 - p_G) & (1 - \lambda_p)p_B
\end{pmatrix}.
\]

\[
A_0 = \begin{pmatrix}
(1 - \lambda_p)(1 - \sum_{i \in I} p_i^1 a_i)p_G & 0 \\
(1 - \lambda_p)(1 - \sum_{i \in I} p_i^1 a_i)(1 - p_G) & 0
\end{pmatrix}.
\]

\[
A_1 = \begin{pmatrix}
[\lambda_p(1 - \sum_{i \in I} p_i^1 a_i) + (1 - \lambda_p)\sum_{i \in I} p_i^1 a_i]p_G & (1 - \lambda_p)(1 - p_B) \\
[\lambda_p(1 - \sum_{i \in I} p_i^1 a_i) + (1 - \lambda_p)\sum_{i \in I} p_i^1 a_i](1 - p_G) & (1 - \lambda_p)p_B
\end{pmatrix}.
\]

\[
A_2 = \begin{pmatrix}
\lambda_p\sum_{i \in I} p_i^1 a_i p_G & \lambda_p(1 - p_B) \\
\lambda_p\sum_{i \in I} p_i^1 a_i(1 - p_G) & \lambda_p p_B
\end{pmatrix}.
\] (2.27)
By solving equation (2.6), but with $A_0$, $A_1$, and $A_2$ listed in equation (2.27), the elements of the matrix $R$ can be calculated as follows:

$$
\begin{align*}
 r_{11} &= \frac{\lambda_p \sum_{i \in I} p_i a_i}{(1 - \lambda_p)(1 - \lambda_p + \lambda_p p_B + \lambda_p p_G - 1)}; \\
 r_{12} &= \frac{\lambda_p \sum_{i \in I} p_i a_i (1 - p_G)}{(1 - \lambda_p)(1 - \lambda_p + \lambda_p p_B - p_B - \lambda_p p_G + 1)}; \\
 r_{21} &= \frac{\lambda_p \sum_{i \in I} p_i a_i (1 - p_G)}{(1 - \lambda_p)(1 - \lambda_p + \lambda_p p_B - p_B - \lambda_p p_G + 1)}; \\
 r_{22} &= \frac{\lambda_p \sum_{i \in I} p_i a_i (1 - p_G)}{(1 - \lambda_p)(1 - \lambda_p + \lambda_p p_B - p_B - \lambda_p p_G + 1)}.
\end{align*}
$$

(2.28)

where $B_5$ is shown in (2.29).

$$
B_5 = \frac{\lambda_p (\lambda_p + \sum_{i \in I} p_i a_i + p_B - \lambda_p \sum_{i \in I} p_i a_i - \lambda_p p_B - \lambda_p p_G - \sum_{i \in I} p_i a_i p_B - \sum_{i \in I} p_i a_i p_B - \sum_{i \in I} p_i a_i p_B - \lambda_p \sum_{i \in I} p_i a_i p_B + \lambda_p \sum_{i \in I} p_i a_i p_G)}{(1 - \lambda_p)(1 - \lambda_p + \lambda_p p_B - p_B - \lambda_p p_G + 1)}.
$$

(2.29)

The balance equation around state $(0, G)$ is given by:

$$
\begin{align*}
\pi_0^G (1 - (\lambda_p + \sum_{i \in I} p_i a_i) p_G + (1 - \lambda_p) p_G)) = \\
\pi_0^B (1 - \lambda_p + \sum_{i \in I} p_i a_i) (1 - p_B) + \pi_1^G (1 - \lambda_p + \sum_{i \in I} p_i a_i) (1 - p_G).
\end{align*}
$$

(2.30)

The balance equation around state $(0, B)$ is given by:

$$
\begin{align*}
\pi_0^B (1 - (1 - \lambda_p) p_B) = \\
\pi_0^G (1 - \lambda_p + \sum_{i \in I} p_i a_i) (1 - p_B) + (1 - \lambda_p) (1 - p_G) + \\
\pi_1^G (1 - \lambda_p + \sum_{i \in I} p_i a_i) (1 - p_G).
\end{align*}
$$

(2.31)

Jointly solving (2.30) and (2.31), we get

$$
\pi_0^B = \frac{(1 - p_G) \pi_0^G}{\lambda_p p_B - p_B - \lambda_p + \lambda_p p_G + 1},
$$

(2.32)

and $\pi_0^G$ is as shown in (2.33).

$$
\pi_0^G = \frac{\lambda_p p_B - \sum_{i \in I} p_i a_i - p_B - 2 \lambda_p + \lambda_p p_B + \sum_{i \in I} p_i a_i p_B + 1}{(1 - \lambda_p)(1 - \sum_{i \in I} p_i a_i)(2 - p_B - p_G)}.
$$

(2.33)
2. **SU throughput analysis**

In this subsection, we derive closed-form expressions for the SU throughput of the proposed system and the baseline systems.

- **The No CQI Feedback- SD based access system**

  In this system, the PU has ARQ feedback only, and the SU does not have access to this feedback. Therefore, the PU always accesses the channel when it has packets in its queue even if the channel is in the bad state; the SU channel access decisions are based only on its soft-sensing outcome.

  In a given time slot, the SU will be able to transmit a packet without collision with the PU if the following conditions are met: i) the PU queue is empty (the MC is in state $(0, G)$ or $(0, B)$), ii) the PU does not receive a new packet in this time slot, iii) and the SU correctly detects the PU’s absence and decides to access the channel (this event has a probability $\sum_{i \in I} p_i^0 a_i$). Hence, the SU throughput in this system, $\mu_{s1}$, is given by

  $$\mu_{s1} = \left( \sum_{i \in I} p_i^0 a_i \right) (1 - \lambda_p) (\pi_0^G + \pi_0^B).$$

  The complete expression for $\mu_{s1}$ is given in (2.34).

  $$\mu_{s1} = \frac{B_6}{(1 - \sum_{i \in I} p_i^1 a_i)(2 - p_B - p_G)(\lambda_p p_B - p_B - \lambda_p + \lambda_p p_G + 1)},$$

  where $B_6$ is shown in the following equation:

  $$B_6 = \sum_{i \in I} p_i^0 a_i (\lambda_p + p_B + p_G - \lambda_p p_B - \lambda_p p_G - 2)(\lambda_p p_B - \sum_{i \in I} p_i^1 a_i - p_B - 2\lambda_p + \lambda_p p_G + \sum_{i \in I} p_i^1 a_i p_B + 1)$$

  (2.35)

- **The PU CQI feedback- SD based access system**

  In this system, the PU has both ARQ and CQI feedback, but the SU does not have access to any of the PU feedback. Given the
CQI feedback, the PU does not transmit packets if the channel is in the bad state. The SU accesses the channel based on its spectrum soft-sensing operation only.

In a given time slot, the SU will succeed in transmitting a packet without collision with the PU in the following two cases. i) The PU channel is in the bad state (the MC is in state \((K, B)\) for some \(K = 0, 1, 2...\) and the SU successfully detects PU’s absence and decides to access the channel (this event has a probability \(\sum_{i \in I} p_i^0 a_i\)). ii) The PU channel is in the good state, the PU queue is empty and the SU successfully detects PU’s absence and decides to access the channel. Therefore, the SU throughput in this system, \(\mu_{s2}\), can be written as

\[
\mu_{s2} = \sum_{i \in I} p_i^0 a_i \left[ (1 - \lambda_p)\pi_0^G + \sum_{k=0}^{\infty} \pi_k^B \right]
\]

\[
= \sum_{i \in I} p_i^0 a_i \left[ (1 - \lambda_p)\pi_0^G + \sum_{k=0}^{\infty} \pi_k^B \right]^{-1} \left( \pi_0^G \right) \left( I_2 - R \right)^{-1} \left( \pi_0^B \right)
\]

(2.36)

- The proposed CQI feedback- SD based access system

In this system, the PU has both ARQ and CQI feedback, and the SU has access to the PU CQI feedback. Therefore, the SU accesses the channel with probability 1 when the PU’s channel is in the bad state since the PU will be backing-off in this case. When the PU channel is in the good state, the SU relies on its spectrum soft-sensing for channel access decisions.

The SU will successfully transmit a packet in the following cases: i) with probability 1 if the PU channel is in the bad state (the MC is in state \((K, B)\) for some \(K = 0, 1, 2...\)), ii) with probability \(\sum_{i \in I} p_i^0 a_i\) when the PU queue is empty (the MC is in state \((0, G)\)). Note that the event where the PU queue is empty and the channel is in the bad state (state \((0, B)\)) is already included in the first case. Therefore, the SU throughput in this system, \(\mu_{s3}\),
can be expressed as

\[
\mu_{s3} = \sum_{i \in I} p_i^0 a_i (1 - \lambda_p) \pi_i^G + \sum_{k=0}^{\infty} \pi_k^B
\]

\[
= \sum_{i \in I} p_i^0 a_i (1 - \lambda_p) \pi_i^G + [0 \quad 1] (I_2 - R)^{-1} \begin{pmatrix} \pi_i^G \\ \pi_i^B \end{pmatrix}
\]

\[
= \frac{B_T}{(1 - \sum_{i \in I} p_i^0 a_i) (2 - p_B - p_G)}.
\]

where \( B_T \) is given in equation (2.38).

\[
B_T = \left( \sum_{i \in I} p_i^0 a_i - \sum_{i \in I} p_i^1 a_i - p_G - 2 \lambda_p \sum_{i \in I} p_i^0 a_i - \sum_{i \in I} p_i^1 a_i - \sum_{i \in I} p_i^0 a_i - \sum_{i \in I} p_i^1 a_i p_B \right)
\]

\[+ \left( \sum_{i \in I} p_i^1 a_i p_G + \lambda_p \sum_{i \in I} p_i^0 a_i p_B + \lambda_p \sum_{i \in I} p_i^1 a_i p_B + \sum_{i \in I} p_i^1 a_i \right). \]

3. PU delay analysis

In this subsection, we derive expressions for the average PU packet delay of our proposed system and the baseline systems using the PU delay equation shown in equation (2.23).

Substituting in equation (2.23) with the \( R \) matrix, \( \pi_0^G \), and \( \pi_0^B \), we get the closed-form expressions of the average PU packet delay of the two baseline systems and the proposed system and they are given in equation (2.39).

\[
D_p = \frac{B_8}{(2 - p_B - p_G)(\lambda_p p_B - \sum_{i \in I} p_i^1 a_i - p_B - 2 \lambda_p + \lambda_p p_G + \sum_{i \in I} p_i^1 a_i p_B + 1)},
\]

where \( B_8 \) is shown in the following equation:

\[
B_8 = \lambda_p \left( 3 \sum_{i \in I} p_i^1 a_i - p_G - 4 \sum_{i \in I} p_i^1 a_i p_B - 3 \sum_{i \in I} p_i^1 a_i p_B + \sum_{i \in I} p_i^1 a_i p_B^2 \right)
\]

\[+ \left( \sum_{i \in I} p_i^1 a_i p_G + 2 \sum_{i \in I} p_i^1 a_i p_B p_G + 1 \right). \]

4. Optimal SU channel access probabilities

The optimal channel access probabilities \( a_i \)'s are obtained by maximizing the SU’s throughput, \( \mu_{s,j} \), \( j = 1, 2, 3 \), while guaranteeing the
PU’s queue stability. Stability of the PU queue is determined by the value of $\pi_0^G$ and $\pi_0^B$. If these probabilities are greater than zero, it means that the probability of the PU queue being empty is also greater than zero. Therefore, the problem can be formulated as follows:

$$\max_{a_i, i \in I} \mu_{xj},$$

subject to

$$\pi_0^G > 0 \text{ and } \pi_0^B > 0.$$ 

For the proposed scheme, it can be shown that the SU throughput given by equation (2.37) is neither convex nor concave. Therefore, we used a grid search \(^3\) to find the optimal access probabilities $a^*_i$'s that satisfy the PU’s queue stability constraint.

For the two baseline systems, it can be shown that the logarithm of their throughput functions given by (2.34) and (2.36) is the difference between two concave functions. The difference between two concave functions is not concave in general, so we rely on the algorithm developed in [10] to find the optimal access probabilities $a^*_i$'s that maximize the SU throughput in these systems. The algorithm in [10] removes the nonconcave term and replace it with a constant term, for which an exhaustive search for a single parameter is applied. This constant parameter has a limited value, for each value of this parameter a convex optimization problem is solved using any standard convex optimization solver.

