Universitat de les Illes Balears

Doctorat en Informàtica

DOCTORAT EN INFORMÀTICA

MAC-PHY CROSS-LAYER ANALYSIS AND DESIGN OF MIMO-OFDM WLANS BASED ON FAST LINK ADAPTATION

GABRIEL MARTORELL LLITERAS

DIRECTORS:

FELIP RIERA PALOU
GUILLEM FEMENIAS NADAL

Departament de Matemàtiques i Informàtica
Universitat de les Illes Balears

Palma, juliol de 2013

Universitat de les
Illes Balears

MAC-PHY CROSS-LAYER ANALYSIS AND DESIGN OF MIMO-OFDM WLANS BASED ON FAST LINK ADAPTATION

GABRIEL MARTORELL LLITERAS

Departament de Ciències Matemàtiques i Informàtica

Palma, juliol 2013

Felip Riera Palou, professor del Departament de Matemàtiques i Informàtica de la Universitat de les Illes Balears

FA CONSTAR:

que la present memòria *MAC-PHY CROSS-LAYER ANALYSIS AND DESIGN OF MIMO-OFDM WLANS BASED ON FAST LINK ADAPTATION* presentada per *Gabriel Martorell Lliteras* per optar al grau de Doctor en Informàtica, ha estat realitzada sota la seva direcció i cumpleix els requisists per a ser considerada com a tesi doctoral.

Firma i data

Guillem Femenias Nadal, professor del Departament de Matemàtiques i Informàtica de la Universitat de les Illes Balears

FA CONSTAR:

que la present memòria *MAC-PHY CROSS-LAYER ANALYSIS AND DESIGN OF MIMO-OFDM WLANS BASED ON FAST LINK ADAPTATION* presentada per *Gabriel Martorell Lliteras* per optar al grau de Doctor en Informàtica, ha estat realitzada sota la seva direcció i cumpleix els requisists per ser considerada com a tesi doctoral.

Firma i data

Per na Paula, en Sebastià i els meus pares.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis advisors Dr. Felip Riera and Prof. Guillem Femenias for their guidance along this research period. I am very thankful for their valuable assistance, help and enthusiasm in this thesis during the entire five years. I would also like to highlight their passion on their research, pursuing the excellence in all their works.

Special thanks to Dr. Jaume Barceló and Dr. Boris Bellalta from the Pompeu Fabra University (UPF) for their recommendations and valuable suggestions.

Then, I would like to extend my gratitude to all past and present colleagues at the Telematics Engineering research laboratory, for the good times we shared over the past five years. Additionally, I would also like to mention all the other members from the mobile communications group (MCG) and security and electronic commerce (SECOM) whose support was excellent during this time.

Last, but not least, I am thankful to my family and friends for their constant support and encouragement. In particular my brother Sebastià who has always supported me in anything that I need and Paula whose patience, support, encouragement, help and so on, was priceless during this thesis.

Funding

I would like to thank to the Conselleria d'Educació, Cultura i Universitats del Govern de les Illes Balears (co-financed by the European Social Fund) who funded my research through a PhD grant. Additionally, this thesis would not have been possible without the assistance of Conselleria d'Economia, Hisenda i Innovació del Govern de les Illes Balears through project XISPES (PROGECIB-23A) and the Spanish government through projects COSMOS(TEC2008-02422) and AM3DIO (TEC2011-25446).

ABSTRACT

The latest WLAN standard, known as IEEE 802.11n, has notably increased the network capacity with respect to its predecessors thanks to the incorporation of the multiple-input multiple-output (MIMO) technology. Nonetheless, the new amendment, as its previous ones, does not specify how crucial configuration mechanisms, most notably the adaptive modulation and coding (AMC) algorithm should be implemented. The AMC process has proved essential to fully exploit the system resources in light of varying channel conditions.

In this dissertation, a closed-loop AMC technique, referred to as fast link adaptation (FLA) algorithm, that effectively selects the modulation and coding scheme (MCS) for multicarrier multiantenna WLAN networks is proposed. The FLA algorithm determines the MCS that maximizes the throughput while satisfying a quality of service (QoS) constraint, usually defined in the form of an objective packet error rate (PER). To this end, FLA uses a packet/bit error rate prediction methodology based on the exponential effective SNR metric (EESM).

The FLA algorithm performance has been evaluated under IEEE 802.11n systems that thanks to the incorporation of a feedback mechanism are able to implement closed-loop AMC mechanisms. Initially, this AMC technique relies only on physical layer information but it is subsequently extended to also take into account the medium access control (MAC) sublayer performance. At the physical layer, the FLA algorithm has demonstrated its effectivity by performing very close to optimality in terms of throughput, while satisfying a prescribed PER constraint. The FLA algorithm has also been evaluated using imperfect channel information. It has been observed that the proposed FLA technique is rather robust against imperfect channel information, and only in highly-frequency selective channels, imperfect channel knowledge causes a noticeable degradation in throughput.

At the MAC sublayer, the FLA algorithm has been complemented with a timeout strategy that weighs down the influence of the available channel information as this becomes outdated. This channel information outdate is caused by the MAC sublayer whose user multiplexing policy potentially results in large delays between acquiring the instant in which the channel state information is acquired and that in which the channel is accessed. Results demonstrate the superiority of FLA when compared to open-loop algorithms under saturated and non-saturated conditions and irrespective of the packet length, number of users, protocol (CSMA/CA or CDMA/E2CA) and access scheme (Basic Access or RTS/CTS).

Additionally, several analytical models have been developed to estimate the system performance at the MAC sublayer. These models account for all operational details of the IEEE 802.11n MAC sublayer, such as finite number of retries, anomalous

slot or channel errors. In particular, a semi-analytical model that assesses the MAC layer throughput under saturated conditions, considering the AMC performance is first introduced. Then, an analytical model that allows the evaluation of the QoS performance under non-saturated conditions is presented. This model focuses on single MCS and it is able to accurately predict very important system performance metrics such as blocking probability, delay, probability of discard or goodput thanks to the consideration of the finite queues on each station. Finally, the previous non-saturated analytical approach is used to define a semi-analytical model in order to estimate the system performance when considering AMC algorithms (i.e. when multiple MCSs are available).

RESUM

La darrera versió de l'estàndard de WLAN, anomenada IEEE 802.11n, ha augmentat la seva capacitat notablement en relació als sistemes anteriors gràcies a la incorporació de la tecnologia de múltiples antenes en transmissió i recepció (MIMO). No obstant això, la nova proposta, al igual que les anteriors, segueix sense especificar com s'han d'implementar els mecanismes de configuració més crucials, un dels quals és l'algoritme de codificació i modulació adaptativa (AMC). Aquests algoritmes ja han demostrat la seva importància a l'hora de maximitzar el rendiment del sistema tenint en compte les condicions canviants del canal.

En aquesta tesis s'ha proposat un algoritme AMC de llaç tancat, anomenat adaptació ràpida de l'enllaç (FLA), que selecciona eficientment l'esquema de modulació i codificació adaptativa per xarxes WLAN basades en arquitectures multiportadora multiantena. L'algoritme FLA determina el mode de transmissió capaç de maximitzar el *throughput* per les condicions de canal actuals, mentre satisfà un requisit de qualitat de servei en forma de taxa d'error per paquet (PER). FLA utilitza una metodologia de predicció de PER basada en l'estimació de la relació senyal renou (SNR) efectiva exponencial (EESM).

El rendiment de l'algoritme FLA ha estat avaluat en sistemes IEEE 802.11n, ja que aquests, gràcies a la incorporació d'un mecanisme de realimentació de modes de transmissió, poden adoptar solucions AMC de llaç tancat. En una primera part, l'estudi s'ha centrat a la capa física i després s'ha estès a la subcapa MAC. A la capa física s'ha demostrat l'efectivitat de l'algoritme FLA aconseguint un rendiment molt proper al que ens proporcionaria un esquema AMC òptim en termes de *throughput*, alhora que es satisfan els requisits de PER objectiu. L'algoritme FLA també ha estat avaluat utilitzant informació imperfecte del canal. S'ha vist que l'algoritme FLA proposat és robust en front dels efectes d'estimació imperfecte del canal, i només en canals altament selectius en freqüència, la informació imperfecte del canal provoca una davallada en el rendiment en termes de *throughput*.

A la subcapa MAC, l'algoritme FLA ha estat complementat amb una estratègia de temps d'espera que disminueix la dependència amb la informació de canal disponible a mesura que aquesta va quedant desfassada respecte de l'estat actual. Aquesta informació de canal desfassada és conseqüència de la subcapa MAC que degut a la multiplexació d'usuaris introdueix grans retards entre que es determina el mode de transmissió més adequat i la seva utilització per a l'accés al canal. Els resultats obtinguts han demostrat la superioritat de FLA respecte d'altres algoritmes de llaç obert en condicions de saturació i de no saturació, i independentment de la longitud de paquet, nombre d'usuaris, protocol (CSMA/CA i CSMA/E2CA) i esquema d'accés (*Basic Access* i RTS/CTS).

A més, s'han desenvolupat diversos models analítics per tal d'estimar el rendiment del sistema a la subcapa MAC. Aquests models consideren tots els detalls de funcionament de la subcapa MAC del 802.11n, com per exemple un nombre finit de retransmissions de cada paquet, l'*slot* anòmal o els errors introduïts pel canal. Inicialment s'ha proposat un model semi-analític que determina el throughtput en condicions de saturació, considerant el rendiment dels algoritmes AMC. Després s'ha presentat un model analític que estima el rendiment del sistema per condicions de no saturació, mitjançat el modelat de cues finites a cada estació. Aquest model considera modes de transmissió fixes i és capaç de determinar de manera molt precisa mètriques de rendiment molt importants com són la probabilitat de bloqueig de cada estació, el retard mitjà del paquets, la probabilitat de descart o la mesura del *goodput*. Finalment, el model analític de no saturació s'ha utilitzat per definir un model semi-analític per tal d'estimar el rendiment del sistema quan es considera l'ús d'algoritmes AMC.

CONTENTS

List of Figures

List of Tables

Acronyms

ACK positive acknowledgement 56

AMC adaptive modulation and coding 2

AoA angle-of-arrival 15

AoD angle-of-departure 15

AP access point 75

ARF AutoRate Fallback 3

AS angular spread 15

ASEL antenna selection 36

AWGN additive white Gaussian noise 34

BICM bit-interleaved coded modulation 43

BEB binary exponential backoff 3

BER bit error rate 4

CARA collision-aware rate adaptation 94

CDD cyclic delay diversity 36

CP cyclic prefix 34

CSI channel state information 42

CSMA/CA carrier sense multiple access with collision avoidance 3

CSMA/E2CA CSMA with enhanced collision avoidance 93

CTS clear-to-send 77

DCF distributed coordination function 3

DIFS distributed frame interval space 76

DVB-T Digital Video Broadcasting -Terrestrial 33

DVB Digital Video Broadcasting 43

EDCA enhanced distributed channel access 76

EESM exponential effective SNR mapping 59

EIFS extended interframe space 76

EQM equal modulation 44

FFT fast Fourier transform 33

FLA fast link adaptation 7

GI guard interval 34

HCF hybrid coordination function 75

HCCA HCF controlled channel access 76

HOL head-of-line 127

HTC High Throughput control field 104

HT High Throughput 44

IEEE Institute of Electrical and Electronics Engineers 1

IFFT inverse fast Fourier transform 33

ISI intersymbol interference 16

JC joint coding 37

LAN local area network 1

LDPC low-density parity-check 44

LOS line of sight 14

LQM link quality metric 57

LTV linear time-invariant 21

MAC medium access control 1

MCM multicarrier modulation 33

MCS modulation and coding scheme 2

SIFS short interval frame space 77

SIMO single-input multiple-output 3

SISO single-input single-output 3

SNR signal-to-noise ratio 38

STA station 75

STBC space-time block coding 38

STC space time coding 35

TDMA time division multiple access 115

Tx transmitter 3

UEQM unequal modulation 44

Wi-Fi Wireless Fidelity 1

WiMAX Worldwide Interoperability for Microwave Access 33

WLAN Wireless local area network 1

ZF zero forcing 41

RMS root mean square 18

CHAPTER 1

INTRODUCTION

1.1 Wireless Local Area Networks

Wireless local area networks (WLANs) are used to provide connectivity to areas ranging from households and offices and up to whole buildings and small open spaces (e.g. campus, airports). They are considered the wireless version of the local area network (LAN), hence satisfying the same requirements such as high capacity, full connectivity among attached stations and broadcast capability. At the same time, they provide mobility and easy deployment to offer connectivity to places where LANs are non-existent, very costly or not affordable due to their temporality or short operation time, hence making the setting-up of a wired network impractical [7].

In the last years, WLANs have witnessed and impressive rise in demand, being massively deployed at homes, offices and public places. This wide acceptance motivates their support by most personal mobile devices, such as, laptops, tablets and smartphones. Over the last decade, WLAN has increasingly become synonymous with the Institute of Electrical and Electronics Engineers (IEEE) 802.11 family of standards (IEEE std 802.11a, b, g or n) and popularised under the Wireless Fidelity (Wi-Fi) denomination [8, 9, 10].

IEEE 802.11, similarly to other standards and protocols, is designed to ensure that communications equipment from different vendors can interoperate, hence defining rules and language conventions. This standard is solely focused on the physical layer (PHY) layer and medium access control (MAC) sublayer specifications, hence defining the available transmission techniques, frame formats and other physical layer parameters at the PHY layer, and also dictating the medium access rules and algorithms at the MAC sublayer. This standard has notably evolved since its first version, particularly, several PHY layers have been proposed in the IEEE 802.11a, b, g and n amendments, and one additional MAC sublayer was defined IEEE 802.11e. It is worth pointing out that crucial issues, such as implementation of particular processing steps, resource allocation algorithms or system administration are left to the vendors. Consequently, although compatible WLAN devices are evaluated under the same channel

conditions, they may perform differently since their particular implementations of the standard are presumably different.

The most recent standard, namely IEEE 802.11n, supports much higher transmission rates than its previous incarnations thanks to the use of multiple-input multiple-output (MIMO) antenna technology, the possibility of operating on a 40 MHz bandwidth (employing more subcarriers) and the availability of transmission modes using a reduced guard interval. Furthermore, and critically important for the work undertaken in this thesis, the MAC sublayer does incorporate mechanisms to feedback information regarding modulation and coding scheme (MCS) selection, thus making closed-loop adaptive algorithms a feasible option. Note that, as in previous standards, IEEE 802.11n only specifies which MCSs are allowed for which types of MAC frames, but not how and when to switch between the permitted rates.

1.2 Link Adaption in Wireless LANs: An Overview.

Compared to a wired transmission media such as copper, coaxial cable or fiber optic, the wireless radio channel is a very hostile medium. High data-rate transmitted signals are significantly impaired by noise, interference, object obstructions, reflections and refractions whose effects change over frequency, time and space, thus causing the undesired frequency-, time- and space-selective channel fading [1].

To deal with these channel variations, current WLANs, as the majority of wireless networks, require adaptive algorithms that properly set their configuration according to these variable conditions. A proper operation of these algorithms notably enhances the system behaviour, hence contributing to the objective of maximising system efficiency in terms of b/s/Hz. Recently, WLAN systems have enjoyed an impressive capacity increase thanks to the latest advances in wireless transmission techniques. Consequently, new or refined adaptive algorithms are required to properly exploit these enlarged capacities in light of the wireless channel variations.

One of the most widely used reconfiguration techniques is adaptive modulation and coding (AMC) (also known as rate adaptation or link adaptation), which selects an appropriate MCS in response to changes in the environment or system behaviour. AMC algorithms can be broadly categorized as closed- or open-loop, and throughput- or quality of service (QoS)-oriented schemes.

1.2.1 Throughput versus QoS oriented AMCs

According to their optimization objective, the AMC mechanisms can be classified as throughput- or QoS- oriented algorithms. The throughput-oriented algorithms maximize the total system throughput, regardless of QoS measures, such as error rate, delay, or jitter and, moreover, they neglect fairness considerations. Conversely, QoS-oriented mechanisms, select the transmission rate that maximizes the throughput while satisfying predefined QoS constraints that typically take into account fairness issues.

1.2.2 Open- versus closed-loop AMCs

AMC can also be classified as closed- or open-loop algorithms, depending on whether an explicit feedback channel between receiver (Rx) and transmitter (Tx) is used or not. Open-loop setups operate in a heuristic manner and their rate of adaptation tends to be slow with respect to channel changes, thus compromising the fulfilment of QoS constraints. In contrast, closed-loop mechanisms track the channel behavior more accurately and are more reactive to rapid channel variations.

In previous amendments of IEEE 802.11, there was no signaling mechanism specified that would allow a receiver to inform the transmitter about the actual link quality or the rate to be used, consequently, only open-loop AMC techniques where feasible. In the absence of Rx-Tx feedback, the AutoRate Fallback (ARF) link adaptation protocol and its modifications (see [11] and references therein) have been widely used in legacy WLANs.

In IEEE 802.11n, the MAC sublayer does incorporate mechanisms to feedback information regarding MCS selection, thus making closed-loop algorithms a potential option. In MIMO systems, in addition to MCS selection, link adaptation algorithms face another challenge: the MIMO mode selection. In this case, a higher transmission rate in the MCS does no longer imply a higher instantaneous packet error rate (PER), as was the case in single-input single-output (SISO) systems, and thus, the traditional link adaptation algorithms used in single-input multiple-output (SIMO) legacy systems become less effective [12].

1.2.3 MAC sublayer analysis

WLAN systems are also seriously affected by the MAC sublayer operation, due to the large overheads introduced and the crucial role user multiplexing has on the overall system performance. Consequently, the MAC and PHY layers should be jointly treated when evaluating the suitability of any AMC algorithm. This motivates the development of MAC-PHY cross-layer AMC algorithms for MIMO-OFDM systems.

The distributed coordination function (DCF) implemented in most IEEE 802.11 networks has been intensively studied[1]. The basic access mechanism of DCF is based on the carrier sense multiple access with collision avoidance (CSMA/CA) protocol, which makes use of binary exponential backoff (BEB) to share the medium between competing stations. As a random access protocol, BEB-based CSMA/CA inherits the advantages of minimum coordination and distributed control. Unfortunately, its random nature also complicates the derivation of suitable models and their corresponding performance evaluation. In fact, as collisions couple the service processes of stations competing for access, the performance evaluation of this multiple access protocol cannot be analytically investigated unless some judicious simplifying approximations are assumed, which have nevertheless been shown to lead to accurate analytical models.

[1] Note that the IEEE 802.11 MAC layer can also be based on the point coordination function (PCF), however this mode of operation is hardly used in practical deployments.

1.3 Problem Formulation, Motivation and Scope

This dissertation has two main goals. The first goal is to provide an AMC algorithm to efficiently exploit the IEEE 802.11n capacity at the PHY and MAC sublayer levels. Ideally, this algorithm should rapidly react to wireless channel changes by providing suitable MCS selections taking into account the effects of outdated MCS feedback information. Furthermore, this algorithm should be valid irrespective of the scenario configuration (number of users, packet sizes, traffic loads, access schemes and network protocols) and allow the performance assessment of metrics such as goodput, PER, fairness, blocking probability and delay.

The second goal is to derive analytical models to estimate the system performance at the MAC level under saturated and/or non-saturated conditions, while considering the use of fixed MCSs or AMC algorithms. To this end, models based on Markov chains are proposed that are capable of modelling the relevant performance metrics at the MAC level.

The first dissertation goal is formulated into two separated problems, P1 and P2, depending on the layer point of view. P1 refers to the PHY and can be defined as:

P1) *How to select the most suitable MCS at the physical layer under given channel conditions that satisfy QoS constraints while maximizing the throughput.*

Specifically, this thesis initially addresses the problem of MCS selection only considering the PHY layer of IEEE 802.11n. The use of a closed-loop algorithm is motivated by the addition of the MCS feedback mechanism into the standard and the promising performance results obtained by novel closed-loop AMCs in other environments such as IEEE 802.16e. In trying to provide an answer to P1, two important issues should be addressed that lead to subproblems:

P1.1) *How does the availability of non-ideal channel state information affect the MCS selection procedure and its resulting performance.*

The proposed scheme to solve P1 decides the most suitable MCS after predicting PER performance of each MCS. To this end, it uses ideal channel estimations, although wireless systems do not always have access to this precise information. For this reason, the AMC algorithm performance is evaluated under imperfect channel conditions.

P1.2) *How can the MCS selection be adapted to be suitable for different packet sizes.*

The proposed AMC algorithm relies on PER estimations, whose performance is linked to a particular packet size, hence complicating its practical implementation since different systems will typically be configured with different packet lengths. Therefore, an equivalent model is proposed that measures the link quality in terms of bit error rate (BER), hence widening the suitability of the adaptive strategy.

The second problem related to the first goal refers to the MAC sublayer and can be posed as:

P2) *How does the AMC algorithm perform at the IEEE 802.11 MAC sublayer level.*

The addition of the MAC sublayer to the system notably lowers the system performance, since large overheads are introduced between transmissions. These overheads may notably affect the operation of closed-loop algorithms, by introducing large delays between successive station transmissions, and rendering the MCS information provided by the feedback channel outdated. Therefore, some additional mechanism is required to counteract this envisaged degradation.

In particular, this goal is fulfilled by tackling the following two subproblems:

P2.1) *How does MCS feedback delay influence the AMC performance under saturated conditions.*

In P2.1, the AMC performance is thoroughly evaluated under saturated conditions, which means that each station has always packets ready for transmission.

P2.2) *How does delayed information influence the AMC performance under non-saturated conditions.*

In P2.2, AMC is evaluated under non-saturated conditions, by modelling finite queue sizes, where the queue length is decreased after a successful transmission or packet discard, and increased after a packet arrival. When tackling P2.1 and P2.2, strategies are proposed that are shown to effectively combat the deleterious effects delayed information has on MCS selection.

Finally, the third problem treated in this thesis refers also to the MAC sublayer and is defined as:

P3) *How to analyze IEEE 802.11 system performance at the MAC sublayer considering saturated and non-saturated conditions for both fixed-and AMC-based configurations.*

The motivation of this analysis is the lack of studies (to the best of author's knowledge) analyzing the influence of AMC techniques at the MAC layer.

The overall motivation of this thesis through the study of the aforementioned problems resides in the fact that the proper use of MCS notably increases the system performance. The AMC selection is a common issue to be tackled by all vendors implementing IEEE 802.11-based products, hence highlighting the practical importance of this problem. Remarkably, the apparition of the new amendment IEEE 802.11n contributes to new AMC possibilities not previously available. Additionally, it is important to have analytical MAC sublayer models that are able to consider all performance details of current WLAN networks, hence also accounting for the performance of the AMC algorithms.

1.4 Adopted Approach

In this dissertation, the adopted approach to solve Problem P1 is a closed-loop, QoS-oriented AMC technique. This AMC selects the MCS that maximizes the throughput while satisfying a predefined QoS constraint in the form of a PER target. To this end,

the AMC algorithm employs PER estimations of each MCS provided by a PER prediction methodology. This method requires of a calibration process to tune the algorithm to the specific operating environment.

For the Problem P2, the system model is extended to consider IEEE 802.11 MAC sublayer, hence accounting for the influence of delayed information on the MCS selection. To tackle this effect, some alternative strategies are proposed upon detection of stale MCS information.

In order to tackle Problem P3, the system modelling of the MAC sublayer relies on Markov chains. Markov chains have been widely used in modelling telecommunications systems in general, and wireless networks in particular. Remarkably, the most popular model to assess the performance of IEEE 802.11 [13] is indeed based on a Markov Chain model approach.

1.5 Previous Works

The MCS selection at the physical layer of WLAN networks based on MIMO-OFDM was initially treated in [14], [15] and [16]. These works propose closed-loop AMC algorithms rooted on the estimation of the PER for instantaneous channel conditions using the methodology presented in [12]. However, practical issues such as, the availability of non-ideal channel estimation or the use of BER instead of PER to estimate current channel performance are not considered.

Regarding problem P2, the performance evaluation of the proposed closed-loop algorithm in the context of the MAC sublayer has not been treated before. We nevertheless note that open-loop algorithms such as ARF [17], CARA [18] or [19] have indeed been analysed in terms of MAC sublayer performance.

Regarding problem P3, one of the most widely adopted modeling paradigms is the mean-field Markov model approach originally developed by Bianchi on his seminal papers [20, 13]. In these papers, a two-dimensional Markov chain model was proposed to characterize the basic access mechanism of DCF. Bianchi's basic approach has been refined by follow-up works that take into account practical assumptions and expand its original range of applicability. In particular, and specially relevant for this thesis, it is worth mentioning [21, 22, 23], which present an improved analytical framework thoroughly accounting for the backoff freezing details of the DCF specifications; [24, 25, 26, 27, 28], which analyze the DCF for an error-prone channel; [29, 30, 31], which treat unsaturated bufferless stations; and [32, 33, 34], which consider the impact of unsaturated stations with infinite buffers.

Bufferless models [29, 30, 31] overestimate system throughput and underestimate collision probability. Moreover, the use of infinite buffers [20, 13, 32, 33, 34], impede the analytical derivation of practical QoS performance measures such as the blocking probability or the queueing delay. In order to capture these QoS performance metrics, Liu *et al.* [35] proposed a finite buffer three-dimensional Markov chain model able to integrate contention resolution procedures and queueing processes. This extra third dimension, used to model the queue dynamics, allows the calculation of the throughput, transmission failure probability, collision probability, queue blocking probability, queue length, and system delay, without resorting to an explicit service time distribution. Unfortunately, Liu's three dimensional model is based on Bianchi's original

approach and, therefore, it does not take into account neither the refinements dealing with the freezing of backoff counter nor those considering the transmission over an error-prone channel.

1.6 Dissertation Contributions

The main contributions to the field provided in this dissertation are briefly listed regarding Problem (P1):

P1.C1) *Definition of a QoS-oriented AMC algorithm for WLANs based on closed-loop operation, namely fast link adaptation (FLA). The proposed scheme employs a PER prediction methodology that estimates PER in light of the available channel information.*

P1.C2) *Performance characterization of FLA over WLAN scenarios based on the IEEE 802.11n physical layer.*

P1.C3) *Definition of a new FLA algorithm version that bases its selections on the bit error rate (BER) instead of PER, hence notably simplifying the calibration procedure required when using different packet lengths.*

P1.C4) *Characterization of FLA performance with imprecise channel state estimation at the physical layer.*

For the second Problem (P2), the contributions are:

P2.C1) *Performance evaluation of FLA strategies under saturated conditions considering IEEE 802.11n MAC and physical layers. Remarkably, the effects induced by outdated MCS information are thoroughly assessed and tackled.*

P2.C2) *Performance comparison of FLA with respect to other open-loop strategies for different access schemes (Basic or RTS/CTS access) and network protocols (CSMA/CA and CSMA/E2CA).*

P2.C3) *Performance evaluation of AMC algorithms under non-saturated conditions, an scenario found to exacerbate the MCS outdate problem.*

Finally, the contributions to the third Problem (P3) are:

P3.C1) *Definition of a semi-analytical model that assesses the MAC sublayer performance of IEEE 802.11n when considering AMC algorithms. This model is validated through the study of different scenario configurations and AMC algorithms.*

P3.C2) *Analysis of IEEE 802.11 MAC sublayer systems considering non-saturated conditions. This model allows the performance estimation of delay, goodput, blocking probability or other performance metrics suitable for network designers.*

P3.C3) *Extension of the study in P3.C2 by means of a a semi-analytic model that allows the assessment of the system performance when considering AMC algorithms.*

1.6.1 Related publications

Journal publications

- G. Martorell, F. Riera-Palou, and G. Femenias. Cross-Layer Fast Link Adaptation for MIMO-OFDM Based WLANs. *Springer Wireless Personal Communications*, 56(3):599-609, 2011.
- G. Martorell, F. Riera-Palou, and G. Femenias. Modelling Fast Link Adaptation-based 802.11n Distributed Coordination Function. *Springer Telecommunication Systems*, to appear in 2013.
- G. Martorell, G. Femenias, and F. Riera-Palou. Non-Saturated IEEE 802.11 DCF-based Networks. Part I: A three Dimensional Markov Model. Submitted to *IEEE Transactions on Wireless Communications*.
- G. Martorell, G. Femenias, and F. Riera-Palou. Non-Saturated IEEE 802.11 DCF-based Networks. Part II: QoS Performance Analysis. Submitted to *IEEE Transactions on Wireless Communications*.

Papers in International Conferences

- G. Martorell, F. Riera-Palou, and G. Femenias. Cross-layer link adaptation for IEEE 802.11n. In *IEEE Second International Workshop on Cross-Layer Design (IWCLD)*, Palma de Mallorca, Spain, June 2009.
- G. Martorell, F. Riera-Palou, and G. Femenias. DCF performance analysis of open- and closed-loop adaptive IEEE 802.11n networks . In *IEEE International Conference on Communications (ICC)*, Kyoto, Japan, June 2011.
- G.Martorell, F. Riera-Palou, and G. Femenias. Closed-loop adaptive IEEE 802.11n with PHY/MAC cross-layer constraints. In *LNCS. 4th International Workshop on Multiple Access Communications (MACOM)*, Trento, Italy, Sept. 2011.
- G.Martorell, F. Riera-Palou, G. Femenias, J. Barceló, and B. Bellalta. On the performance evaluation of CSMA/- E2CA protocol with open loop ARF-based adaptive modulation and coding. In *18th European Wireless Conference (EW)*, Poznań, Poland, April 2012.
- G. Martorell, F. Riera-Palou, and G. Femenias. Tuning fast link adaptation algorithms for CSMA/CA- and CSMA/E2CA-based WLANs networks. In *IEEE 21st International Conference on Computer Communications and Networks (ICCCN)*, Munich, Germany, July 2012.
- G. Martorell, G. Femenias, and F. Riera-Palou. A Refined 3D Markov model for QoS Performance Analysis of unsaturated IEEE 802.11 DCF Networks. Submitted to *Wireless Days, 2013*.

Papers in national conferences

- G. Martorell, F. Riera-Palou, and G. Femenias. Adaptabilidad de enlace en Sistemas IEEE 802.11n. In *Jornadas de Ingeniería Telemática (JITEL 2009)*, Cartagena, Spain, Sept. 2009.

- G. Martorell, F. Riera-Palou, and G. Femenias. Análisis del goodput para sistemas IEEE 802.11n basados en AMC de lazo abierto y cerrado. In *Jornadas de Ingeniería Telemática (JITEL)*, Valladolid, Spain, Sept. 2010.

- G. Martorell, F. Riera-Palou, and G. Femenias. Performance Analysis of Fast Link Adaptation-based 802.11n Basic and RTS/CTS Access Schemes. In *Jornadas de Ingeniería Telemática (JITEL)*, Santander, Spain, Sep. 2011.

1.7 Dissertation Outline

This dissertation is divided into two parts. Part I, entitled *Link Adaptation Performance on IEEE 802.11n Physical layer* refers to the physical layer performance. This part is made of three chapters: Chapter 2 describes the wireless LAN channel model used in the dissertation, Chapter 3 fully describes and characterizes the IEEE 802.11n physical layer at the transmitter and receiver sides and, finally, Chapter 4 provides an adaptive modulation and coding algorithm for IEEE 802.11n, evaluating its impact on the system performance. Part II, entitled *Link Adaptation Performance on the MAC sublayer*, is focused on the system performance at the MAC sublayer, hence presenting analytical models and evaluating the AMC algorithms provided in the previous part within the MAC context. To this end, Chapter 5 describes the MAC sublayer operation and characterizes its performance under saturated conditions when using fixed MCSs. Chapter 6 provides a semi-analytic model to evaluate the system performance with AMC algorithms under saturated conditions, while providing improvements to the proposed AMC algorithms. Chapter 7 defines a new modelling approach to evaluate the QoS performance of IEEE 802.11n under non-saturated conditions. Chapter 8 incorporates AMC to the non-saturated analytical approach to evaluate the performance of AMC algorithms. Finally, Chapter 9 recaps the dissertation main results and it also describes what are though to be the most promising avenues for further research.

This introduction concludes with a notational remark that applies to the entire dissertation: vectors and matrices are denoted by lower- and upper-case bold letters, respectively. The superscripts ^{T} and ^{H} are used to denote the transpose and conjugate transpose of the corresponding vector/matrix, respectively. The symbol $\boldsymbol{I}_Q$ represents the Q-dimensional identity matrix, whereas $E\{\cdot\}$ is used to denote the expectation operator. $\lceil z \rceil$ denotes the smallest integer greater than or equal to z. Finally, given a matrix $\boldsymbol{A}$, the notation $[\boldsymbol{A}]_{j,i}$ is used to identify the element situated on the ith column and jth row.

Part I

Link Adaptation Performance on IEEE 802.11n Physical Layer

CHAPTER 2

WIRELESS CHANNEL

The wireless radio channel is a very hostile medium for high data-rate transmissions compared to other transmission mediums, such as fiber optics, pair of copper or coax. The transmitted signal is significantly impaired by noise, interference, object obstructions, reflections and refractions whose effects change over frequency, time and space, thus causing the undesired frequency-, time- and space-selective channel fadings [1]. In this chapter, the main characteristics of WLAN wireless channels are presented. Firstly, a general overview of the wireless channel characterization is provided in Section 2.1, presenting the large-scale fading, the small-scale fading and the impulse response model for single antenna systems (SISO). Secondly, in Section 2.2, the TGn channel model for WLAN environments is detailed, thus describing the MIMO channel impulse response and the different types of TGn channels. Finally, a brief description of the TGn channel generation tool is provided as well as a collection of simulation results.

2.1 General Overview

2.1.1 Large-scale fading

Large-scale fadings are slow power variations of the received signal as a function of distance from the transmitter and position of the transmitter and receiver, which are caused by the path loss and shadow fading. The path loss is the signal attenuation attributed to the signal propagation through the channel, while shadow fading is the loss of power due to obstructions situated in between transmitter and receiver [1, 36, 37]. In Fig. 2.1, a realization of the path loss and shadow fading is depicted as a function of the distance between the transmitter and the receiver.