### 2.3.4 Multi-user case

In this section, we extend our model to incorporate multiple PUs and multiple SUs. We consider a primary network with $M_p$ PUs employing Time

\(^3\)Grid search is a simple exhaustive numerical search. The optimization domain is divided into small steps and exhaustive search is done over these points. This is applicable for the case of small number of optimization parameters. For bigger problems, we can use any other heuristic approach to find a solution for our optimization problem.
2.3. CQI feedback- soft detection (SD) based access system

<table>
<thead>
<tr>
<th>Number</th>
<th>Transition Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\left(1 - \lambda_p\right) + \lambda_p (1 - \sum_{i \in I} P_i a_i) M_s) p_G$</td>
</tr>
<tr>
<td>2</td>
<td>$\left(1 - \lambda_p\right)(1 - p_B)$</td>
</tr>
<tr>
<td>3</td>
<td>$(1 - \lambda_p)p_B$</td>
</tr>
<tr>
<td>4</td>
<td>$\left((1 - \lambda_p) + \lambda_p (1 - \sum_{i \in I} P_i a_i) M_s) (1 - p_G)\right)$</td>
</tr>
<tr>
<td>5</td>
<td>$\left(1 - \lambda_p\right)(1 - \sum_{i \in I} P_i a_i) M_s p_G$</td>
</tr>
<tr>
<td>6</td>
<td>$\left(1 - \lambda_p\right)(1 - \sum_{i \in I} P_i a_i) M_s (1 - p_G)$</td>
</tr>
<tr>
<td>7</td>
<td>$\lambda_p (1 - \sum_{i \in I} P_i a_i) M_s + (1 - \lambda_p)(1 - (1 - \sum_{i \in I} P_i a_i) M_s) p_B$</td>
</tr>
<tr>
<td>8</td>
<td>$\left(1 - \lambda_p\right)(1 - p_B)$</td>
</tr>
<tr>
<td>9</td>
<td>$(1 - \lambda_p)p_B$</td>
</tr>
<tr>
<td>10</td>
<td>$\lambda_p (1 - \sum_{i \in I} P_i a_i) M_s p_B$</td>
</tr>
<tr>
<td>11</td>
<td>$\lambda_p (1 - \sum_{i \in I} P_i a_i) M_s (1 - p_B)$</td>
</tr>
<tr>
<td>12</td>
<td>$\lambda_p (1 - p_B)$</td>
</tr>
<tr>
<td>13</td>
<td>$\lambda_p (1 - p_B)$</td>
</tr>
<tr>
<td>14</td>
<td>$\lambda_p p_B$</td>
</tr>
</tbody>
</table>

Division Multiple Access (TDMA) as a multiple access protocol, and a secondary network having $M_s$ SUs employing random access scheme (slotted-ALOHA) \(^4\) to access the PUs’ channel. The use of TDMA in the primary network ensures that PUs will not interfere with each other since every time slot will have at most one active PU. Therefore, the queue of any PU can be modeled using the two-dimensional MC model shown in Fig. 2.3, where the transitions correspond to the time slots allocated to this PU. We assume that the PU arrival process is modeled as a Bernoulli process between the time slots allocated to the user, i.e., a PU can have zero or one packet arrival per frame. However, the transition probabilities of the MC in the case of multiple SUs will be different and are given in Table 2.3. For simplicity of presentation, we have assumed symmetry among the SUs. Following the same analysis of the single SU case, the steady-state distribution of the MC can be derived. The probability $\pi_G^0$ is given in (2.41), and $\pi_B^0$ can be

\(^4\)ALOHA is a random multiple access scheme to organize the transmissions of two or more nodes over a shared medium. In ALOHA, each node attempts transmission (with some access probability) whenever it has packets in its queue. If more than one node transmit simultaneously a collision occurs, and the colliding packets are assumed to be lost. Each node, however, has the ability to determine which packet is lost and to attempt to re-transmit it after some random back-off time.
obtained using equation (2.32).

\[ \pi G \frac{G_0}{0} = \lambda p p B + (1 - \sum_{i \in I}^1 p_i a_i)^{M_s} - 2 \lambda p + \lambda p p G + (1 - \sum_{i \in I}^1 p_i a_i)^{M_s} p B . \]

(2.41)

1. SUs throughput analysis

In this subsection, we characterize the throughput of a typical SU. From the symmetry assumption, the throughput will be the same for all SUs. In the following systems, each SU in the secondary network employs soft spectrum sensing to detect the presence of PUs and take the channel access decisions.

- **The No CQI feedback- SD based access system**

In this system, we assume that the primary network does not support CQI feedback. A SU will successfully transmit a packet in a given time slot if i) the queue of PU owning this time slot is empty, and a new packet does not arrive at this queue by the start of the time slot (these two events have a joint probability \((1 - \lambda p)(\pi G_0^0 + \pi B_0^0)\)), ii) the SU successfully detects that the time slot is idle and gains access to the channel (the probability of this event is \(\sum_{i \in I}^1 p_i^0 a_i\)), and iii) all other SUs do not access the channel (this event has a probability \((1 - \sum_{i \in I}^1 p_i^0 a_i)^{M_s} - 1\)). The per-user SU throughput is given by

\[ \mu s_1 = M p \left( \sum_{i \in I}^1 p_i^0 a_i \right)(1 - \sum_{i \in I}^1 p_i^0 a_i)^{M_s - 1} (1 - \lambda p)(\pi G_0^0 + \pi B_0^0). \]

(2.42)

- **The PU CQI feedback- SD based access system**

In this system, the primary network has CQI feedback but the secondary network does not have access to this feedback. PUs back-off from transmission in case of bad channels. Therefore, whenever the PUs are backing-off the SUs will be able to access the channel when it is sensed idle. The PU owning the current
2.3. CQI feedback- soft detection (SD) based access system

time slot will be backing-off with probability \( \sum_{k=0}^{\infty} \pi_k^B \). Therefore, the SU throughput in this system is given by

\[
\mu_{s2} = M_p \left( \sum_{i \in I} p_i^0 a_i \right) (1 - \sum_{i \in I} p_i^0 a_i)^{M_s-1} \left( (1 - \lambda_p) \pi_0^G + \sum_{k=0}^{\infty} \pi_k^B \right).
\]

(2.43)

- The proposed CQI feedback- SD based access system

In this system, the SUs have access to the PUs CQI feedback. In this case, the SUs will know when a PU is backing-off in his assigned time slots. If the SUs decide to access the PU channel with probability 1, they will collide with each other. Therefore, when a PU is backing-off, the SUs try to access the channel in this PU’s time slot with random access probability \( p_s \) to limit the collisions between SUs. The throughput of a typical SU, in this case, is given by

\[
\mu_{s3} = M_p \left( \sum_{i \in I} p_i^0 a_i \right) (1 - \sum_{i \in I} p_i^0 a_i)^{M_s-1} (1 - \lambda_p) \pi_0^G
\]

\[
+ p_s (1 - p_s)^{M_s-1} \sum_{k=0}^{\infty} \pi_k^B.
\]

(2.44)

The SU random access probability \( p_s \) is selected to maximize the SU throughput. It can be readily proved that the SU optimum access probability equals \( \frac{1}{M_s} \).

2. Access Probabilities Calculation

In the system with multiple PUs and SUs, the SUs access probabilities \( a_i \)'s are chosen to maximize the throughput of a typical SU under the constraint of keeping all PUs queues stable. This optimization problem is formulated as follows:

\[
\max_{a_i, i \in I} \mu_{s,j},
\]

subject to

\( \pi_{0j}^G > 0 \) and \( \pi_{0j}^B > 0 \), for \( j = 1, 2, ..., M_p \).
Chapter 2. CQI feedback based access system

This optimization problem can be solved using the same method described in Section 2.3.3. From the symmetry of SUs assumption, all SUs will have the same optimal access probabilities $a_i$s obtained from the solution of the optimization problem above. Note that, as it will become clear in the simulation section, the per-user SU throughput in the multiuser case will be less than the throughput of the single SU case. This stems from the fact that multiple users are now contending for the channel resources. Furthermore, the sum throughput of the SUs will still be less than the throughput in the single SU case because some of the spectrum opportunities will be lost due to collisions between SUs.

2.3.5 Effect of CQI erasures at the SUs

In this section, we consider the effect of imperfect feedback channel at the SU. To isolate the effects of imperfect feedback channel, we are assuming a single PU and a single SU. The CQI feedback channel is modeled as an erasure channel\(^5\). The PU’s CQI feedback channel is still assumed to be error-free. The CQI feedback erasure channel model is shown in Fig. 2.12. Any CQI feedback can be erased by an erasure probability $p_e$. Whenever the CQI feedback is not received at the SU due to an erasure, it is interpreted as a good CQI. This means that the SU will sense the PU activity before

\(^5\)The feedback erasure channel can be used to model many practical channels. For example, for a general fading channel model, the receiver can set a certain SNR threshold below which an erasure is declared and no decoding is attempted; above this threshold, the receiver is assumed to always decode correctly. Another practical system that can be modeled by an erasure channel is a system where the feedback information is protected by a strong error detection code (e.g., a strong cyclic redundancy check (CRC)). If the received CQI packet fails CRC check an erasure is declared and this CQI packet is discarded.
2.3. CQI feedback- soft detection (SD) based access system

attempting to access the channel. Of course, this makes sense because if the SU has no access to the PU CQI information the system should fall back to the sensing-based access scheme.

The PU’s queue, in this case, can be still modeled as the two dimensional MC shown in Fig. 2.3 with the transition probabilities shown in Table 2.2. Since erasures are interpreted as good CQI at the SU, a good CQI feedback will always be received correctly. However, erased bad CQI’s will be interpreted as good CQI’s feedback. This will not affect the PU’s performance since, in the bad channel state, the PU will not be able to deliver any packets to its destination irrespective of the SU access decisions.

At the SU, erasure of bad CQI’s will result in the loss of some of the spectrum opportunities. This is because in case of a bad CQI erasure SU will have to sense the channel before accessing it, and sensing errors will result in lost opportunities. Clearly, the CQI feedback erasure channel at the SU has no effect on the performance of the two baseline systems as the SU does not have access to the CQI feedback in these systems. However, the performance of our proposed will be affected by CQI erasures. By considering the effect of the CQI erasure events, the SU’s throughput of the proposed system with CQI erasure channel, $\mu_{se}$, is given by:

$$
\mu_{se} = \left( \sum_{i \in I} p^0_i a_i \right) (1 - \lambda_p) \pi^G_0 + (1 - p_e) \sum_{k=0}^{\infty} \pi^B_k \right) + p_e \left( \sum_{i \in I} p^0_i a_i \right) \left( \sum_{k=0}^{\infty} \pi^B_k \right).
$$

(2.45)

2.3.6 Performance results

In this section, we will provide a comparative study in terms of the SU throughput and the PU delay of the proposed system, the two baseline systems, the CQI Feedback- HD based access system and the perfect sensing CQI feedback based access system. The system parameters are as follows; the transmit power is set to 30 mW, the path loss exponent $\gamma = 3.7$, and the noise power spectral density $N_0 = 6 \times 10^{-11} W/Hz$. The energy detector’s threshold $\eta$ is chosen by the Neyman-Pearson (N-P) detector design
Chapter 2. CQI feedback based access system

Figure 2.13: The SU throughput for different access schemes

\[ M_p = 1, M_s = 1, p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1 \].

Figure 2.14: The SU throughput for different access schemes

\[ M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_g = 0.6, p_d = 0.9, p_f = 0.1 \].
2.3. CQI feedback- soft detection (SD) based access system

![CQI Feedback SD System](image)

**Figure 2.15:** The SU throughput for different access schemes

\[(M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_G = 0.6, p_d = 0.7, p_f = 5.8717 \times 10^{-4})\]

rule, which maximizes the detection probability under a constraint on the false-alarm probability [38]. For the soft-sensing case, the region below the threshold is divided into \( n = 4 \) regions each having a different access probability \( a_i, i = 1, \ldots, 4 \).

First, we consider the system with one PU and one SU. In Fig. 2.13, Fig. 2.14, and Fig. 2.15, the SU throughput is plotted against the PU arrival rate for the different access schemes. It can be seen that, the SU throughput for the two baseline systems equals to \((1 - p_f)\), at zero PU arrival rate. Moreover, when the PU CQI feedback is available at the SU, the proposed scheme and the CQI Feedback- HD based access system achieve SU throughput equals to \( \zeta_B + \zeta_G (1 - p_f) \) at zero PU arrival rate.

In all the considered systems, the achievable SU’s throughput decreases linearly with the increase of the PU’s arrival rate. This is because with the increase of the PU arrival rate, the PU will be accessing the channel more often resulting in less spectrum opportunities for the SU. For the first baseline system where there is no CQI feedback at all, the SU throughput reaches zero for high PU arrival rate. It can be seen that the maximum PU arrival
rate beyond which the SU throughput is zero is equal to \( \zeta_g p_d \), which is the product of the probability that the channel in the good state and the probability that the SU correctly detects the PU’s presence. For all the other systems which have PU CQI feedback, the SU throughput reaches a floor not equal to zero. Since the PU is backing-off in the bad channel state, there will always be idle time slots for the SU to utilize. In the baseline system 2, where the SU does not have access to the PU CQI, this floor is equal to \( \zeta_B (1 - p_f) \). However, in the proposed system and the CQI Feedback- HD based access system where the SU uses the PU CQI feedback, the floor is equal to \( \zeta_B \) since the SU access the channel with probability 1 in case of bad CQI in these two systems.

In Fig. 2.13 and Fig. 2.14, it is noted that the proposed scheme has comparable performance to the CQI Feedback- HD based access system. When the system has a relatively high probability of detection (0.9 in Fig. 2.13 and Fig. 2.14), soft-sensing has no or little performance gain over hard-sensing. As the probability of detection increases, the value of the decision threshold \( \eta \) required to achieve this probability of detection decreases, hence, the
2.3. CQI feedback- soft detection (SD) based access system

Figure 2.17: The SU optimal access probabilities for the PU CQI Feedback- SD based access system 
\((M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_G = 0.6, p_d = 0.9, p_f = 0.1)\).

width of interval \([0, \eta]\) decreases. In this case, the differences between the access probability in the different subinterval of \([0, \eta]\) become very small. Therefore, the effect of soft-sensing is less prominent. Fig. 2.16 and Fig. 2.17 show the optimal SU access probability in the proposed access scheme with soft-sensing for the same system parameters of Fig. 2.13 and Fig. 2.14, respectively. It can be noted that for most of the PU arrival rate values, the access probabilities in all the subintervals are all equal to one. As the PU arrival rate increases the values of the access probabilities in the intervals closer to the threshold \(\eta\) start to decrease until they reach zero. For high values of the PU arrival rate, where the PU queue stability cannot be maintained, all the access probabilities are equal to zero. It can also be noted that the access probability \(a_1\) has the highest value compared to the other access probabilities; this is because the energy received at the SU in the \(a_1\) region lies far from the threshold \(\eta\), therefore, the SU becomes more aggressive in accessing the PU channel, as the PU is more likely to be idle.