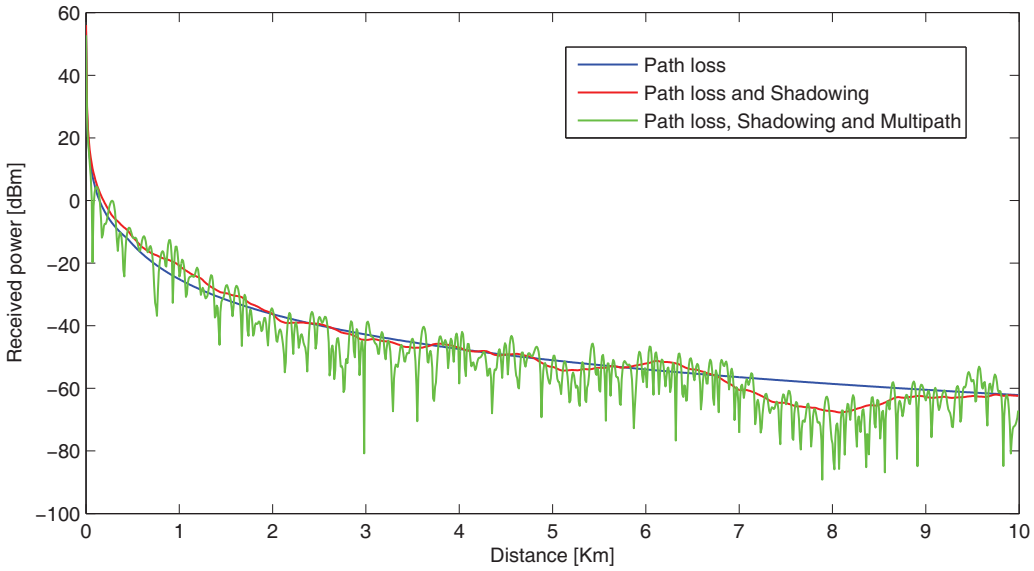


Figure 2.1: Path loss, Shadow and Multipath fading versus distance.

Path loss modeling

In the literature, different path loss models have been proposed. Some of these models rely on making use of an accurate description of the scenario configuration and/or empirical measurements, thus achieving high precision results, i.e., the Okumura model [38], the Hata model [39], or the COST 231 model [40]. Alternative simpler models are based on theoretical approximations that capture the essence of the propagation loss, without requiring complex measurements or scenario characterizations. These models usually consist of slight modifications of the free-space path loss model.

The free-space path loss model estimates the power loss of a signal propagated through free-space to a receiver located at a distance d from the transmitter. As its name suggests, in free-space the path is clear from obstructions and the signal propagates between receiver and transmitter through a straight line whose path is referred to as line of sight (LOS). The free-space path loss ($G_{FS}(d)$) referred to WLAN environments (slope of 2) expressed in decibels (dB) is defined as [41, 2]

$$G_{FS} = 20\log\frac{4\pi d_0}{\lambda_0} + 20\log\frac{d}{d_0} \tag{2.1}$$

where d_0 is fixed to 1 meter and λ_0 denotes the carrier wave length.

Shadow fading modeling

Apart from the path loss, the received signal is usually affected by objects encountered through the propagation path, thus suffering signal blockage, reflection or scattering. The combination of these effects results in a random signal attenuation referred to as shadow fading (see Fig. 2.1). Since the location, size and dielectric properties of the blocking objects as well as the changes in reflecting surfaces and scattering objects that cause the random attenuation are generally unknown, statistical models must

be used to characterize the shadow fading. Empirical measurements over different indoor and outdoor channel environments demonstrate that a model based on the Log-normal distribution is an accurate mechanism to characterize shadow fading [42, 43].

Log-normal shadow fading, $G_{\text{Shadow fading}}$, is typically parameterized by the log mean $\mu_{G_{FS}}$, that is, the average path loss, and its standard deviation $\sigma_{G_{FS}}$. Note that this implies that the distribution of G_{FS} on a dB scale is Gaussian with mean $\mu_{G_{FS}}$ and standard deviation $\sigma_{G_{FS}}$, that is,

$$G_{\text{Shadow Fading}}\left(G_{FS}(d)\right) = \frac{1}{\sqrt{2\pi}\sigma_{G_{FS}}} exp\left[-\frac{\left(G_{FS}-\mu_{G_{FS}}\right)^2}{2\sigma^2_{G_{FS}}}\right]. \tag{2.2}$$

These models are generally known as combined path loss and shadow fading models.

2.1.2 Small-scale fading

The small-scale fading is a fast variation of amplitude and phase on the received signal in response to small changes in the environment (i.e. position of the Tx, Rx or scatterers), giving rise to large power variations (see Fig 2.1). Small-scale fading is originated by two effects: the time-variant behavior of the channel and the time spreading of the signal. The time-variant behavior is mainly produced by the movement of any of the objects involved in the configuration of the propagation path (Tx, Rx or any of the objects around or in between), hence resulting in the Doppler effects. While the time-spreading is provoked by the reception of multiple signal copies spread in time due to the multipath propagation. In the next paragraphs, the multipath propagation and the Doppler effects are studied in the domains of time and frequency. Most parts of this subsection were extracted from [1, 36].

Multipath Propagation

The electromagnetic waves of the transmitted signal propagate through the media experiencing reflections, scattering or refractions on the environmental elements (buildings, walls, furniture, trees, cars, etc.), thus enabling multiple signal propagation paths to reach the receiver. In Fig. 2.2, these paths are depicted for a generic single transmit and receive antenna scenario. Due to this variety of propagation paths, multiple copies of the transmitted signal arrive at the receiver and their addition constitutes the received signal. Note that those replicas may have opposite phase shifts and their sum might imply a reduction in the received signal strength. This random power variation is generally referred to as multipath fading.

Although multipath propagation can be characterized using ray-tracing models for deterministic channels, in practice, these models are rarely available, since they require stringent knowledge of the scenario's features. Most wireless channel models are modelled using statistical approximations where the channel response is presented as a random time-varying impulse response. This channel response usually consists of the sum of channel responses from each propagation path.

As stated by Saleh and Valenzuela [44], propagation paths can be grouped in clusters characterized by an angle-of-departure (AoD) from the Tx, an angle-of-arrival (AoA) at the Rx, and angular spread (AS) at both Tx and Rx (one AS value for each).

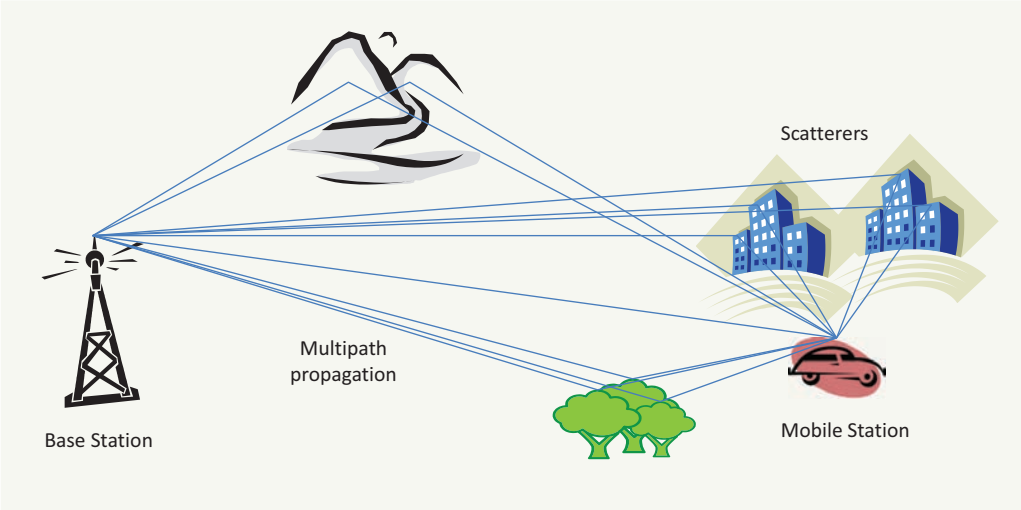


Figure 2.2: Transmitter and receiver in an area with scatterers.

For each cluster, the number of multipath components is determined by the receiver's ability to distinguish received signal components separated within a given time-span. Each multipath component can model the performance of a single ray or a combination of rays which, due to their proximity in time, cannot be separated and treated as single rays. More specifically, two multipath components with delays $\eth_1$ and $\eth_2$, are resolvable if their delay difference significantly exceeds the inverse signal bandwidth B_s, that is, ($|\eth_1 - \eth_2| >> B_s^{-1}$). If the components do not meet this criterion, they cannot be separated out at the receiver, since these components are non-resolvable. Therefore, the received signal can be expressed as the sum of propagation paths coming from $\mathfrak{L}$ clusters, where each cluster l, $1 \leq l \leq \mathfrak{L}$, has N_l multipath components.

The amplitude $\alpha_{\hat{n},l}(t)$, for each multipath component $\hat{n}$ ($1 \leq \hat{n} \leq N_l$) can be characterized statistically by a Ricean or a Rayleigh distribution, depending on the presence or absence of the line of sight (LOS) component. If the considered multipath component is only composed of non-line-of-sight (NLOS) paths, $\alpha_{\hat{n},l}(t)$ can be approximated by a Rayleigh distribution. Otherwise, when a LOS component is present, $\alpha_{\hat{n},l}(t)$ can be modelled by a Ricean distribution.

At the receiver, the sum of all the multipath components spreads the received signal in time with respect to the original transmitted signal, causing the appearance of intersymbol interference (ISI) among successively transmitted symbols. Several techniques can be used in order to counteract ISI and time spreading, such as, equalization, multicarrier modulation and spread spectrum.

In channel modeling, the delay spread T_m represents the signal's propagation delay that exceeds the arrival of the first component at the receiver. In practice, the last multipath signal is the one whose power falls to some threshold level below that of the strongest component. An example of a threshold value could be a power value of 10 or 20 dB below the power of the strongest multipath component. Therefore, T_m is defined as the time lapsed between the first and last received multipath components above the predefined threshold.

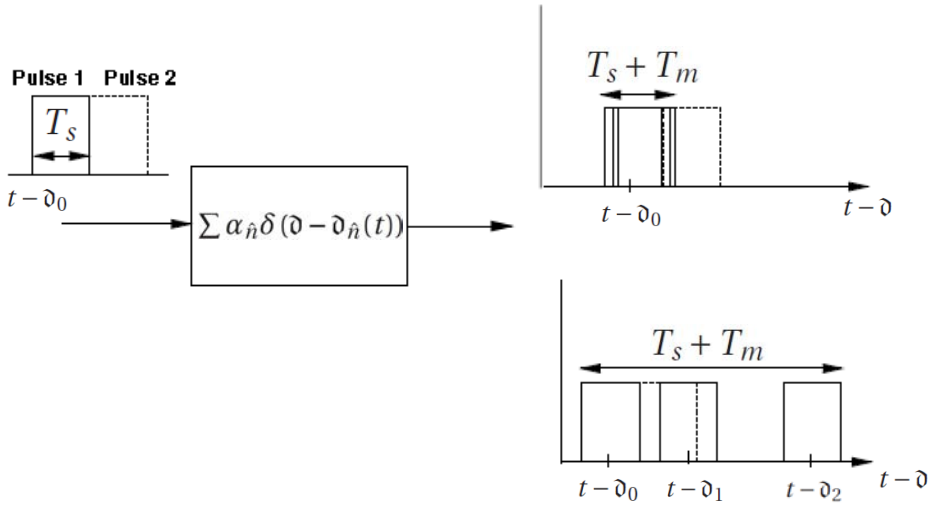


Figure 2.3: Multipath resolution (extracted from [1]).

Time domain characterization. Depending on the relation of delay spread (T_m) and the inverse of the signal bandwidth ($1/B_s = T_s$, where T_s is the symbol rate), the propagation channels can be classified as narrowband or wideband channels. If $T_m << T_s$, the LOS and multipath components are non-resolvable, thus leading to a narrowband model. Otherwise, the received multipath components are resolvable, thus spreading the received signal in the time domain.

In narrowband channels, all the multipath components from a transmitted pulse arrive at the receiver simultaneously, thus not extending the signal duration at the receiver side. This effect can be observed in Fig. 2.3 where the transmission of a train of pulses of duration T_s takes place over two different wireless channels with distinct delay spread values. If we consider the narrowband case where the delay spread is $T_m << T_s$ then the multipath components are received roughly on top of one another, (see Fig. 2.3). The resulting constructive and destructive interference causes narrowband fading of the pulse, but there is little time-spreading and therefore little interference with a subsequently transmitted pulse. Note that ISI mitigation is not necessary if $T_s >> T_m$, but this can place significant constraints on data rate.

By contrast, in wideband channels, the received signal suffers an important distortion due to the multipath delay spread. A single pulse transmission of duration T_s results in a received signal of duration $T_s + T_m$ (see Fig. 2.3). Therefore, the duration of the received pulse can increase severely due to the channel delay spread. Consider now a train of pulses where each pulse contains information in its amplitude or phase corresponding to a data bit or symbol. If $T_m >> T_s$, each multipath component can be resolved. However, multipath components from different symbols are received at the same moment, thus causing ISI.

The delay spread, T_m, is not necessarily the best indicator of how any given system will perform on a channel because different channels with the same value of T_m can exhibit very different profiles of signal intensity over the delay span. The power delay profile (PDP) represents the average power associated with a given multipath delay and is easily measured empirically. It is commonly used in channel type

definitions, such as the IEEE channel models, where the channel model is obtained through the PDP definition of each cluster. Mathematically, the PDP as a function of delay is defined as the channel response autocorrelation with time difference $\Delta t = 0$: $A_c(\mathfrak{d}) \triangleq A_c(\mathfrak{d}, 0)$. Note that $A_c(\mathfrak{d}, \Delta t)$ is expressed as

$$A_c(\mathfrak{d}, \Delta t) = E[h^*(\mathfrak{d}; t)h(\mathfrak{d}; t + \Delta t)]. \tag{2.3}$$

where $c(\mathfrak{d}, t)$ is the channel impulse response and $h^*(\mathfrak{d}, t)$ denotes the complex conjugate of $h(\mathfrak{d}, t)$.

Two more valuable and descriptive channel delay spread measures are the average and root mean square (RMS) delay spread. The average and RMS delay spread are defined in terms of $A_c(\mathfrak{d})$ as

$$\mu_{T_m} = \frac{\int_0^\infty \mathfrak{d} A_c(\mathfrak{d}) d\mathfrak{d}}{\int_0^\infty A_c(\mathfrak{d}) d\mathfrak{d}}, \tag{2.4}$$

and

$$\sigma_{T_m} = \sqrt{\frac{\int_0^\infty (\mathfrak{d} - \mu_{T_m})^2 A_c(\mathfrak{d}) d\mathfrak{d}}{\int_0^\infty A_c(\mathfrak{d}) d\mathfrak{d}}}. \tag{2.5}$$

Defining the mean and RMS delay spread by (2.4) and (2.5), respectively, weights the delay associated with a given multipath component by its relative power, so that weak multipath components contribute less to delay spread than strong ones. In particular, those components below the noise floor will not significantly impact these delay spread characterizations. Practical system designs are based on σ_{T_m} since those based on the strict delay spread definition lead to over-pessimistic solutions.

Frequency domain characterization. An analogous characterization of signal dispersion can be outlined in the frequency domain. It is obtained by evaluating the Fourier transform of the channel PDP with respect to $\mathfrak{d}$ ($A_C(\Delta f; \Delta t) = \mathfrak{F}(A_c(\mathfrak{d}, \Delta t))$, where $\mathfrak{F}(x)$ denotes the Fourier transform of x. Note that $A_C(\Delta f; \Delta t)$ represents the correlation between the channel's response to two signals as a function of the frequency and time separation.

A typical channel performance metric in the frequency domain is the coherence bandwidth, B_c, defined as the rank of frequencies over which the channel response can be considered flat. That is, two transmitted signals with a frequency separation smaller than B_c will experience the same fading or gain over the propagation channel. Its value is defined as the Δf for $\Delta t = 0$ where $A_C(\Delta f) \approx 0$ for all $\Delta f > B_c$. To unwind this definition, other authors define B_c as the rank of frequencies whose $A_C(\Delta f)$ is higher than 0.9. Similarly, using the $A_C(\Delta f)$ relation with the PDP, B_c can also be related to the RMS delay spread as $B_c \approx 0.2/\sigma_{T_m}$, that corresponds to a channel correlation exceeding the 0.9 value in the frequency domain.

In general, if a narrowband signal with a bandwidth $B_s << B_c$ is transmitted, then fading across the entire signal bandwidth is highly correlated, i.e., the fading is roughly equal across the entire signal bandwidth. Therefore, all of the signal's spectral components will be affected by the channel in a similar manner. This is usually referred to as flat fading or frequency non-selective. Otherwise, when the signal bandwidth $B_s >> B_c$, then the channel amplitude varies widely across the signal bandwidth. In this case the channel is said to be frequency selective.

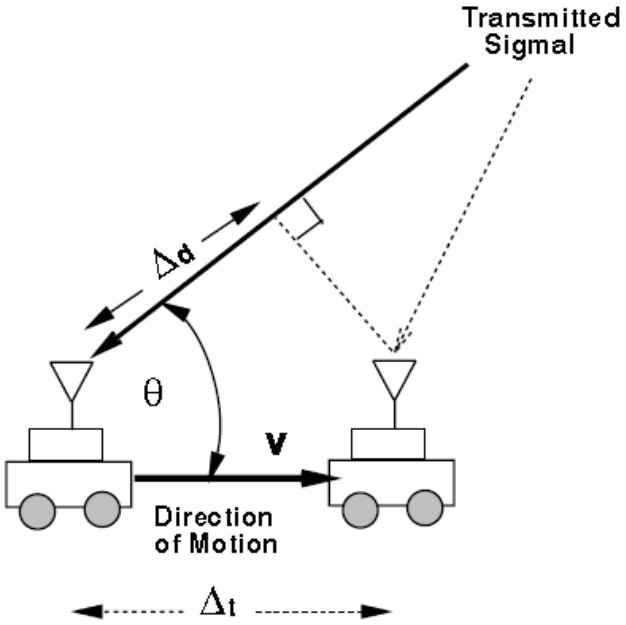


Figure 2.4: Geometry associated with Doppler shift (extracted from [1]).

Doppler Effect

Apart from the multipath fading, channel models must consider the time-varying channel response derived from the movements of any element involved in the transmission, such as transmitter, receiver or any of the objects whose presence affects the propagation path. As a result of such motion, for a transmitted continuous wave signal, the receiver suffers variations in the signal's amplitude and phase.

These movements produce a frequency shift on the received signal, known as Doppler shift. The Doppler shift, denoted by B_d, results from the fact that transmitter's or receiver's movements over a short time interval Δt cause a slight change in distance Δd on the propagation path. The phase change due to this path length difference is $\nu_d = 2\pi V \Delta t \cos\theta/\lambda_0$, where θ is the angle of arrival of the received signal relative to the direction of motion, $\lambda_0 = 3 \times 10^8/f_0$ is the signal wavelength (f_0 denotes the carrier frequency), and $\Delta d = V\Delta t \cos\theta$ with V being the receiver velocity with respect to the transmitter in the direction of motion. The geometry associated with the Doppler shift is shown in Fig. 2.4. Doppler effects can also be characterized by the Doppler frequency B_d that is obtained from the relationship between signal frequency and phase:

$$B_d = \frac{1}{2\pi}\frac{\Delta\nu_d}{\Delta t} = V\cos\theta/\lambda_0 \tag{2.6}$$

If the receiver is moving towards the transmitter, i.e. $0 \le \theta \le \pi/2$, then the Doppler frequency is positive ($B_d > 0$). Otherwise, it is negative ($B_d < 0$).

Time domain characterization The time-variant nature of the channel response can be easily characterized by measuring its autocorrelation in the time domain given by $A_c(\Delta t, f)$ function. $A_c(\Delta t, f)$ specifies the extent to which there is correlation between the channel's response to a signal sent at time t_1 and the response to a similar signal

sent at time t_2 ($\Delta t = t_2 - t_1$). The correlation function for a given Δt and frequency f can be obtained empirically by measuring the received signal correlation between two sinusoids delayed in time Δt seconds. In dense-scatterer channel models, with constant velocity of motion, and an unmodulated continuous wave signal, the normalized function $A_c(\Delta, \eth)$ is typically modelled as [45]:

$$A_c(\Delta t, f) = J_0(\frac{2\pi V \Delta t}{\lambda_0}) \tag{2.7}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind.

The coherence time, T_c, is a measure of the expected time duration over which the channel's response remains invariant. Therefore, for $t < T_c$ the channel response is highly correlated and the $A_c(\Delta t)$ is approximately constant. Note that T_c gives information about the rapidity of channel fading changes.

Suppose the transmission of a frame of N_f symbols, whose duration is $T_f = N_f T_s$. According to the relation between T_c and T_f, the channel degradation can be classified into fast or slow fading. A fast fading channel corresponds to the situation where $T_c < T_f$, thus implying that the transmitted frame suffers a variation of the channel response during its transmission. This effect can lead to several problems in the communication systems such as synchronization problems (failure of phase-locked-loop receivers), in addition to difficulties to adequately define a matched filter. When $T_c > T_f$ the channel is referred to as slow fading channel and the frame transmission is affected by an invariant channel response during the whole frame.

Frequency domain characterization. In the frequency domain, the channel impulse response, $S_c(\nu_d)$, can be characterized as function of the Doppler shift ν_d. $S_c(\nu_d)$ or Doppler power spectrum is highly dependent on the scenario configuration, thus presenting different profile spectrums according to the scenario configuration. For example, $S_c(\nu_d)$ of denser-scatterer models are characterized by a bowl shape spectrum, while, $S_c(\nu_d)$ of indoor radio channel models are commonly defined by a flat shape spectrum. Despite different spectrum shapes, all the $S_c(\nu_d)$ functions are usually centered on the carrier frequency f_0 and defined over the rank $[f_0 - B_d, f_0 + B_d]$.

$S_c(\nu_d)$ allows to characterize the spectral broadening as a function of the rate of change in the channel state. In a typical multipath environment, the received signal arrives from several reflected paths with different path distances and different angles of arrival, and the Doppler shift of each arriving path is generally different from any other path. The effect on the received signal is seen as a Doppler spreading or spectral broadening of the transmitter signal frequency, rather than a shift. Note that the Doppler spread, B_d, and the coherence time, T_c, are reciprocally related (within a multiplicative constant). Therefore, the approximate relationship between the two parameters is

$$T_c \approx \frac{1}{B_d}. \tag{2.8}$$

If, as commonly done, T_c is defined as the time duration over which the channel response of two spared signals in the time-domain experience a correlation factor higher than 0.5, then

$$T_c \approx \frac{9}{16\pi B_d}. \tag{2.9}$$

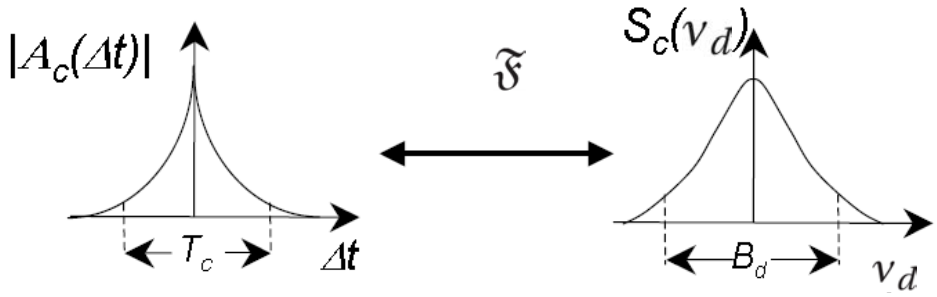


Figure 2.5: Doppler Power Spectrum, Doppler Spread, and Coherence Time (extracted from [1]).

A popular rule of thumb is to define T_c as the geometric mean of (2.8) and (2.9). This yields

$$T_c \approx \sqrt{\frac{9}{16\pi B_d^2}} = \frac{0.423}{B_d} \tag{2.10}$$

In Fig. 2.5, the Doppler power spectrum $S_C(\nu_d)$ and its inverse Fourier transform

$$A_C(\Delta t) \ = A_C(\Delta f = 0, \Delta t) \tag{2.11}$$

are illustrated.

Using the relation between B_d and T_c, a channel will be fast fading if the signal frame rate or bandwidth B_f is less than the Doppler frequency shift, $B_f < B_d$ or $T_f > T_c$. Conversely, the channel will be referred to as slow fading if the signalling rate is greater than the fading rate. That is, $B_f > B_d$ or $T_f < T_c$.

2.1.3 SISO channel modeling

Considering the previous fading descriptions, the channel impulse response can be mathematically expressed as a sum of signal copies arriving at different instants from different paths and affected by a common large-scale fading.

Denoting the transmitted signal as $\check{x}(t)$, the received signal for a single input single output (SISO) system considering the cluster modeling approach is given by

$$\check{u}(t) = \sqrt{G_{LS}(d)} \sum_{l=1}^{\mathfrak{L}} \sum_{\hat{n}=1}^{N_l} \alpha_{l,\hat{n}} \check{x}\left(t - \mathfrak{d}_{l,\hat{n}}\right) \tag{2.12}$$

where $G_{LS}(d)$ are the large scale fadings (comprising path loss and shadowing), $\alpha_{l,\hat{n}}$ and $\mathfrak{d}_{l,\hat{n}}$ are, respectively, the small-scale fading gain and propagation delay at time t on the $\hat{n}^{th}$ path from the cluster l. Since the propagation channel is linear time-invariant (LTV) [46], it can be described by the baseband equivalent channel impulse response

$$h(\mathfrak{d}, t) = \sqrt{G_{LS}(d)} \sum_{l=1}^{\mathfrak{L}} \sum_{\hat{n}=1}^{N_l} \alpha_{l,\hat{n}} e^{-j2\pi f_0 \mathfrak{d}_{l,\hat{n}}(t) + \nu_{d_{l,\hat{n}}}} \delta\left(t - \mathfrak{d}_{l,\hat{n}}\right). \tag{2.13}$$

Note that $\nu_{d_{l,\hat{n}}}$ is the Doppler shift of the $\hat{n}^{th}$ propagation path of cluster l introduced in the previous subsection.

2.2 TGn Channel Model

The TGn Channel Models are a set of WLAN indoor channel model types defined by the IEEE Task Group n (TGn) [2]. These models, suitable for multiple antenna systems, are based on the cluster modeling approach presented by Saleh and Valenzuela in [44] and further elaborated by Spencer et al. [47], Cramer et al. [48] and Poon and Ho [49]. Through the definition of different channel configuration parameters, such as, RMS delay spread, number of clusters, LOS/NLOS components or cluster configurations, these models are able to characterize the majority of indoor scenarios where WLAN technologies can be deployed (small rooms, medium or large offices,...).

Depending on the TGn channel model type, the number of clusters and the overall RMS delay spread varies from 2 to 6, and 0 (flat fading) to 150 ns, respectively. This is consistent with numerous experimentally determined results reported in the literature [44, 47, 48, 49, 50, 51] and also using ray-tracing methods [52]. The power of each tap in a particular cluster is determined so that the sum of the powers of overlapping taps corresponding to different clusters corresponds to the powers of the original power delay profiles. Next, AS, AoA, and AoD values are assigned to each tap and cluster (using statistical methods) so that they match experimentally determined values reported in the literature. Cluster AS was experimentally found to be in the 20^o to 40^o range [47, 51], and the mean AoA was found to be random with a uniform distribution. With the knowledge of each tap power, AS, and AoA (AoD), for a given antenna configuration, the channel matrix $\mathbf{H}$ is determined by using the MIMO channel impulse response described in next subsection.

Time-variations in the model are intended to emulate moving environmental scatterers. The prescribed Doppler spectrum consists of a bell-shaped part with low Doppler frequency and an optional additional peak at a larger Doppler frequency that corresponds to vehicles passing by [53]. A special feature of the model are channel time-variations caused by fluorescent lights. This is taken into account by modulating several channel taps to artificially produce an amplitude modulation. As an additional option, polarization can be included. For further details about the channel model description, please refer to the channel model's document in [2].

2.2.1 The MIMO channel modeling

As previously stated in the introductory chapter, the IEEE 802.11n physical layer has incorporated multiple antennas at the transmitter and receiver sides, through the use of the MIMO technology. MIMO notably increases the diversity gain by creating new propagation paths between the transmitter and the receiver [54]. This can be easily observed at the channel impulse response modeling of MIMO systems, consisting of a set of SISO channel impulse responses generated in parallel [3]. Each following the general characteristics previously described in section 2.1.

Consider a general MIMO system with N_T transmitter and N_R receiver antennas. The corresponding MIMO channel impulse response for the v^{th} tap can be expressed

as follows,

$$\boldsymbol{H}(\upsilon) = \sqrt{G_{LS}(d)} \begin{bmatrix} \alpha_{1,1}(\upsilon)e^{\phi_{1,1}(\upsilon)} & \alpha_{1,2}(\upsilon)e^{\phi_{1,2}(\upsilon)} & \cdots & \alpha_{1,N_T}(\upsilon)e^{\phi_{1,N_T}(\upsilon)} \\ \alpha_{2,1}(\upsilon)e^{\phi_{2,1}(\upsilon)} & \alpha_{2,2}(\upsilon)e^{\phi_{2,2}(\upsilon)} & \cdots & \alpha_{2,N_T}(\upsilon)e^{\phi_{2,N_T}(\upsilon)} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{N_R,1}(\upsilon)e^{\phi_{N_R,1}(\upsilon)} & \alpha_{N_R,2}(\upsilon)e^{\phi_{N_R,2}(\upsilon)} & \cdots & \alpha_{N_R,N_T}(\upsilon)e^{\phi_{N_R,N_T}(\upsilon)} \end{bmatrix} \tag{2.14}$$

where $\alpha_{n_R,n_T}(\upsilon) = \sum_{l=1}^{\mathfrak{L}} \alpha_{n_R,n_T,l}(\upsilon)$ and $\phi_{n_R,n_T}(\upsilon) = \sum_{l=1}^{\mathfrak{L}} \phi_{n_R,n_T,l}(\upsilon)$ are the signal amplitude and the phase shift, respectively, of the path linking the n_T transmitter and n_R receiver antennas for the υ^{th} tap. The $\alpha_{n_R,n_T,l}(\upsilon)$ and $\nu_{n_R,n_T,l}(\upsilon)$ components are the equivalent to the $\alpha_l(\upsilon)$ and $\nu_l(\upsilon)$ terms for the SISO signal model previously presented in (2.13), respectively.

In practical MIMO systems, there is some correlation among the signals received at different antennas [2, 3]. This correlation can be decreased by increasing the distance between antenna array elements; however, the space constraints in the majority of current systems does not allow an excessive array span. In any case, a small separation of $\lambda_0/2$ is enough to benefit from the diversity gain benefits of MIMO systems.

LOS and NLOS components, and antenna correlation in MIMO systems

Following the approximation of [55, 56], the channel response can be modelled as a MIMO matrix where the receive and transmit correlation are taken into account. The MIMO channel matrix $\boldsymbol{H}$ for each tap υ, at one instance of time, can be separated into a fixed (constant, LOS) matrix ($\boldsymbol{H}_F(\upsilon)$) and a Rayleigh (variable, NLOS) matrix ($\boldsymbol{H}_V(\upsilon)$),

$$\boldsymbol{H}(\upsilon) = \sqrt{G_{LS}(d)} \left(\sqrt{\frac{K}{K+1}} \boldsymbol{H}_F(\upsilon) + \sqrt{\frac{1}{K+1}} \boldsymbol{H}_V(\upsilon) \right) \tag{2.15}$$

with

$$\boldsymbol{H}_F(\upsilon) = \begin{bmatrix} e^{j\phi_{1,1,\upsilon}} & e^{j\phi_{1,2,\upsilon}} & \cdots & e^{j\phi_{1,N_T,\upsilon}} \\ e^{j\phi_{2,1,\upsilon}} & e^{j\phi_{2,2,\upsilon}} & \cdots & e^{j\phi_{2,N_T,\upsilon}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j\phi_{N_R,1,\upsilon}} & e^{j\phi_{N_R,2,\upsilon}} & \cdots & e^{j\phi_{N_R,N_T,\upsilon}} \end{bmatrix} \tag{2.16}$$

and

$$\boldsymbol{H}_V(\upsilon) = \begin{bmatrix} \chi_{1,1,\upsilon} & \chi_{1,2,\upsilon} & \cdots & \chi_{1,N_T,\upsilon} \\ \chi_{2,1,\upsilon} & \chi_{2,2,\upsilon} & \cdots & \chi_{2,N_T,\upsilon} \\ \vdots & \vdots & \ddots & \vdots \\ \chi_{N_R,1,\upsilon} & \chi_{N_R,2,\upsilon} & \cdots & \chi_{N_R,N_T,\upsilon} \end{bmatrix}. \tag{2.17}$$

In (2.17), the variable NLOS components $\chi_{n_R,n_T,\upsilon}$ (n_R^{th} receiving, n_T^{th} transmitting antenna and υ^{th} tap) are correlated zero-mean, unit variance complex Gaussian random variables. In (2.16), K is the Ricean factor representing the relative strength of the LOS component. It is assumed that each tap consists of a number of individual rays so that the complex Gaussian assumption is valid.

Before continuing, it should be pointed out that most of the channel parameters, such as, the number of taps, the average power of each tap or the AS, among others, are defined by the specific channel model employed [2].

Fixed LOS contribution The fixed LOS contribution is only applied to the first tap, as the LOS path clearly arrives first. Assuming that transmitter and receiver use uniform linear antenna arrays (ULA) with elements separated by a distance d_T and d_R, respectively, the fixed LOS matrix for the v^{th} tap can be written as [2, 3]

$$\boldsymbol{H}_F(v) = \begin{bmatrix} e^{j\phi_{1,1,v}} & \dots & e^{j\phi_{1,N_T,v}} \\ e^{j\phi_{2,1,v}} & \dots & e^{j\phi_{2,N_T,v}} \\ \vdots & & \vdots \\ e^{j\phi_{N_R,1,v}} & \dots & e^{j\phi_{N_R,N_T,v}} \end{bmatrix} = \exp\left(j2\pi\frac{V}{\lambda_0}t\cos(\pi/4)\right)\mathfrak{S} \tag{2.18}$$

where $\mathfrak{S}$ is the Rice Steering matrix:

$$\mathfrak{S} = \begin{bmatrix} 1 \\ e^{j2\pi(d_R/\lambda_0)\sin AoA} \\ \vdots \\ e^{j2\pi(d_R/\lambda_0)\sin((N_R-1)AoA)} \end{bmatrix} \begin{bmatrix} 1 \\ e^{j2\pi(d_T/\lambda_0)\sin AoD} \\ \vdots \\ e^{j2\pi(d_T/\lambda_0)\sin((N_T-1)AoD)} \end{bmatrix}^T. \tag{2.19}$$

NLOS matrix A correlated NLOS Rayleigh component can be expressed as [55],

$$\boldsymbol{H}_V(v) = \sqrt{\boldsymbol{R}_R}\boldsymbol{H}_{iid}(t)\sqrt{\boldsymbol{R}_T}^T \tag{2.20}$$

where $\boldsymbol{H}_{iid}(t)$ is an $(N_R \times N_T)$ matrix of independent, unit variance, zero-mean, complex random processes, and the transmit ($\boldsymbol{R}_T$) correlation matrix is defined as

$$\boldsymbol{R}_T = \begin{bmatrix} 1 & \rho_{T_{1,2}} & \cdots & \rho_{T_{1,N_T}} \\ \rho_{T_{2,1}} & \rho_{T_{2,2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{X_{N_T-1,N_T}} \\ \rho_{T_{N_T,1}} & \cdots & \rho_{T_{N_T,N_T-1}} & 1 \end{bmatrix} \tag{2.21}$$

with $\rho_{T_{i,j}}$ denoting the complex correlation coefficient between the i^{th} and j^{th} transmit antennas. Similarly, the receive correlation matrix, R_R, is analogously defined, replacing $\rho_{T_{i,j}}$ by $\rho_{R_{i,j}}$ that is the complex correlation coefficient between the i^{th} and j^{th} receive antennas. In the next paragraphs only the transmitter correlation coefficients $\rho_{T_{i,j}}$ are calculated. Their receiver correlation coefficient counterparts, i.e., $\rho_{R_{i,j}}$, are correspondingly obtained by replacing the transmitter variables by the receiver ones.

Assuming uniform linear array configurations, each complex correlation coefficient, can be obtained as

$$\rho_{T_{i,j}} = R^{\mathsf{RR}}_{T_{i,j}} + R^{\mathsf{RI}}_{T_{i,j}} \tag{2.22}$$

where $R^{\mathsf{RR}}_{T_{i,j}}$ is the cross-correlation between the real parts of the NLOS channel coefficients corresponding to the i^{th} and j^{th} antennas, and $R^{\mathsf{RI}}_{T_{i,j}}$ is the cross-correlation between the real and the imaginary parts. These parameters can be calculated as

$$\begin{aligned} R^{\mathsf{RR}}_{T_{i,j}} &= \int_{-\pi}^{\pi} \cos(d_{T_{i,j}}\sin\phi)\mathrm{PAS}_T(\phi)d\phi, \\ R^{\mathsf{RI}}_{T_{i,j}} &= \int_{-\pi}^{\pi} \sin(d_{T_{i,j}}\sin\phi)\mathrm{PAS}_T(\phi)d\phi, \end{aligned} \tag{2.23}$$

Channel model	σ_{T_m} [ns]	$\mathfrak{L}$	K LOS/NLOS	Description
A	0	1 tap	$0/-\infty$	Flat fading with (one tap at 0 ns delay model)
B	15	2	$0/-\infty$	Typical residential environment
C	30	2	$0/-\infty$	Typical residential or small office environment
D	50	3	$3/-\infty$	Typical office environment
E	100	4	$6/-\infty$	Typical large open space and office environments
F	150	6	$6/-\infty$	Large open space (indoor and outdoor)

Table 2.1: TGn channel models main characteristics (extracted from [2]).

where $d_{T_{i,j}}$ are the distance between the i^{th} and j^{th} transmit antennas and $\mathrm{PAS}_T(\phi)d\phi$ denotes the power angular spectrum (distribution of the signal power over angle) of the transmitter. Some examples of typical PAS distributions are uniform, truncated Gaussian or truncated Laplacian distributions.

As an example, the IEEE 802.11 TGn channel models use truncated Laplacian PAS shape:

$$\mathrm{PAS}_T(\phi) = \begin{cases} \sum_{l=1}^{\mathfrak{L}} \frac{Q_l e^{-\sqrt{2}|\phi-\phi_{0,l}|/\sigma_l}}{\sigma_l\sqrt{2}} & \phi_{0,l}-\Delta\phi_l \le \phi\phi_{0,l}+\Delta\phi_l \\ 0 & \text{otherwise} \end{cases}, \tag{2.24}$$

where $\phi_{0,l}$ is the AoA (or AoD) for cluster l, σ_l is the angular spread for cluster l, and Q_l are normalization constants, which satisfy

$$\sum_{l=1}^{\mathfrak{L}} Q_l \left[1 - e^{-\sqrt{2}\Delta\phi_l/\sigma_l}\right] = 1. \tag{2.25}$$

2.2.2 TGn channel types

The TGn model types are made of 6 profiles, named A to F, modeling different scenarios. Some of these model types are an extension of the SISO WLAN models proposed by Medbo et al. [57, 58]. In particular, channel models D, E and F correspond to the A, B, C channel models from Medbo [57], respectively. The other models have been obtained from empirical measurement campaigns. Each channel model defines different channel features, which are summarised in Table 2.1. Moreover, the cluster distribution for each profile is also defined in [2]. As an example, the PDP of TGn Channel D for each cluster is presented in Fig. 2.6. Note that clusters are presented using different colors, and for each path of the cluster, its delay path, average power, AoD, AS and AoD are specified, thus fully describing each path component.

	Tap index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	Excess delay [ns]	0	10	20	30	40	50	60	70	80	90	110	140	170	200	240	290	340	390
Cluster 1	Power [dB]	0	-0.9	-1.7	-2.6	-3.5	-4.3	-5.2	-6.1	-6.9	-7.8	-9.0	-11.1	-13.7	-16.3	-19.3	-23.2		
AoA	AoA [°]	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9	158.9		
AS (receiver)	AS [°]	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7	27.7		
AoD	AoD [°]	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1	332.1		
AS (transmitter)	AS [°]	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4	27.4		
Cluster 2	Power [dB]											-6.6	-9.5	-12.1	-14.7	-17.4	-21.9	-25.5	
AoA	AoA [°]											320.2	320.2	320.2	320.2	320.2	320.2	320.2	
AS	AS [°]											31.4	31.4	31.4	31.4	31.4	31.4	31.4	
AoD	AoD [°]											49.3	49.3	49.3	49.3	49.3	49.3	49.3	
AS	AS [°]											32.1	32.1	32.1	32.1	32.1	32.1	32.1	
Cluster 3	Power [dB]															-18.8	-23.2	-25.2	-26.7
AoA	AoA [°]															276.1	276.1	276.1	276.1
AS	AS [°]															37.4	37.4	37.4	37.4
AoD	AoD [°]															275.9	275.9	275.9	275.9
AS	AS [°]															36.8	36.8	36.8	36.8

Figure 2.6: Channel D cluster model (extracted from [2]).

New model	$d_{BP}(m)$	Slope before d_{BP}	Slope after d_{BP}	σ_{BP} (dB) before d_{BP} (LOS)	σ_{BP} (dB) after d_{BP} (NLOS)
A (optional)	5	2	3.5	3	4
B	5	2	3.5	3	4
C	5	2	3.5	3	5
D	10	2	3.5	3	5
E	20	2	3.5	3	6
F	30	2	3.5	3	6

Table 2.2: Path loss model parameters according to the IEEE 802.11n channel models (extracted from [2]).

Large-scale fading for TGn models.

In the IEEE 802.11n channel models, a combined path loss and shadow fading model is employed [2]. Throughout this work, path loss models using the configuration parameters presented in Table 2.2 have been applied. Note that different parameter values are taken depending on whether the distance (d) exceeds the break point distance (d_{PB}) or not, thus obtaining the path loss equation as follows

$$G_{\text{Path Loss}}(d) = \begin{cases} G_{FS}(d) & d \leq d_{BP} \\ G_{FS}(d) + 35log_{10}\left(\frac{d}{d_{BP}}\right) & d > d_{BP}. \end{cases} \tag{2.26}$$

For the shadow fading, a log-normal distribution with $\mu_{BP} = 0$ and σ_{BP} obtained from Table 2.2 is used. Shadow fading and path loss are combined by a simple addition,

$$G_{LS}(d) = G_{\text{Path Loss}}(d) + G_{\text{Shadow Fading}}\left(G_{\text{Path Loss}}(d)\right). \tag{2.27}$$

2.3 Channel Simulation Tool

In addition to the TGn channel models, the IEEE TGn committee also supported the development of its channel model emulator by AAU (Aalborg University)-CSys (Cellular systems division)[3]. In this Matlab simulation tool[1], a time-variant channel impulse response is provided for the user-defined Tx-Rx link configuration. These responses are obtained for each tap and they are spaced in time according to the output sampling rate; they correspond to time samples of the tap-channel matrix $\boldsymbol{H}_v$ presented in eq. (2.14). The link configuration, the number of Tx and Rx antennas, the antenna array separation at the Tx and Rx, the Tx-Rx distance and the channel model, among others, are user-configurable values in the simulator. Similarly, the output sample rate and the simulation duration are also setup parameters that can be easily customized. This configurability allows the simulation of lots of different scenarios and configurations, thus providing a convenient testbed for the IEEE 802.11n system performance evaluation.

2.3.1 Some simulation performance results

Without loss of generality, this subsection studies the simulator performance for a given scenario configuration. The selected scenario is a general one with similar conditions as those proposed in further chapters. In addition, the results are contrasted with the wireless channel features described in Sect. 2.1 and 2.2.

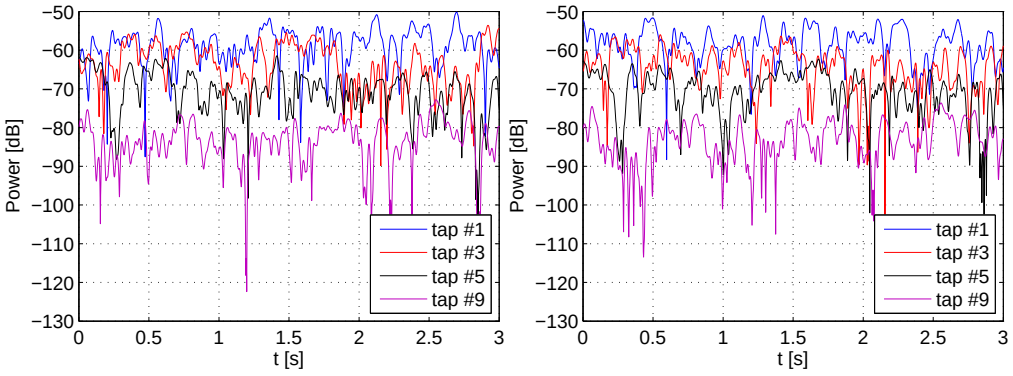
The setup used consists of a 2×2 MIMO system, i.e., $N_T = 2$ Tx antennas and $N_R = 2$ Rx antennas, with $d_T = \lambda_0$ and $d_R = \lambda_0/2$. The transmitter and receiver separation is 3 m and the carrier frequency f_0 is 5 GHz. For the simulation output configuration, the sampling rate is fixed to 125000 Hz with a downsampling factor of 125, thus leading to one channel sample for each millisecond. The simulation duration is set to 50 coherence times and, finally, only Channel model B and E are simulated.

In Figs. 2.7 and 2.8, the instantaneous power of several taps is presented for channel model B and E, respectively. To ease understanding, and without loss of generality, only several channel taps where depicted for each model. Particularly, the $1,3,5$ and 9, and the $1,3,5$ and 18 path indexes are shown for the Channel model B and E, respectively. Note that each figure frame is divided into four subfigures, one for each Tx-Rx propagation path.

In each subfigure of Figs. 2.7 and 2.8, each tap experiences fast and high power variations, as expected from the channel model descriptions given in Sect. 2.1. There, it was pointed out that the power variations of a given tap can be modelled following random distributions, particularly, a Rician or Rayleigh distribution depending on the presence or the lack of the LOS component, respectively. Apart from that, note that the fadings are independent among taps belonging to the same Tx-Rx propagation path, as it is expected, because they model different multipath components.

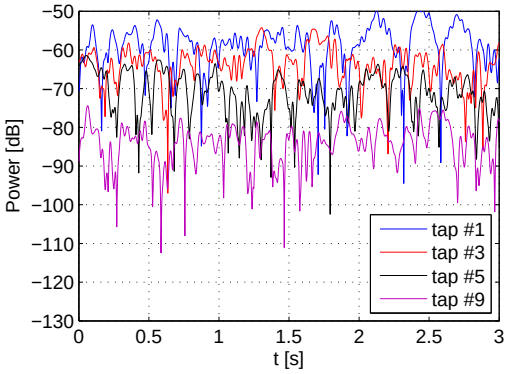
In Figs. 2.9 and 2.10, the correlation of several tap amplitudes is presented for channel models B and E, respectively. In these figures, the $\tilde{m}$ propagation path is referred to the path through which a hypothetical signal is transmitted from the n_T Tx antenna to the n_R Rx antenna. The $\tilde{m}^{th}$ value is set to $\tilde{m} = 2(n_T - 1) + n_R$. Note

[1] This simulation tool can be accessed by any person agreeing to its terms of use by emailing Laurent Schumacher at *lsc@info.fundp.ac.be*.

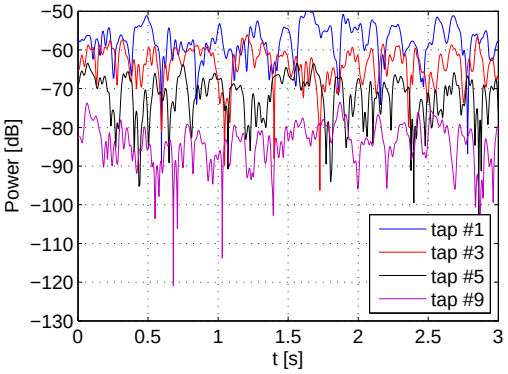


(a) $n_T = 1$, $n_R = 1$

(b) $n_T = 1$, $n_R = 2$



(c) $n_T = 2$, $n_R = 1$



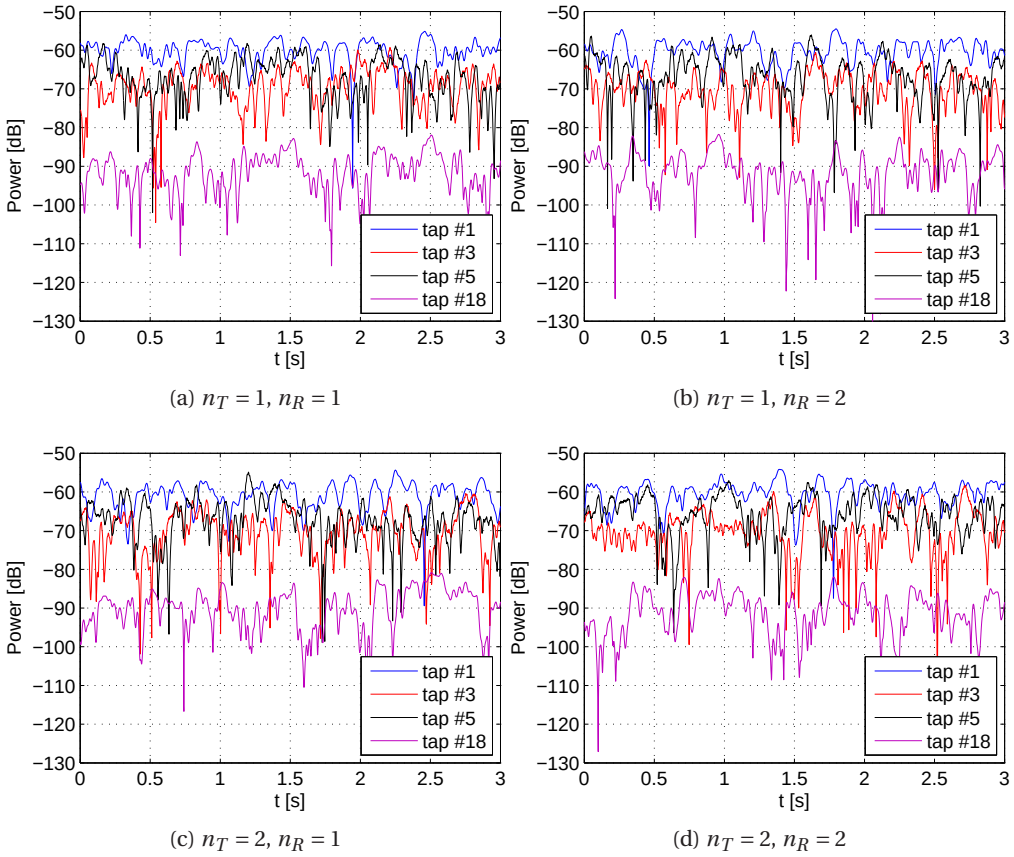
(d) $n_T = 2$, $n_R = 2$

Figure 2.7: Instantaneous power for several Channel model B taps.

that, in each subfigure, two graphs are presented. The graph above corresponds to the three-dimensional representation of the correlation factor between paths, while the frame below depicts the correlation of the first path ($\tilde{m} = 1$) with all the remaining paths ($\tilde{m} \in 2,3,4$). In Figs. 2.9a 2.9b, 2.10a and 2.10b, an important path correlation can be noticed for those paths where the transmitter or receiver antenna is shared. Specially, those cases where the reception antennas are shared, having lower separation distance ($\lambda_0/2$) of the array elements than in the transmitter side (λ_0). As it is expected, in most of the cases, the correlation between paths with non-shared antennas is negligible; for instance, see the correlation among paths $\{1,4\}$ or $\{2,3\}$. Similar conclusions can be observed from the correlation paths of other taps not presented in Figs. 2.9 and 2.10.

In Figs. 2.11 and 2.12, the time-variant transfer function, referred to as $\boldsymbol{H}_t[k]$, for path $\tilde{m} = 1$ of channel models B and E are presented, respectively. These transfer functions have been normalized dividing each amplitude by $\max_{t,k}\{|\boldsymbol{H}_t[k]|^2\}$. Note that on the frequency domain, the fading variations of the Channel model E are more pronounced than in Channel model B, thus implying a lower coherence bandwidth than Channel model B. This is mainly caused by the different density of scatterers.

Lastly, the channel coherence time has been empirically measured for a correlation threshold of 0.5, considering a velocity of the scatters of $v_0 = 1.2$ Km/h. The ob-



(a) $n_T = 1$, $n_R = 1$

(b) $n_T = 1$, $n_R = 2$

(c) $n_T = 2$, $n_R = 1$

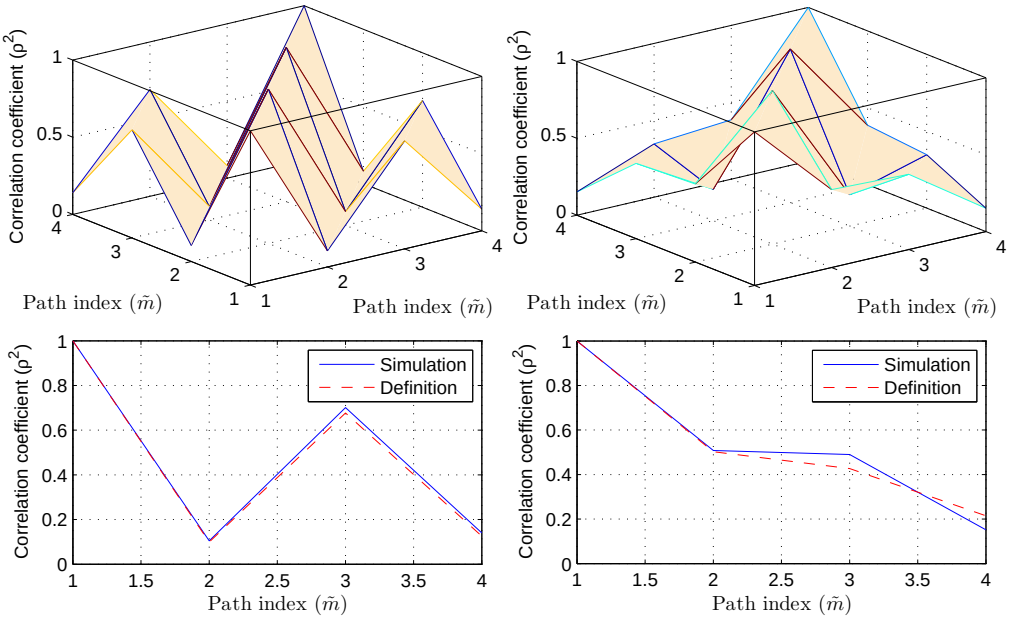
(d) $n_T = 2$, $n_R = 2$

Figure 2.8: Instantaneous power for several Channel model E taps.

tained coherence time is 33 ms.

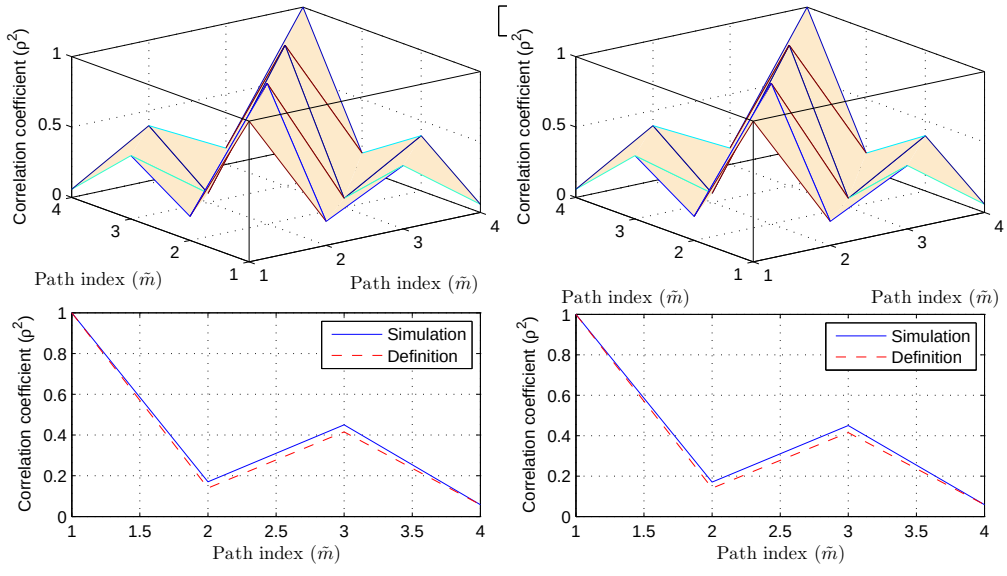
2.4 Chapter summary

In this chapter, the channel response modeling of the wireless LAN environments has been addressed. First, a general overview of the propagation impairments has been reviewed by detailing the large and small-scale fadings, as the two most important fading types inherent of any wireless communication system. Firstly, as large scale fading, the loss of power caused by path loss and shadow fadings have been detailed. Secondly, for the small-scale fadings, the multipath propagation and the Doppler effects have been studied in time- and frequency-domains. After that, and considering these channel fadings, the SISO channel model impulse response has been shown. A particular channel model specially designed for WLAN environments, namely the TGn Channel model, has also been presented. This model's main features have been described, putting special emphasis on the MIMO channel impulse response modeling. To conclude this chapter, several performance results of the channel simulation tool have been depicted and contrasted with the channel descriptions provided in previous sections.



(a) Tap #1 (b) Tap #3

Figure 2.9: Correlation coefficients for taps #1 and #3 of channel model B.



(a) Tap #2 (b) Tap #4

Figure 2.10: Correlation coefficients for taps #2 and #4 of Channel model E.

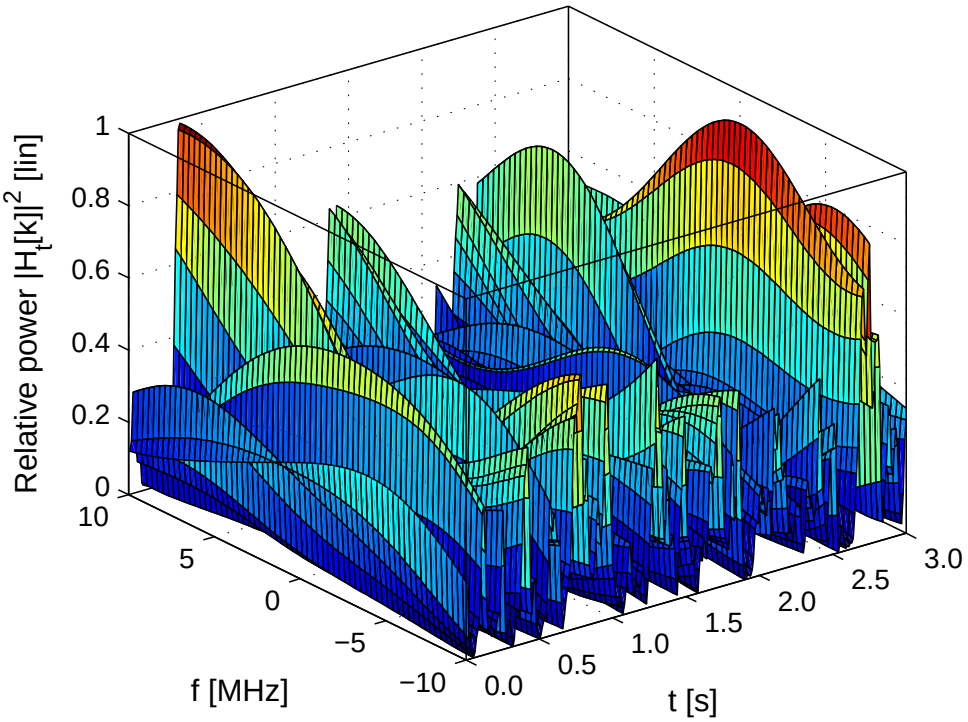


Figure 2.11: Time-variant transfer function for Channel model B.

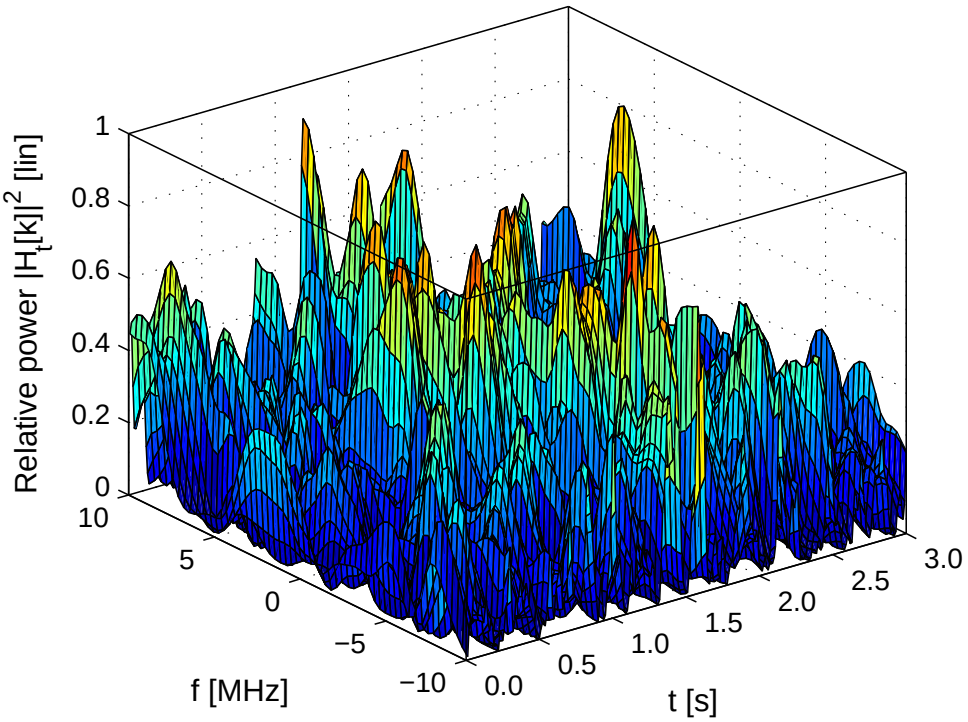


Figure 2.12: Time-variant transfer function for Channel model E.

CHAPTER 3

THE IEEE 802.11N PHYSICAL LAYER

This chapter presents the physical layer of the IEEE 802.11n. Firstly, in Sect. 3.1, the main IEEE 802.11n PHY transmission technologies are individually described, introducing their main characteristics, objectives and operation. Subsequently, in Sect. 3.2, these techniques are applied to review the set of allowed configurations included in the IEEE 802.11n standard. Next, the IEEE 802.11n system model is detailed in Sect. 3.3. Finally, the PHY performance evaluating different channel models and considering an imperfect channel estimator is presented in Sect. 3.4.

3.1 PHY Technology Overview

3.1.1 Orthogonal frequency division multiplexing (OFDM)

The orthogonal frequency division multiplexing (OFDM) has enjoyed wide acceptance in current standards such as IEEE 802.16 (Worldwide Interoperability for Microwave Access (WiMAX)), IEEE 802.11a/g/n or Digital Video Broadcasting -Terrestrial (DVB-T), among others [59]. OFDM is a particular type of multicarrier modulation (MCM) which, besides the parallel symbol transmission through multiple subcarriers, has the peculiarity of using orthogonal subcarriers with overlapping frequency bands. Remarkably, modulation and demodulation can be efficiently implemented using fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT) [60].

OFDM modulator

The main elements of a generic OFDM modulator are presented in Fig. 3.1. In the modulation process, the quadrature amplitude modulation (QAM) symbols that form the input stream are serial to parallel converted into as many substreams as available subcarriers. Then, the fist symbol of each substream ($\hat{s}[1] \dots \hat{s}[k] \dots \hat{s}[N_c - 1]$) is modu-

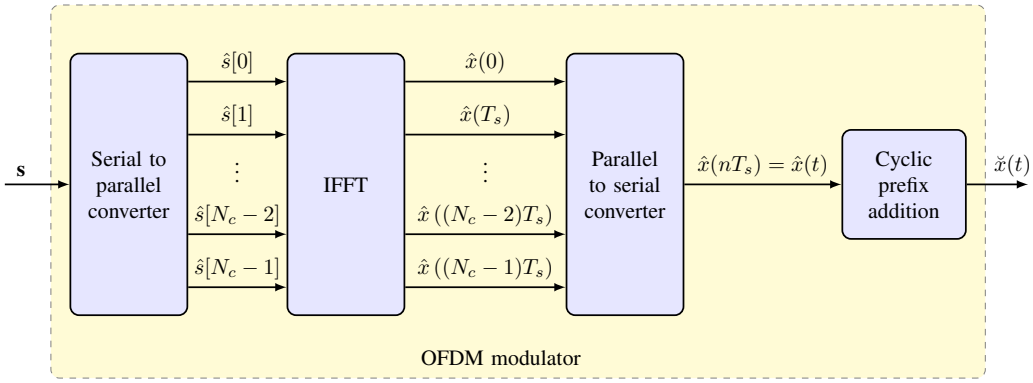


Figure 3.1: OFDM Tx system blocks.

lated by means of the IFFT processing block, obtaining the following sampled signal,

$$\hat{x}(nT_s) = \frac{1}{\sqrt{N_c}} \sum_{k=0}^{N_c-1} \hat{s}[k]\, e^{j2\pi f_k nT_s},\ n = 0,\ldots,N_c-1, \tag{3.1}$$

where N_c denotes the number of subcarriers, T_s the original symbol rate of $\hat{\boldsymbol{s}}$ and $f_k = \frac{kB_T}{N_c}$ is the k^{th} subchannel frequency with B_T denoting the total system bandwidth. Generally, N_c does not match the number of data substreams (N_d), since some subcarriers are used for channel estimation or synchronization purposes. Due to the FFT and IFFT design constraints, N_c is fixed to a 2^i, where i denotes a natural number.

After that, $\hat{x}(nT_s)$ is parallel-to-serial converted and then extended by adding a signal prefix, referred to as cyclic prefix (CP), consisting of the last μ signal samples. The CP usage has mainly two objectives, ISI removal and complexity reduction. The CP can be interpreted as a guard interval (GI) among consecutive OFDM symbols. It is important to notice that the CP reduces the system transmission rate by a factor of $N_c/(N_c+\mu)$, since each OFDM symbol duration is enlarged by μT_s time.

OFDM symbol through the channel

When $\check{x}(t)$ propagates though a LTV SISO channel, the received signal $\check{u}(t)$ is obtained as the convolution of $\check{x}(t)$ and the channel impulse response $\check{h}(t)$, plus the additive white Gaussian noise (AWGN) component ($\eta(t)$):

$$\check{u}(t) = \check{h}(t) * \check{x}(t) + \eta(t), \tag{3.2}$$

where $\check{h}(t) = \int_0^{T_m} h(\mathfrak{d},t)d\mathfrak{d}$, using $h(\mathfrak{d},t)$ as defined in (2.13). Note that due to the cyclic prefix addition, $\check{x}(nTs)$ is a periodic signal for $-\mu \le n \le N_c-1$ and its convolution with $\check{h}(t)$ results in a circular convolution that offers a reduced equalization complexity after applying the FFT [1].