Fig. 2.15 shows the SU throughput when the threshold \(\eta\) is designed for a lower probability of detection. Here, the gain of the soft-sensing approach
over the hard-sensing one is clear. This is because as shown in Fig. 2.18, the access probabilities are optimally selected to counter the effect of the low probability of detection and protect the PU from possible collisions with SU transmissions. This protection can be seen when we look at the value of the access probabilities $a_3$ and $a_4$, which correspond to the two subintervals closer the threshold $\eta$. The optimal values for the access probabilities $a_3$ and $a_4$ are equal to zero for almost all values of the PU arrival rate to avoid collisions with the PU. Since measurements that fall in these subintervals are those which most probably cause miss-detection events. Finally, we can always think of the hard-sensing scheme as a special case of the soft-sensing scheme, where all access probabilities are assumed to be the same. This is why we see gains from applying the soft-sensing scheme whenever the optimal access probabilities are not supposed to be equal.

In Fig. 2.19, Fig. 2.20 and Fig. 2.21, the PU packet delay is plotted against the PU arrival rate for the different access schemes. It can be seen that with more feedback information at the SU, it will have better predictions of the PU activity, which leads to fewer collisions resulting in lower
2.3. CQI feedback- soft detection (SD) based access system

PU delays. As clear from these figures, the two baseline systems have the highest PU delay. The SU does not have extra information in the second baseline system compared to the first baseline system; therefore, the PU delays in the two baseline systems are equal. It can also be readily seen that the proposed soft-sensing scheme improves the PU delay compared to the hard-sensing scheme, especially in the case of lower probability of detection. Finally, the perfect sensing system has the lowest PU delay among all systems since in this system, the SU presence is completely transparent to the PU with no collisions.

The SU throughput in the single user case with CQI feedback erasures at the SU is depicted in Fig. 2.22 and Fig. 2.23 with the probability of erasure $p_e$ set to 0.1 and 0.3, respectively. It is noted that, for zero PU arrival rate, the SU throughput is $(1 - p_f)\zeta_G + (1 - p_f)p_e\zeta_B + (1 - p_e)\zeta_B$ in case of erasures compared to $(1 - p_f)\zeta_G + \zeta_B$ in case of no erasures. And the SU arrival rate floor for high PU arrival rates is $(1 - p_e)\zeta_B$ in case of erasures compared to $\zeta_B$ in case of no erasures. Comparing the SU throughput in Fig. 2.22 with the no erasure case in Fig. 2.13, it can be noted that, for the
selected set of parameters, the performance of the CQI Feedback- HD based access system and CQI Feedback- SD based access system is not affected by the relatively small erasure probability of 0.1. However, when the erasure probability is increased to 0.3, the performance of these two systems is degraded slightly, with the soft-sensing system being less susceptible to the feedback erasures. However, exploiting the available, limited CQI feedback information will always result in performance gains if compared to systems that do not exploit them.
2.3. CQI feedback- soft detection (SD) based access system

\[ (M_p = 1, M_s = 1, p_B = \zeta_B = 0.4, p_G = \zeta_G = 0.6, p_d = 0.7, p_f = 5.8717 \times 10^{-4}). \]

\[ (M_p = 1, M_s = 1, p_B = \zeta_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_g = 0.875, p_d = 0.9, p_f = 0.1, p_e = 0.1). \]
Figure 2.23: The SU throughput for erasure effect
\((M_p = 1, M_s = 1, p_B = 0.3, \zeta_B = 0.125, p_G = 0.9, \zeta_G = 0.875, p_d = 0.9, p_f = 0.1, p_e = 0.3)\).
Chapter 3

Hybrid feedback-based access scheme

3.1 Introduction

In this chapter, CR system is studied in which the SU leverages the PU’s CQI and the PU’s ARQ. The SU randomly accesses the PU channel with access probabilities based on its spectrum sensing outcome and the PU feedbacks. We consider hard sensing and soft sensing. SU’s access probabilities are selected through an optimization problem with the objective to maximize the SU’s throughput while ensuring the stability of the PU’s packet queue. This system is modeled using a three-dimensional MC. This model enabled us to derive a closed-form expression for the SU’s throughput, which is used in the throughput maximization problem.

3.2 ARQ-CQI feedback-hard detection (HD) based access scheme

3.2.1 System model

We consider a CR system has the system model, shown in Fig. 3.1.

Channel model

The same channel model that were described in Section 2.2.1 is considered.
3.2.2 SU spectrum access schemes

The No feedback system

In this system, the PU has ACK/NACK ARQ feedback but has no CQI feedback information. Therefore, the PU transmits its packets regardless of the channel state. In this system it is assumed that the SU does not access the PU feedback. Therefore, it accesses the channel with an access probability \( a_s \) in every time slot based on its spectrum sensing.

The ARQ feedback-based system

In this system, the SU has access to the PU’s ARQ feedback. Observing a NACK, the SU backs-off since it knows that the PU will re-transmit its undelivered packet from the previous time slot, thus avoiding sure collisions with the PU. However, upon hearing an ACK, the SU accesses the channel with an access probability \( a_s \) based on its spectrum sensing.

The CQI feedback-based system

In this system, the PU has ARQ feedback, and CQI feedback of the channel state, which is an indicator of how good/bad the channel between the PU transmitter and receiver is. If a good CQI feedback is observed, the PU transmits whenever it has packets in its queue. Observing a bad PU CQI feedback, the PU backs-off since it knows that the packet will not be received correctly. The SU monitors only the PU CQI feedback. Hearing a bad PU CQI feedback, the SU accesses the channel with probability 1. If
3.2. ARQ-CQI feedback- hard detection (HD) based access scheme

The SU observes a good PU CQI feedback, it accesses the channel with an access probability $a_s$ based on its spectrum sensing.

Proposed hybrid feedback HD-based access system

The proposed system model is shown in Fig. 3.1, in which the PU has both CQI and ARQ feedbacks; the SU listens to these two types of feedback. The SU accesses the channel depending on the hard-decision sensing scheme and the PU feedbacks. If a good CQI feedback and ACK message is observed, and the SU does not detect the PU’s existence, the SU accesses the channel with access probability $a_s$. If a bad CQI feedback is observed, the SU exploits the knowledge that the PU will be in “back-off” state during the next time slot to transmit with probability 1 irrespective of the ARQ feedback. In the case of PU NACK with a good CQI, the SU backs-off to allow for collision-free transmission for the PU. In the case of PU NACK with bad CQI, the SU accesses the channel with probability 1, as the SU knows for sure that the PU will be in the “back-off” state, although it has packets to transmit.
### 3.2.3 Performance analysis

The PU’s queue in the proposed system is modeled by the three-dimensional MC \( \{X(n), n = 0, 1, 2, \ldots\} \) shown in Fig. 3.2, whose state space is given by \( S=\{(K,D,T) : K = 0, 1, 2, \ldots; D \in \{F,R\}; T \in \{G,B\}\} \), where \( K \) is the number of PU packets in the queue, \( D \) is the ARQ feedback, where \( F \) means that the packet at the head of the PU queue is being transmitted for the first time, while \( R \) means that the packet at the head of the PU is being re-transmitted due to failure in the previous time slot. Finally, \( T \) is the CQI feedback, where \( G \) means that the PU channel is in the good state and \( B \) means that it is in the bad state.

The transitions between states are as follows:

- From \((K, F, G)\) to \((K - 1, F, G), K > 0\): the transition in this case occurs according to the following equation:
  \[
  \Pr(X(n+1) = (K - 1, F, G) \mid X(n) = (K, F, G)) = \Pr((\text{no new packet arrives at the PU queue}) \cap (\text{SU does not detect the PU presence and decides not to access the channel}) \cap (\text{the channel in the next time slot remains in the good state})) \cup \Pr(\text{no new packet arrives at the PU queue}) \cap (\text{SU detects the PU presence}) = (1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)p_g.
  \]

- From \((K, F, G)\) to \((K - 1, F, B), K > 0\): it is the same as the previous transition but \( p_g \) is replaced by \( 1 - p_g \). Therefore the transition probability equals to \((1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)(1 - p_g)\).

- From \((K, R, G)\) to \((K, F, G), K > 0\): as the MC is in \((K, R, G)\), the packet at the head of the queue will be transmitted successfully with probability 1. The transition in this case occurs according to this equation:
  \[
  \Pr(X(n+1) = (K, F, G) \mid X(n) = (K, R, G)) = \Pr((\text{new packet arrives at the PU queue}) \cap (\text{the channel in the next time slot remains in the good state})) = \lambda_pp_g.
  \]

- The complete transition probabilities of the MC are shown in Table 3.1.

#### 1. The steady state distribution calculation

To get an expression for the SU throughput of the proposed system,
we start by calculating the steady state distribution of the MC shown in Fig. 3.2.

The steady state distribution vector is given by

$$\mathbf{v} = [\pi_0^{FG}, \pi_0^{FB}, 0, 0, \pi_1^{FG}, \pi_1^{FB}, \pi_1^{RG}, \pi_1^{RB}, \ldots].$$

It is clear that the PU can not be in the re-transmission state if it has no packets. Therefore, $\pi_0^{RG} = 0$ and $\pi_0^{RB} = 0$. Define the vector

$$\mathbf{v}_k = \begin{pmatrix} \pi_k^{FG} \\ \pi_k^{FB} \\ \pi_k^{RG} \\ \pi_k^{RB} \end{pmatrix},$$

note that $\mathbf{v}_0 = \begin{pmatrix} \pi_0^{FG} \\ \pi_0^{FB} \\ 0 \\ 0 \end{pmatrix}$.

The state transition matrix of the MC shown in Fig. 3.2 can be written as

$$\Phi = \begin{pmatrix} B & A_0 & 0 & 0 & \ldots \\ A_2 & A_1 & A_0 & 0 & \ldots \\ 0 & A_2 & A_1 & A_0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

(3.1)
where \( B, A_0, A_1, A_2 \) are shown in equation (3.2).

\[
B = \begin{pmatrix}
(1 - \lambda_p) + \lambda_p((1 - a_s)(1 - p_d) + p_d)p_g & (1 - \lambda_p)(1 - p_B) & 0 & 0 \\
(1 - \lambda_p) + \lambda_p((1 - a_s)(1 - p_d) + p_d)(1 - p_g) & (1 - \lambda_p)p_B & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}.
\]

\[
A_0 = \begin{pmatrix}
(1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)p_g & 0 & (1 - \lambda_p)p_g & 0 \\
(1 - \lambda_p)((1 - a_s)(1 - p_d) + p_d)(1 - p_g) & 0 & (1 - \lambda_p)(1 - p_g) & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}.
\]

\[
A_1 = \begin{pmatrix}
\lambda_p((1 - a_s)(1 - p_d) + p_d)p_g & (1 - \lambda_p)(1 - p_B) & \lambda_p p_g & 0 \\
\lambda_p((1 - a_s)(1 - p_d) + p_d)(1 - p_g) & (1 - \lambda_p)p_B & \lambda_p(1 - p_g) & 0 \\
(1 - \lambda_p)a_s(1 - p_d)p_g & 0 & 0 & (1 - \lambda_p)(1 - p_B) \\
(1 - \lambda_p)a_s(1 - p_d)(1 - p_g) & 0 & 0 & (1 - \lambda_p)p_B
\end{pmatrix}.
\]

\[
A_2 = \begin{pmatrix}
0 & \lambda_p(1 - p_B) & 0 & 0 \\
0 & \lambda_p p_B & 0 & 0 \\
\lambda_p a_s(1 - p_d)p_g & 0 & 0 & \lambda_p(1 - p_B) \\
\lambda_p a_s(1 - p_d)(1 - p_g) & 0 & 0 & \lambda_p p_B
\end{pmatrix}.
\]

The state transition matrix \( \Phi \) is a block-tridiagonal matrix; therefore the MC shown in Fig. 3.2 is a homogeneous quasi birth-and-death (QBD) MC. To make the state transition matrix a block-tridiagonal matrix, a transition from \( \pi_{RB0} \) to \( \pi_{RG1} \) and a transition from \( \pi_{RB0} \) to \( \pi_{RB1} \) are added. Adding these transitions do not affect the MC as the probabilities of being in \( \pi_{RG0} \) and \( \pi_{RB0} \) are equal to zero.

The steady state distribution of the MC shown in Fig. 3.2 satisfies the following equation[36]:

\[
v_k = R^k v_0, \quad k > 0,
\]

(3.3)
3.2. ARQ-CQI feedback- hard detection (HD) based access scheme

where the rate matrix $R$:

$$R = \begin{pmatrix}
  r_{11} & r_{12} & r_{13} & r_{14} \\
  r_{21} & r_{22} & r_{23} & r_{24} \\
  r_{31} & r_{32} & r_{33} & r_{34} \\
  r_{41} & r_{42} & r_{43} & r_{44}
\end{pmatrix},$$

is the solution of

$$A_2 + (A_1 - I_4)R + A_0R^2 = 0_{4 \times 4}, \quad (3.4)$$

where $I_4$ is the $4 \times 4$ identity matrix, which can obtained by substituting equation (3.3) in the next equation

$$v_k = A_2v_{k-1} + A_1v_k + A_0v_{k+1}, \quad k \geq 1. \quad (3.5)$$

Equation (3.5) can be easily derived using the states balance equations.