OFDM demodulator

The general OFDM demodulator is illustrated in Fig. 3.2. It basically performs the inverse operation of the transmitter but incorporates several extra functions, such as

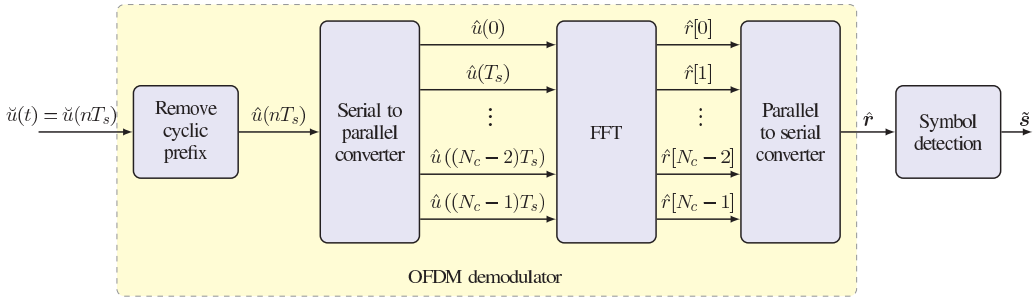


Figure 3.2: OFDM Rx system blocks.

the addition of synchronization and frequency offset estimation. In the first block, for each OFDM symbol, the received signal part corresponding to the CP, i.e., $\breve{u}(nT_s)$ for $-\mu < n < -1$, is removed, obtaining $\hat{u}(t) = \hat{u}(nT_s)$. The remaining N_c samples are then serial-to-parallel converted and processed though the FFT block. Assuming perfect synchronisation, the resulting signal, $\hat{r}$, consists of scaled versions of $\hat{s}[k]$ plus an AWGN component,

$$\hat{r}[k] = h[k]\hat{s}[k] + \zeta[k], \tag{3.3}$$

where $\zeta[k]$ is the AWGN noise and $h[k]$ is the flat fading channel gain associated with the k^{th} subchannel,

$$h[k] = \sum_{n=1}^{N_c} \breve{h}(nT_s)e^{-j2\pi f_k nT_s}. \tag{3.4}$$

Note that, thanks to the OFDM operation, each $h[k]$ models a flat fading response.

Thereafter, the FFT output is serialized, obtaining $\hat{\boldsymbol{s}}$ filtered by $h[k]$ and corrupted by the AWGN noise. To recover $\boldsymbol{s}$, each symbol $\hat{r}[k]$ is equalized by the inverse of the channel gain $1/h[k]$, and its output is passed through a M-ary quadrature amplitude modulation (M-QAM) symbol detector.

3.1.2 Multiple-input multiple-output (MIMO)

The initial excitement about MIMO was sparked by the pioneering work of Winters [61], Foschini [54], Gans [62], and Telatar [63] forecasting remarkable spectral efficiencies for wireless systems with multiple transmit and receive antennas.

In MIMO systems, the multiple antennas at the transmitter and receiver can be employed either to increase the data rate by multiplexing multiple data streams or to improve the transmission robustness through diversity techniques, such as, space time coding (STC), delay diversity or beamforming. Combinations of multiplexing and diversity are also possible. The costs of these enhancements range from the added complexity of deploying multiple antennas and their corresponding radio frequency (RF) chains, and the space and power requirements of these extra antennas (especially critical in small handheld units), to the requirement of more sophisticated signal processing [1]. Moreover, they demand accurate knowledge of the channel at the receiver, and, in some schemes, this is also needed at the transmitter.

The proposed MIMO schemes can be split into two groups depending on their performance objective: STC [64] and spatial division multiplexing (SDM) [62, 65]. The

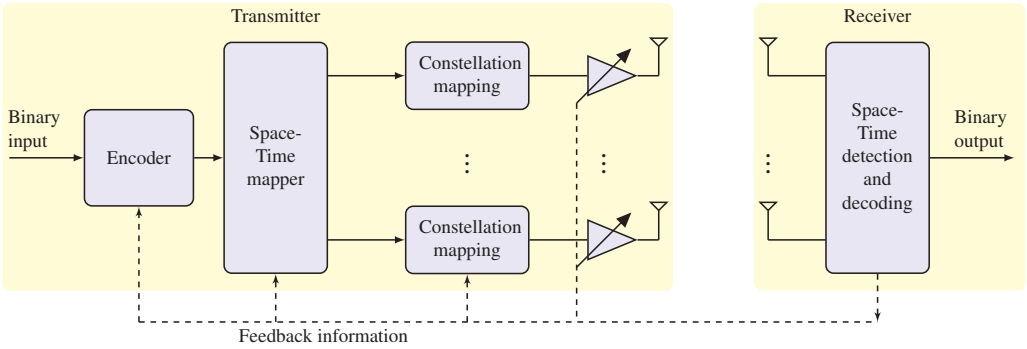


Figure 3.3: A general MIMO architecture.

STC schemes increase the robustness of the communication system by coding over the different transmitter branches, while SDM techniques achieve a higher data rate by simultaneously transmitting independent data streams on the different transmitter branches at the same carrier frequency. These systems offer a variety of trade-offs between capacity, error-rate performance, computational complexity and sensitivity to channel/interference estimation mismatch [66].

The following paragraphs introduce the main MIMO schemes, by referring to a general MIMO framework. Subsequently, the general single-carrier MIMO signal model and the STC and SDM techniques, are presented. In SDM, the MIMO detection techniques are covered by introducing the three main detection families: the linear, the successive interference cancellation (SIC) and the maximum likelihood (ML) detectors. Finally, other MIMO techniques adopted in the IEEE 802.11n, including e.g., beamforming, cyclic delay diversity (CDD) or antenna selection (ASEL), are briefly described.

General Structure

A MIMO structure that could form the basis of a general MIMO architecture is presented in Fig. 3.3 [66]. A Tx signal for a MIMO system with N_T transmit antennas is generated by performing the following tasks on the incoming bit stream:

- channel encoding
- mapping of the encoded bits on the spatial and/or temporal dimensions
- mapping of the (coded) bits into modulated symbols (and, if necessary, weighting of the resulting streams).

On the receiver side, detection is performed jointly over the spatial and temporal dimension. The complexity of this procedure strongly depends on the Tx signal design. In principle, the number of codewords grows exponentially with the size of the spatial and temporal dimensions thus compromising the design of optimal detection schemes. Some particular designs of the Tx signal, as will be shown later, allow receivers to achieve (near) optimal performance at a very low computational complexity.

Other two important common classification criteria are:

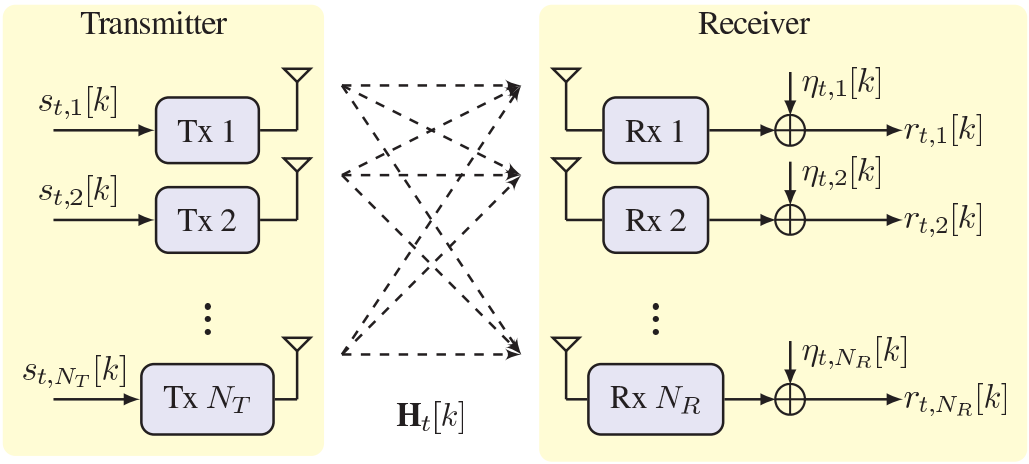


Figure 3.4: The MIMO system signal model.

- *Open-loop versus closed-loop techniques.* The distinction is made between systems that do not rely on knowledge of the channel responses at the transmitter -i.e., open-loop schemes- and systems that do assume partial or full availability of channel information at the Tx through some form of feedback mechanism -i.e., closed-loop schemes. In general, the feedback loop is designed to provide information for selection of the coding rate, constellation size, precoder design/selection, type of space-time mapping, and/or per antenna Tx power (see Fig. 3.3).

- *Joint coding (JC) versus per antenna coding (PAC).* Under the scheme known as JC (or vertical encoding, [67]), the original bit stream is first encoded and then demultiplexed into coded substreams, each of which is modulated and mapped onto the corresponding transmit antenna. In contrast, with PAC (or horizontal encoding, [67]), the original bit stream is first demultiplexed into a number of uncoded bit substreams which are then individually encoded, modulated and mapped onto the transmit antennas ([68]). The advantage of joint coding (JC) is that the coding is performed over the time and spatial dimensions, which could result in a better performance than the per antenna coding (PAC). The advantage of PAC, however, is that the receiver architecture might be less complex, since the encoding over the time and spatial dimensions is separated.

The MIMO signal model

Let us consider a MIMO system with N_T transmitter and N_R receiver antennas, such as the one presented in Fig. 3.4. As previously stated, note that the use of OFDM will ensure a flat fading channel response for all the symbols propagated through any subcarrier, allowing the application of MIMO techniques on a per-subcarrier basis, allowing us to focus on the MIMO performance over a single OFDM subcarrier.

For a given subcarrier k, the received signal vector at time instant t can be expressed as

$$\boldsymbol{r}_t[k] = \boldsymbol{H}_t[k]\boldsymbol{s}_t[k] + \boldsymbol{\eta}_t[k] \tag{3.5}$$

where $\boldsymbol{r}_t[k] = \left[r_{t,1}[k], r_{t,2}[k], \ldots, r_{t,N_R}[k]\right]$ is an N_R-element column vector containing the received signals on the N_R Rx antennas, $\boldsymbol{s}_t[k] = \left[s_{1,t}[k], s_{2,t}[k], \ldots, s_{N_T,t}[k]\right]^T$ denotes the transmitted symbols on the N_T Tx antennas, $\boldsymbol{\eta}_t[k] \sim C\mathcal{N}\left(0, \sigma_\eta^2 I_{N_R}\right)$ forms a $N_R \times 1$-dimensional vector representing the thermal noise, and $\boldsymbol{H}_t[k]$ is the $N_R \times N_T$ MIMO channel matrix for the k^{th} subcarrier,

$$\boldsymbol{H}_t[k] = \begin{bmatrix} h_{t,1,1}[k] & h_{t,1,2}[k] & \ldots & h_{t,1,N_T}[k] \\ h_{t,2,1}[k] & h_{t,2,2}[k] & \ldots & h_{t,2,N_T}[k] \\ \vdots & \vdots & \ddots & \vdots \\ h_{t,N_R,1}[k] & h_{t,N_R,2}[k] & \ldots & h_{t,N_R,N_T}[k] \end{bmatrix}, \tag{3.6}$$

with $h_{t,n_R,n_T}[k]$ denoting the channel coefficient linking Tx antenna n_T with Rx antenna n_R at time instant t and it is obtained from (3.4). In this case, each element of $\boldsymbol{H}_t[k]$ models a single frequency flat fading channel, assumed to be known at the receiver (e.g. via the prior transmission of training sequences). In (3.5), the elements of $\boldsymbol{H}_t[k]$ are properly defined to have unit variance, i.e., the average channel gain is normalized to one.

Unless mentioned otherwise, the total transmit power P_T is assumed to be uniformly distributed among the different transmit antennas. More precisely, the vector $\boldsymbol{s}_t[k]$ is assumed to have zero-mean and covariance matrix

$$\boldsymbol{Q}_s = E\left[\boldsymbol{s}_t[k]\boldsymbol{s}_t[k]^H\right] = \sigma_s^2 \boldsymbol{I}_{N_T} = \frac{P_T}{N_T}\boldsymbol{I}_{N_T}. \tag{3.7}$$

For a given MIMO channel realisation, $\boldsymbol{H}_t[k]$, the covariance matrix of $r_t[k]$ can then be written as

$$\begin{aligned} \boldsymbol{Q}_r &= E[\boldsymbol{r}_t[k]\boldsymbol{r}_t[k]^H] = E\left[(\boldsymbol{H}_t[k]\boldsymbol{s}_t[k] + \boldsymbol{\eta}_t[k])(\boldsymbol{H}_t[k]\boldsymbol{s}_t[k] + \boldsymbol{\eta}_t[k])^H\right] \\ &= \boldsymbol{H}_t[k]\boldsymbol{Q}_s\boldsymbol{H}_t[k]^H + \boldsymbol{Q}_\eta = \sigma_s^2\boldsymbol{H}_t[k]\boldsymbol{H}_t[k]^H + \sigma_\eta^2\boldsymbol{I}_{N_R}. \end{aligned} \tag{3.8}$$

Therefore, the average signal-to-noise ratio (SNR) at the output of the receive antenna n_R can be obtained as

$$\rho_{n_R} = \frac{E_s}{N_0} = \frac{\sigma_s^2\left(E\left[\boldsymbol{H}_t[k]^H\boldsymbol{H}_t[k]\right]\right)}{\left(\boldsymbol{Q}_\eta\right)} = \frac{N_T\sigma_s^2}{\sigma_\eta^2} = \frac{P_T}{\sigma_\eta^2} \tag{3.9}$$

where E_s represents the average transmitted energy per-symbol and N_0 denotes the noise power spectral density.

Space Time Coding (STC)

Different techniques have been used to efficiently code over both space and time dimensions. In this thesis we are going to focus on a particular space-time block coding (STBC) technique known as Alamouti STC [69], which has been included in the IEEE 802.11n standard.

Space Time Block Coding (STBC) The STBC input signal is assumed to be a stream of modulated symbols from a real or complex constellation. The STBC encoder maps

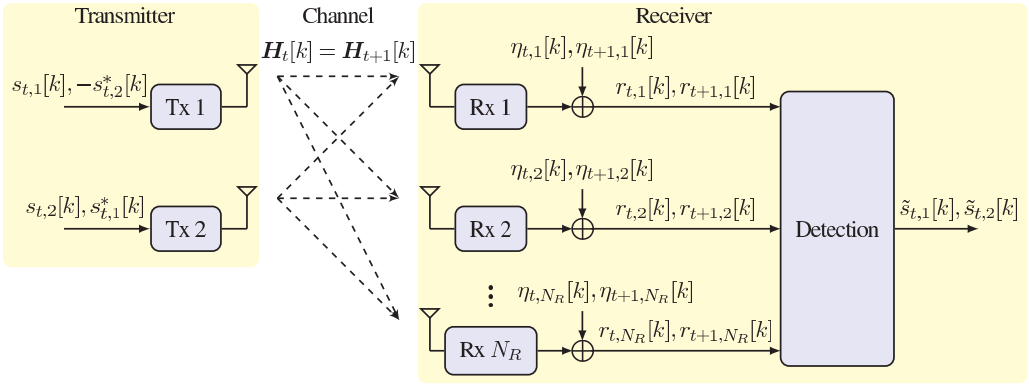


Figure 3.5: The Alamouti scheme.

these symbols onto codewords that span a block in both the spatial and temporal dimensions. The encoder operates on a block of input symbols producing an $N_T \times N_s$ encoding matrix $\boldsymbol{C}_{\text{STBC}}$, where N_s denotes the number of symbol periods over which the block encoding process is performed [66]:

$$\boldsymbol{C}_{\text{STBC}} = \begin{pmatrix} c_{1,1} & \cdots & c_{1,N_s} \\ \vdots & \ldots & \vdots \\ c_{N_T,1} & \cdots & c_{N_T,N_s} \end{pmatrix}. \tag{3.10}$$

One of the most common forms of STBC is the Alamouti coding scheme, shown schematically in Fig. 3.5.

Alamouti scheme The Alamouti scheme encodes one spatial stream into two space-time streams. It employs 2 transmit and N_R receive antennas, and performs the following three functions:

- *Information encoding and transmission*

  At time instant t, two symbols are simultaneously transmitted from the two Tx antennas on the k^{th} subcarrier, $s_{t,1}[k]$ and $s_{t,2}[k]$, respectively (see Fig. 3.5). The next transmission, at time instant $t+1$, consists of an encoded version of the previous ones, in particular, $s_{t+1,1}[k] = -s^*_{t,2}[k]$ and $s_{t+1,2}[k] = s^*_{t,1}[k]$. Note that this codification can be expressed as a specific case of STBC code with the following STBC encoding matrix:

  $$\boldsymbol{C}_{\text{Alamouti}} = \begin{pmatrix} c_1 & c_2 \\ -c_2^* & c_1^* \end{pmatrix}. \tag{3.11}$$

  Assuming the transmission over a flat fading channel, the fading can be modelled as a complex multiplicative distortion $h_{t,n_R,n_T}[k]$, where n_T and n_R are used to denote the transmit and receive antennas, respectively, and t is used to denote the time index. Assuming a constant fading across two consecutive symbols, then

  $$h_{t+1,n_R,n_T}[k] = h_{t,n_R,n_T}[k] = h_{n_R,n_T}[k]. \tag{3.12}$$

The signal at the output of receive antenna n_R during symbol periods t and $t+1$ can be expressed as,

$$\begin{aligned} r_{t,n_R}[k] &= h_{n_R,1}[k]s_{t,1}[k] + h_{n_R,2}[k]s_{t,2}[k] + \eta_{t,n_R}[k] \\ r_{t+1,n_R}[k] &= h_{n_R,1}[k](-s^*_{t,2}[k]) + h_{n_R,1}[k]s^*_{t,1}[k] + \eta_{t+1,n_R}[k] \end{aligned} \tag{3.13}$$

where $\eta_{t,n_R}[k]$ and $\eta_{t+1,n_R}[k]$ are complex random variables representing receiver noise.

- *Combining scheme of the received signals*

  At the receiver side, the received signals are linearly combined in order to obtain estimations of the original transmitted symbols, namely, $\tilde{s}_{t,1}[k]$ and $\tilde{s}_{t,2}[k]$.

  In a single Rx antenna system ($N_R = 1$, the combiner produces the following estimations

$$\begin{aligned} \tilde{s}_{t,1}[k] &= h^*_{1,1}[k]r_{t,1}[k] + h_{1,2}[k]r^*_{t+1,1}[k] \\ &= \left(\left|h_{1,1}[k]\right|^2 + \left|h_{1,2}[k]\right|^2\right)s_{t,1}[k] + h^*_{1,1}[k]\eta_{t,1}[k] + h_{1,2}[k]\eta^*_{t+1,1}[k] \\ \tilde{s}_{t,2}[k] &= h^*_{1,2}[k]r_{t,2}[k] - h_{1,1}[k]r^*_{t+1,1}[k] \\ &= \left(\left|h_{1,1}[k]\right|^2 + \left|h_{1,2}[k]\right|^2\right)s_{t,2}[k] - h_{1,1}[k]\eta^*_{t+1,1}[k] + h^*_{1,2}[k]\eta_{t,1}[k] \end{aligned} \tag{3.14}$$

  In case of employing a system with $N_R = 2$, the combination scheme leads to

$$\begin{aligned} \tilde{s}_{t,1}[k] =& h^*_{1,1}[k]r_{t,1}[k] + h_{1,2}[k]r^*_{t+1,1}[k] + h^*_{2,1}[k]r_{t,2}[k] + h_{2,2}[k]r^*_{t+1,2}[k] \\ =& \left(\left|h_{1,1}[k]\right|^2 + \left|h_{1,2}[k]\right|^2 + \left|h_{2,1}[k]\right|^2 + \left|h_{2,2}[k]\right|^2\right)s_{t,1}[k] + h^*_{1,1}[k]\eta_{t,1}[k] \\ &+ h_{1,2}[k]\eta^*_{t+1,1}[k] + h^*_{2,1}[k]\eta_{t,2} + h_{2,2}\eta^*_{t+1,2}[k] \\ \tilde{s}_{t,2}[k] =& h^*_{1,2}[k]r_{t,1}[k] - h_{1,1}[k]r^*_{t+1,1}[k] + h^*_{2,2}[k]r_{t,2}[k] - h_{2,1}[k]r^*_{t+1,2}[k] \\ =& \left(\left|h_{1,1}[k]\right|^2 + \left|h_{1,2}[k]\right|^2 + \left|h_{2,1}[k]\right|^2 + \left|h_{2,2}[k]\right|^2\right)s_{t,2}[k] - h_{1,1}[k]\eta^*_{t+1,1}[k] \\ &+ h^*_{1,2}[k]\eta_{t,1}[k] - h_{2,1}[k]\eta^*_{t+1,2}[k] + h^*_{2,2}[k]\eta_{t,2}[k] \end{aligned} \tag{3.15}$$

  Obviously, the Alamouti scheme can be extended to be used in systems with higher N_R values [69], however its main limitation is that it can only be applied to systems with two transmit antennas.

  Note that the combination of (3.14) and (3.15) leads to an expression very similar to that obtained when using 1×2 maximal ratio combining (MRC). The most notable difference being the power factors associated to the fact that the energy is distributed among the available transmit antennas.

- *The decision rule for maximum likelihood detection*

  Finally, the $\tilde{s}_{t,1}[k]$ and $\tilde{s}_{t,2}[k]$ symbol estimations are supplied to an ML detector in order to determine their respective transmitted symbols. Since as seen in eq. (3.15), the detection of a symbol does not include any interfering term caused by the other symbol, detection can be conducted on a symbol-by-symbol basis.

Spatial-Division Multiplexing (SDM)

In contrast to STC, the multiple antennas can also be used to increase the data rate. This can be achieved by simultaneously transmitting different streams of data on the different Tx antennas (over the same carrier frequency), using a transmission technique usually known as SDM. Although these parallel data streams are mixed up by the wireless channel, they can be recovered at the receiver by using multiple receive antennas and suitable signal-processing algorithms. Unfortunately, the main disadvantage is that no redundancy is added and, then, it might suffer from poor link reliability. To overcome this problem, additional channel coding can be introduced at the cost of reducing the data rate.

MIMO detection In SDM, MIMO detection refers to the process of determining the transmitted data symbols from the received signal vector. This requires both separating each transmitted data stream from the other transmitted streams, as well as performing channel equalization [70]. Broadly speaking, three main detection strategies exist: linear detectors, interference cancellation detectors and ML detectors.

- *Linear detectors.*

  The linear detectors are low-complexity MIMO detection schemes, which obtain symbol estimates by performing a linear transformation on the received data vector $\boldsymbol{r}_t[k]$. Mathematically, this reduces to estimate the symbol vector as,

  $$\tilde{\boldsymbol{s}}_t[k] = \boldsymbol{W}_t[k]\boldsymbol{r}_t[k]. \tag{3.16}$$

  Each linear detector scheme defines a particular matrix $\boldsymbol{W}_[k]$ according to its detection strategy, being the zero forcing (ZF) and the minimum mean squared error (MMSE) the most common linear detectors.

  Zero-forcing

  In this case, $\mathbf{W}_t[k]$ is employed to completely eliminate the interference caused by the other antenna symbols (interference cancellation), that is,

  $$\mathbf{W}_t[k] = \left(\mathbf{H}_t[k]^H\mathbf{H}_t[k]\right)^{-1}\mathbf{H}_t[k]^H. \tag{3.17}$$

  A drawback of the ZF linear detector is that it focuses solely on interference cancellation and this can remove signal energy of symbol components. The zero forcing is ideal when the channel is noiseless. However, when the channel is noisy, the zero-forzing equalizer will amplify the noise at those components where the small channel response magnitude [70].This problem is usually referred to as noise enhancement.

  MMSE

  The MMSE linear detector chooses $\mathbf{W}_t[k]$ to minimize the mean squared error (MSE), where $MSE = E\left[\|\mathbf{W}_t[k]\mathbf{r}_t[k] - \mathbf{s}_t[k]\|^2\right]$ and then the $\boldsymbol{W}_t[k]$ matrix for the MMSE is obtained as

  $$\mathbf{W}_t[k] = \left(\mathbf{H}_t[k]^H\mathbf{H}_t[k] + N_0\mathbf{I}\right)^{-1}\mathbf{H}_t[k]^H. \tag{3.18}$$

- *Successive interference cancellation (SIC).*

  This is a non-linear scheme that estimates the transmitted symbols iteratively from the received signal. Usually, these schemes consist of detecting the strongest symbol and then removing its contribution on the received signal, thus facilitating the later detection of the remaining symbols. This process is repeated until there are no more symbols to be estimated. For the symbol detection, linear techniques, such as, ZF or MMSE can be employed. Note that an erroneous detection during one of the first symbols leads to a non-desired modification of the received signal, and this will hinder the proper estimation of any of the remaining symbols. In general, an error at the first symbol produces several symbol errors and this effect is generally known as the error propagation problem, usual in SIC detection.

- *Maximum-likelihood detector.*

  The ML detector consists of an exhaustive search of the constellation symbol ($\mathbf{s}_t[k]$) in such a way that $\boldsymbol{H}_t[k]\mathbf{s}_t[k]$ has the minimum euclidian distance with respect to the received signal. Mathematically, this is expressed as,

$$\tilde{\boldsymbol{s}}_t[k] = \arg \min_{\mathbf{s}_t[k] \in \mathscr{S}} \|\boldsymbol{r}_t[k] - \boldsymbol{H}_t[k]\boldsymbol{s}_t[k]\|^2, \tag{3.19}$$

  thus improving the performance with respect to linear or SIC detectors [66]. Unfortunately, its complexity is notably higher than that of linear or SIC detection. A full complexity analysis in terms of number of operations (sums and multiplications) for each detector can be found in [66]. From there it can be concluded that the complexity of ML detectors increases exponentially with the number of transmitter antennas and also depends on the size of set $\mathscr{S}$. Meanwhile, the ZF and MMSE linear detector complexity grows only polynomially with N_T (N_T^3), irrespective of the size of set $\mathscr{S}$.

Other MIMO transmission techniques

Antenna selection (ASEL) The ASEL is a MIMO technique where among all the available antennas, only some of them are selected to perform the transmission and reception. The ASEL can be applied at the transmitter and/or receiver sides [71], encompassing a wide range of transmission and reception techniques. The ASEL is a low-cost low-complexity alternative that captures many of the advantages of MIMO systems. This process can reduce hardware complexity and cost, it can achieve full diversity and, in case of transmit ASEL, it can gain rate (or capacity) [71].

Note that, in case of relying on ASEL, a feedback path from the receiver to the transmitter is required in order to inform the transmitter about which antennas should be activated. The antenna selection process employs the channel state information (CSI) corresponding to all pairs of transmit and receive antennas, on all OFDM subcarriers [72]. Usually, the selection process consists of an exhaustive search to find the combination which offers the highest capacity or robustness against errors.

Beamforming Transmit beamforming is a technique where, in light of CSI at the transmitter, precoding matrices are designed with the objective of approaching the

channel capacity. Typically, beamforming implies a decomposition of channel coefficient matrix using singular value decomposition (SVD) and uses these decomposed unitary matrices as pre- and post-processing filters at the transmitter and the receiver to achieve near capacity gains [73]. Note that in beamforming, the feedback channel is required to provide CSI to the transmitter.

Cyclic delay diversity (CDD) The delay diversity is a simple transmit diversity technique that sends signal copies from multiple antennas with a different delay factor [74]. It can be shown that this results in an equivalent single-input channel with increased frequency-selectivity- and therefore, increased frequency diversity in comparison to a single antenna system. Interestingly, in case of using OFDM, this frequency diversity can be exploited by a FEC decoder. In contrast, the introduction of the additional delay requires a longer GI and consequently, the bandwidth efficiency is reduced. If the GI is not sufficiently long, the additional delay will cause inter-symbol and inter-carrier interference [75].

An elegant way to produce an increased frequency selectivity without exceeding the GI is the CDD scheme proposed in [76, 77]. Here, the delay applied on each antenna is introduced in a cyclic manner, so that the part of the delayed signal that should be transmitted after the non-delayed signal is appended at the beginning of the delayed signal. As a result, each antenna's signal is a cyclic shift of the original signal, elapsing the same transmission time than non-CDD systems and not exceeding the GI.

CDD is attractive as a low cost transmit diversity scheme. Furthermore, it can be applied to an arbitrary number of transmit antennas without suffering from a rate loss as in most STBC cases [75].

3.1.3 Bit interleaved coded modulation (BICM)

The flat-fading experienced on each subcarrier will typically lead to BER degradations. In order to combat this phenomenon, bit-interleaved coded modulation (BICM) is the most common solution, and it has been used in many standards such as WiMAX, Digital Video Broadcasting (DVB), and IEEE 802.11n.

The basic idea behind BICM over time and frequency is to encode data bits into codewords, interleave the resulting coded bits over both time and frequency, and then transmit the coded bits over the corresponding subchannel so that the coded bits within a given codeword all experience independent fading [78]. If most of the subchannels have a high SNR, the codeword will receive most coded bits correctly, and the errors associated with the few bad subchannels could be corrected by applying a Viterbi decoder. The Viterbi decoder performs a ML detection and is able to correct errors taking place in bits far apart in the stream decoded. Note that thanks to the interleaving, consecutive bits in the original stream suffer different subcarrier responses, and then the erroneous bits associated to a single subcarrier are scattered on the recomposition of the original bit-stream. This technique only works well if there is sufficient frequency diversity across the total system bandwidth. If the coherence bandwidth of the channel is large, then the fading across subchannels will be highly correlated, thus significantly reducing the effect of coding/interleaving.

MCS index (m)	Modulation	R_m
0	BPSK	1/2
1	QPSK	1/2
2	QPSK	3/4
3	16 QAM	1/2
4	16 QAM	3/4
5	64 QAM	2/3
6	64 QAM	3/4
7	64 QAM	5/6

Table 3.1: Modulation and coding schemes (MCS) for single-stream configurations.

3.2 High Throughput (HT) physical layer

In this section, the High Throughput (HT) PHY layer of the IEEE 802.11n is described by defining the supported configurations of MIMO, OFDM and BICM technologies. The IEEE 802.11n MIMO setup is designed to incorporate up to 4 antennas at each communication end. The supported MIMO techniques are SDM, ASEL, STBC, hybrid schemes, CDD and beamforming.

In IEEE 802.11n, the OFDM block configuration varies according to the selected bandwidth mode. In sharp contrast to previous amendments, the standard allows the 20 MHz and the 40 MHz bandwidth modes. In the 20 MHz mode, the total number of subcarriers is fixed to $N_c = 64$, out of which 52 are used for data transmission ($N_d = 52$) while the rest correspond to pilots and guard intervals. In the 40 MHz mode, N_c is set to 128 and N_d to 108. Regardless of the bandwidth mode, a variable CP can be employed, which extends the fixed OFDM symbol duration of 3.2 μs. The CP or GI, can be set to 0.4 μs, 0.8 μs or 1.6 μs. Obviously, a reduced GI provides higher transmission rates, however, in some channel models, this leads to a performance degradation caused by ISI since the GI duration is shorter than T_m.

Regarding BICM, the IEEE 802.11n employs a $R_{cc} = 1/2$ convolutional encoder. The main difference compared to the previous amendments (IEEE 802.11a/g) is the incorporation of the $R_m = \frac{5}{6}$ coding rate. As shown in Table 3.1, the MCS set includes $R_m = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}\}$ code rates and the BPSK, 4 QAM, 16 QAM, and 64 QAM modulation schemes. At the same time, IEEE 802.11n has also introduced the possibility to employ low-density parity-check (LDPC) codes [79] instead of the traditional $R_{cc} = 1/2$ convolutional encoder. However, their performance is not studied in this thesis.

In IEEE 802.11n, the number of streams (N_s) and the MCS employed on each of these streams is defined by the MCS index. In particular, the indices ranging from MCS0 to MCS32 correspond to equal modulation (EQM) MCSs that employ the same modulation on each stream, while the MCS33 to MCS76 indices specify unequal modulation (UEQM) MCSs applying different MCSs per stream. Note that although the MCS index is provided, the transmission rate cannot be specified yet, since some additional information, such as the bandwidth mode and the CP duration, is required.

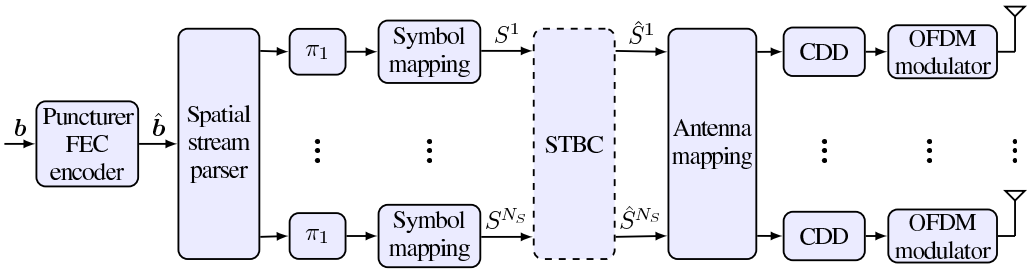


Figure 3.6: IEEE 802.11n Transmitter block diagram.

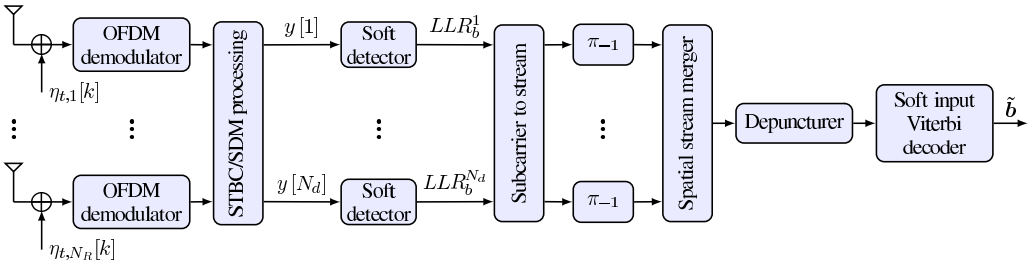


Figure 3.7: Generic IEEE 802.11n receiver block diagram.