By solving equation (3.4), the matrix $R$ can be obtained. To get the steady state distribution of the MC, the following normalization requirement is applied:

$$\sum_{k=0}^{\infty} (\pi_k^{FG} + \pi_k^{FB} + \pi_k^{RG} + \pi_k^{RB}) = 1,$$

and using equation (3.3), we have

$$\mathbf{I} \left( \sum_{k=0}^{\infty} R^k \right) v_0 = 1, \text{ where } \mathbf{I} = [1 \quad 1 \quad 1 \quad 1].$$

So, $\mathbf{I} \left( \sum_{k=0}^{\infty} R^k \right) v_0 = \mathbf{I}(I_4 - R)^{-1} \begin{pmatrix}
  \pi_0^{FG} \\
  \pi_0^{FB} \\
  0 \\
  0
\end{pmatrix} = 1.$

To get the relationship between $\pi_0^{FG}$ and $\pi_0^{FB}$, the balance equations
around \((0, F, G)\) and \((0, F, B)\) are solved.

The balance equation around state \((0, F, G)\) is given by:

\[
\left[ a_s \lambda_p p_g - p_g - a_s \lambda_p p_d p_g + 1 \right] \pi_0^{FG} - (1 - \lambda_p)(1 - p_B) \pi_0^{FB} \\
+ \left[ p_g (p_d + (1 - a_s)(1 - p_d))(1 - \lambda_p) \right] \pi_1^{FG} \\
+ (1 - \lambda_p) p_g \pi_1^{RG},
\]

(3.6)

The balance equation around state \((0, F, B)\) is given by:

\[
\left[ \lambda_p p_B - p_B + 1 \right] \pi_0^{FB} = \\
\left[ (1 - \lambda_p)(1 - p_g) + \lambda_p ((1 - a_s)(1 - p_d) + p_d)(1 - p_g) \right] \pi_0^{FG} \\
+ \left[ (1 - \lambda_p)(1 - a_s)(1 - p_d) + p_d(1 - p_g) \right] \pi_1^{FG} \\
+ (1 - \lambda_p)(1 - p_g) \pi_1^{RG}.
\]

(3.7)

Solving equation (3.6) and equation (3.7), we get

\[
\pi_0^{FB} = \frac{(1 - p_g) \pi_0^{FG}}{\lambda_p p_B - p_B - \lambda_p + \lambda_p p_g + 1},
\]

(3.8)

and \(\pi_0^{FG}\) is obtained as shown in equation (3.9).

\[
\pi_0^{FG} = \frac{B_8}{(\lambda_p - 1)(p_B + p_g - 2)},
\]

(3.9)

where \(B_8\) is shown in the following equation:

\[
B_8 = \lambda_p p_B - p_B - 2a_s \lambda_p - 2\lambda_p + \lambda_p p_g + 2a_s \lambda_p p_d + a_s \lambda_p p_B + a_s \lambda_p p_g \\
- a_s \lambda_p p_d p_B - a_s \lambda_p p_d p_g + 1.
\]

(3.10)

2. The secondary throughput analysis

- The No feedback system

In this system, the PU has ACK/NACK ARQ feedback but has no CQI feedback information. Therefore, the PU transmits its packets regardless of the channel state. In this system it is assumed that the SU does not access the PU feedback. Therefore, it
accesses the channel with an access probability $a_s$ in every time slot based on its spectrum sensing. The analysis of this system was studied in Section 2.2.3. A closed form expression of the SU throughput for the system with no feedback was shown in equation (2.14).

- The ARQ system

In [10], the authors have done the analysis of ARQ system utilizing the soft sensing scheme. In this work, the soft sensing scheme is converted to the hard sensing scheme so the proposed scheme can be compared with the ARQ system. After modifying the results in [10] to match our model the SU throughput is given by,

$$\mu_s = (1 - \frac{\lambda_p}{1 - (1 - p_d)(1 - \xi_B)a_s})(1 - p_f)a_s. \quad (3.11)$$

The access probability that maximizes the SU service rate $a_s^*$ is obtained by differentiating the SU service rate in equation (3.11) with respect to $a_s$ and equating the result to zero, so

$$a_s^* = \left(\sqrt{\lambda_p} - 1\right)\left(p_B + p_a - 2\right) \frac{(1 - p_f)}{(1 - p_d)(1 - p_B)}. \quad (3.12)$$

- The CQI feedback HD-based access system

In this system, the PU has ARQ feedback, and CQI feedback of the channel state, which is an indicator of how good/bad the channel between the PU transmitter and receiver is. If a good CQI feedback is observed, the PU transmits whenever it has packets in its queue. Observing a bad PU CQI feedback, the PU backs-off since it knows that the packet will not be received correctly. The SU monitors only the PU CQI feedback. Hearing a bad PU CQI feedback, the SU accesses the channel with probability 1. If the SU observes a good PU CQI feedback, it accesses the
Chapter 3. Hybrid feedback-based access scheme

channel with an access probability $a_s$ based on its spectrum sensing. The analysis of this system was studied in Section 2.2.3 and the closed form expression for the system was shown in equation (2.17).

- **The SU Perfect sensing with PU CQI feedback based-access system**

  In this system, which is an upper bound system, the PU has a CQI feedback. The PU accesses the channel if there is a new arrival based on the CQI feedback. The SU accesses the PU channel in the bad channel states with probability 1. When the PU channel is in the good state and the PU’s queue is empty, the SU accesses the channel with probability 1 as well (because of perfect sensing). The analysis of the PU’s queue for the SU perfect sensing with the PU CQI feedback based access system was provided in Section 2.2.3. The same expression for the SU service rate, shown in equation (2.22), is used in this scheme as the ACK does not add information to the SU in the perfect sensing system.

- **The proposed hybrid feedback HD-based access system**

  The closed-form expressions for the SU throughput of the proposed system are derived as follows. In this system, the SU has access to the PU CQI and ARQ feedback messages; therefore, the SU accesses the channel with probability 1 under bad PU CQI feedback irrespective of the PU ARQ messages, since the PU is backing off due to the bad channel. However, under good PU CQI feedback and PU first transmission state the SU accesses the channel with probability $a_s$ if the SU decides that the PU is absent through sensing. Therefore, the SU transmits its packets collision-free in the bad states $(K, F, B)$ and $(K, R, B)$ with probability 1 and in the empty first transmission good state, $(0, F, G)$, with probability $a_s(1 - p_f)$. Hence, the SU throughput in this
The access probability \( a_s \) has to be selected to maximize the secondary throughput, \( \mu_{s,i} \), \( i = 1, 2, 3 \), while keeping the PU queue stable. Stability of the PU queue is determined by the value of \( \pi_{0}^{FG} \) and \( \pi_{0}^{FB} \). If these probabilities are greater than zero, it means that the probability of the PU queue being empty is also greater than zero. Therefore, the problem can be formulated as follows:

\[
\max_{a_s} \mu_{s,i}
\]

subject to

\[
\pi_{0}^{FG} > 0 \text{ and } \pi_{0}^{FB} > 0.
\]

By differentiating the expression of \( \mu_{s,i} \) with respect to \( a_s \) and equating the derivative to zero, the optimal access probability \( a_s^* \) can be obtained. The closed-form expressions of the access probabilities to
maximize the secondary throughput of the proposed system are as follows,

\[ a^*_s = \frac{\lambda_p p_B - p_B - 2\lambda_p p_g + 1}{4\lambda_p - 4\lambda_p p_d - 2\lambda_p p_B - 2\lambda_p p_g + 2\lambda_p p_d p_B + 2\lambda_p p_d p_g}. \] (3.14)

### 3.2.4 Performance results

In this section, we compare the performance of the proposed hybrid feedback SU access scheme with the no-FB scheme, the ARQ FB-based scheme and the CQI FB-based scheme. We also compare it with the performance of the perfect-sensing system where at each time slot the SU is able to sense the channel occupancy without error. Therefore, the perfect sensing system provides an upper bound on the performance of any access scheme.

Fig. 3.3 and Fig. 3.4 depict the SU throughput as a function of the PU arrival rate for the different access schemes. The figures differ in the steady state probability of the channel being in the bad or good states, the probability of detection and the probability of false alarm. At zero PU arrival rate, the perfect sensing scheme achieves a SU throughput of 1 since it can access all the time slots. The hybrid and CQI FB-based schemes benefit from the CQI feedback to access the channel without sensing when the PU channel is in the bad state. Therefore, achieving a throughput of \( \zeta_B + \zeta_G (1 - p_f) \), which is equal to 0.81 in Fig. 3.4. The ARQ FB-based and No-FB schemes are limited by the false alarm rate and will start at the value of 0.7, which is \( 1 - p_f \). At high PU arrival rates, the hybrid and the perfect sensing scheme achieve a minimum SU throughput equal to the steady state probability of the PU channel being in the bad state. This minimum value is guaranteed as the PU is backing off under bad channel conditions allowing the SU to access the channel with probability 1. By combining the information of the CQI and ARQ feedbacks, the proposed hybrid scheme outperforms the other schemes for all values of the PU arrival rate. In Fig. 3.3, the SU throughput for the ARQ FB-based scheme is slightly better than that of the CQI FB-based one for relatively low PU arrival rates. This is because in Fig. 3.3 the probability of the PU channel being in the bad state is 0.125.
3.2. ARQ-CQI feedback- hard detection (HD) based access scheme

Thus, the gain achieved by preventing the collisions with the PU in the ARQ FB-based scheme outweigh the gain of accessing the channel when the PU refrains from accessing it when it is in the bad state in the CQI FB-based scheme. At high PU arrival rates, the only opportunity for the SU to access the channel is when the PU refrains from using it in the bad state, hence, the CQI FB-based schemes performing better than the ARQ FB-based one. In Fig. 3.4 when the the probability of the PU channel being in the bad state is 0.3636, the gain of exploiting the channel when it is in the bad state outweigh that of preventing the collisions with the PU. Therefore, the CQI FB-based scheme outperforms the ARQ FB-based scheme for all values of the PU arrival rates.

In Fig. 3.5 and Fig. 3.6, the SU access probabilities are plotted against the PU arrival rate for different access schemes. The steady state probability of the channel being in the bad state in Fig. 3.5 is 0.125 and in Fig. 3.6 is 0.3636. The results in Fig. 3.5 and Fig. 3.6 show very interesting insight. Comparing the ARQ FB-based scheme with the CQI FB-based, we can see that the access probability of the ARQ FB-based system is higher. This can
be attributed to the fact that in the ARQ FB-based system the SU can be more aggressive in accessing the channel, and under collision, it will go to a back-off state to allow for collision-free transmission from the PU user; this is not the case for the CQI FB-based system, since under collision, there is no guarantee for the PU in the next time-slot. Therefore, the SU should limit its collisions with the PU in the CQI FB-based system as much as it can. Also, in the CQI FB-based system, the SU is always guaranteed to receive an access to the channel whenever the PU channel becomes in the bad state, therefore, it can limit its access probability in the good PU states. As clear from these two figures, the gap between the ARQ FB-based and the CQI FB-based access probabilities is bigger when the probability of the PU channel being in the bad state becomes higher. Again this is expected, since as the probability of the PU channel being in the bad state becomes higher, the SU will get a higher service in the bad state in the CQI FB-based system, so it will be less aggressive in accessing the PU channel under good PU channel state. Finally, comparing the hybrid FB-based system and the ARQ FB-based system access probabilities, we can see that none of them...
3.3 ARQ-CQI Feedback- Soft detection (SD) based access scheme

In this system, we have proposed a SUs access scheme in a CR network in which SUs leverage the primary network feedback signals in the form of ARQ and CQI feedback messages. Moreover, SUs employ soft spectrum sensing for PUs activity detection and channel access probability selection.

will be always higher than the other. For the case of high probability of the PU being in the bad channel, the access probability of the ARQ FB-based system will always be higher since in the hybrid FB system the SU benefits from the PU bad channel states, so it can be less aggressive in accessing the channel when the PU channel becomes good (cf. Fig. 3.6). For the other case of having lower probability of the PU channel being in the bad state, knowing more about the PU activity through the PU feedback might mean more aggressive access under small PU arrival rates and less aggressive access under higher PU arrival rates (cf. Fig 3.5).

3.3 ARQ-CQI Feedback- Soft detection (SD) based access scheme

In this system, we have proposed a SUs access scheme in a CR network in which SUs leverage the primary network feedback signals in the form of ARQ and CQI feedback messages. Moreover, SUs employ soft spectrum sensing for PUs activity detection and channel access probability selection.
Chapter 3. Hybrid feedback-based access scheme

The exploitation of the PUs feedback signals allows the SUs to better estimate the PUs activities.

### 3.3.1 System model

We consider a CR network with $M_p$ PUs in a TDMA based primary network. The TDMA access scheme of the PUs allows them to access the channel without interference among themselves. The secondary network consists of $M_s$ SUs, which compete for PUs’ channel opportunities through slotted ALOHA random access scheme. Each PU and each SU has an infinite length queue to store their packets waiting for transmission as shown in Fig. 3.7.