3.3 System model

3.3.1 Transmitter Block diagram

Our study focuses on the IEEE 802.11n [5] whose physical layer transmitter structure is shown in Fig. 3.6. Information bits $\{b_1, b_2, \ldots, b_L\}$ are first encoded with a rate $R_{cc} = \frac{1}{2}$ convolutional encoder with generator polynomials $\mathbf{g} = [133, 171]$ (in octal form) and then punctured to one of the possible coding rates $R_m \in \{1/2, 2/3, 3/4, 5/6\}$. Depending on the selected MIMO configuration, the resulting bits are demultiplexed into N_s spatial streams. For each stream, the coded bits are interleaved and then mapped to symbols from one of the allowed constellations (BPSK, QPSK, 16-QAM or 64-QAM). In accordance with the selected MIMO configuration, the symbols are then either STBC encoded or antenna mapped on the available N_T transmit antennas. The resulting symbols are finally supplied to a conventional OFDM modulator consisting of an IFFT and the addition of a GI.

For simplicity of exhibition, this work focuses on a 2×2 MIMO system ($N_T = 2$ and $N_R = 2$), implying that MCSs with $N_s = 1$ and $N_s = 2$ spatial streams employ STBC [80] and SDM [54], respectively[1].

[1] Extension of this work to the use of CDD is straightforward as the CDD processing can be effectively modeled as a simple modification of the channel delay profiles [75].

3.3.2 Receiver

A generic block diagram of the receiver is depicted in Fig. 3.7. Reception begins by inverting the OFDM modulation (i.e. GI removal and FFT processing) to obtain the received baseband samples, which for the k^{th} subcarrier at time instant t can be expressed as

- SDM

$$\boldsymbol{r}_t[k] = \boldsymbol{H}_t[k]\boldsymbol{s}_t[k] + \boldsymbol{\eta}_t[k] \tag{3.20}$$

- STBC

$$\begin{aligned} \boldsymbol{r}_t[k] &= \boldsymbol{H}_t[k]\begin{bmatrix} s_{t,1}[k] \\ s_{t,2}[k] \end{bmatrix} + \boldsymbol{\eta}_t[k] \\ \boldsymbol{r}_{t+1}[k] &= \boldsymbol{H}_{t+1}[k]\begin{bmatrix} -s^*_{t,2}[k] \\ s^*_{t,1}[k] \end{bmatrix} + \boldsymbol{\eta}_{t+1}[k]. \end{aligned} \tag{3.21}$$

The STBC/SDM processing unit takes care of performing the appropriate spatial processing. Particularly, if $N_s = 1$, STBC has been applied at the transmitter side and usual Alamouti decoding is conducted. As it has been shown in section 3.1.2, the symbol estimation can be obtained as[2]

$$\boldsymbol{y}^{STBC}[k] = \Lambda_k \boldsymbol{s}[k] + \tilde{\boldsymbol{\eta}}[k] \tag{3.22}$$

where $\Lambda_k = \sum_{n_R=1}^{N_R}\left(\left|h_{1,n_R}[k]\right|^2 + \left|h_{2,n_R}[k]\right|^2\right)$. The output SNR corresponding to transmitted symbol j can then be calculated as

$$SNR_j[k] = \frac{P_T \Lambda_k}{N_T \sigma^2_\eta}. \tag{3.23}$$

Alternatively, if $N_s = 2$, a linear MMSE detector is applied on the received samples in order to decouple the two streams. That is, the MMSE symbol estimation can be written as

$$\boldsymbol{y}^{SDM}[k] = \boldsymbol{W}[k]\boldsymbol{r}[k]. \tag{3.24}$$

In this case, the post-MMSE equalizer SNR corresponding to transmitted symbol j is given by [81]

$$SNR_j[k] = \frac{1}{\left[\left(\frac{P_T}{N_T\sigma^2_\eta}\boldsymbol{H}^H[k]\,\boldsymbol{H}[k] + \boldsymbol{I}_{N_T}\right)^{-1}\right]_{j,j}} - 1. \tag{3.25}$$

It is well-known that soft decoding yields important benefits over hard-decision decoding. To this end, soft information in the form of likelihood ratios (LLRs) should be computed. Using (3.23), for the STBC case, or (3.25), for the SDM case, the LLR for the in-phase bit on the p^{th} position of the transmitted symbol j can be obtained as [82]

$$LLR\left(s_j[k], b_{I,p}[k]\right) = SNR_j[k] D_{I,p}[k] \tag{3.26}$$

where $D_{I,p}[k]$ is defined in Table 3.2 with $y_I[k] = \mathrm{Re}\{y_j[k]\}$. The LLRs for the quadrature bits are computed using an analogous procedure.

[2] In order to simplify the notation and due to independence among the transmitted/received symbol blocks, the time subindex t is dropped from this point onwards.

Modulation	$D_{I,1}[k]$	$D_{I,2}[k]$	$D_{I,3}[k]$
BPSK	$y_I[k]$		
QPSK	$y_I[k]$		
16QAM	$y_I[k]$	$-\left\|y_I[k]\right\|+2$	
64QAM	$y_I[k]$	$-\left\|y_I[k]\right\|+4$	$-\left\|\left\|y_I[k]\right\|-4\right\|+2$

Table 3.2: Auxiliary table for LLR computation.

MCS index (m)	N_s	Modulation	R_m	MIMO technique	$R_b^{(m)}$ [Mbps]
0	1	BPSK	1/2	STBC	6.5
1	1	QPSK	1/2	STBC	13.0
2	1	QPSK	3/4	STBC	19.5
3	1	16 QAM	1/2	STBC	26.0
4	1	16 QAM	3/4	STBC	39.0
5	1	64 QAM	2/3	STBC	52.0
6	1	64 QAM	3/4	STBC	58.5
7	1	64 QAM	5/6	STBC	65.0
8	2	BPSK	1/2	SDM	13.0
9	2	QPSK	1/2	SDM	26.0
10	2	QPSK	3/4	SDM	39.0
11	2	16 QAM	1/2	SDM	52.0
12	2	16 QAM	3/4	SDM	78.0
13	2	64 QAM	2/3	SDM	104.0
14	2	64 QAM	3/4	SDM	117.0
15	2	64 QAM	5/6	SDM	130.0

Table 3.3: Modulation and coding schemes (MCS) for one and two spatial streams using 0.8 μs cyclic prefix and 20 MHz bandwidth.

The resulting LLRs are reformatted into streams and deinterleaved. Finally, spatial deparsing is applied and the output LLR sequence is supplied to a soft Viterbi decoder after suitable depuncturing.

3.4 Physical layer performance results

In this section, the performance of the IEEE 802.11n physical layer using a 2×2 MIMO antenna system is evaluated. The system transmits on the 5.25 GHz carrier frequency and uses the 20 MHz bandwidth mode, thus implying the utilization of $N_c = 64$ and $N_d = 52$ subcarriers. Additionally, the CP is fixed to 0.8 μs and the packet size is fixed to 1664 bits. The corresponding MCS configurations with their respective rates and MIMO techniques are presented in Table 3.3. For the channel modelling, the channel realizations are obtained from the channel generation tool of [3], already described in Chapter 2. To this end, the system simulator has been configured with a Tx and Rx antenna spacing of λ and 0.5 λ, respectively.

In Fig. 3.8, the system PER performance of one stream MCS (or mode) is evaluated

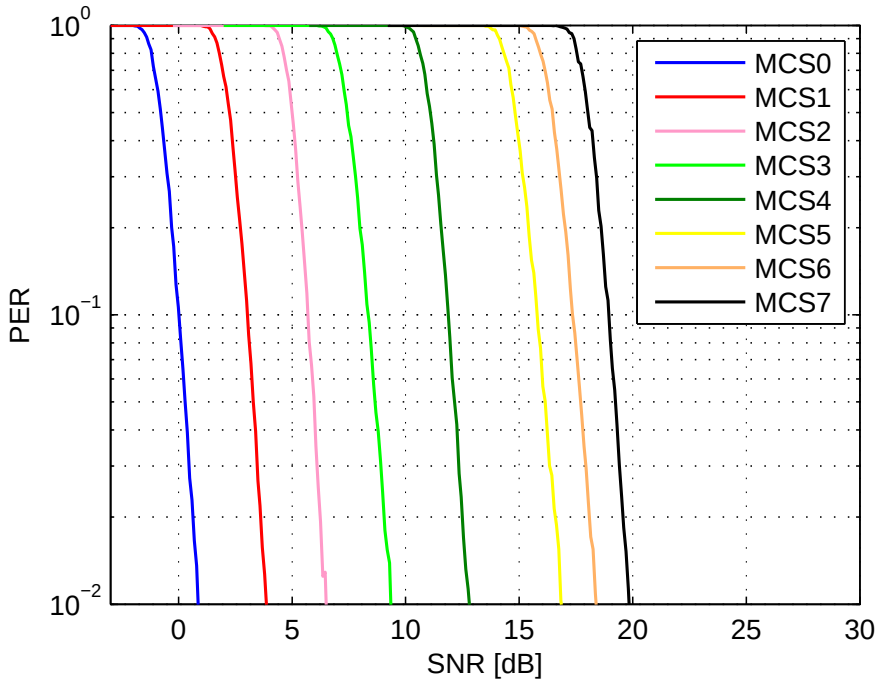


Figure 3.8: AWGN PER performance of one stream MCSs in a SISO system.

on the AWGN channel. Obviously, the best PER performance results are obtained by using the most robust MCS that combines the lowest modulation order and the highest coding rate. However, it provides the smallest transmission rate and usually its utilization implies an important waste of the available system resources, specially, in terms of system throughput. Therefore, there exists a clear trade-off between throughput and PER performance that must be considered in order to exploit efficiently the system's capabilities. In the next chapter, this trade-off is studied in detail by the introduction of the adaptive modulation and coding algorithms.

The system PER for $N_s = 1$ and $N_s = 2$ stream MCSs when using Channel model B is depicted in Figs. 3.9 and 3.10, respectively. Note that the fading effects introduced in Channel model B negatively affect the PER performance by reducing the curve slope of each MCS compared to that in the AWGN case (see Figs. 3.8 and 3.9).

From Figs. 3.9 and 3.10, it can be noticed an important difference in PER performance between $N_s = 2$ and $N_s = 1$ MCSs which employ the same modulation order and coding rate combination. In particular, higher PER values are obtained for the same SNR values and PER curve slope is also less pronounced than that of one stream MCSs. Obviously, these performance differences are motivated by the different usage of the spatial diversity gain of the 2×2 MIMO system. The two streams MCSs make use of SDM to achieve higher transmission rates while sacrificing the robustness offered by STBC, which is employed in one stream MCSs.

The system's throughput performance of $N_s = 1$ and $N_s = 2$ MCSs evaluated under Channel model B conditions is depicted in Fig. 3.11 and 3.12, respectively. The

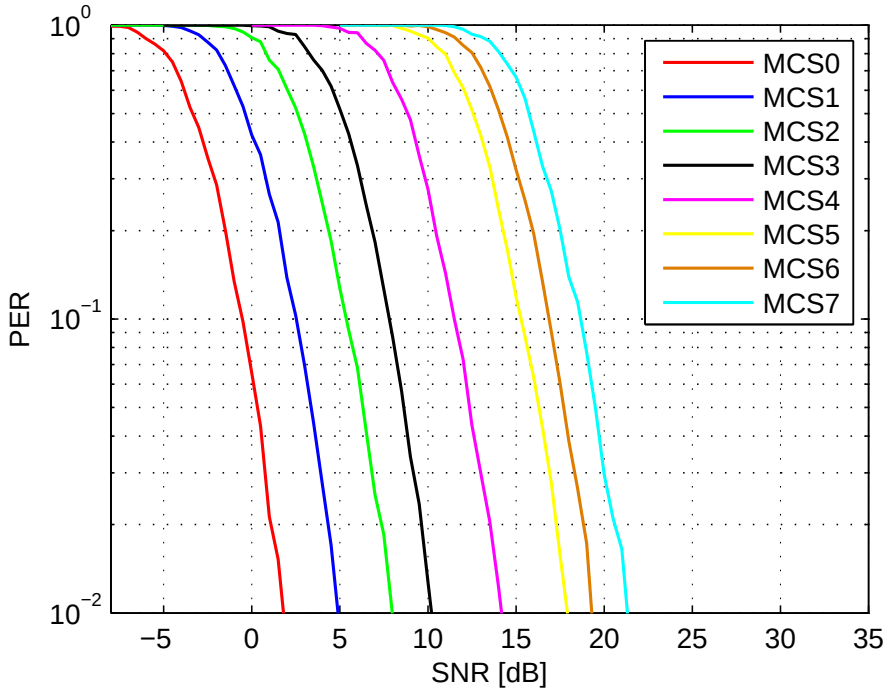


Figure 3.9: PER performance for 1 streams MCSs in Channel model B using STBC over a 2×2 MIMO system.

throughput has been defined as [83]:

$$\Gamma_m(\text{SNR}) = (1 - \text{PER}(m, \text{SNR}))\, R_b^{(m)}, \tag{3.27}$$

where m is the MCS index, PER$(m,$SNR$)$ denotes the PER for MCS m and a signal to noise ratio equal to SNR shown in Figs. 3.9 and 3.10, respectively, and $R_b^{(m)}$ denotes the transmission rate for MCS m (see Fig. 3.11 and 3.12). Note that similar throughput performance is obtained by those MCSs that share the same transmission rate having different number of streams (see also Table 3.3). Only small throughput improvements can be observed at low SNR values when using the one stream MCSs, since they usually offer better PER performance than that of the two stream MCSs (see Figs. 3.9 and 3.10).

The system PER performance comparison between Channel model B and E for $N_s = 1$ and $N_s = 2$ modes is depicted in Figs. 3.13 and 3.14, respectively. The higher frequential selectivity of Channel model E is exploited by the BICM to improve the PER performance gain with respect to Channel model B. This gain is more pronounced for two streams MCSs, because the worse detection performance of $N_s = 2$ MCSs is partially solved by the BICM scheme thanks to the higher frequency diversity provided by Channel model E.

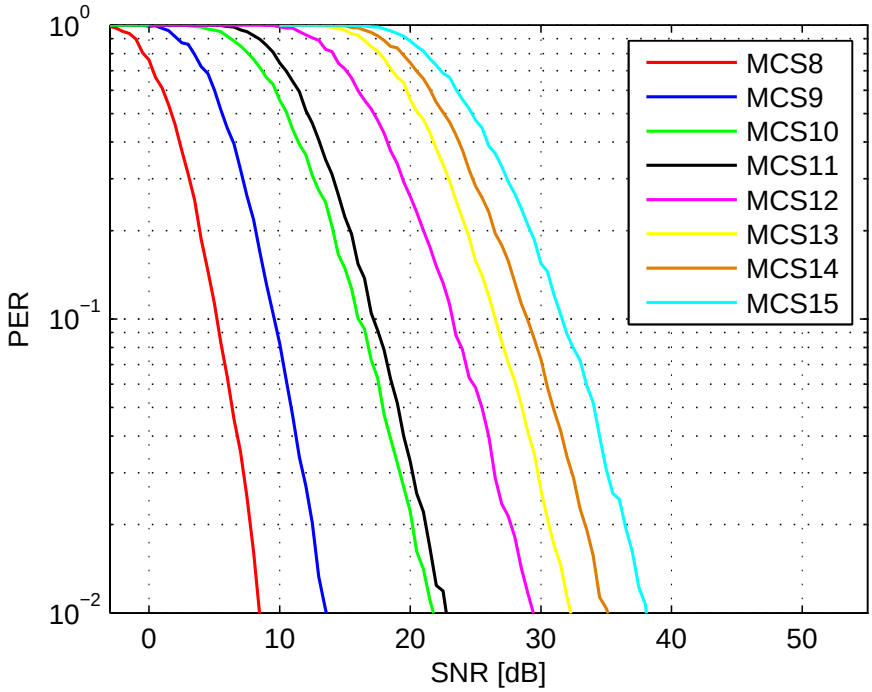


Figure 3.10: PER performance for 2 streams MCSs in Channel model B using SDM over a 2×2 MIMO system.

3.4.1 Imperfect channel estimation performance

In order to account for non-ideal CSI at the receiver side, the estimated channel is modelled as $\hat{\boldsymbol{H}}[k] = \boldsymbol{H}[k] + \boldsymbol{\xi}$ where $\boldsymbol{\xi}$ is an $N_R \times N_T$ matrix representing the channel estimation error with each element being Gaussian distributed with zero-mean and variance $\sigma_E^2 = \sigma_\eta^2 T_m / T_{OFDM}$, where T_{OFDM} is the OFDM symbol period. This variance corresponds to the MSE of the maximum performance bound of the optimal MIMO-multicarrier channel estimator proposed in [84].

The PER results when experiencing an imperfect channel estimation of Channel model B are presented in Fig. 3.15. From this figure, it can be observed that the PER performance suffers a slight degradation that translates to a non-noticeable system throughput loss. Note that in Channel model B, the error introduced by the estimation is negligible, since the T_m value of Channel model B is notably lower than $T_{OFDM} = 4\ \mu s$, and this fact notably reduces the error estimation. These observations are also valid for $N_s = 2$ modes.

In Figs. 3.16 and 3.17, the PER performance using imperfect channel estimation is evaluated on Channel model E for $N_s = 1$ and $N_s = 2$ streams, respectively. In contrast to Channel model B, the PER performance is seriously affected as can be inferred from the SNR losses of around 1.2, 1.4 and 1.6 dBs, for one and two streams MCS sets, respectively. In Channel model E, the T_m value is comparable to the CP duration, hence increasing the estimation error.

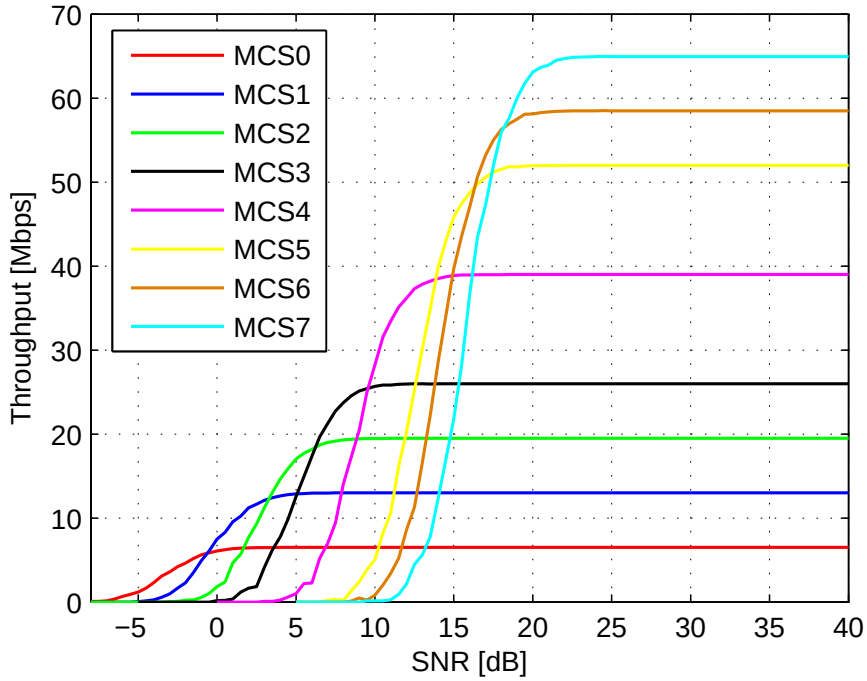


Figure 3.11: Throughput performance for 1 stream MCSs in Channel model B using STBC over a 2×2 MIMO system.

3.5 Chapter Summary

In this chapter the HT PHY as the new physical layer definition of the IEEE 802.11n amendment, has been fully presented, studying several transmission techniques and its performance under different channel conditions. Firstly, a general description of the main IEEE 802.11n PHY technologies has been introduced, including MIMO transmission techniques, OFDM schemes and BICM strategies. Later, the HT PHY is described by introducing the enabled configurations of MIMO, OFDM and BICM techniques. Additionally, the whole IEEE 802.11n signal model for the transmitter and the receiver has been provided using a general system configuration. Finally, the performance of IEEE 802.11n, employing fixed MCSs, has been evaluated for different channels and assuming imperfect channel estimation.

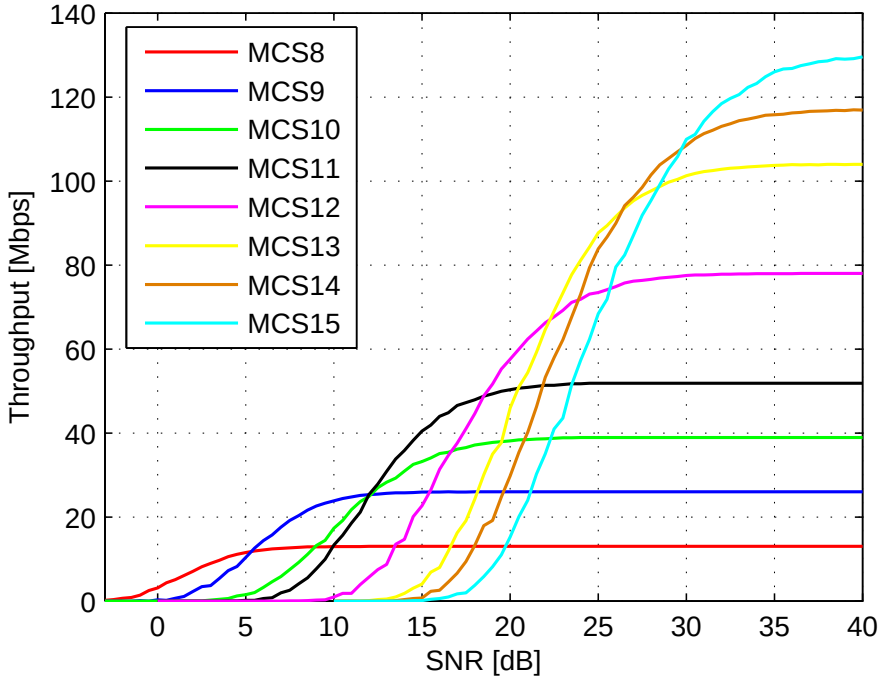


Figure 3.12: Throughput performance for 2 stream MCSs in Channel model B using SDM over a 2×2 MIMO system.

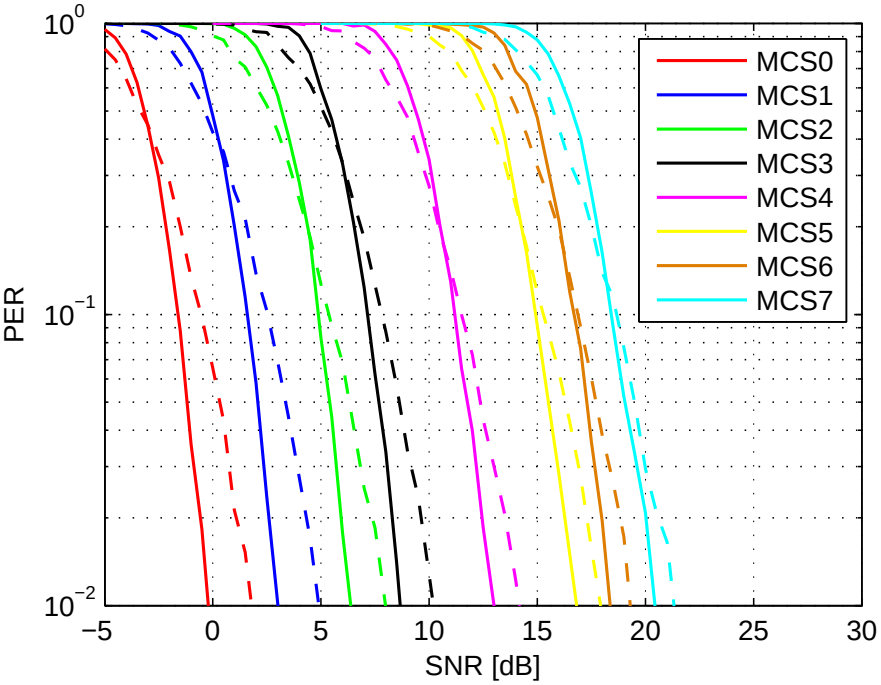


Figure 3.13: PER performance comparison between Channel model B (dashed lines) and E (continuous lines) cases for 1 stream MCSs.

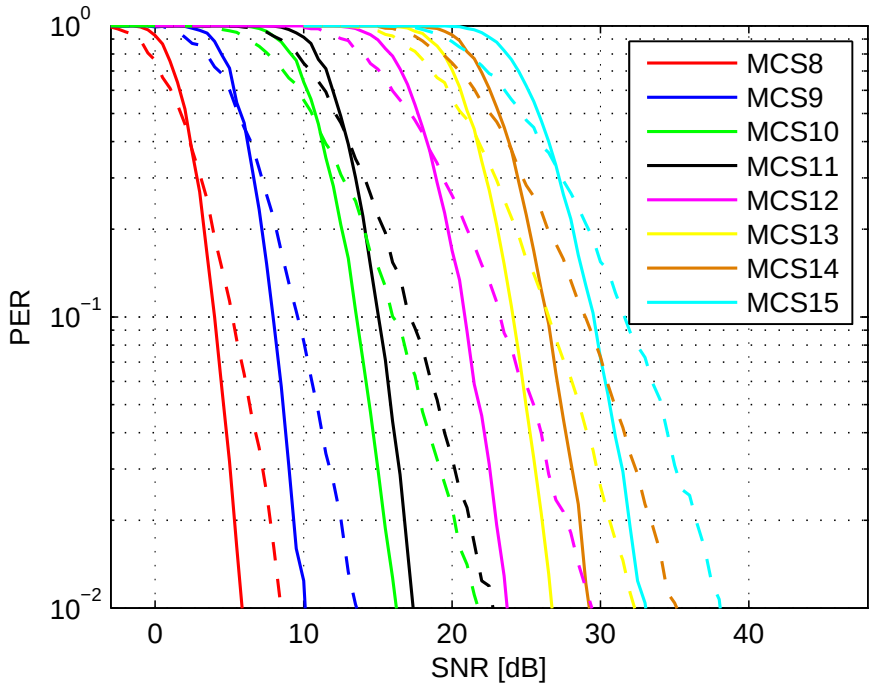


Figure 3.14: PER performance comparison between Channel model B (dashed lines) and E (continuous lines) cases for 2 streams MCSs.

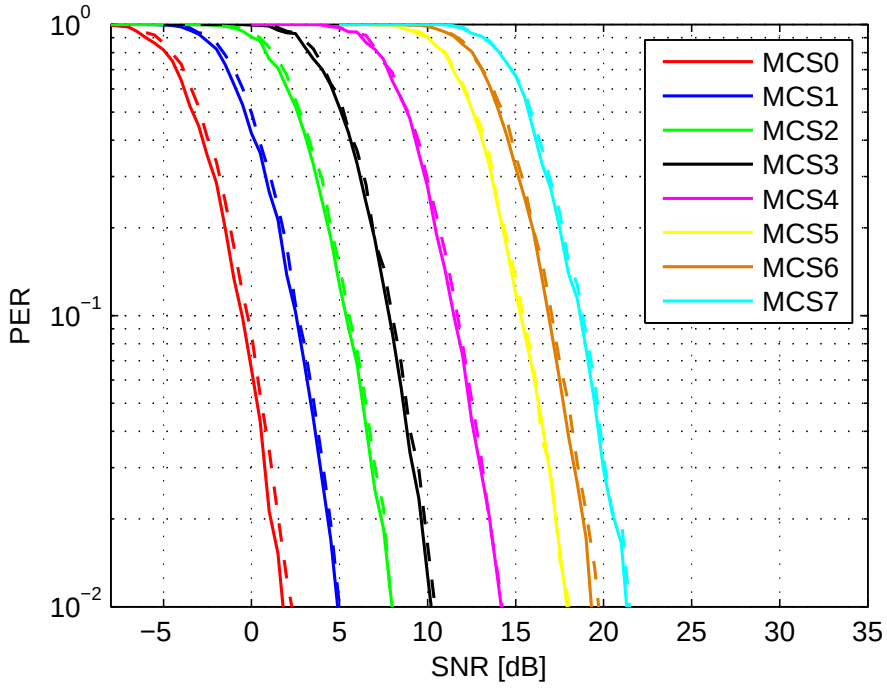


Figure 3.15: PER performance comparison using ideal (continuous lines) and imperfect (dashed lines) channel estimations for 1 stream MCSs in Channel model B.

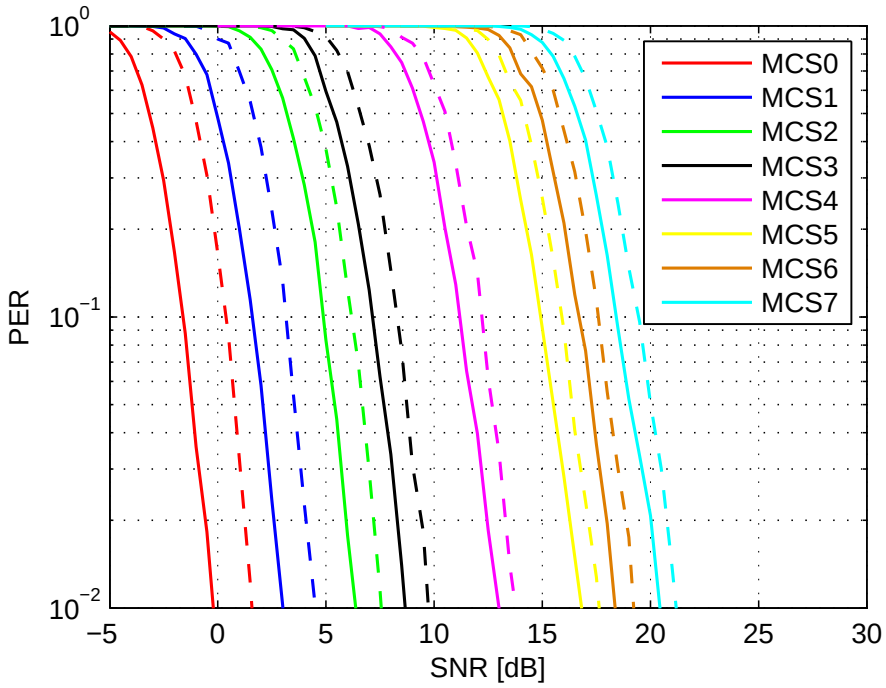


Figure 3.16: PER performance comparison using ideal (continuous lines) and imperfect (dashed lines) channel estimations for 1 stream MCSs in Channel model E.

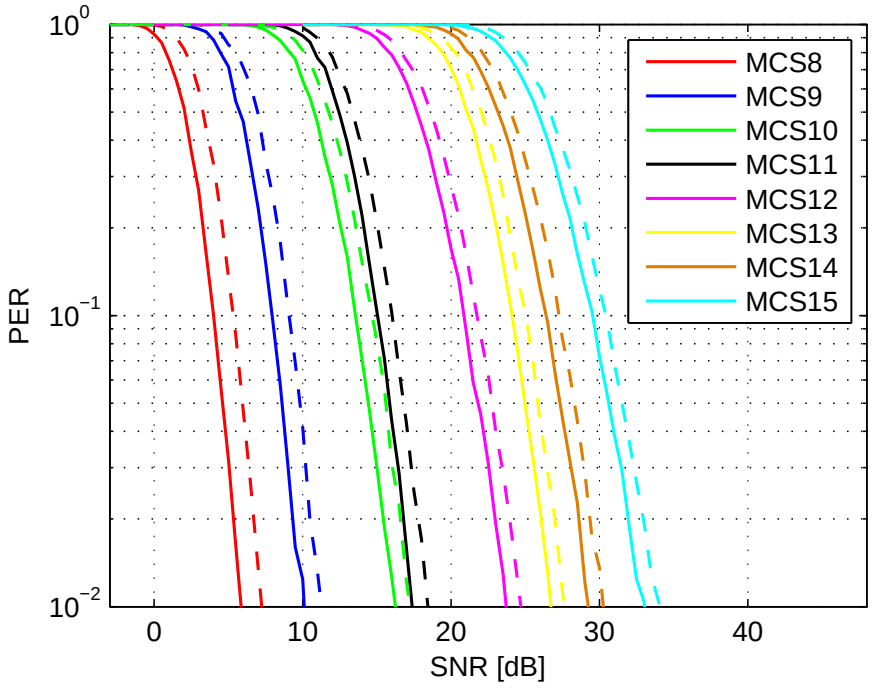


Figure 3.17: PER performance comparison using ideal (continuous lines) and imperfect (dashed lines) channel estimations for 2 streams MCSs in Channel model E.

CHAPTER 4

FAST LINK ADAPTATION IN IEEE 802.11N

This chapter considers the use of cross-layer FLA for WLANs employing a MIMO-OFDM physical layer. The FLA algorithm is a QoS-oriented adaptive modulation and coding (AMC) algorithm that selects the modulation and coding scheme (MCS) with the aim of maximizing the throughput while satisfying prescribed QoS constraints. To this end, FLA uses a packet error rate (PER) prediction mechanism that, without loss of generality, makes use of the exponential effective SNR mapping (EESM) and a selection algorithm. Both of these PER-based FLA mechanisms are duly presented in this chapter.

Additionally, a FLA scheme relying on bit error rate (BER) metrics is introduced that simplifies the link adaptation procedure without any significant performance degradation compared to the PER-based one. Results show that both PER- and BER-based FLA techniques perform in a close to optimum fashion in the sense of maximising the data throughput while satisfying prescribed quality of service constraints. Channel estimation errors have also been considered, revealing the importance of good channel estimators in order for FLA strategies to work satisfactorily.