**Channel model**

The system has a channel model same as in Section 2.2.1.
3.3.2 SU spectrum access schemes

- **The No feedback- SD based access system**
  In this baseline system, each PU has ARQ feedback and no CQI feedback information. Based on the ARQ feedback, the PU either retransmits its last transmitted packet (in the case of a NACK feedback) or transmits a new packet from the head of its queue (in the case of ACK feedback). Since the PU does not have CQI feedback, it transmits its packets irrespective of the channel condition. In this system, we assume that the SUs do not have access to the PU’s ARQ feedback and employ soft-sensing scheme to access the PU channel.

- **The CQI feedback- SD based access system**
  In this system, PUs have both ARQ feedback and CQI feedback information. Hearing a good CQI feedback, the PU will transmit a packet (whether a new packet or retransmit the last packet in case of failure in the previous transmission attempt). In the case of a bad channel state indicated by the CQI feedback, the PU will refrain from transmitting since it knows that its packet will be lost due to the bad channel condition. In this system, we assume that the SUs use soft-sensing scheme, and only have access to the CQI feedback.

  When the CQI feedback overheard by SUs indicates a good channel state, SUs sense the channel and access it based on an energy level dependent access probabilities as discussed above. When the CQI feedback indicates a bad channel condition (PUs refraining from transmission), SUs will compete for the available channel in a slotted ALOHA fashion with some access probabilities that are selected to maximize the SUs’ service rate.

- **The hybrid feedback- HD based access system**
  In this system, each PU has ARQ feedback and CQI feedback, as well. SUs can overhear the two PU feedback signals and sense the PU activity using the hard-sensing scheme. The PU refrains from transmitting any packets in time slots in which the channel is in its bad state and
SUs access the PU channel with probability $p_s$ upon overhearing a bad CQI report. SUs do not transmit any packets when overhearing a good CQI and a NACK message as the PU will be retransmitting its previously lost packet with probability 1. Upon hearing a good CQI feedback and an ACK message, each SU senses the PU channel, and based on the received energy level makes a hard decision of whether the PU is present or not. The SUs that decide that the PU is idle will access the channel with probability $p_s$. The value of $p_s$ will be optimized later to maximize the SUs’ service rate while guaranteeing some QoS for the PUs.

- **The proposed hybrid feedback-SD based access system**

In this scheme, PUs have both ARQ feedback and CQI feedback. SUs can overhear both PUs feedback signals and they use energy soft-sensing for detecting the PUs’ activities. The SUs operate in the same manner as in the previous system, except that in the case of an ACK message and a good CQI feedback, SUs decide to access the channel using access probabilities that are dependent on the sensed PU energy level (which is the main difference between soft-sensing and hard-sensing).
3.3.3 Performance analysis

To analyze the SUs’ throughput and PUs’ packet delay and how they are affected by the SUs’ access scheme, the queue of any PU in the system under consideration is modeled by the three-dimensional MC shown in Fig. 3.2 but with transitions shown in Table 3.2.

The transition probabilities between the different states of this MC are given in Table 3.2. Below, we will give two examples of how these transition probabilities are calculated; the other transition probabilities can be readily derived following the same approach.

- The transition from state \((K, F, G)\) to state \((K - 1, F, G)\), \(K > 0\), occurs when the PU packet is transmitted successfully, no new packets arrive at the PU queue and the PU’s channel remains in the good state. Taking into consideration that these three events are independent and that for a successful PU transmission all SUs should not access the busy PU channel, the probability of this transition can be calculated as 
  \[(1 - \lambda_p)(1 - \sum_{i \in I} p_i a_i)^{M_s} p_G.\]

- Considering the transition from state \((K, R, G)\) to state \((K, F, G)\), \(K > 0\): Recall that when the PU is in the retransmission mode (the \(R\) states), the SUs do not access the channel. Then, the PU transmission will be successful with probability 1 (under our collision channel model where packet losses can only occur due to packet collisions) and this transition happens when a new PU packet arrives and the channel remains in the good state. Hence, the transition probability is given by \(\lambda_p p_G\).

1. The steady state distribution calculation

We start by calculating the steady state distribution of the MC shown in Fig. 3.2 with transition probabilities given in Table 3.2 so that we can get an expression for the SU throughput of the proposed system. The matrices \(B, A_0, A_1, A_2\) of the transition matrix, shown in equation
Table 3.2: MC Transition probabilities for the ARQ-CQI feedback SD system.

<table>
<thead>
<tr>
<th>Transition Number</th>
<th>Transition Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$[(1 - \lambda_p) + \lambda_p(1 - \sum_{i \in I} p_i ^1 a_i ^M_s)] p_G$</td>
</tr>
<tr>
<td>2</td>
<td>$[(1 - \lambda_p) + \lambda_p(1 - \sum_{i \in I} p_i ^1 a_i ^M_s)](1 - p_G)$</td>
</tr>
<tr>
<td>3</td>
<td>$\lambda_p(1 - (1 - \sum_{i \in I} p_i ^1 a_i ^M_s)) p_G$</td>
</tr>
<tr>
<td>4</td>
<td>$\lambda_p(1 - (1 - \sum_{i \in I} p_i ^1 a_i ^M_s))(1 - p_G)$</td>
</tr>
<tr>
<td>5</td>
<td>$(1 - \lambda_p)(1 - p_B)$</td>
</tr>
<tr>
<td>6</td>
<td>$(1 - \lambda_p)p_B$</td>
</tr>
<tr>
<td>7</td>
<td>$\lambda_p(1 - p_B)$</td>
</tr>
<tr>
<td>8</td>
<td>$\lambda_p p_B$</td>
</tr>
<tr>
<td>9</td>
<td>$(1 - \lambda_p)(1 - \sum_{i \in I} p_i ^1 a_i ^M_s) p_G$</td>
</tr>
<tr>
<td>10</td>
<td>$(1 - \lambda_p)(1 - \sum_{i \in I} p_i ^1 a_i ^M_s)(1 - p_G)$</td>
</tr>
<tr>
<td>11</td>
<td>$\lambda_p(1 - \sum_{i \in I} p_i ^1 a_i ^M_s) p_G$</td>
</tr>
<tr>
<td>12</td>
<td>$\lambda_p(1 - \sum_{i \in I} p_i ^1 a_i ^M_s)(1 - p_G)$</td>
</tr>
<tr>
<td>13</td>
<td>$(1 - \lambda_p)(1 - (1 - \sum_{i \in I} p_i ^1 a_i ^M_s)) p_G$</td>
</tr>
<tr>
<td>14</td>
<td>$(1 - \lambda_p)(1 - (1 - \sum_{i \in I} p_i ^1 a_i ^M_s))(1 - p_G)$</td>
</tr>
<tr>
<td>15</td>
<td>$(1 - \lambda_p)p_G$</td>
</tr>
<tr>
<td>16</td>
<td>$(1 - \lambda_p)(1 - p_G)$</td>
</tr>
<tr>
<td>17</td>
<td>$\lambda_p p_G$</td>
</tr>
<tr>
<td>18</td>
<td>$\lambda_p(1 - p_G)$</td>
</tr>
</tbody>
</table>

(3.1) can be written as:

$$
A_0 = \begin{pmatrix}
(1 - \lambda_p)(1 - \sum_{i \in I} p_i ^1 a_i ^M_s) p_G & 0 & (1 - \lambda_p)(1 - p_G) \\
(1 - \lambda_p)(1 - \sum_{i \in I} p_i ^1 a_i ^M_s)(1 - p_G) & 0 & (1 - \lambda_p)(1 - p_G) \\
0 & 0 & 0
\end{pmatrix},
$$

$$
A_1 = \begin{pmatrix}
\lambda_p(1 - \sum_{i \in I} p_i ^1 a_i ^M_s) p_G & (1 - \lambda_p)(1 - p_B) & \lambda_p p_G \\
\lambda_p(1 - \sum_{i \in I} p_i ^1 a_i ^M_s)(1 - p_G) & (1 - \lambda_p)p_B & \lambda_p(1 - p_G) \\
(1 - \lambda_p)(1 - \sum_{i \in I} p_i ^1 a_i ^M_s)p_G & 0 & 0 \\
(1 - \lambda_p)(1 - (1 - \sum_{i \in I} p_i ^1 a_i ^M_s))p_G & 0 & 0 \\
(1 - \lambda_p)(1 - \sum_{i \in I} p_i ^1 a_i ^M_s)(1 - p_G) & 0 & 0 \\
\end{pmatrix}.
$$
3.3. ARQ-CQI Feedback- Soft detection (SD) based access scheme

\[
\mathbf{A}_2 = \begin{pmatrix}
0 & \lambda_p (1 - p_B) & 0 & 0 \\
0 & \lambda_p p_B & 0 & 0 \\
\lambda_p (1 - (1 - \sum_{i \in I} p_i^1 a_i)^M_s) p_G & 0 & 0 & \lambda_p (1 - p_B) \\
\lambda_p (1 - (1 - \sum_{i \in I} p_i^1 a_i)^M_s) (1 - p_G) & 0 & 0 & \lambda_p p_B
\end{pmatrix}.
\]  

(3.15)

The matrix \( \mathbf{R} \) is obtained by solving equation (3.4) but using \( A_0, A_1, \) and \( A_2 \), shown in equation (3.15).

The balance equation around state \((0, F, G)\) is given by:

\[
[(1 - p_G)(1 - \lambda_p) + \lambda_p (1 - p_G)(1 - \sum_{i \in I} p_i^1 a_i)^M_s + \\
\lambda_p (1 - (1 - \sum_{i \in I} p_i^1 a_i)^M_s)] \pi_0^{FG} = (1 - \lambda_p)(1 - p_B) \pi_0^{FB} + \\
p_G (1 - \sum_{i \in I} p_i^1 a_i)^M_s (1 - \lambda_p) \pi_1^{FG} \\
+ (1 - \lambda_p)p_G \pi_1^{RG}.
\]

(3.16)

The balance equation around state \((0, F, B)\) is given by:

\[
[\lambda_p p_B - p_B + 1] \pi_0^{FB} = [(1 - \lambda_p) + \lambda_p (1 - \sum_{i \in I} p_i^1 a_i)^M_s] \\
(1 - p_G) \pi_0^{FG} + (1 - \lambda_p)(1 - \sum_{i \in I} p_i^1 a_i)^M_s (1 - p_G) \pi_1^{FG} \\
+ (1 - \lambda_p)(1 - p_G) \pi_1^{RG}.
\]

(3.17)

Solving equation (3.16) and equation (3.17), we get

\[
\pi_0^{FB} = \frac{(1 - p_G) \pi_0^{FG}}{\lambda_p p_B - p_B - \lambda_p + \lambda_p p_G + 1},
\]

(3.18)

and \( \pi_0^{FG} \) is obtained as shown in the following equation:

\[
\pi_0^{FG} = \frac{B_{10}}{(1 - \lambda_p)(2 - p_B - p_G)},
\]

(3.19)

where \( B_{10} \) is shown in the following equation:

\[
B_{10} = 4 \lambda_p + p_B - 2 \lambda_p p_B - 2 \lambda_p p_G - 2 \lambda_p (1 - \sum_{i \in I} p_i^1 a_i)^M_s \\
+ \lambda_p p_B (1 - \sum_{i \in I} p_i^1 a_i)^M_s + \lambda_p p_G (1 - \sum_{i \in I} p_i^1 a_i)^M_s - 1.
\]

(3.20)
2. Secondary throughput analysis

**The No feedback- SD based access system**

In a previous Section 2.3.3, this system was studied, in which PUs have ARQ feedback but have no CQI feedback. SUs do not monitor the PU feedback. SUs access the PU channel based on soft decision scheme. The PU’s queue in this system was modeled by the two-dimensional MC, whose state space is given by $S=\{(K,T) : K = 0, 1, 2, \ldots, T \in \{G, B\}\}$. Where $K$ is the number of PU packets in the queue, and $T$ is the CQI feedback, where $G$ means that the PU channel is in the good state and $B$ means that it is in the bad state. The closed form of the SU service rate derived for this system to consider the multi user case is shown in equation (3.21):

$$\mu_{s1} = M_p \sum_{i \in I} p_i^0 a_i (1 - \sum_{i \in I} p_i^0 a_i)^{M_i - 1}(1 - \lambda_p)(\pi_0^G + \pi_0^B)$$

$$= M_p B_{11}$$

$$= (1 - \sum_{i \in I} p_i^1 a_i)^{M_i} (2 - p_B - p_G)(\lambda_p p_B - p_B - \lambda_p + \lambda_p p_G + 1),$$

(3.21)

where $B_{11}$ is given by:

$$B_{11} = \sum_{i \in I} p_i^0 a_i (1 - \sum_{i \in I} p_i^0 a_i)^{M_i - 1}(\lambda_p + p_B + p_G - \lambda_p p_B - \lambda_p p_G - 2)(\lambda_p p_B$$

$$+ (1 - \sum_{i \in I} p_i^1 a_i)^{M_i} - 2\lambda_p + \lambda_p p_G - p_B (1 - \sum_{i \in I} p_i^1 a_i)^{M_i}).$$

(3.22)

**The CQI feedback- SD based access system**

In this system which was studied in Section 2.3.3, PU has ARQ feedback and CQI feedback. SU employs soft sensing scheme to detect the PU existence. SU also hears the PU CQI feedback. The multi user case for this system was studied and the SU service rate expression was obtained in equation (2.44).