4.1 Motivation and Problem Statement

Most current WLAN networks are based on one of the flavors of the IEEE 802.11 family of standards, notably, IEEE 802.11a/g and n. The physical layer of the IEEE 802.11 family of standards (see Chapter 3), typically employs OFDM and AMC based on BICM. Ideally, and based on some form of CSI, AMC strategies are used to select a combination of MCS aiming at the optimization of the spectral efficiency subject to QoS constraints such as, for instance, a maximum average PER outage probability, P_{out}, for a given target PER, PER_0. Mathematically, the *instantaneous* PER of MCS $m \in \mathscr{M}$, where $\mathscr{M}$ denotes the set of MCSs, can be expressed as

$$P_m = \mathscr{F}_m(\text{SNR}, L_M, \boldsymbol{H}), \tag{4.1}$$

where SNR is the received average signal to noise ratio, L_M is the MAC layer data unit (MPDU) packet length in number of information bits, $\boldsymbol{H}$ denotes a given channel realization and $\mathscr{F}_m(\cdot)$ represents an MCS-dependent mapping function. Thus, the optimum MCS selection process can be formulated as

$$m^* = \arg\max_{m \in \boldsymbol{\mathcal{M}}} \Gamma_m = \arg\max_{m \in \boldsymbol{\mathcal{M}}} R_b^{(m)} (1 - P_m) \tag{4.2}$$

subject to

$$Pr\{\text{PER} > PER_0\} \leq P_{\text{out}}, \tag{4.3}$$

where Γ_m and $R_b^{(m)}$ denote the instantaneous throughput and transmission rate of MCS m, respectively. This optimization process, usually known as FLA, clearly shows a close interaction between the PHY and the MAC sublayer and claims for a PHY-MAC cross-layer design.

Despite this claim, legacy WLAN standards only specify which MCSs are allowed for which types of MAC frames, but not how and when to switch between the permitted rates. Furthermore, there is no signaling mechanism specified that would allow a receiver to inform the transmitter about the actual link quality or the rate to be used. In order to overcome the lack of Rx-Tx feedback, the ARF link adaptation protocol and its modifications (see [11] and references therein) have been widely used in legacy WLANs. In this method, the PHY layer automatically switches to a lower rate after two consecutive transmission errors (missed positive acknowledgements (ACKs)) and switches to a higher rate either after ten successful transmissions (ACK reception) or after a time out. The rationale behind this approach is that for the MCS set used in legacy systems and for a given SNR, a higher transmission rate in the MCS implies a higher instantaneous PER.

IEEE 802.11n supports much higher transmission rates thanks to the use of MIMO antenna technology, the possibility of operating on a 40 MHz bandwidth (employing more subcarriers) and transmission modes using a reduced guard interval, as previouly stated in Chapter 2. Furthermore, the MAC sublayer does incorporate mechanisms to feedback information regarding MCS selection, thus making FLA a feasible option. In MIMO systems, in addition to MCS selection, link adaptation algorithms face another challenge: the MIMO mode selection. In this case, a higher transmission rate in the MCS does no longer imply a higher instantaneous PER and thus, the traditional link adaptation algorithms used in SIMO legacy systems become hardly effective. This motivates the development of cross-layer FLA algorithms for MIMO-OFDM systems. As it can be inferred from (4.2), the key elements of the FLA optimization process are, on one hand, a high quality instantaneous PER prediction tool at the PHY layer for all possible MCS/MIMO modes, packet lengths and channel realizations; and on the other hand, an MCS/MIMO mode selection methodology at the MAC sublayer that ensures the fulfilment of QoS constraints.

In contrast to SISO systems, where approximate PER closed-form expressions are available [1], in the MIMO-OFDM case there is no simple and systematic approach for predicting PER assuming arbitrary MCS/MIMO modes, packet sizes, and channel realizations in frequency selective channels with arbitrary channel correlations. Unlike previous related works such as [14, 15, 16], this chapter proposes PHY abstraction techniques which enable accurate PER prediction based on frame-by-frame BER prediction. These techniques are based on a common approach that maps system param-

eters like the selected MCS/MIMO operation mode and channel realization onto a link quality metric (LQM) that can be associated to the PER by means of simple look-up-tables, which are independent from the packet length. Consequently, the approach presented in this chapter greatly simplifies the calibration process of these look-up tables without significantly affecting system performance. Furthermore, appropriate LQMs are derived for different MIMO detection strategies used at the receiver side which allow the use of SDM- and STBC-based transmission modes. An MCS selection approach that fulfils the optimization constraint on the PER outage probability is introduced. Additionally, the impact of channel estimation errors on the performance of FLA techniques for IEEE 802.11n networks is considered.

This chapter is organized as follows: Sect. 4.2 describes the FLA algorithm, detailing the PER prediction methodology, the calibration procedure and the MCS selection process. Sect. 4.3 introduces the new BER-based FLA technique that simplifies the calibration process. The performance bounds algorithm (PBA) algorithm, an ideal AMC algorithm whose performance provides the maximum achievable throughput, is defined in Sect. 4.4. Sect. 4.5 deals with the FLA algorithm performance under Channel models B and E, for perfect and imperfect CSI. Finally, the chapter summary is presented in Sect. 4.6.

4.2 Fast link adaptation

As stated in Sect. 4.1, FLA algorithms are used to select the MCS that maximizes the instantaneous throughput subject to a constraint on the maximum average PER outage probability, P_{out}, for a given target PER, PER_0. Thus, a high quality instantaneous PER prediction tool is required at the PHY layer for all possible MCS/MIMO modes, packet lengths and channel realizations. Using these performance predictions, an MCS/-MIMO mode selection methodology can then be implemented at the MAC sublayer in order to ensure the fulfilment of QoS constraints.

As shown in (4.1), the PER can be expressed as a function of MCS $m \in \mathscr{M}$, the received SNR, the packet length L_M and the channel realization $\boldsymbol{H}$. It is important to remark that in MIMO-OFDM systems, each subcarrier on each spatial channel experiences a different channel response [85]. Consequently, the function $\mathscr{F}_m$ in (4.1) must consider each individual subcarrier/spatial response in order to predict the PER for a given system configuration. It is difficult to find a closed-form expression that accurately relates all the channel parameters into a single PER value. Due to these difficulties, several alternatives based on MCS-dependent tunable expressions have been proposed. These expressions, commonly known as link quality metric (LQM) functions, determine for each channel realization a scalar quality value that can be easily associated to a PER value. They are MCS dependent and must then be calibrated for each MCS in order to be effective.

4.2.1 PER prediction methodology

Different approaches have been proposed to map the instantaneous channel parameters onto a single LQM that could be associated to the PER by means of look-up tables obtained from off-line simulations (see, e.g., [12, 14] and references therein). Among

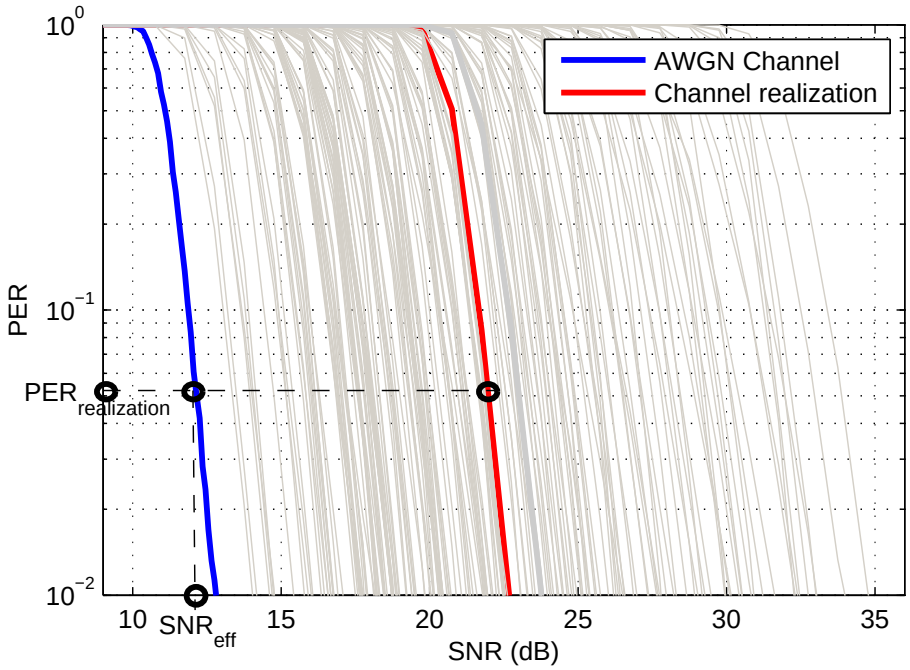


Figure 4.1: Effective SNR for a given channel realization and PER value ($PER_{realization}$) when using MCS12. Grey curves are PER performance results of other channel realizations.

all the proposed prediction strategies, those based on a one-dimensional mapping between a particular LQM, called effective SNR, and the PER are particularly interesting as demonstrated in [12] and references therein. Given an MCS $m \in \mathcal{M}$, the corresponding effective SNR can be defined as the SNR that would be required by this MCS on an AWGN channel to attain the same PER obtained over the frequency selective fading channel realization (see Fig. 4.1 where the effective SNR value for a given PER value and channel realization is presented). For a particular channel realization, $\boldsymbol{H}$, from which its post processing SNR can be computed, the effective SNR for a given MCS can be expressed as

$$SNR_{eff}^{(m)} = \psi_1^{(m)} J^{-1}\left(\frac{1}{N_S N_d}\sum_{j=1}^{N_s}\sum_{k=1}^{N_d} J\left(\frac{SNR_j[k]}{\psi_2^{(m)}}\right)\right) \tag{4.4}$$

where $SNR_j[k]$ is the post processing SNR on subcarrier k, N_d denotes the number of data subcarriers per OFDM symbol, N_s is the number of streams of the MCS m, $J(\cdot)$ is a model-specific LQM function and $J^{-1}(\cdot)$ is its inverse. Note that $SNR_j[k]$ was previously provided in Sect. 3.3. The parameters $\psi_1^{(m)}$ and $\psi_2^{(m)}$ allow the model to be adapted to the characteristics of the corresponding MCS and they are calculated off-line using a calibration procedure. As it can be deducted from the effective SNR definition, the look-up tables are the AWGN channel curves that take care of mapping the SNR to the PER system performance over the AWGN channel.

Different LQM functions can be employed to perform the effective SNR mapping, for example:

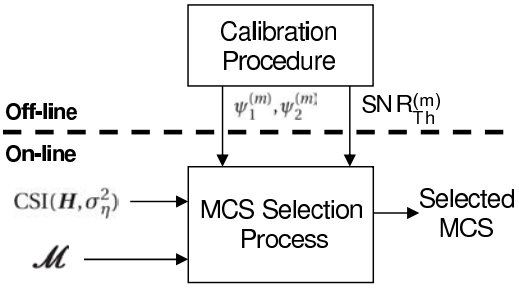


Figure 4.2: FLA block diagram.

- Capacity effective SNR metric (CESM)

$$J(\gamma) = \log_2(1+\gamma).$$

- Exponential effective SNR metric (EESM)

$$J(\gamma) = \exp(-\gamma).$$

- Mutual information effective SNR metric (MIESM)

$$J(\gamma) = \mathscr{I}_m(\gamma),$$

  where $\mathscr{I}_m(\cdot)$ is the mutual information of MCS m.

- Logarithmic effective SNR metric (LESM)

$$J(\gamma) = \log_{10}(\gamma).$$

These functions use the post processing SNR for each transmitted symbol and determine the link quality for each OFDM symbol. As it can be observed, most of these expressions are simple to compute and have similar complexity (only MIESM methods are more computationally intensive to implement). In this thesis, and without loss of generality, we concentrate on the exponential effective SNR mapping (EESM)-based PER prediction strategy.

As shown in Fig. 4.2, the proposed FLA scheme consists of an off-line calibration procedure and an on-line MCS selection process. In the off-line phase, a large set of channel realizations is used to obtain system performance results in terms of PER. These results then serve to calibrate the parameters of LQM functions which can later be used during the on-line operation in order to determine the best MCS for the current channel realization, packet length and SNR. The following subsections explain in greater detail the steps conducted during the off-line and on-line phases.

4.2.2 Calibration procedure (off-line)

The optimum values for the fitting parameters $\psi_1^{(m)}$ and $\psi_2^{(m)}$ in (4.4) are obtained off-line by an exhaustive search aimed at minimizing the RMS $SNR_{eff}^{(m)}$ estimation error

MCS (m)	N_s	Modulation	R_m	$\psi_1^{(m)}$	$\psi_2^{(m)}$	MSE	$SNR_{th}^{(m)}$ [dB]
0	1	BPSK	1/2	0.6865	0.6862	0.0008	0.0093
1	1	QPSK	1/2	1.1469	1.1472	0.0007	3.0269
2	1	QPSK	3/4	1.3598	1.3612	0.0007	5.7037
3	1	16QAM	1/2	3.1790	3.1801	0.0012	8.3704
4	1	16QAM	3/4	5.5062	5.5075	0.0008	11.8879
5	1	64QAM	2/3	15.0853	15.1442	0.0017	15.8023
6	1	64QAM	3/4	20.2015	20.2077	0.0014	17.3917
7	1	64QAM	5/6	25.2049	25.2330	0.0015	18.9456
8	2	BPSK	1/2	1.1507	1.0663	0.4753	0.9964
9	2	QPSK	1/2	2.7012	2.6926	0.0326	3.1574
10	2	QPSK	3/4	2.0991	2.0792	0.0125	5.7615
11	2	16QAM	1/2	8.6236	8.9311	0.0300	8.4177
12	2	16QAM	3/4	9.5287	9.5111	0.0246	11.9284
13	2	64QAM	2/3	31.6432	32.8464	0.0429	15.8454
14	2	64QAM	3/4	33.6721	32.9344	0.0337	17.4610
15	2	64QAM	5/6	38.5250	37.9968	0.0285	19.0199

Table 4.1: Numeric calibration results $L_M = 1664$.

averaged over a large set $\mathscr{H}$ of independent channel realizations and average packet error rate interval $\mathscr{P} = [PER_{\min}, PER_{\max}]$, that is,

$$\left(\psi_{1_{\text{opt}}}^{(m)}, \psi_{2_{\text{opt}}}^{(m)}\right) = \underset{\psi_1,\psi_2}{\arg\min} E_{\mathscr{H},\mathscr{P}} \left\{ \left| SNR_{eff}^{(m)} - SNR_{AWGN}^{(m)} \right|^2 \right\} \tag{4.5}$$

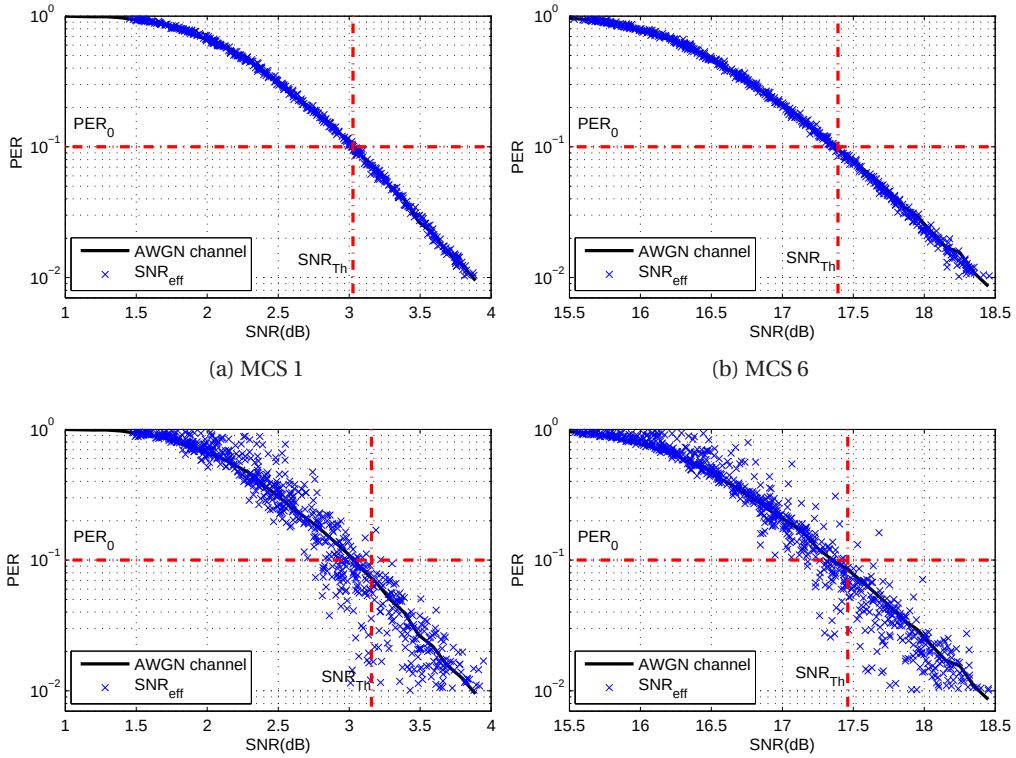
for all $m \in \mathscr{M}$, where $SNR_{AWGN}^{(m)}$ is the required SNR for mode m to obtain the same PER on the AWGN channel as on the current channel realization.

In order to demonstrate the performance results of the PER prediction method once calibrated, several figures for single and double stream MCSs are presented. As a measure of the predicted PER accuracy the MSE as given in (4.5) has been computed. For single stream MCSs (see Fig. 4.3a and Fig. 4.3b), a high PER prediction accuracy can be appreciated (see MSE column in Table 4.1 for numerical results). In contrast, for two-stream MCSs (see Fig. 4.3c and Fig. 4.3d), the prediction is not as tight as for single stream MCSs, however it will be shown later that these inaccuracies do not significantly affect the performance of the FLA system. It is noticeable that, as pointed out in [12] and as it can be appreciated in Table 4.1, for any $m \in \mathscr{M}$ it holds that optimum parameters, $\psi_1^{(m)}$ and $\psi_2^{(m)}$, are very similar (i.e. $\psi_1^{(m)} \sim \psi_2^{(m)} \ \forall \ m$).

In order to fulfill the optimization constraint on the average PER outage, an effective SNR threshold $SNR_{Th}^{(m)}$ is obtained for each $m \in \mathscr{M}$ such as

$$\Pr\left\{ PER_{AWGN}^{(m)}\left(SNR_{eff}^{(m)}\right) > PER_0 \right\} \leq P_{\text{out}}, \tag{4.6}$$

whenever $SNR_{eff}^{(m)} \geq SNR_{Th}^{(m)}$ and $PER_{AWGN}^{(m)}(\cdot)$ represents the PER for mode m over the AWGN channel. The $SNR_{Th}^{(m)}$ is computed as the $SNR_{eff}^{(m)}$ satisfying (4.6) with equality. This probability is found numerically using all realizations over the set $\mathscr{H}$. For this



(a) MCS 1

(b) MCS 6

(c) MCS 9

(d) MCS 14

Figure 4.3: Effective SNR mapping accuracy compared to the ideal AWGN channel reference curve for different MCS: a) MCS1, b) MCS6 (1 stream MCSs) c) MCS9 and d) MCS14 (2 streams MCSs). The system has been calibrated over a PER interval set to $\mathscr{P} = [0.01, 0.95]$ with 200 channel realizations from Channel models B and E using the channel generator tool described in [3]. The PER target has been set to $PER_0 = 0.1$, and the maximum outage probability has been set to $P_{out} = 0.05$

configuration, the $SNR_{Th}^{(m)}$ values presented are those $SNR_{eff}^{(m)}$ ensuring that 95% of the channel realizations used in the $\mathscr{H}$ set have a PER value lower than PER_0. The $SNR_{Th}^{(m)}$ values for several MCSs are represented using a vertical line in Figs. 4.3a, 4.3b, 4.3c and 4.3d, and in table 4.1 for all studied MCSs.

4.2.3 MCS selection process (on-line)

Figure 4.4 shows a flow diagram of the iterative process used to select the most suitable MCS for a given channel realization. The FLA algorithm evaluates the effective SNR for the different MCSs in descendent throughput order. During the evaluation of a given MCS, two situations may occur: if the effective SNR of the MCS is above the corresponding $SNR_{Th}^{(m)}$, the evaluated MCS is selected as a possible transmission mode; otherwise, the considered MCS is deemed unsuitable. This iterative procedure continues until either one MCS is selected or all MCSs have been discarded, in which case, the no transmission mode is selected. Some throughput values may be achieved using either SDM or STBC, in such cases, both MCSs are evaluated and if both are found

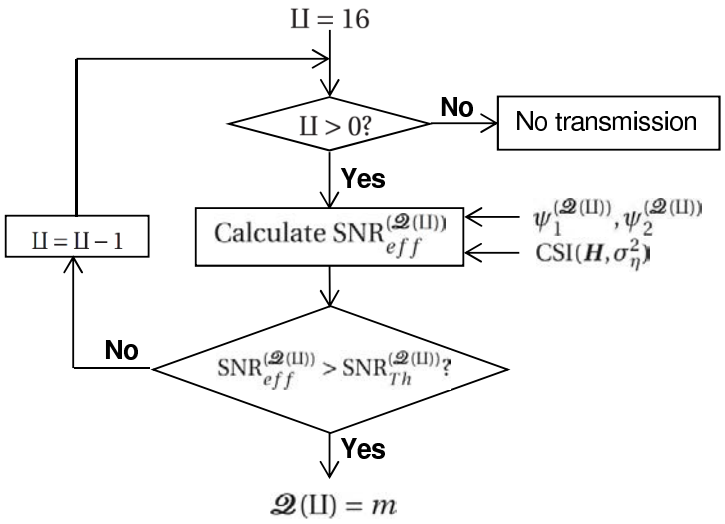


Figure 4.4: MCS Selection Process flow diagram, where q denotes the index position of the $\overline{\mathscr{M}}$ set. $\overline{\mathscr{M}}$ is the $\mathscr{M}$ set ordered in low to high throughput order. The $\mathscr{Q}(q)$ permutation operation returns the $m \in \mathscr{M}$ MCS from the q index.

suitable for transmission, the STBC-based one is selected due to its higher spectral efficiency at low SNR regimes [80].

4.3 BER-based PER prediction methodology

Previously, it has been stated that PER prediction methods depend on the particular packet length L_M, leading to an individual calibration/prediction procedure for each potential packet length. That is, a set of look-up tables would be required for each packet length. The current set of specifications for IEEE 802.11n does not specify a fixed packet length. Instead, the system determines dynamically the packet size in an attempt to efficiently trade system performance and data throughput depending on the operating scenario [86].

Consequently, in order for FLA PER-based algorithms to be usable in IEEE 802.11n, a tedious calibration procedure would need to be conducted taking into account every possible packet length (or a representative subset of them). In order to avoid this situation, a novel solution is presented here that relies on BER, rather than PER, thus making the FLA strategy independent of the packet length L_M. The rational behind this strategy is the assumption that for long-enough packets, the BER is independent of L_M and consequently, if BER and PER can be related by means of a closed-form expression, PER prediction can be made on the basis of BER.

The error event probability for a convolutionally encoded packet using MCS m can be approximated by

$$P_\epsilon^{(m)} \approx P_\varpi^{(m)} / d_f, \tag{4.7}$$

where d_f is the free distance of the convolutional code and $P_\varpi^{(m)}$ is the coded bit error probability of MCS m. This approximation is based on the assumption that, in the medium-to-high SNR regime, the number of bit errors per error event is approx-

imately equal to the free distance of the convolutional code. Note that for low SNRs, where this approximation could compromise the adaptation procedure, the no transmission mode is typically selected and thus it does not affect the performance of the adaptive scheme. Using $P_\epsilon^{(m)}$, and given a fixed channel realization, then

$$PER^{(m)} \approx 1 - \left(1 - P_\epsilon^{(m)}\right)^{\left(\frac{L_M}{R_m}\right)} \tag{4.8}$$

where R_m is the MCS code rate. This approximation is based on the assumption that an error-free packet is due to the absence of error events in each possible transition along the convolutional code trellis.

The BER-based estimation relies on the EESM technique introduced earlier although modifying some of its characteristics. Using (4.4), this new method determines the $SNR_{eff}^{(m)}$ for each channel realization with $\psi_1^{(m)}$ and $\psi_2^{(m)}$ obtained from a calibration phase, where PER curves have been replaced by BER curves. Correspondingly, an $SNR_{Th}^{(m)}$ is determined for each MCS $m \in \mathscr{M}$ in such a way that

$$\Pr\left\{BER_{AWGN}^{(m)}\left(SNR_{Th}^{(m)}\right) > \left(1 - (1 - PER_0)^{\frac{R_m}{L_M}}\right) d_f\right\} = P_{\text{out}}, \tag{4.9}$$

where $BER_{AWGN}^{(m)}(\cdot)$ represents the BER for mode m over the AWGN channel. The search algorithm is not modified.

4.4 Performance bounds algorithm (PBA)

The PBA is an ideal AMC algorithm that achieves the maximum throughput performance irrespective of the environment conditions. This algorithm has the ability to predict the MCS with the highest transmission rate that will successfully transmit the packet. To this end, the algorithm is provided by a perfect knowledge of the environment and receiver state, i.e., having all the information about the channel state and each of the noise components at the receiver. Note that such conditions are not possible in real systems, so this PBA algorithm only provides an ideal performance bound of the AMC performance. Making use of this information, the transmitter can simulate the packet transmission using each of the enabled MCSs and then select the MCS which achieves the successful packet transmission and offers the highest rate.

4.5 Results

The same system model of Chapter 3 has been used in this chapter. The system has been equipped with $N_T = N_R = 2$ and defined according to the specifications of IEEE 802.11n [5]. It has been configured to use full GI (0.8 μs) and $N_c = 64$ subcarriers over a 20 MHz bandwidth on the 5.25 GHz carrier frequency with $N_d = 52$ data subcarriers (remaining subcarriers are pilots or nulls). Frequency-time selective fading channel realizations, fully compliant with the IEEE Channel models [2], have been generated using the MIMO channel model generator tool described in [3]. Denoting by λ the operating wavelength, the system has been configured with a Tx antenna spacing of λ and receive antenna spacing of 0.5λ. In order to determine $\psi_1^{(m)}$ and $\psi_2^{(m)}$, the calibration set $\mathscr{H}$ has been defined by a mixture of 200 channel realizations from Channel models B and E over the PER interval $\mathscr{P} = [0.01, 0.95]$ using $PER_0 = 0.1$ and $P_{out} = 0.05$.

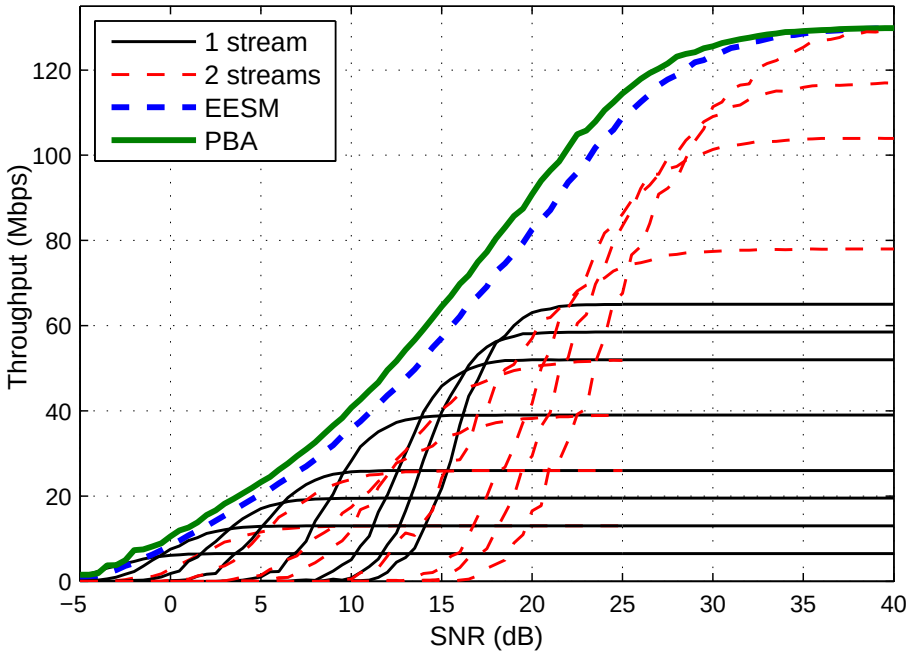


Figure 4.5: Throughput performance for Channel model B with perfect CSI.

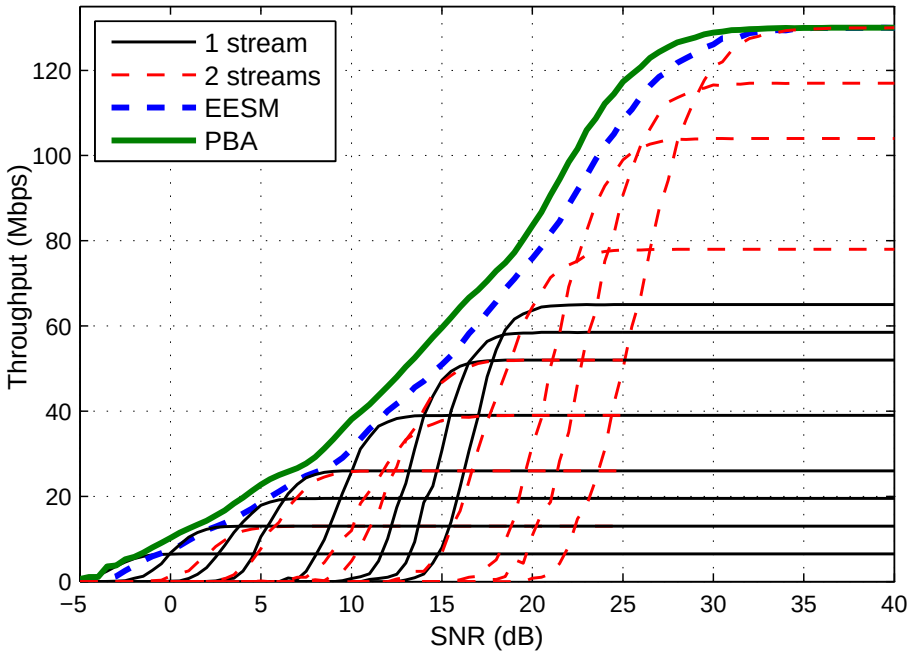


Figure 4.6: Throughput performance for Channel model E with perfect CSI.

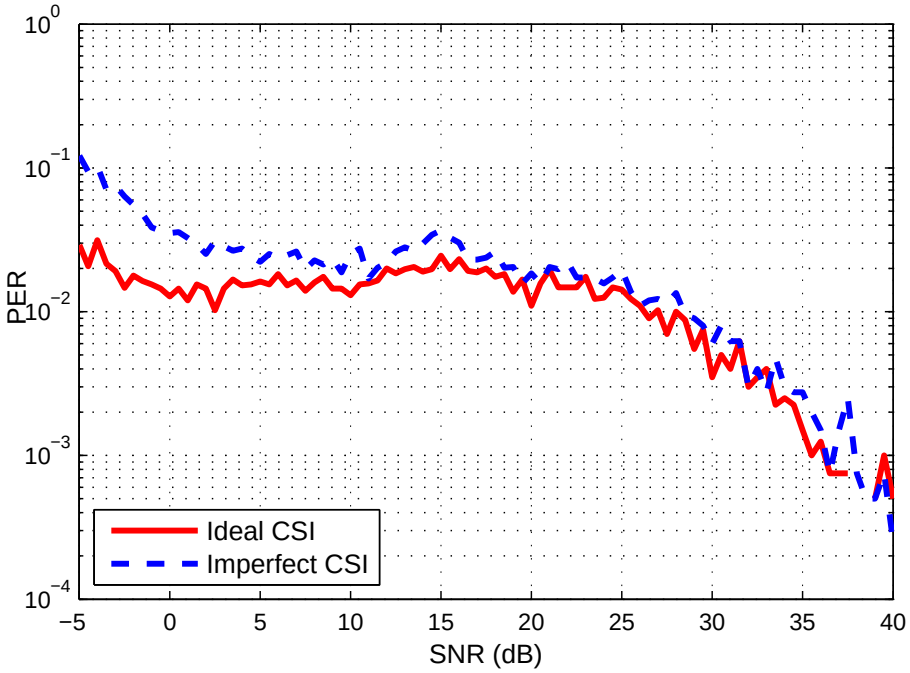


Figure 4.7: System PER performance for Channel model B.

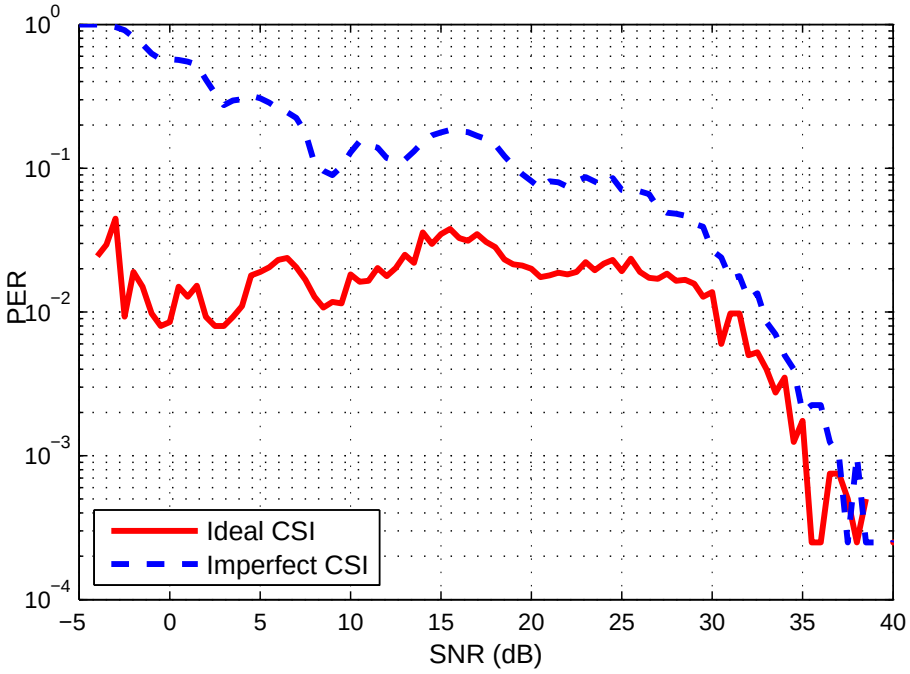


Figure 4.8: System PER performance for Channel model E.

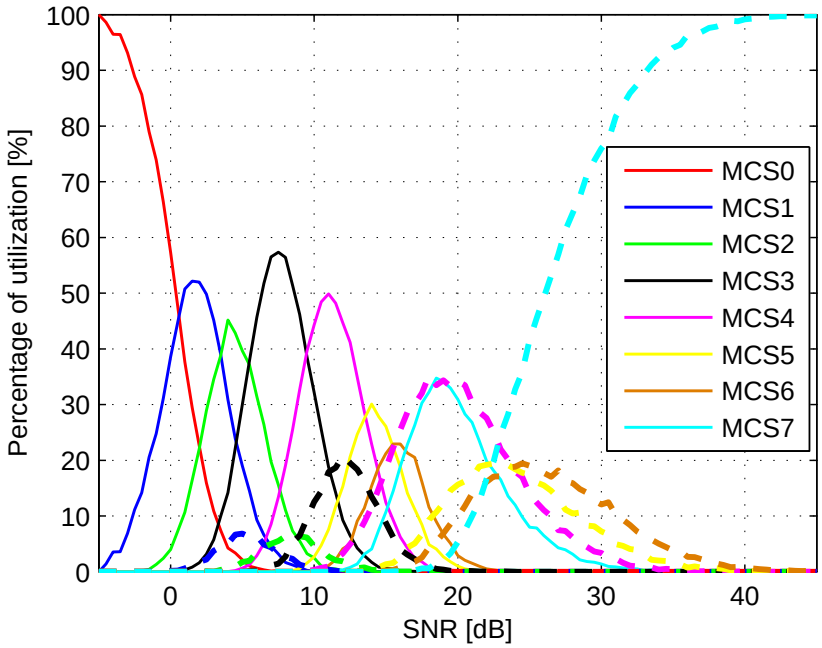


Figure 4.9: Percentage of utilization for one-stream MCSs (lines) and two-streams MCSs (dashes) in Channel model B.

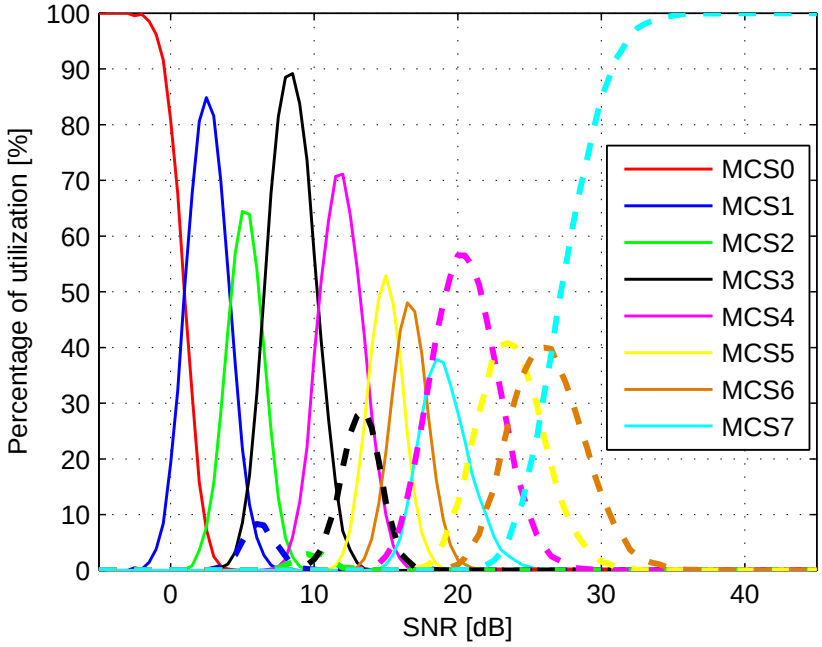


Figure 4.10: Percentage of utilization for one-stream MCSs (lines) and two-streams MCSs (dashes) in Channel model E.

Figures 4.5 and 4.6 show throughput results for Channel models B and E, respectively, obtained under the assumption of ideal CSI for both fixed and adaptive transmission strategies when using packets of length L_M = 1664 bits. For comparison purposes, the results obtained using the PBA are also shown. As previously stated, the PBA is an ideal FLA algorithm that for any channel realization is able to select, in a genie-aided fashion, the MCS with maximum throughput while ensuring zero transmission errors [16]. The EESM FLA clearly outperforms the fixed MCS and remains within 1.5 dB of the PBA. Furthermore, EESM fulfills the QoS constraints by keeping the PER well below PER_0 (see Figs. 4.7 and 4.8 where the system PER of EESM FLA from Figs. 4.5 and 4.6 are depicted, respectively). The large difference in actual PER and PER_0, resulting in an overly pessimistic system, is basically due to the absence of power control.

To complete the FLA performance study, the usage of each MCS per SNR value of Channel models B and E is presented in Figs. 4.9 and 4.10, respectively. In both channel models, one and two stream MCSs are mainly employed by FLA at low and high SNR values, respectively, because the one stream MCSs offer robustness against channel errors and the two stream MCSs provide higher transmission rates when having proper channel conditions. Therefore, these MCS selections clearly respond to the main objective of FLA, by selecting the best MCS for the current channel conditions, in order to maximize the throughput while satisfying the target PER. Furthermore, some differences on the MCS utilization can be noticed in Channel models B and E. Note that in Channel model B when using medium SNR values (i.e. from 0 to 25dBs), the quantity of employed MCS per SNR value is notably higher than that of Channel model E, presenting in some cases a usage of up to six MCSs per SNR value. These divergencies in the MCS utilization are mainly produced by the different diversity gain experienced on each channel. As it was stated in Sect. 3.4, the PER performance slopes in Channel model B are less steep than those in Channel model E and then several modes are available with similar PER performance for the same SNR value, thus making possible a higher number of MCSs candidates per SNR value.

Figs. 4.11 and 4.12 present the throughput performance of EESM FLA using imperfect CSI for channels models B and E, respectively. It should be stated that the imperfect channel performance is obtained by following the model described in Sect. 3.4.1. As it can be observed in Fig. 4.12, when transmitting over Channel model E, imperfect CSI produces a non negligible throughput degradation. Additionally, notice that for low SNRs, the system PER constraint is not fulfilled (see Fig. 4.8). The reason for this misbehavior is the large temporal dispersion of Channel model E, which causes a large MSE in the channel estimation process. The fulfillment of the QoS constraints could be achieved by selecting more conservative SNR threshold values at the expense of an additional decrease in throughput performance. Note that for Channel model B, and owing to its small temporal dispersion, imperfect CSI barely affects performance (see Figs. 4.5, 4.7 and 4.11).

Figures 4.13, 4.14, 4.15 and 4.16 show results obtained when using BER-based PER prediction methods with L_M as parameter. It has been experimentally found that, in the considered scenarios, the BER performance is barely sensitive to packet length values over several hundred bits. Therefore, BER curves with L_M = 832 bits can serve as a reference to provide calibration parameters for any longer packet length. As an illustrative example, results for L_M = 832 bits have been extrapolated to predict the

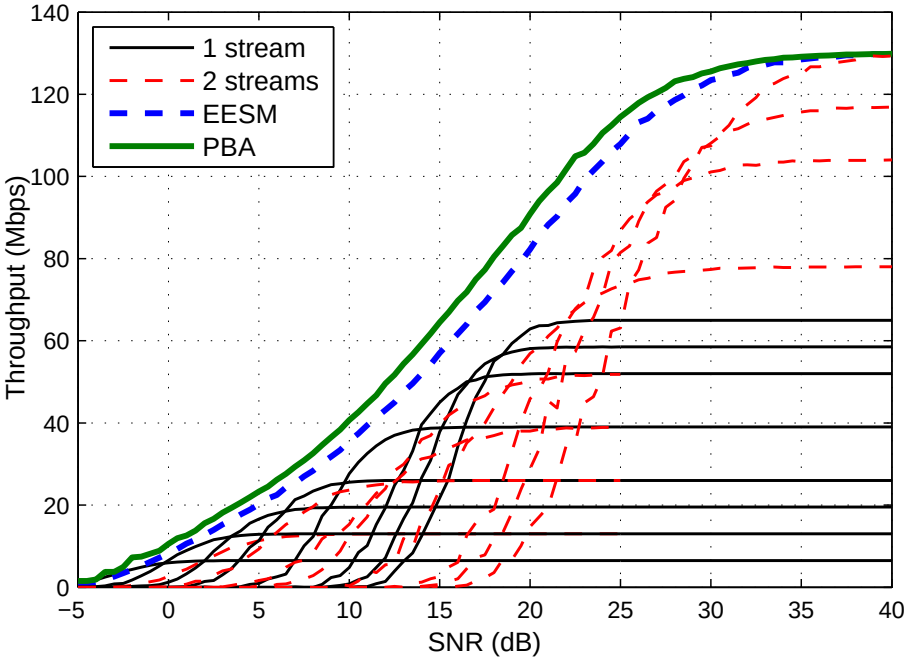


Figure 4.11: Throughput performance for Channel model B with imperfect CSI.

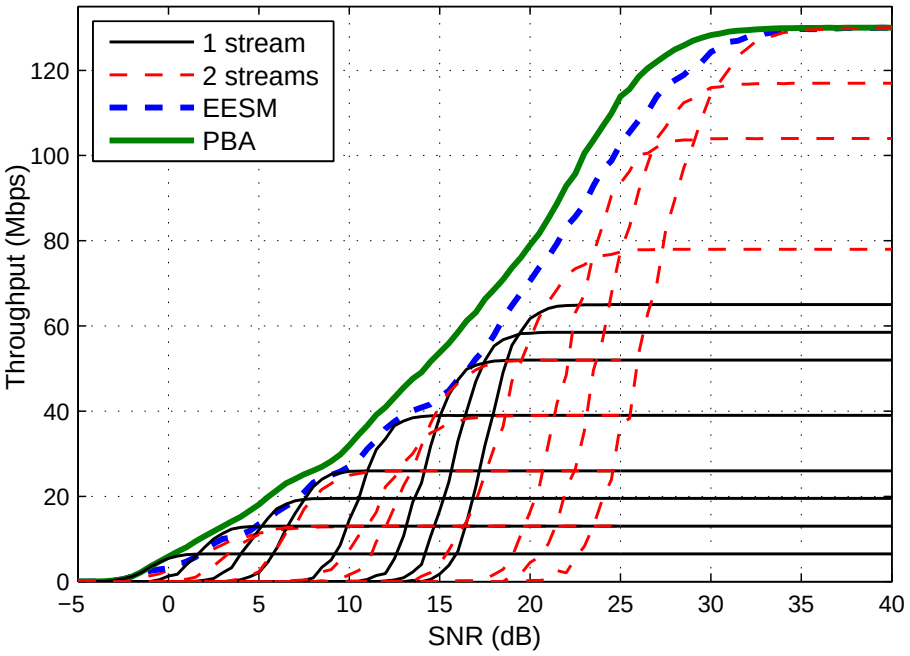


Figure 4.12: Throughput performance for Channel model E with imperfect CSI.

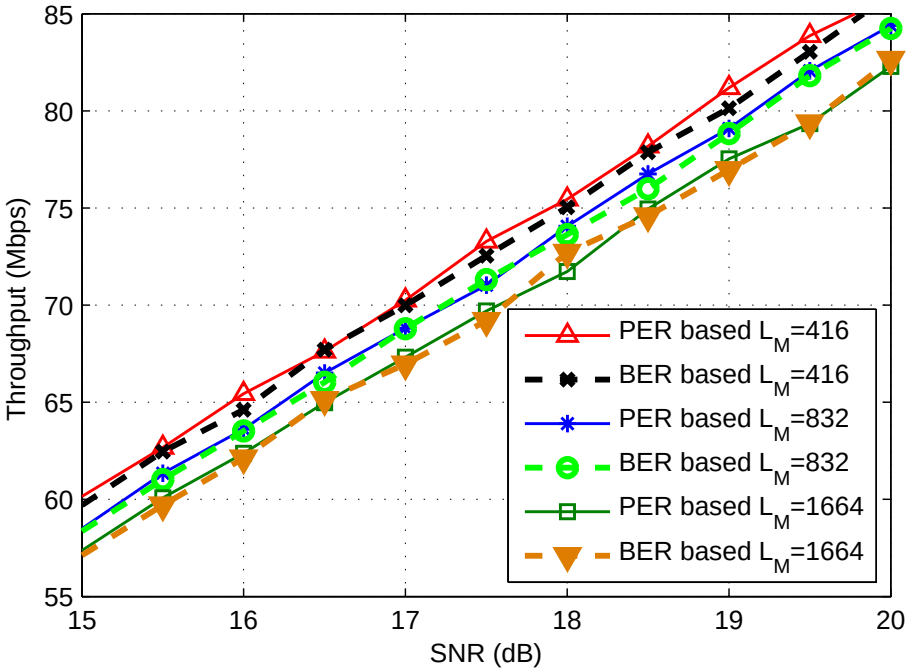


Figure 4.13: Channel model B ideal estimation throughput using BER estimation.

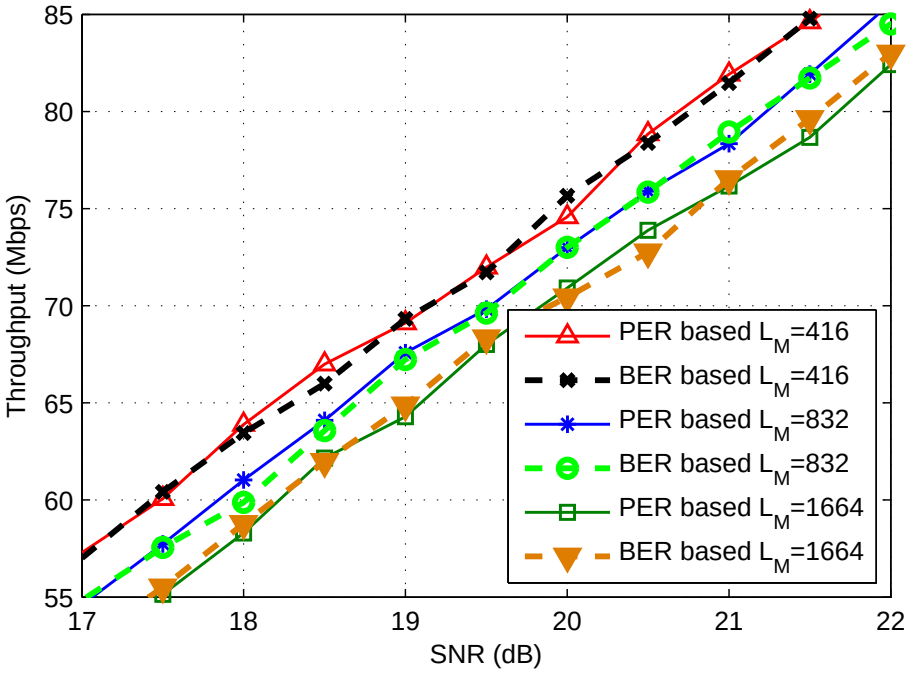


Figure 4.14: Channel model E imperfect estimation using BER estimation.

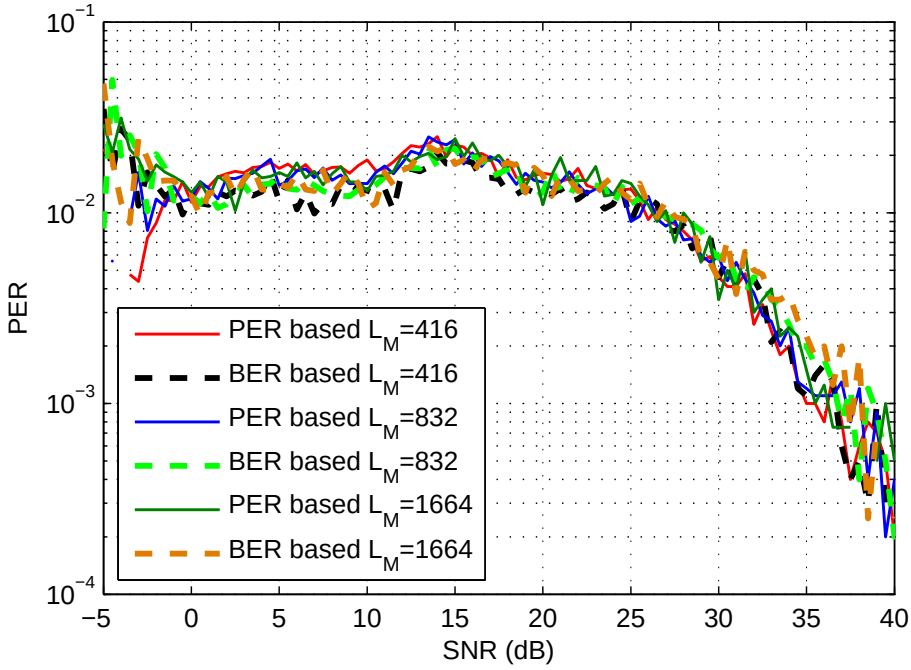


Figure 4.15: PER performance in Channel model B using the BER estimation with ideal CSI.

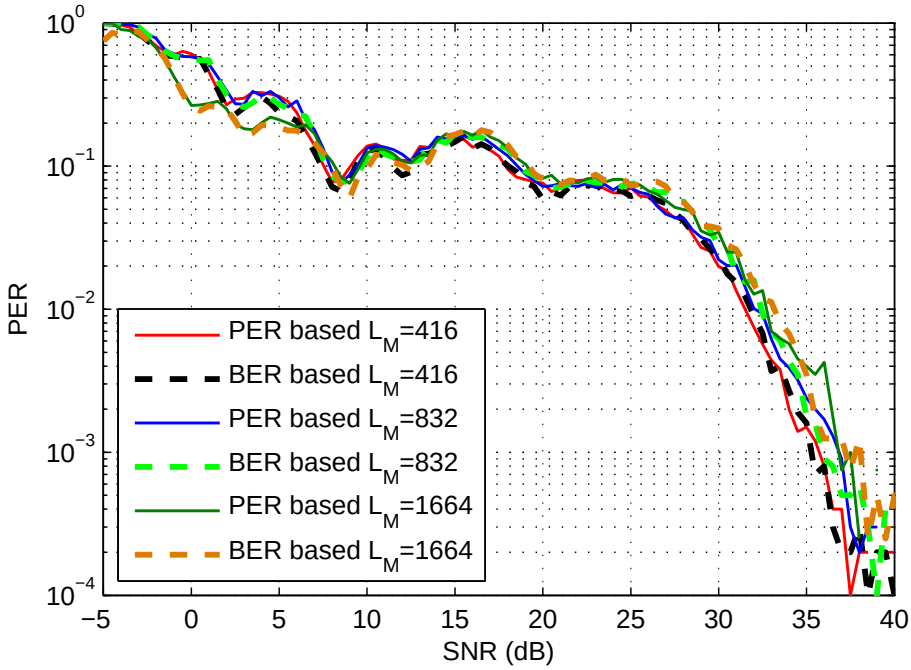


Figure 4.16: PER performance in Channel model E using the BER estimation with imperfect CSI.

PER performance for L_M = 416 and L_M = 1664 bits. As it can be observed, this approach leads to almost identical performance results (Fig. 4.13 and Fig. 4.14) as the PER-based approach, thus validating the BER-prediction accuracy for practical values of L_M. There are no significant differences between PER-based FLA and BER-based FLA for ideal channel estimation (see Fig. 4.13 and 4.15) and imperfect channel estimation (Fig. 4.14 and 4.16). Notice that, as shown in Figs. 4.15 and 4.16, even for short packet lengths (L_M = 416 bits), the degradation due to the use of the BER-based approach is negligible.

4.6 Chapter Summary

This chapter has addressed cross-layer FLA techniques within the framework of IEEE 802.11n networks. Single and double stream MCSs using STBC and SDM, respectively, have been considered. It has been shown that PER prediction EESM-based FLA with ideal channel estimation meets the prescribed quality of service constraints while performing very close ($\sim$ 1.5 dB) to the PBA in terms of throughput. When channel estimation errors are considered in highly frequency-selective channels (i.e., Channel model E), the throughput performance of the proposed FLA algorithms is significantly affected and therefore, it is important to take these effects into consideration during the system design phase. A new variant of FLA has been introduced that is based on BER prediction (rather than PER) using EESM. This technique performs almost identically to its PER-based counterpart while, due to its independence from the packet length, it simplifies the costly calibration/prediction procedure.

Part II

Link Adaptation Performance on the MAC sublayer

CHAPTER 5

MAC SUBLAYER PERFORMANCE UNDER SATURATED CONDITIONS

This chapter analyses the system performance of the IEEE 802.11 MAC sublayer when using the DCF mechanism under saturated conditions. To this end, the MAC sublayer is first described in Sect. 5.1, focusing on the operation of a single station (STA) and providing the expected time spent inside each of its possible states. A short literature review of the most important DCF modeling approaches and the detailed model slot of IEEE 802.11 standard are presented in Sect. 5.2. Subsequently, the latest modelling approach is then fully described and detailed in Sect. 5.3. In Sect. 5.4, the analytic performance results are validated by comparing them to a custom simulation tool fully compliant with the standard specifications and developed within the context of this thesis. Finally, the chapter summary is presented in Sect. 5.5.

5.1 MAC sublayer description

The IEEE 802.11 standard specifies three different MAC mechanisms for WLANs, namely, the DCF, the PCF and the hybrid coordination function (HCF). The DCF is the mandatory MAC mechanism for the IEEE 802.11 standard [5]. It is a random access scheme based on the CSMA/CA protocol that incorporates a BEB algorithm to manage the retransmission of collided and erroneous packets.

Apart from the DCF, the standard also defines the PCF as an optional MAC mechanism, which is only usable on infrastructure network configurations. It is a centralized MAC protocol where a point coordinator (PC), usually the access point (AP), indicates which STA has currently the right to access the medium. This function achieves the collision free operation using a carrier sense mechanism aided by an access priority technique. Its performance evaluation will not be addressed in this thesis since it is not commonly supported by current devices. Obviously, PCF can be considered as an interesting research topic for future studies.

Lastly, the HCF mechanism, incorporated into the standard in the IEEE 802.11e amendment, improves DCF and PCF mechanisms with enhanced QoS support, by defining the HCF controlled channel access (HCCA) and the enhanced distributed channel access (EDCA) for PCF and DCF, respectively. Both HCCA and EDCA mechanisms classify the traffic under different categories, thus providing different access priorities to the medium. The EDCA protocol is regarded as an enhanced version of DCF and, although it is out of the scope of this thesis, the application of the AMC techniques developed in this thesis in connection with EDCA, constitutes an interesting thread for future research.

5.1.1 Distributed coordination function (DCF)

An STA with a pending data-frame, i.e., an MPDU, to transmit shall sense the medium to determine whether it is busy or idle. If the medium is sensed busy, the STA has to defer transmission until the medium is sensed idle without interruption for a prescribed time lapse. This time interval is equal to either the distributed frame interval space (DIFS), after detecting a correctly received frame, or equal to the extended interframe space (EIFS), after detecting an incorrectly received frame. By doing this, it avoids the foreseeable collision with the on-going transmission. Following this DIFS or EIFS medium idle time, the STA shall generate a random backoff interval for an additional deferral period before transmitting, unless the backoff timer already contains a nonzero value. The utilization of this random backoff instead of directly transmitting the packet reduces the collision probability between those STAs that started their contention phase during the current busy period, since the random backoff selection of each STA will probably schedule their transmission on different slots. The backoff interval is slotted and expressed as an integer number of backoff slots, which is known as the backoff counter. Such a number is decremented as long as the medium is sensed idle during a backoff slot, *frozen* as soon as the channel is sensed busy in a backoff slot, and reactivated when the medium is sensed idle again for more than a DIFS or EIFS lapse. Note that during the backoff interval, other contending STAs can access the medium and transmit their packets, since each STA pauses its counter and avoids the collision with the transmitting STA. When the backoff counter reaches zero the STA is allowed to transmit. In order to avoid channel capture by a station[1], a random backoff interval is also generated between two consecutive new packet transmissions from the same STA, even if the medium is sensed idle during the whole DIFS time.

For each MPDU transmission, the backoff counter is set to a pseudo-random integer drawn from a uniform distribution over the interval $[0, CW-1]$, where CW denotes the contention window size for the current backoff stage. The first time an STA tries to transmit a given MPDU, CW must be initialized to $W_0 = CW_{min} + 1$, where CW_{min} denotes the minimum contention window size. After each unsuccessful transmission it is updated to $CW = \min\{W_i, CW_{\max} + 1\}$, where $W_i = 2^i W_0$ denotes the backoff window size for the i^{th} packet retransmission and $CW_{\max}$ defines the maximum contention window size. The value of CW shall be reset to W_0 after successful transmission, or when the retransmission counter reaches a predefined retry limit R causing

[1] One station occupies the medium continuously without allowing the transmissions of the other contending STAs

the MPDU to be dropped. The backoff adjustment procedure described above is generally referred to as the BEB algorithm.

After each MPDU transmission the STA activates an ACK timeout, which has to be set to a time lapse equal to a short interval frame space (SIFS) plus the time it takes to transmit an ACK control frame at the lowest mandatory PHY rate. The successful reception of an MPDU requires the receiving STA to respond with an ACK control frame (i.e., positive acknowledgement). This STA must wait for a SIFS and then respond with an ACK frame using the most appropriate transmission mode, as decided by the adaptive modulation and coding strategy implemented at the receiver side. As the SIFS is shorter than the DIFS, no other active STA will be able to detect the channel to be idle for a DIFS before the end of the ACK. Lack of reception of an expected ACK frame during the ACK timeout indicates to the STA initiating the frame exchange that an error has occurred and that it should schedule a retransmission.

The two-way handshaking technique just described is called the *Basic Access mechanism* and is the most extensively used [18] in current deployments. However, the IEEE 802.11 DCF describes another four-way handshaking mechanism known as request-to-send/clear-to-send (RTS/CTS). The RTS/CTS access method relies on a virtual carrier sensing mechanism achieved by distributing reservation information announcing the impending use of the medium. Before transmitting an MPDU, a short request-to-send (RTS) frame is transmitted. If the receiving STA is able to correctly decode this control frame (i.e., idle channel), it responds with a short clear-to-send (CTS) frame, allowing the STA that initiated the handshaking procedure to transmit the data-frame. Finally, if the MPDU transmission succeeds, an ACK control frame will be sent by the receiving STA. All four frames must be separated by a SIFS. In case of detecting an erroneous frame, the transmitting STA shall defer transmission until the medium is sensed idle without interruption for a time lapse equal to EIFS.

Note that collisions can only occur during the RTS/CTS frame exchange duration, thus drastically reducing the collision time caused by long packets in the Basic Access technique. However, RTS/CTS induces a considerable increase of the system overhead, lowering in this way the system performance when short packet lengths or high transmission rates are used. Nevertheless, thanks to its reduced collision duration, it outperforms Basic Access in dense user scenarios where the collision probability is high [18, 13]. Furthermore, the use of RTS/CTS frame exchange avoids the hidden terminal problem[2].

In the standard document [5], the Basic Access is defined as the mandatory access scheme employed in the DCF mechanism. In contrast, the implementation of RTS/CTS is optional, yet it is an advisable feature whenever the system operates with packets whose length exceeds a predefined threshold (S_{thres}). Typically S_{thres} is defined by a value comprised in the $0-2347$ bytes interval [18].

[2] A pair of stations in the range of the AP are hidden from each other because the transmission from one of them cannot be heard by the other.

5.1.2 Timing of DCF events

Basic Access technique

According to the Basic Access scheme of DCF, the elapsed time for a successful transmission of an L_M-bit MPDU using MCS m is

$$T_{T_s}^{bas}(m, L_M) = T_{\text{MPDU}}(m, L_M) + \text{SIFS} + \text{ACK}(m) + \text{DIFS}, \tag{5.1}$$

where SIFS, DIFS and σ are 802.11n time constants defined in [5]. The time elapsed in the MPDU transmission, $T_{\text{MPDU}}(m, L_M)$, is defined as

$$T_{\text{MPDU}}(m, L_M) = t_{\text{Preamble}} + N_{\text{Sym}}(m, L_M)\ t_{\text{Sym}}, \tag{5.2}$$

with t_{Preamble} representing the PHY layer convergence procedure (PLCP) preamble duration, t_{Sym} denoting the OFDM symbol period and

$$N_{\text{Sym}}(m, L_M) = m_{\text{STBC}} \left\lceil \frac{L_M + 22}{m_{\text{STBC}} N_{\text{DBPS}}(m)} \right\rceil, \tag{5.3}$$

being the number of OFDM symbols required for the transmission of a complete MPDU, where $N_{\text{DBPS}}(m)$ is the number of bits forming each OFDM symbol as defined by MCS m and $m_{\text{STBC}} = 2$ if STBC is used and $m_{\text{STBC}} = 1$ otherwise. Similarly, the time required for the transmission of an ACK frame using PHY mode m is given by

$$T_{\text{ACK}}(m) = t_{\text{Preamble}} + N_{\text{Sym}}\left(m, 14\text{bytes} \times \frac{8\text{bits}}{\text{byte}}\right) t_{\text{Sym}}. \tag{5.4}$$

A collision occurs whenever two or more STAs transmit on the same slot, finishing an EIFS period after the end of the longest transmission of the collided STAs. That is, its duration depends on the MCS and MPDU length corresponding to the longest transmission, denoted by m^* and L_M^*, respectively. Mathematically, the collision duration can be expressed as

$$T_{T_c}^{bas}(m^*, L_M^*) = T_{\text{MPDU}}(m^*, L_M^*) + \text{EIFS}, \tag{5.5}$$

where

$$\text{EIFS} = \text{SIFS} + T_{\text{ACK}}(\text{m=0}) + \text{DIFS}. \tag{5.6}$$

Finally, the MPDU error transmission duration, defined as $T_{T_e}(m, L_M)$, is the time elapsed in a transmission that experiences channel errors, and it can be expressed as

$$T_{T_e}^{bas}(m, L_M) = T_{\text{MPDU}}(m, L_M) + \text{EIFS}. \tag{5.7}$$

RTS/CTS access technique

Unlike Basic Access, the RTS/CTS mechanism incorporates a frame exchange prior to data transmission. One of the consequences of this handshake procedure is that when a collision takes place, its duration is minimized. This is because collisions can only occur during the RTS/CTS exchange, which only involves very short frames. In this work, a short retry limit (R=4) has been applied to both the RTS frames and data packets.

The RTS/CTS time elapsed for a successful transmission of an L_M-bit MPDU using MCS m is

$$T_{T_s}^{rts}(m, L_M) = T_{\text{RTS}}(m) + T_{\text{CTS}}(m) + T_{\text{MPDU}}(m, L_M) + T_{\text{ACK}}(m) + 3\,\text{SIFS} + \text{DIFS}, \quad (5.8)$$

where

$$T_{\text{RTS}}(m) = t_{\text{Preamble}} + N_{\text{Sym}}\left(m, 14\text{bytes} \times \frac{8\text{bits}}{\text{byte}}\right) t_{\text{Sym}}, \quad (5.9)$$

and

$$T_{\text{CTS}}(m) = t_{\text{Preamble}} + N_{\text{Sym}}\left(m, 20\text{bytes} \times \frac{8\text{bits}}{\text{byte}}\right) t_{\text{Sym}}, \quad (5.10)$$

are the time durations of RTS and CTS control frames, respectively.

Similarly, the RTS/CTS elapsed time in a transmission error of an L_M-bit MPDU using MCS m is

$$T_{T_e}^{rts}(m, L_M) = T_{\text{RTS}}(m) + T_{\text{CTS}}(m) + T_{\text{MPDU}}(m, L_M) + 2\,\text{SIFS} + \text{EIFS}. \quad (5.11)$$

Finally, the collision duration can be defined as

$$T_{T_c}^{rts}(m') = T_{\text{RTS}}(m') + \text{EIFS}, \quad (5.12)$$

where m' is the lowest MCS value that causes the longest RTS transmission duration from the collided users.

In this model, due to its negligible probability of occurrence, we have not considered the possibility of an error in the ACK, RTS and/or CTS transmissions. The ACK transmission takes place under the same system conditions than the packet being acknowledged (i.e., using the same MCS and suffering similar channel conditions) however, its packet size is considerably smaller than that of the information packets and therefore, its error probability can be safely considered insignificant. Similarly, packet sizes for RTS and CTS transmissions are also small and the use of the most reliable MCS warrants a negligible error probability.

5.2 Review of analytical models

A simple but accurate modelling approach for evaluating the goodput performance of the IEEE 802.11 DCF was proposed by Bianchi in [20] and further detailed in [13]. Since then, a number of papers have enhanced this basic model by either adapting it to different backoff freezing assumptions [21], [22], or accounting for supplementary modeling details such as finite retransmission attempts [87], k-ary exponential backoff and multiple traffic classes [88], error-prone channel conditions [89], channel capture [90], and hidden terminals [91]. In addition to the goodput analysis, a derivation of the average delay performance is described in [21], whereas a derivation for the case of finite retry limits is given in [92].

The model proposed by Bianchi in [13], as well as many of its extensions, implicitly assume that the backoff counter is deterministically decremented at the beginning of a slot time. As it is recognized in several papers [21], [22], [29], this assumption does not conform to the original DCF specification [6]. To face this issue, Tinnirello et. al in [23] proposed a new backoff model that retains the simplicity of traditional DCF models

while offering a better accuracy when compared to other similar studies. This model considers the presence of anomalous slots in which the channel access probability is much lower than the average probability, and, moreover, it also considers a finite number of retransmissions for each packet, discarding those packets whose number of retransmissions exceeds the maximum allowed value.