**The hybrid feedback- HD based access system**

In Section 3.2.3, the analysis of this system was studied with single PU and single SU. In this system, the SU hear the PU CQI feedback and the ARQ feedback. The difference between this
3.3. ARQ-CQI Feedback- Soft detection (SD) based access scheme

In this system, the SU utilizes hard decision scheme and the PU feedbacks unlike the proposed system which depends on the soft sensing scheme. The SU accesses the channel with probability $a_s$ if good CQI and ACK are observed. We can extend the SU throughput driven in Section 3.2.3 to consider multi user case. For the multi user system, the SU throughput is given by,

$$
\mu_s = M_p a_s (1 - a_s)^{M_s - 1} (1 - \lambda_p) (1 - p_f) \pi_0^{FG} + M_p p_s (1 - p_s)^{M_s - 1} \sum_{k=0}^{\infty} [\pi_k^{FB} + \pi_k^{RB}] = \frac{M_p B_{12}}{p_B + p_G - 2},
$$

(3.23)

where $B_{12}$ is shown in equation (3.24).

$$
B_{12} = p_G - a_s + 2a_s \lambda_p + a_s p_f + a_s p_B + 2a_s^2 \lambda_p - 2a_s \lambda_p p_f - a_s \lambda_p p_B - a_s \lambda_p p_G - a_s p_f p_B - 2a_s^2 \lambda_p p_d - 2a_s^2 \lambda_p p_f - a_s^2 \lambda_p p_B - a_s \lambda_p p_f p_B + a_s \lambda_p p_f p_B + a_s \lambda_p p_d p_B + a_s^2 \lambda_p p_d p_B + a_s^2 \lambda_p p_f p_B + a_s^2 \lambda_p p_d p_B - a_s^2 \lambda_p p_d p_f p_B - a s^2 \lambda_p p_d p_f p_G - 1.
$$

(3.24)

- **The hybrid feedback- SD based access system**

The individual SU throughput expression of this system is derived as follows. In this system, the SU leverages the two PU feedbacks, the ACK feedback and the CQI feedback. Therefore, the SU has the information about the PU transmission modes. The SU will transmit its packet successfully without interfering with the PU in two cases: i) The PU’s queue in the bad states $(k, F, B)$ and $(K, R, B)$ for some $K = 0, 1, 2,...$. Therefore, SU transmits its packets with probability $p_s$ as it has information about the PU back-off mode. ii) The PU’s queue in state $(0, F, G)$, in which the good CQI and ACK message are observed at the SU, the PU has no packets to transmit, and the SU detects PU’s absence and decides to access the channel and other SUs refrains from accessing the PU channel in this time slot. The probability of this event is $\sum_{i \in I} P_i^0 a_i (1 - \sum_{i \in I} P_i^0 a_i)^{M_s - 1}$. The SU throughput
in this system, \( \mu_{s4} \) can be written as shown in equation (3.25).

The SU random access probability \( p_s \) is selected to achieve the maximum SU throughput. It can be proved that \( \frac{1}{M_s} \) is the optimum SU access probability.

\[
\mu_{s4} = M_p \sum_{i \in I} p_i^0 a_i (1 - \sum_{i \in I} p_i^0 a_i)M_s^{-1}(1 - \lambda_p)\pi_0^{FG} + M_p p_s (1 - p_s)^{M_s-1} \sum_{k=0}^{\infty} [\pi_k^{FB} + \pi_k^{RB}] \\
= M_p \sum_{i \in I} p_i^0 a_i (1 - \sum_{i \in I} p_i^0 a_i)M_s^{-1}(1 - \lambda_p)\pi_0^{FG} + \\
M_p \begin{pmatrix} 0 & p_s (1 - p_s)^{M_s-1} & 0 \end{pmatrix} \left( \begin{array}{c} \pi_0^{FG} \\ \pi_0^{FB} \\ \pi_0^{RG} \\ \pi_0^{RB} \end{array} \right) \\
= M_p \sum_{i \in I} p_i^0 a_i (1 - \sum_{i \in I} p_i^0 a_i)M_s^{-1}(1 - \lambda_p)\pi_0^{FG} + \\
M_p \begin{pmatrix} 0 & 1 - \frac{1}{M_s} \end{pmatrix} \left( \begin{array}{c} \pi_0^{FG} \\ \pi_0^{FB} \\ \pi_0^{RG} \\ \pi_0^{RB} \end{array} \right). 
\]

- **The SU perfect sensing with PU CQI feedback based-access system**

This system can be considered an upper bound system, where we assume perfect sensing of the PUs activities. In this upper bound system, the PU is assumed to have a CQI feedback. PU transmits its packets if the CQI is good, otherwise, it does not transmit. The SU accesses the PU channel in two cases: 1) If the MC is in the bad states \((K, B)\) for \( K = 0, 1, 2, \cdots \). 2) If the MC is in the good state and the PU queue is empty \((0, G)\). In both cases, a successful transmission from any SU means that other SUs have refrained from transmitting in this time slot. Therefore, any SU can successfully transmit its packet over the PU channel, in both of the above cases, with probability \( \frac{1}{M_s}(1 - \frac{1}{M_s})^{M_s-1} \). The
3.3. ARQ-CQI Feedback- Soft detection (SD) based access scheme

The overall SU service rate is given by equation (3.26).

\[ \mu_{sp} = \begin{cases} \frac{M_p}{M_s} (1 - \frac{1}{M_s})^{M_s-1} (1 - \lambda_p), & \text{if } \lambda_p < 1 - \zeta_B, \\ \frac{M_p}{M_s} (1 - \frac{1}{M_s})^{M_s-1} \zeta_B, & \text{otherwise.} \end{cases} \]  

(3.26)

3. Primary delay analysis

In this subsection, the average PU packet delay expression for the proposed system are derived using Little’s law as follows:

\[ D_p = \frac{E(Q_p)}{\lambda_p} = \frac{1}{\lambda_p} \sum_{k=0}^{\infty} k (\pi_k^{FG} + \pi_k^{FB} + \pi_k^{RG} + \pi_k^{RB}) \]

\[ = \frac{1}{\lambda_p} |1 1 1 1| R (I_2 - R)^{-2} \begin{pmatrix} \pi_0^{FG} \\ \pi_0^{FB} \\ 0 \\ 0 \end{pmatrix}. \]  

(3.27)

In equation (3.27), the average PU packet delay expression of the proposed system is shown. The same PU delay equation can be used to calculate the PU delay for the proposed systems with soft sensing scheme and hard sensing scheme. However, different schemes will have different values for the PU average delays as each system has different values for the optimum SU access probabilities.

4. Access probabilities calculation

The SU access probabilities have to be selected to maximize the SU throughput \( \mu_s \) and maintain the PU queue stability. Therefore, the problem can be formulated as follows:

\[ \max_{a_i, i \in I} \mu_s \]

subject to

\[ \pi_0^{FG} > 0 \text{ and } \pi_0^{FB} > 0. \]

In the simulation section, we have used a simple grid search to find the access probabilities, \( a_i \)'s, for our proposed schemes.
3.3.4 Performance results

In this section, we compare the per-user SU throughput and PU delay of the proposed system using soft sensing scheme and hard sensing scheme, the No Feedback-SD based access system, the CQI Feedback-SD based access system and the perfect sensing CQI feedback based access system. We consider a system with $M_p = 2$ PUs. We simulate the system with different values of SUs ($M_s$). The system parameters are as follows; the transmit power is set to 30 mW, the path loss exponent $\gamma = 3.7$, and the noise power spectral density $N_0 = 6 \times 10^{-11} W/Hz$. The Neyman-Pearson (N-P) detector design rule is used to choose the energy detector’s threshold $\eta$. The region below the threshold in soft sensing case is divided into $n = 4$ regions each having a different access probability $a_i, i = 1, \cdots, 4$.

In Fig. 3.8, Fig. 3.9, Fig. 3.10, Fig. 3.11, and Fig. 3.12, the SU throughput is plotted against the PU arrival rate for the different access schemes. In Fig. 3.8, we set the distance between any PU transmitter and any SU transmitter $r_{PS} = 100m$. However, in Fig. 3.9, Fig. 3.11, and Fig. 3.12, we set $r_{PS} =$
3.3. ARQ-CQI Feedback-Soft detection (SD) based access scheme

![Graph 1](image1)

**Figure 3.9:** The SU throughput for different access schemes

\( p_B = \zeta_B = 0.3, \ p_G = \zeta_G = 0.7, \ p_d = 0.7, \ p_f = 0.6712, \ r_{PS} = 400, \ M_s = 2, \ M_p = 2 \).

![Graph 2](image2)

**Figure 3.10:** The SU throughput for different access schemes

\( p_B = \zeta_B = 0.3, \ p_G = \zeta_G = 0.7, \ p_d = 0.6, \ p_f = 0.5649, \ r_{PS} = 400, \ M_s = 2, \ M_p = 2 \).
Chapter 3. Hybrid feedback-based access scheme

Figure 3.11: The SU throughput for different access schemes
\( (p_B = \zeta_B = 0.3, p_G = \zeta_G = 0.7, p_d = 0.7, p_f = 0.6712, r_{PS} = 400, M_s = 3, M_p = 2) \).

Figure 3.12: The SU throughput for different access schemes
\( (p_B = \zeta_B = 0.3, p_G = \zeta_G = 0.7, p_d = 0.7, p_f = 0.6712, r_{PS} = 400, M_s = 4, M_p = 2) \).
3.3. ARQ-CQI Feedback- Soft detection (SD) based access scheme

We simulate the system with three values of $M_s$ SUs. In Fig. 3.9, Fig. 3.11, and Fig. 3.12, we set $M_s$ to 2, 3, and 4 SUs, respectively. In Fig. 3.10, the parameters of sensing algorithm are changed. We set $p_d = 0.6$, $r_{PS} = 400m$, and so $p_f = 0.5649$.

It can be easily shown that the SU throughput at zero PU arrival rate equals to $\sum_{i \in I} p_i^0 a_i (1 - \sum_{i \in I} p_i^0 a_i)^{M_s - 1}$ for the no FB system, and $\frac{1}{M_s} (1 - \frac{1}{M_s})^{M_s - 1}$ for the perfect sensing system. This is equal to the probability of only one SU accessing the channel and the remaining SUs not accessing, thus, no collision between SUs. For the FB systems, the SU throughput equals to $\frac{1}{M_s} (1 - \frac{1}{M_s})^{M_s - 1} \zeta_B + \sum_{i \in I} p_i^0 a_i (1 - \sum_{i \in I} p_i^0 a_i)^{M_s - 1} \zeta_G$ when the PU has no packets to transmit.

It can also be shown that the SU throughput converges to $\frac{1}{M_s} (1 - \frac{1}{M_s})^{M_s - 1} \zeta_B$ at high PU arrival rate for all systems except the No Feedback-SD based access system. This value is equal to the probability of only one SU accessing the channel when the PU channel is in the bad state (thus, the PU is not
transmitting) and there is no false alarm. The No Feedback- SD based access system converges to zero SU throughput at high PU arrival rate. Since there is no CQI feedback in this system, at high PU arrival rates the PU will always be transmitting packets and the SU has no chance to transmit its packets.

In Fig. 3.8, it is noted that the proposed system with soft sensing has a very close performance to the CQI Feedback- SD based access system (this behavior is interpreted below when discussing Fig. 3.14. It is also noted that the proposed system with soft sensing scheme achieves better performance than the proposed system with hard sensing scheme. This gain is attributed to the use of soft energy sensing since it allows the SU to have a reliability measure for its sensing outcome and a higher granularity in selecting its access probabilities as will be shown in the discussion of Fig. 3.13 and 3.14.

In Fig. 3.9, Fig. 3.11, and Fig. 3.12, the distance between the PUs and SUs is set to \( r_{PS} = 400 \text{m} \). Increasing the distance from 100m to 400m results in a decreased reliability of the sensing process, which is evidenced by
the increase in the false alarm probability. In this case, the information provided to the SUs through the PU feedback channels has significant importance. Therefore, our proposed system, which leverages both ARQ and CQI feedback information, shows a remarkable gain over the CQI Feedback-SD based access system. The increase in the number of SUs in the system results only in a decrease of the SUs per user throughput since the available channel resources are shared among the increased number of SUs. It is also noted that the proposed systems with hard sensing and soft sensing have the same performance, which will be clarified below. In Fig. 3.10, at $\lambda_p = 0.1$, the percentage increase in throughput of the proposed scheme over the CQI system is 15.79% and the percentage increase of the proposed scheme over the CQI system in Fig. 3.9 is 13.08%. This increase is due to the fact that the sensing process in Fig. 3.10 is less reliable than the sensing process in Fig. 3.9. Therefore, the proposed scheme results in more gain in the system with unreliable sensing decisions.

In Fig. 3.13 and Fig. 3.14, the optimal SU access probabilities for the
proposed access scheme with soft-sensing are plotted for different PU arrival rates. The network is assumed to have $M_p = 2$ PUs and $M_s = 2$ SUs.

In Fig. 3.13, the parameters are same as the parameter used in Fig. 3.8. The distance between SUs and PUs is set to $r_{PS} = 100m$, which results in a fairly reliable energy sensing process. When the energy level sensed by an SU falls in the first region (which is the closest to zero), the SU accesses the channel with access probability $a_1 = 0.56$ for low PUs arrival rate. It does not access the channel with a higher probability in order to avoid colliding with the second SU which because of the relatively reliable sensing process is most likely to make a similar decision.

In Fig. 3.14, which has the same parameters we set in Fig 3.9, the distance between PUs and SUs is set to $r_{PS} = 400m$. In this case, all access probabilities are equal to one for most of the PU arrival rates. This can be resorted to the fact that at 400m the false alarm probability is high and the probability of the sensed energy level falling below the energy threshold $\eta$ is low. Therefore, the probability that the energy level received by both SUs fall below the energy threshold is very low. Thus, SUs can increase their
channel access probability knowing that a collision with the second SU is less likely. It is also noted that because all access probabilities are 1 for a wide range of PU arrival rates, the proposed system with soft-sensing has no gain over the proposed system with the hard sensing as shown in Fig. 3.9. It is noted that the hard sensing scheme is a special case of the soft sensing scheme with equal access probabilities in all regions.

Finally, Fig. 3.15 and Fig. 3.16 depict the average PU queuing delay for different access scheme. Fig. 3.15 and Fig. 3.16 have the same parameters as we used in Fig. 3.8 and Fig. 3.9, respectively. It is shown that the perfect sensing system has the lowest PU delay among all systems since in this system, the SU presence is completely transparent to the PU with no collisions. All the other schemes result in comparable PU average delay performance as maximizing the SUs throughput without PU delay provisioning can result in significant queuing delays in the primary queue. However, for any given PU average delay requirement, the schemes presented are always guaranteed to achieve better SUs throughput performance than the baseline schemes as they can make more informed channel access decisions.
Chapter 4

Learning to communicate with multi-agent reinforcement learning

4.1 Introduction

In this chapter, we consider a reinforcement learning (RL) based approach to design an access scheme for SUs in a CR network. In the proposed system, we implement a Q-network using multilayer perceptron (MLP) to enable SUs to access the PU channel based on their past experience and the PU network’s ARQ feedback. Each SU tries to cooperate with other SUs to avoid collisions with other SUs and with the PU. The SUs can partially observe the state of the environment as they cannot observe the number of packets in the PU’s queue; therefore, a Partially Observable Markov Decision Process (POMDP) is used to model the SUs’ access decisions and an RL Q-network using MLP is used to learn the best actions. A comparative study between the proposed system and other systems is presented. We also compare the proposed system with the perfect sensing system and the system exploiting only the last ARQ feedback. Our results show that the proposed RL based access scheme can result in comparable performance to other baseline ARQ based access schemes. Our proposed scheme achieves this performance with little knowledge compared to other ARQ based access schemes, which assume some perfect system parameters knowledge,
e.g., primary user arrival rate. Our proposed system can implicitly learn these parameters and can adapt to their variation, which is not the case for other baseline systems.

4.2 Background

4.2.1 Reinforcement learning (RL)

RL is a machine learning algorithm, in which the agent learns itself about the surrounding environment. On the contrast, supervised and unsupervised machine learning algorithms provide data to the agent to learn about the environment.

Markov Decision Processes (MDPs), which are described by the tuple $(S, A, P(s'|s, a), R(s, a))$, are used to model the RL, where $S$ is the possible states of the RL environment and $A$ is the possible actions the agent can do. $P(s'|s, a)$ is the transition probability from state $s$ to $s'$ given the action $a$ and $R(s, a)$ is the reward that the agent receives corresponding to state $s$ and action $a$ [39]. The interaction between the agent and the environment at time $t$ can be described as follows:

1. At time $t$, the agent observes the state $s_t \in S$ of the environment.

2. The agent determines action $a_t \in A$ based on $s_t$.

3. Based on the transition probability $P(s'|s, a)$, the system makes a transition from $s$ to $s_{t+1}$, where $s_{t+1} \in S$.

4. The agent receives the reward $r_t \sim R(s_t, a_t)$ corresponding to the state $s_t$ and action $a_t$ and the above procedures are repeated.

The agent aims at maximizing the discounted reward $R_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + ..., $ where $r_t$ is the reward function at time $t$. $\gamma \in (0, 1)$ is the discount factor, which reflects the effect of future reward on the decisions at each time.

The main goal of the agent is to find an optimal policy $\pi^*(s),$, where $\pi(s)$ is the mapping from the set of states to the set of actions, that the agent...
4.2. Background

follows at each state $s \in S$ in order to maximize the expected discounted reward over an infinite time [39]:

$$V^\pi(s) = E\{\gamma r_{t+1} + \gamma^2 r_{t+2} + ... | s_t = s\} = E\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s\},$$

(4.1)

where $V^\pi(s)$ is the value function of state $s$, which represents the expected discounted infinite reward when the state is $s_t$, and a policy $\pi$ is followed. In MDP, an optimal stationary and deterministic policy always exists and so there is only optimal action for a certain state. Equation (4.1) can be expressed as follows [39]:

$$V^\pi(s) = E\{r_{t+1} + \sum_{k=0}^{\infty} \gamma^k r_{t+k+2} | s_t = s\}$$

$$= E\{r_{t+1} | s_t = s\} + \gamma \sum_{s' \in S} P(s' | s, a) E\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+2} | s_{t+1} = s'\}$$

(4.2)

So the optimal value function $V^*(s)$ can be written as follows [39]:

$$V^*(s) = \max_{a \in A \{E\{r_{t+1} | s_t = s\} + \gamma \sum_{s' \in S} P(s' | s, a) V^*(s')\}}.$$  

(4.3)

The transition probability is not usually known in advance as the case in most of CR systems. Q-learning algorithm is a model free RL algorithm, in which the agent finds the optimal policy $\pi^*(s)$ that corresponds to $V^*(s)$ without having any prior knowledge about the transition probabilities $P(s' | s, a)$. The agent performs that by defining a new value called Q-value for all pairs of $(s, a)$. The Q-values are stored in table with size $m \times n$, where $m$ is the number of states and $n$ is the number of actions. The optimal Q-value is defined as follows [39]:

$$Q^*(s, a) = E\{r_{t+1} | s_t = s\} + \gamma \sum_{s' \in S} P(s' | s, a) \max_{a' \in A} Q^*(s', a').$$

(4.4)
Comparing equation (4.3) and equation (4.4), so $V^*(s') = \max_{a' \in A} Q^*(s', a')$.

By finding the optimal $Q$-value at each pair of $(s, a)$, the optimal policy $\pi^*(s)$ can be determined by $\pi^*(s) = \arg \max_{a \in A} Q^*(s, a)$. $Q^*(s, a)$ can be calculated using the following equation [39]:

$$Q(s, a) = Q(s, a) + \alpha (r(s, a) + \gamma \max_{a' \in A} Q(s', a')) - Q(s, a),$$

where $\alpha \in (0, 1)$ is a learning rate and $\max_{a' \in A} Q(s', a')$ is the maximum expected reward for all possible actions at next state $s'$.

The agent selects its action using $\epsilon$ greedy policy. In this policy, at the beginning of the interaction between the agent and the environment, $\epsilon$ has a high value and the agent starts to explore the environment by taking a random action. Then, the value of $\epsilon$ decreases and the agent starts to exploit the environment by taking an action, which results in the maximum $Q$-value.

The steps of the Q-learning algorithm can be summarized as follows:

1. Initialize the Q-table with the shape (state, action) to zero.
2. The agent starts at state $s_t \in S$.
3. $\epsilon$ greedy policy is considered to select the action which results in higher Q-value with probability $1 - \epsilon$ and a random action with probability $\epsilon$.
4. Apply the selected action $a_t \in A$ to the environment to determine the instantaneous reward $r_t \sim R(s_t, a_t)$ and the next state $s_{t+1}$ of $P(s' | s, a)$.
5. The Q-table is updated using equation (4.5).

It was proven that the Q-learning algorithm converges to the optimal Q function if all state-action pairs can be visited infinitely often [40] [41].

Instead of the look up table in Q-learning, Q-Learning is combined with the neural network. In this algorithm, two neural networks are used, the
predicted network and the target network, and both networks are parameterized by $\theta$ to represent $Q(s, a; \theta)$. The input to the predicted network is the observation $s$ and the output is the $Q(s, a, \theta_i)$, where $\theta_i$ are the parameters of the predicted network. $\epsilon$ greedy policy is considered to select the action which maximizes $Q(s, a, \theta_i)$ with probability $1 - \epsilon$ and a random action with probability $\epsilon$. The target network aims to minimize the loss:

$$L_i(\theta_i) = E[(r_t + \gamma \max_{a'}Q(s', a'; \theta^-_i) - Q(s, a, \theta_i))^2], \quad (4.6)$$

where $\theta^-_i$ are the target network parameters that are used to update the parameters for the predicted network [42].

Fully observable single agent RL (SARL) can be readily extended to the case of independent multiple agent RL (MARL). Each agent independently learns its own Q-function and selects its action. Multi agent can also be cooperative by sharing information needed to select the optimum action for agents. We use the cooperative Q-learning algorithm in [43]. The cooperative Q-learning algorithm can be described as follows:

1. Agent $n$ observes the state $s_t$ of the environment.

2. Agent $n$ shares the row of its Q-table that corresponds to the current state $s_t$, $Q_n(s_t)$, with all other cooperating agents.

3. Agent $n$ determines its action $\hat{a}_n$ based on

$$\hat{a}_n = \arg \max_{a'_{n'}} \left( \sum_{1 < n' < N} Q_{n'}^t(s_t) \right) \quad (4.7)$$

4. Agent $n$ receives a reward $r_n$ corresponding to action $\hat{a}_n$ and state $s_t$.

5. Agent $n$ updates the Q-value ($Q(s_t, \hat{a}_n)$) using equation 4.5, and the process is then repeated
4.2.2 Partial observability

It is not common that the agent has full information about the system environment. The agent however can only observe partial information about the environment state. For the partially observable environment, the agent does not observe the environment state $s_t$ but receives observation $o_t$ which is related to $s_t$ in a certain manner. POMDP is used to model the system dynamics instead of MDP in full observability.

POMDP can be characterized by the tuple $(\mathcal{S}, \mathcal{A}, P(s'|s, a), R(s, a), \Omega, O)$. The only difference between the MDP tuple and POMDP tuple are $\Omega$, $O$. In POMDP, the agent does not have complete observability of the environment state $s$ but only receives an observation $o \in \Omega$. $O$ denotes the set of conditional observation probabilities $O(o|s', a)$, which is the probability of observing $o$ when the state is $s'$ after taking the action $a$. To enable the partial observability agent to learn about the unknown environment, the agent has to use all the previous observations to try to capture the environment state [39].

4.3 System model

4.3.1 Channel Access Model

We consider $M_s$ SUs and $M_p$ PUs. PUs employ a TDMA access scheme so only one PU accesses the channel at one time slot. SUs access the PU channel by a random access scheme. The PU has an infinite buffer for storing its incoming packets. The packet arrival process is assumed to be Bernoulli i.i.d. with an average arrival rate of $\lambda_p$ packets per time slot. A slot duration is equal to the packet transmission time, and therefore, we assume $0 \leq \lambda_p \leq 1$ or else the queues will not be stable. Finally, we consider the case where SUs are backlogged which means SUs always have packets to send. We assume a collision channel model, i.e. packets are lost only when more than one transmission occurs simultaneously.
4.3. System model

4.3.2 The POMDP model

We consider the case of cooperative multi agent (SU) and partial observability. Agents cooperate in the PU channel utilization and the state of the PU queue is hidden from the agents (i.e., SUs). The SUs can only observe the PU ARQ which is NO-FB, ACK, and NACK so that each SU can compute its belief vector. The belief vector is the probability of the PU queue being in a certain state. The PU queue is modelled using the MC shown in Fig. 4.1 and the SU access decisions are modelled using a POMDP due to our system model assumptions.

The MC shown in Fig. 4.1 is described by the state space given by $S = \{K_D : K=0, 1, 2, \cdots \text{ and } D \in \{F, R\}\}$, where $k$ is the number of the PU queue packets, $D$ depends on the PU ARQ feedback as follows: $F$ means that the packet on the head of the PU queue is being transmitted for the first time, while $R$ means that the packet on the head of the PU is being retransmitted due to failure in the previous time slot.

The POMDP is characterized by the tuple $(S, A, P(s'|s, a), R(s, a), \Omega, O)$, where $S$ denotes the states of the PU queue’s MC, $S = \{\{i_F\}, \{j_R\}\}$, $i = 0, 1, \cdots$ and $j = 1, 2, \cdots$. The set $A$ denotes the set of each SU actions, i.e., $A = \{\text{access, no access}\}$. $P(s'|s, a)$ is the probability of the queue to move into state $s'$ if the action $a$ is taken while the queue is in state $s$. The
transition probabilities are given by:

\[ P(i_R|j_F, \text{no access}) = 0, \forall i, j \]
\[ P(i_F|j_F, \text{access}) = 0, \forall i, j \neq 0 \]
\[ P(1_F|0_F, \text{access}) = \lambda_p, \]
\[ P(1_F|0_F, \text{no access}) = \lambda_p, \]
\[ P(i_R|j_F, \text{access}) = \begin{cases} 
1 - \lambda_p & \text{if } i = j, j \neq 0 \\
\lambda_p & \text{if } i = j + 1, j \neq 0 \\
0 & \text{otherwise,}
\end{cases} \quad (4.8) \]
\[ P(i_F|j_R, \text{no access}) = \begin{cases} 
1 - \lambda_p & \text{if } i = j - 1 \\
\lambda_p & \text{if } i = j \\
0 & \text{otherwise.}
\end{cases} \]

The reward function for each SU, \( R \), is defined as follows.

\[ R(s, a) = \begin{cases} 
r_1 & a = \text{access}, s = 0_F \\
0 & a = \text{no access}, \forall s \quad , \quad (4.9) \\
-1 & a = \text{access}, s \neq 0_F
\end{cases} \]

where we set \( r_1 \) to be 10 in our simulation based on many trials\(^1\).

The value of reward function is obtained by trial and error. If only one SU accesses the PU channel while the PU queue is empty, the reward will be 10. If all SUs do not access the PU channel, the reward will be 0. In the case of collision between SUs or between any SU and the PU, the reward will be -1.

The set \( O \) denotes the set of conditional observation probabilities. \( O(o|s', a) \) is the probability of observing \( o \) when the state is \( s' \) after taking the action

---

\(^1\)The value of \( r_1 \) is a control parameter in our system that can control the access decisions of the SU and how aggressive these access decisions are.
and can be calculated as follows.

\[
O(o|s' = i_F, \text{no access}) = \begin{cases} 
0 & o = \text{NACK and } \forall i_F \\
F_{\text{ACK}}(0_F) & o = \text{ACK, } i_F = 0 \\
1 - F_{\text{ACK}}(0_F) & o = \text{No-FB, } i_F = 0 \\
F_{\text{ACK}}(1_F) & o = \text{ACK, } i_F = 1 \\
1 - F_{\text{ACK}}(1_F) & o = \text{No-FB, } i_F = 1 \\
1 & o = \text{ACK, } i_F \geq 2 \\
0 & o = \text{No-FB, } i_F \geq 2
\end{cases}
\] (4.10)

Note that the values of \(F_{\text{ACK}}(0_F), F_{\text{No-FB}}(0_F), F_{\text{ACK}}(1_F)\) and \(F_{\text{No-FB}}(1_F)\) will not affect our formulation and are mentioned here for the completeness of the POMPD analysis.

\[
O(o|i_F, \text{access}) = \begin{cases} 
0 & \forall o \text{ and } i_F \geq 2 \\
1 & o = \text{No-FB, } i_F = 0, 1 \\
0 & o = \text{ACK or NACK, } i_F = 0, 1
\end{cases}
\] (4.11)

\[
O(o|i_R, \text{no access}) = 0 \quad \forall o
\] (4.12)

\[
O(o|i_R, \text{access}) = \begin{cases} 
1 & o = \text{NACK} \\
0 & \text{otherwise}
\end{cases}
\] (4.13)

The last set \(\Omega\) denotes the set of observations which is given by \(\Omega = \{\text{ACK, NACK, No-FB}\}.

\textsuperscript{2}By abuse of notations, we set \(Pr(A|B) = 0\) if \(Pr(B) = 0\) (for example we set \(\Omega(o|i_R, \text{no access}) = 0\) since \(Pr(i_R, \text{no access}) = 0\) under our collision system model assumption).
4.4 RL architecture and implementation using MLP

In this section, the RL model using MLP employed for SUs operation will be described. As shown in Fig. 4.2, the SU (agent) is a neural network that has the PU’s ARQ feedback as input (observation) from the environment. The output of the RL model using MLP is the Q-value for each action (an action is the probability with which the SU decides to access the PU channel in the current time slot). The SU then chooses the action with the highest Q-value. In case of the $\epsilon$-greedy policy, it chooses the action with the highest Q-value or a random action with probability $1 - \epsilon$ and $\epsilon$, respectively.

The structure of the agent’s neural network model is shown in Fig. 4.2. The agent’s model is implemented using the Keras functional API [45] and using TensorFlow back-end [46]. As shown in Fig. 4.3, it consists of an input layer (input_1) and two fully connected layers (dense_1 and dense_2). The inner layers (dense_1) use the sigmoid activation function, and the final output layer (dense_2) uses a linear activation function.

The overall RL environment is implemented using the OpenAI Gym toolkit [47]. Algorithm 1 summarizes the operation of the SUs’ Q-learning model using MLP. Table 4.1 states the values of the parameters used in our implementation. It is noted that the values were selected based on trial and error in order to give the best possible performance.
4.4. RL architecture and implementation using MLP

**Algorithm 1** SU’s Q-learning algorithm using MLP

**Input:** History of PU ARQ state

**Output:** The optimum SU action

1. Initialize predicted model weights with random weights $\theta$.
2. Initialize target model weights with random weights $\theta^-$. 
3. Reset the environment to its initial state (No FB state).

   for episode=0:M do
   
4. Select a random action $a_t$ with probability $\epsilon$.

   otherwise $a_t = \arg \max_a \sum_{1<n'<N} Q^t_{n'}(s_t)$

5. Apply action $a_t$ to the PU environment to observe the reward $r_t$ and next state $s_{t+1}$.

6. set $z_j=\begin{cases} r_j & \text{if } j+1 = M \\ r_j + \gamma \max_a Q(s, a; \theta^-) & \text{otherwise} \end{cases}$

7. Perform a gradient descent step on $(z_j - Q(s, u, \theta_i))^2$ with respect to the network $\theta$.

8. Set $\theta = \theta^-$

**FIGURE 4.2:** RL architecture using MLP.
Chapter 4. Learning to communicate with multi-agent reinforcement learning

Figure 4.3: RL model using MLP summary.

Table 4.1: List of network parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>The number of episodes (M)</td>
<td>10000</td>
</tr>
<tr>
<td>The discount factor ($\gamma$)</td>
<td>0.85</td>
</tr>
<tr>
<td>Initial value in $\epsilon$-greedy exploration ($\epsilon_{initial}$)</td>
<td>1</td>
</tr>
<tr>
<td>Final value in $\epsilon$-greedy exploration ($\epsilon_{final}$)</td>
<td>0.01</td>
</tr>
<tr>
<td>Decay factor of $\epsilon$-greedy exploration ($\epsilon_{decay}$)</td>
<td>0.995</td>
</tr>
<tr>
<td>Learning rate ($\alpha$)</td>
<td>0.001</td>
</tr>
<tr>
<td>Agent history length (L)</td>
<td>1-5</td>
</tr>
</tbody>
</table>

4.5 Numerical results

In this section, we present the performance results of our proposed feedback-based RL channel access scheme. We consider the SUs throughput as the performance metric. For our RL based scheme, we consider the case where the system relies only on the information collected from the PUs’ ARQ feedback without sensing the PUs’ channel status and compare it to the works in [22] and [44]. We also consider the case where the system uses spectrum sensing in addition to the PUs’ ARQ feedback and compare it with the scheme presented in [10]. Moreover, all systems are compared with the ideal case (which constitutes an upper bound on performance) where SUs can perfectly sense the PUs presence, hence, no collisions between PUs and SUs can take place, and SUs are able to exploit all idle time slots.

In the following results, we consider different numbers of PUs $M_p$ and SUs $M_s$. Finally, it is assumed that the SUs action space has five actions. That is, a SU can access the PU’s channel in a given time slot with one of
4.5. Numerical results

**Figure 4.4:** SU throughput for different PU feedback history length ($L$) without a spectrum sensing.

($M_s = 1, M_p = 1$)

**Figure 4.5:** SU throughput (per user) for different spectrum access schemes

($M_s = 2, M_p = 2, p_d = 0.9, p_f = 0.1, L = 1$).
Chapter 4. Learning to communicate with multi-agent reinforcement learning

In Fig. 4.4, we investigate the effect of the length of the PU feedback (FB) history \( L \) collected by the SU on the achievable SU throughput. In the shown results, we consider only one SU and one PU, and do not employ spectrum sensing. The SU bases its access decision based on the last \( L \in \{1, 3, 5\} \) ARQ feedback packets received from the PU. The results reveal that increasing the length of the feedback history used has a negligible effect on the SU’s throughput. Therefore, it can be concluded that the last feedback packet received from the PU has the predominate effect on the SU’s throughput.

Fig. 4.4 also compares the performance of our RL-based scheme with the systems proposed in [22] and in [44]. In [22], authors introduced an access scheme for the SU in cognitive radio systems based on PU’s ARQ feedback. In [22], the SU only utilizes the last PU ARQ feedback to take its access decision to the PU channel. In [44], the authors proposed a protocol based on PU ARQ feedback in cognitive radio systems. The SU observes the history of the PU ARQ feedback. A POMDP models the SU access policy, an exact solution for the POMPD was not derived due to its computational complexity. Therefore, the authors in [44] proposed a greedy algorithm that
enables the SU to use the PU feedback history to decide its actions. It is noted that the three systems have comparable performance. However, it should be noted that in both [22] and in [44] the SU is assumed to know the number of PUs and their arrival rates in addition to observing the ARQ feedback messages, where in our scheme the SU is assumed to observe only the PUs’ ARQ feedback without any prior information about the PUs.

In Fig. 4.5 and Fig. 4.6 we investigate the performance of our RL-based access scheme using hard-decision energy detection for spectrum sensing. Our scheme is compared with the access scheme proposed in [10] which exploits the last PU’s ARQ feedback and employs a soft-sensing energy detector. Since our system is based on hard-decision spectrum sensing, we had to rework the soft-sensing based SUs’ throughput analysis of [10]. The hard-decision based SUs’ throughput of the system proposed in [10] for the case of two SUs is given by

\[
\mu_s = a_s (1 - p_f) (p_f + (1 - a_s)(1 - p_f)) \\
\left( \lambda_p a_s^2 p_d^2 - 2 \lambda_p a_s^2 p_p + \lambda_p a_s^2 + 2 \lambda_p a_s p_d - 2 \lambda_p a_s - \lambda_p + 1 \right),
\]

where \(p_d\) is the detection probability of the spectrum sensor and \(p_f\) is the probability of false alarm.

The access probability that maximizes the SU service rate \(a_s^*\) is obtained by differentiating the SU service rate in equation (4.14) with respect to \(a_s\) and equating the result to zero.

Fig. 4.5 and Fig. 4.6 depict the per SU throughput for our RL-based access scheme using the hard-decision spectrum sensing (RL sensing) using only the last PU’s ARQ feedback \((L = 1)\). Results are compared with the case of not using spectrum sensing (RL no sensing), the system proposed in [10] (FB sensing), and the ideal system (perfect sensing). The system is composed of \(M_s = 2\) SUs and \(M_p = 2\) PUs. The effect of the performance of the spectrum sensing scheme on the SUs throughput is investigated by setting the probability of detection \(p_d = 0.9\) and probability of false alarm
$p_f = 0.1$ in Fig. 4.5, and setting $p_d = 0.6$ and $p_f = 0.3$ in Fig. 4.6.

From Fig. 4.5 and Fig. 4.6, it can be seen that there is a significant improvement in SUs throughput by adding the hard-decision spectrum sensing to our RL-based access scheme. It is clear that by exploiting the information embedded in the PUs ARQ feedback alongside spectrum sensing, both our RL-based scheme and the scheme of [10] have almost the same performance, which is very close to the ideal case that assumes perfect sensing. It should be noted that our proposed RL-based access scheme does not have any information about the PUs except the received ARQ feedback. However, in [10] it is assumed that the SUs know the PUs arrival rates in addition to the received feedback messages. The information about the PUs arrival rate is not expected to be always available to the SUs. Additionally, this arrival rate can vary with time, which will force the system in [10] to recalculate its access probabilities while our RL-based scheme will be able to adapt seamlessly to any change in PUs arrival rates.
Chapter 5

Conclusions and future work

5.1 Conclusions

In this work, we have proposed secondary users (SUs) access schemes in a cognitive radio network in which SUs leverage the primary network feedback signals in the form of ARQ and CQI feedback messages. Moreover, SUs employ hard/soft spectrum sensing for PUs activity detection and channel access probability selection. The exploitation of the PUs feedback signals allows the SUs to better estimate the PUs activities. SUs are able to avoid clear collisions with PUs transmissions upon the reception of a NACK feedback. Also, SUs are able to access the PUs channel in the case of a bad channel CQI report (which means that a PU will not be attempting any transmissions). The system is analyzed from a queuing-theoretic point of view. The PU queue is modeled as a multi-dimensional Markov chain (MC). The steady-state distribution of this MC is derived. We derive expressions for the SUs throughput and the PUs average delay. Our results demonstrate the performance improvement of our proposed schemes in terms of the SUs throughput compared to baseline schemes where feedback information is not, or partially, utilized or available. In conclusion, allowing the SUs to access the primary network feedback information will always result in enhanced network performance as it allows the SUs to have better estimates of the PUs activities. We also propose a novel random access scheme for ARQ-based cognitive radio networks using RL. RL enables the SU to explore the environment and adapt its actions to best exploit the ARQ feedback on the PU channel, without prior knowledge about the PU
network, the system model or parameters. In the proposed system model, the SU exploits the PU ARQ feedback and its past experience to optimize its access decision to the PU channel.

5.2 Future directions

In the following we list some possible directions for extending our current work

- The system performance can be investigated using other spectrum sensing techniques as we have only considered the energy detection technique in this work.

- We assume a single carrier system so CQI is used to report this single carrier channel quality. Our model can be extended to multi-carrier systems and explore the effect of using wideband CQI values. Moreover, we can extend the work by considering more CQI values, not just the two-state channel model we have considered in this thesis.

- Another direction for future work is to extend the RL algorithm to consider other PU feedback information. We only considered the use of ARQ feedback in the RL based access scheme.

- Another direction for future work is to study the effect of combining the RL algorithm with the soft sensing detection scheme.
Bibliography