Recently, Tinnirello's model has been extended by Chen in [28] by considering non-ideal channel conditions. This modelling approach accounts for the errors produced inside the anomalous slots, hence offering precise analytic results. This model is thoroughly discussed in this chapter and its performance compared to that of previous channel modelling approaches.

5.2.1 Some notes about the backoff counter and the anomalous slot definition

From [23] and references therein, it can be deducted that the backoff counter compliant with the standard document [6] is only decreased at the end of each idle slot, i.e., in slots free of transmissions. This end-of-slot backoff decrement rule has some implications on the DCF modelling framework that are not obvious. The first implication is that the slot immediately following a successful transmission cannot be used for transmissions by any other STA, except the transmitting one. Hence, the transmission placed inside the slot just after a successful transmission is free of collisions, since only one STA can access this slot. This special slot is referred to as the anomalous slot. The second implication is that the slot just after a packet error or collision is always empty, because no STA can proceed with transmission.

To understand these special probabilities for the anomalous slot, let us consider an scenario with multiple contending STAs where a target STA has just ended a successful packet transmission. The target STA will transmit on the next slot only if the selected backoff value from the BEB algorithm is zero, an event that has probability $1/W_0$. In this case, the target STA immediately transmits the packet at the beginning of the slot, thus avoiding any idle slots between its two consecutive transmissions. Therefore, the other STAs cannot decrease their counters and remain in the same slot contending to access the medium, thus avoiding the collision with the target STA. Conversely, if the backoff value is greater than zero, then the first slot after the transmission is idle, because all STAs have nonzero backoff values. As a result, the backoff counter of each STA is decreased at the end of this slot.

The behaviour of an STA in anomalous slot can be observed in the backoff operation example of Fig. 5.1, where two STAs (A and B) are contending for the medium. In this figure, note that two consecutive transmissions of STA A take place within one slot of STA B.

Consider now the collision case. We assume that all STAs in the scenario are able to detect a collision when this happens and then resume the backoff process after an EIFS time from the previous channel activity. In this case, neither the transmitting STAs nor other STAs can use the first slot after an EIFS time from the previous transmission. In fact, the EIFS duration is equal to $ACK_{timeout} + DIFS - \sigma = T_{ACK}(m = 0) + DIFS$. Thus, after an EIFS from a previous collision, any STA involved in the collision and waiting for the ACK timeout has to hold on for a further backoff slot before extracting a new backoff value. Moreover, STAs sensing the collision resume the back-

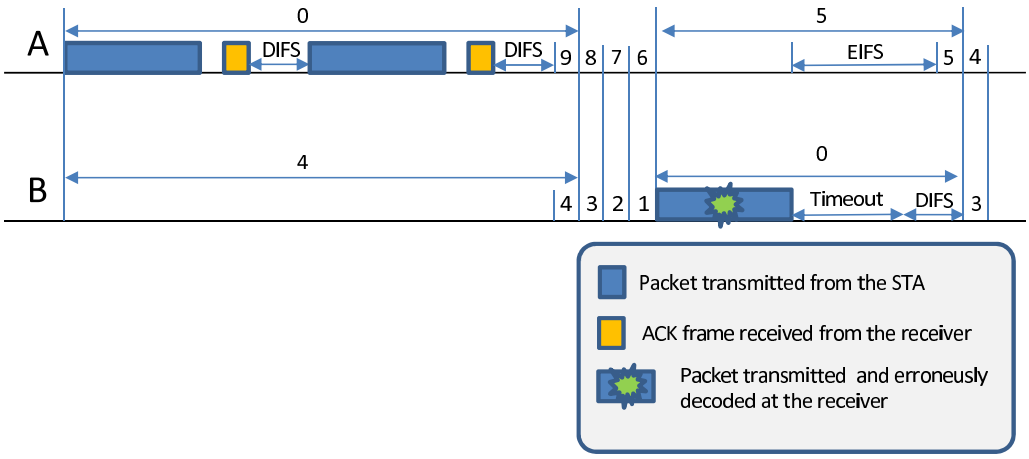


Figure 5.1: Example of backoff counter operation for two contending STAs.

off counter to the frozen value, which is different from zero, because they were not transmitting, and consequently, they cannot use the first slot after an EIFS time either. We conclude that the extra slot after the end of an EIFS will not be used by any STA (neither those involved in a collision nor the other STAs monitoring the channel). This case is again represented in the backoff counter example of Fig. 5.1. In this illustration, STA B experiences a collision and sets its backoff counter after $ACK_{timeout} + DIFS$ time; meanwhile, STA A, as a contending STA, decreases one unit its backoff counter. Therefore, the slot unit after a collision is always empty, since no STA can transmit on this slot.

In short, as stated by Tinnirello [23], it is concluded that the slots right after a channel busy period are anomalous slots since their probability of transmitting and colliding is different from that in other slots.

5.2.2 The slot model

Note that the anomalous slot and its error modelling have some effects on the model slot presented by Bianchi. In Chen [28], the new model slot time is defined as the time interval between two consecutive backoff counter decrements of a non-transmitting STA, as presented in Fig. 5.1. During this time, one of the following five distinct events might occur:

1. All the STAs are experiencing a backoff decrement. This is also known as an idle slot.

2. There are multiple successful transmissions.

3. One or more successful transmissions are followed by a failed transmission (MAC packet error or acknowledgement packet error).

4. One collision occurs.

5. A failed transmission happens.

In both the second and third cases, more than one transmission occurs due to the occurrence of anomalous slots. Any non-transmitting STA shall not decrease its backoff counter when multiple packets are transmitted without any separate idle backoff slots. Thus, the model slot time remains the same. This is exemplified in STA B of Fig. 5.1 that keeps the backoff counter to 4 during two successful transmissions of STA A within the anomalous slot. Multiple transmissions will be terminated by either a nonzero backoff counter or a failed transmission.

In Fig. 5.1, it can also be observed the system operation when a packet is erroneously detected. During this slot, the sender updates the backoff counter, while the other STAs decrease the backoff counter at the end of the slot. In contrast to the successful transmission, the backoff counter is updated in all STAs.

For each of the previous events, the slot duration considering the use of MCS m is then defined as follows:

1. σ corresponds to an idle slot duration. No STAs are transmitting during this period.
2. $j_s T_{T_s}(m) + \sigma$ is the anomalous slot duration having j_s consecutive successful transmissions when using MCS m.
3. $j_s T_{T_s}(m) + T_{T_e}(m) + \sigma$ denotes the time lapsed in an anomalous slot that ends with an erroneous transmission after j_s successful transmissions.
4. $\hat{T}_{T_c}(m) = T_{T_c}(m) + \sigma$ is the duration of a colliding slot.
5. $\hat{T}_{T_e}(m) = T_{T_e}(m) + \sigma$ is the slot duration of an erroneous transmission.

5.3 System analysis for saturated conditions

Before starting the system analysis, some initial considerations should be taken into account. In this model, the number of STAs (or users) per scenario is fixed, consequently, the complex modelling approaches that determine the number of current users per scenario can be avoided. Furthermore, it is supposed that each STA has always a packet ready for its transmission (saturated condition). Additionally, the scenario conditions are idealized by avoiding the hidden terminal problem or channel capture by any station.

The analytic model introduced in this section evaluates the MAC sublayer performance under saturated conditions. The main performance measure is the saturation throughput, that is, the quantity of information transmitted at the MAC sublayer per unit of time. This is a fundamental performance metric defined as the limit reached by the system throughput as the offered load increases and represents the maximum load that the system can carry under stable conditions.

To undertake the analytical evaluation of the saturation throughput, the analysis is divided into two distinct parts. First, the behavior of a single station is tackled employing a Markov chain, hence obtaining the transmission probability of each station per slot unit, referred to as τ. Note that the model slot used follows the slot definition presented in the previous subsection from [28]. Subsequently, by examining the possible events that can occur within each slot, the system saturation goodput can be obtained as a function of the number of users, τ and the time duration of each event.

5.3.1 The Markov model

Given an STA, a bidimensional Markov chain whose states are denoted by $(s(t), b(t))$ where $b(t)$ and $s(t)$ represent the value of the backoff counter and the backoff stage at model slot time t, respectively. This Markov chain is shown in Fig. 5.2. Based on the backoff procedure previously described, let us define the contention window size for the i^{th} retransmission as

$$W_i = \begin{cases} 2^i W_0 = 2^i (CW_{\min} + 1) & \text{for } i \in [0, a-1], R > a \\ 2^a W_0 = 2^a (CW_{\min} + 1) & \text{for } i \in [a, R], R > a \\ 2^i W_0 = 2^i (CW_{\min} + 1) & \text{for } i \in [0, R], R \leq a \end{cases} \tag{5.13}$$

where $a = log_2 (CW_{\max} + 1) / (CW_{\min} + 1)$. In Fig. 5.2, p stands for the collision and/or error probability that a transmitted packet experiences and p_e is the probability that the MPDU suffers only errors from channel transmissions, hence excluding errors caused by collisions. It is assumed that p is constant and independent of the number of collisions or errors a packet has suffered in the past. Let τ be the probability that an STA transmits in a randomly chosen model slot and p_c be the corresponding packet collision probability. Then, it follows that in a symmetric network with n users

$$\begin{cases} p_c = 1 - (1-\tau)^{n-1} \\ p = 1 - (1-p_c)(1-p_e) \end{cases} \tag{5.14}$$

As inferred from Fig. 5.2, the zero backoff stage can only be accessed after a successful transmission or after dropping a packet. Note that each of these events trigger different counter intervals although they belong to the same backoff stage. In particular, the backoff counter interval is set to $[0, W_0 - 2]$ after a successful transmission and $[0, W_0 - 1]$ when the packet was discarded. In order to differentiate how the zero backoff stage is accessed, the 0^+ and 0^- notations are used. The 0^+ and 0^- stages define the 0-stage triggered by a successful transmission and a packet drop, respectively.

To simplify notation let us define $\Omega = p_e/(W_0 + p_e - 1)$ and $P\{i,k|i,k+1\}$ as the transition probability to reach the (i,k) state from $(i,k+1)$ state. In the Markov chain, the non-zero one-step transition probabilities are:

$$\begin{cases} P\{i,k|i,k+1\} = 1, & k \in [0, W_i - 2], i \in [1, R] \\ P\{0^-,k|0^-,k+1\} = 1, & k \in [0, W_0 - 2] \\ P\{0^+,k|0^+,k+1\} = 1, & k \in [0, W_0 - 3] \\ P\{0^-,k|R,0\} = p/W_0, & k \in [0, W_0 - 1] \\ P\{i,k|i-1,0\} = p/W_i, & k \in [0, W_i - 1], i \in [1, R] \\ P\{0^+,k|i,0\} = (1-p)\Omega/pe, & k \in [0, W_0 - 2], \forall i \\ P\{1,k|i,0\} = (1-p)\Omega/W_1, & k \in [0, W_1 - 1], \forall i \end{cases} \tag{5.15}$$

In (5.15), the first three sub-equations can be derived because the backoff counter must be decreased by one in the next model slot. Either a collision or a failed transmission can result in an increase of backoff stage (with probability p). The fourth and fifth sub-equations are easily obtained since they are the probabilities of reaching a given backoff state, triggered by a packet drop or an erroneous transmission, respectively.

After a successful transmission in one model slot, there may be multiple consecutive transmissions in the same model slot. Within the multiple transmissions, if the

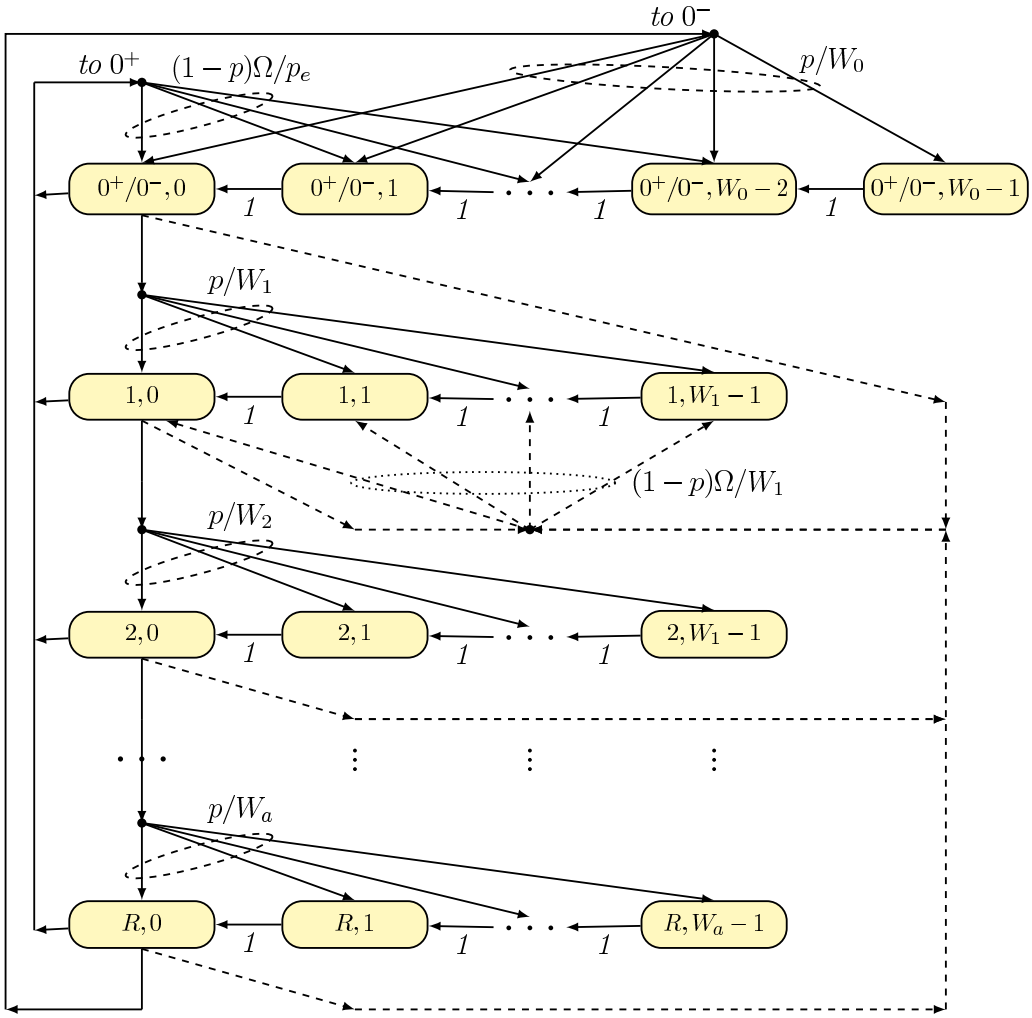


Figure 5.2: Markov chain model states of a given DCF STA.

backoff counter is set to a nonzero value after a successful transmission, the current model slot shall be ended and the backoff stage shall go into stage 0^+. On the other hand, if a failed transmission follows multiple successful transmissions, the backoff stage shall go into into 1-stage because a transmission error results in an increase of the backoff stage. The sixth and seventh sub-equations in (5.15) correspond to the above two cases, respectively.

Let us consider the case where there are ($j_s > 0$) consecutive successful transmissions in one model slot. After the first transmission (with success probability $1-p$), the next $j_s - 1$ transmissions are collision-free. Thus, the $j_s - 1$ transmissions are successful with probability $(1-p_e)^{j_s-1}$. Furthermore, after each transmission except for the j_s^{th} transmission, the backoff counter is set to 0 with probability $1/W_0$. The j_s^{th} transmission ends the model slot with a nonzero backoff counter value with probability $(W_0-1)/W_0$.

For the case when $j_s - 1 (> 0)$ consecutive successful transmissions are followed by a failed transmission in one model slot, all j_s transmissions are successful with proba-

bility $(1-p)(1-p_e)^{j_s-1}$. After each successful transmission, the backoff counter is set to 0. A failed transmission (with error probability p_e) ends the model slot regardless of the value of the backoff counter. Based on the above discussion, equation (5.16) can be obtained

$$\begin{cases} P\{0^+,k|i,0\} = \frac{1-p}{W_0-1}\sum_{j_s}^{\infty}\left(\frac{1-p_e}{W_0}\right)^{j_s-1}\left(\frac{W_0-1}{W_0}\right) \\ P\{1,k|i,0\} = \frac{1-p}{W_1-1}\sum_{j_s}^{\infty}\frac{1}{W_0^{j_s}}\left(1-p_e\right)^{j_s-1}p_e \end{cases} \tag{5.16}$$

Note that the expressions in (5.15) can be derived from (5.16).

Let $b_{i,k} = \lim_{t\to\infty} P\{s(t)=j, b(t)=k\}$ be the stationary probability of the Markov chain. Owing to the chain regularities, and by imposing the normalization condition, it is found that

$$\tau = \sum_{i=1}^{R} b_{i,0} + b_{0^+,0} + b_{0^-,0} = \frac{2(1-p^{R+1})}{(1-p)\Theta} \tag{5.17}$$

where

$$\Theta = \sum_{i=1}^{R}(W_i+1)\left[\Omega p^{i-1} + (1-\Omega)\,p^i\right] + W_0 + 1 - (1-\Omega)\left(1-p^{R+1}\right). \tag{5.18}$$

5.3.2 The saturation goodput

The system saturation goodput G can be defined as

$$G = \frac{E\{\text{payload information in a slot}\}}{E\{\text{duration of a slot}\}}. \tag{5.19}$$

As previously stated in the slot model, each slot can be occupied by different events of different durations. Each of these events has also a different probability of occurrence to be considered when calculating the average slot duration.

Let $P_b^{(n)}$ denote the probability that the channel is busy in an scenario with n users and $P_s^{(n)}$ define the probability that a single STA successfully transmits in a slot time when having symmetric n users in the scenario. In this case, as in [13], it readily follows that

$$\begin{aligned} P_b^{(n)} &= 1-(1-\tau)^n \\ P_s^{(n)} &= n\tau(1-\tau)^{n-1}(1-p_e) \end{aligned} \tag{5.20}$$

According to the model proposed, we have to recall that the anomalous slot may cause multiple transmissions inside the same slot and this must be accounted for when calculating the average duration of successful slots. Note that $\hat{T}_{T_c}(m) = \hat{T}_{T_e}(m)$ when using the same MCS m for all STAs, then the goodput for this MCS can be defined as

$$G(m) = \frac{P_s^{(n)}\overline{L_s}}{\left(1-P_b^{(n)}\right)\sigma + P_s^{(n)}\overline{T}_s(m) + (P_b^{(n)}-P_s^{(n)})\hat{T}_{T_c}(m)} \tag{5.21}$$

where $\overline{L_s}$ is the average payload size of successful transmissions and $\overline{T}_s(m)$ is the average slot duration, which includes at least one successful transmission. Given the condition that there is a successful transmission, for a model slot that consists of

j_s transmissions (with probability $\left(\frac{1-p_e}{W_0}\right)^{j_s-1}\left(1-\frac{1}{W_0}\right)$), then $\overline{L_s} = j_s L_p$ and $\overline{T}_s(m) = j_s T_{T_s}(m)+\sigma$. Note that $L_p = L_M - L_h$ denotes the payload information transmitted inside a packet with L_h defining the MAC sublayer overhead. If the slot contains j_s successful transmissions and a failed transmission (with probability $\left(\frac{(1-p_e)^{j_s-1}p_e}{W_0^{j_s}}\right)$), then $\overline{L_s} = j_s L_p$ and $\overline{T}_s(m) = (j_s+1)T_{T_s}(m)+\sigma$. Taking into account all cases with $j_s > 0$ and then their corresponding probabilities of occurrence, results in

$$\overline{L_s} = \frac{W_0}{W_0 + p_e - 1} L_p \tag{5.22}$$

$$\overline{T}_s(m) = \frac{W_0 + p_e}{W_0 + p_e - 1} T_{T_s}(m) + \sigma \tag{5.23}$$

.

5.4 MAC sublayer performance results

In order to validate the analytical model just presented, an IEEE 802.11n MAC sublayer simulator has been implemented using MATLAB. This simulator mainly focuses on the MAC sublayer, strictly following the descriptions presented in Sect. 5.1. The physical layer is modeled by a fixed MCS-special channel error probability p_e and the corresponding overhead. An uplink scenario is considered with a fixed number of users using $CW_{min} = 15$. This simulator has been duly validated by reproducing results from [13] and [28].

Figures 5.3 and 5.4 show the τ and p performance, respectively, as a function of the number of users when using $R = 7$ and $p_e = \{0.05, 0.1, 0.2\}$. Note that these results are also compared to those obtained using the analytical models proposed by Bianchi [13], Tinnirello et. al [23] and Chen [28]. It should be mentioned that Tinnirello and Bianchi models are modified in order to account for the channel errors, by defining p as a function of p_e using (5.14). It can be observed that the modelling of the anomalous slot (Tinnirello, Chen) leads to more accurate results than the ones provided by original Bianchi's model. Regarding p and τ performance, note that higher p values are obtained as the number of contending users increases, in contrast, τ decreases since the backoff algorithm doubles the contention window more often in response to the collisions increase. It is worth mentioning that the p and τ performance is identical irrespective of the selected MCS, since each MCS suffers the same p_e value.

In case of $R = 4$, p and τ performance is shown in Figs. 5.5 and 5.6, respectively. Note that these two figures are the Figs. 5.3 and 5.4 counterparts. Remarkably, a slight probability increase is noticeable in both τ and p in comparison to the $R = 7$ case, since the number of backoff stages is lower, thus implying a reduction in the average backoff value. Notice also that similar accuracy performance is achieved by the whole set of analytical models.

The system goodput performance when using MCS0, MCS4, MCS12 and MCS15 for $R = 7$ is presented in Figs. 5.7, 5.8, 5.9 and 5.10, respectively. In these figures, the system performance is compared with that of the analytical models by Bianchi, Tinnirello and Chen. As before, Tinnirello and Bianchi goodput expressions were slightly modified to account for the channel error cases by using the P_s expression defined in

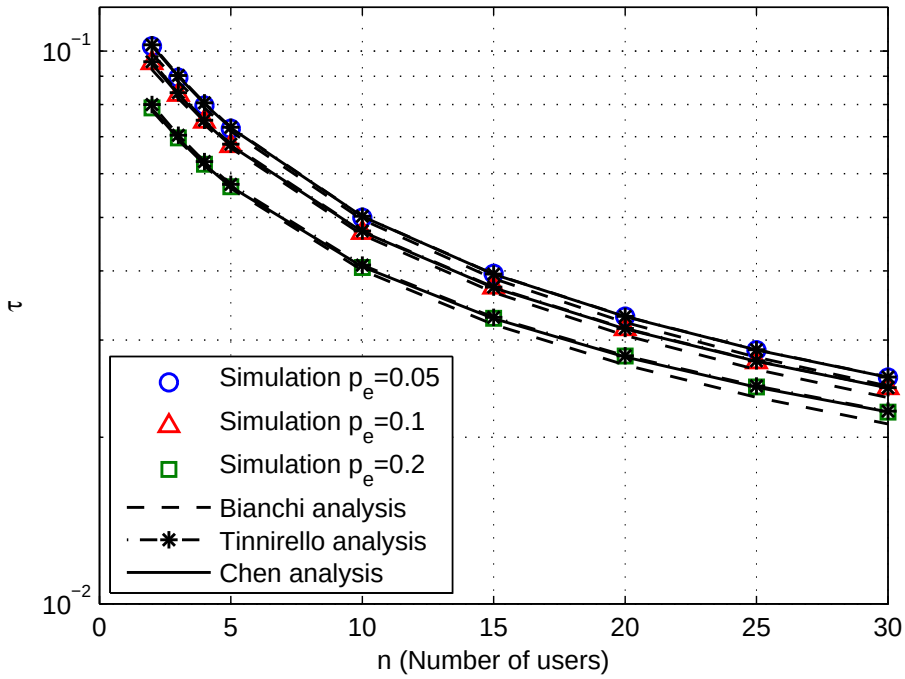


Figure 5.3: Probability of transmission of each station (τ) when using a system configuration with $R = 7$, $CW_{min} = 15$ and $a = 6$.

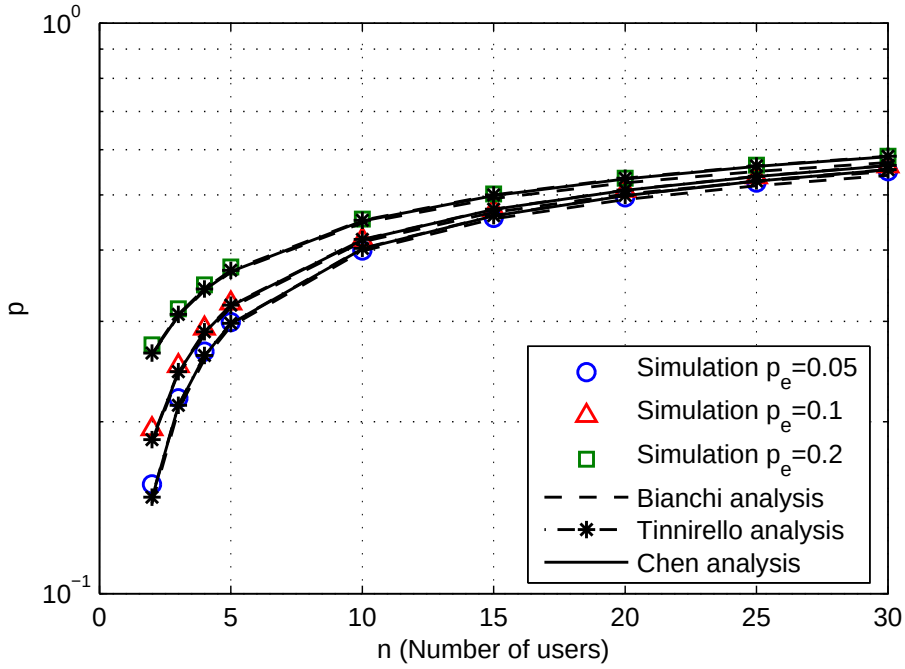


Figure 5.4: Probability of collision and error of each station (p) when using a system configuration with $R = 7$, $CW_{min} = 15$ and $a = 6$.

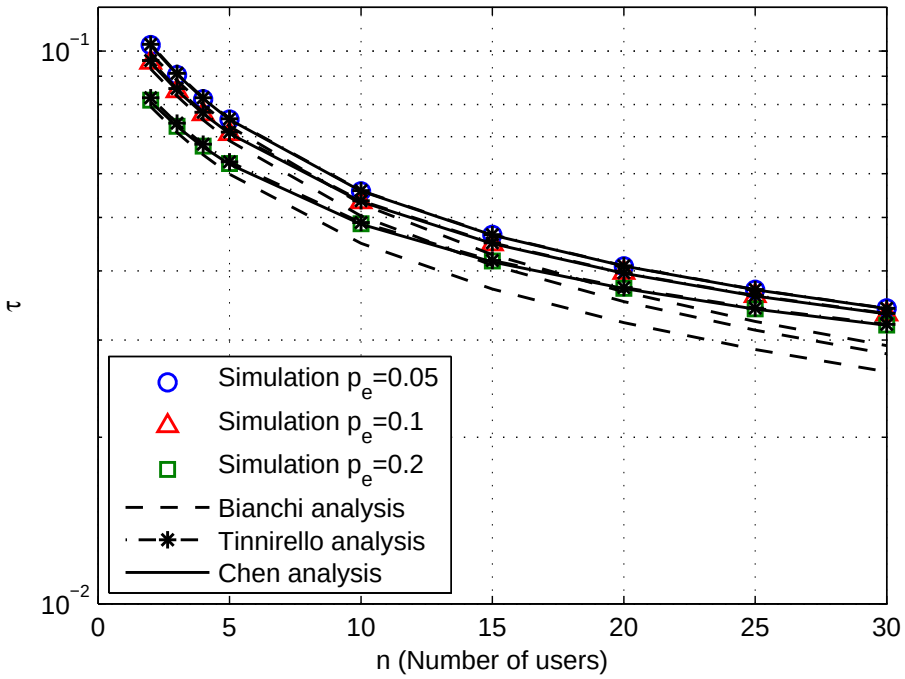


Figure 5.5: Probability of transmission of each station (τ) when using a system configuration with $R = 4$, $CW_{min} = 15$ and $a = 4$.

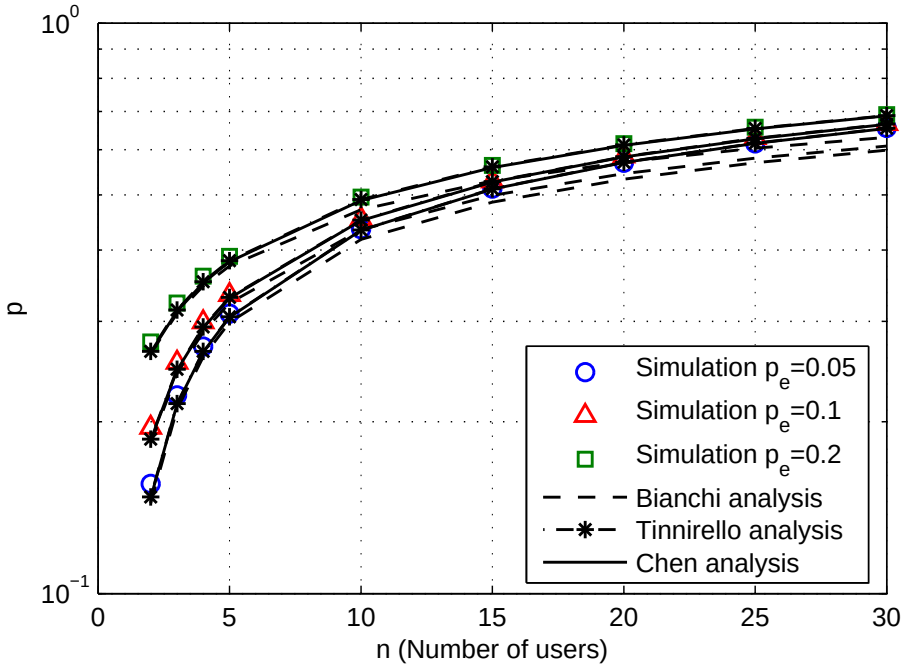


Figure 5.6: Probability of collision and error of each station (p) when using a system configuration with $R = 4$, $CW_{min} = 15$ and $a = 4$.

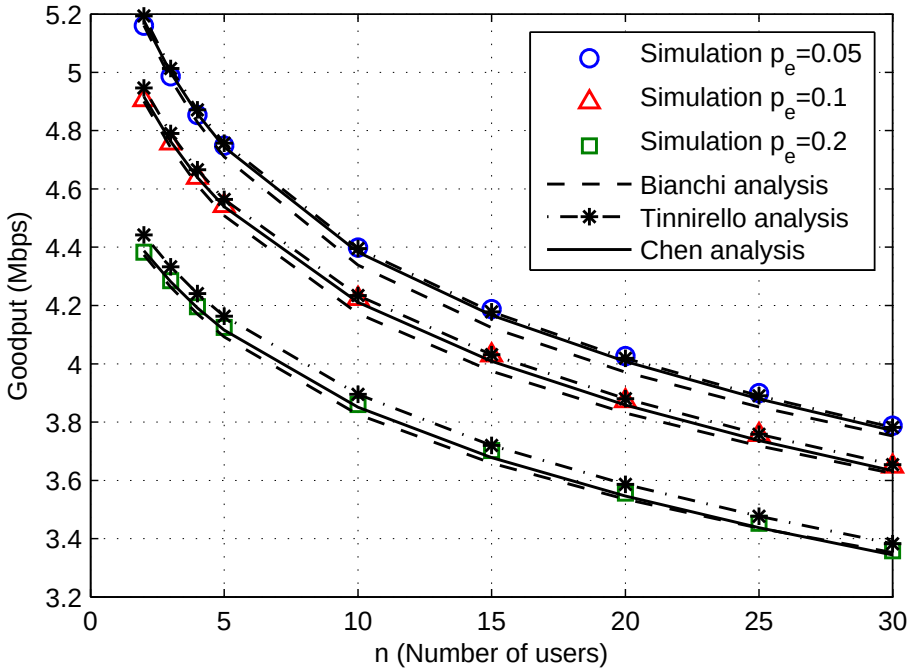


Figure 5.7: System goodput for MCS0 when using $R = 7$, $CW_{min} = 15$ and $a = 6$.

(5.20). Note that in each figure, three packet error probabilities are evaluated, $p_e = \{0.05, 0.1, 0.2\}$. From these results, we conclude that accurate goodput predictions are achieved when using Tinnirello and Chen models. Conversely, as it was previously stated in the τ and p modelling approaches, important deviations are appreciated when using the Bianchi model. Regarding the employed MCS, goodput improvements are achieved when using higher MCS values. However, this does not avoid the performance degradation caused by collisions. Furthermore, as expected, the goodput performance is lowered as p_e increases.

Finally, the goodput performance using $R = 4$ for MCS0 and MCS15 is depicted in Figs. 5.11 and 5.12, respectively, for different number of users and for $p_e = \{0.05, 0.1, 0.2\}$. In this case, the best performance results are also obtained using the Tinnirello and Chen models (specially Chen's one). Comparing the goodput performance with the results shown in Figs. 5.7 and 5.10, note that an important goodput reduction is observed due to a greater collision probability.

5.5 Chapter Summary

In this chapter, the DCF mechanism of the IEEE 802.11 has been thoroughly studied. Firstly, a description of the DCF MAC sublayer mechanism has been provided, hence detailing the access protocol operation, the binary exponential backoff algorithm, the time duration of each event and the available DCF access mechanisms. Secondly, a short literature review of analytic models has been presented, detailing the model slot that accurately follows the standard definition. Subsequently, the analysis model for saturation conditions has been introduced, allowing the evaluation of system perfor-

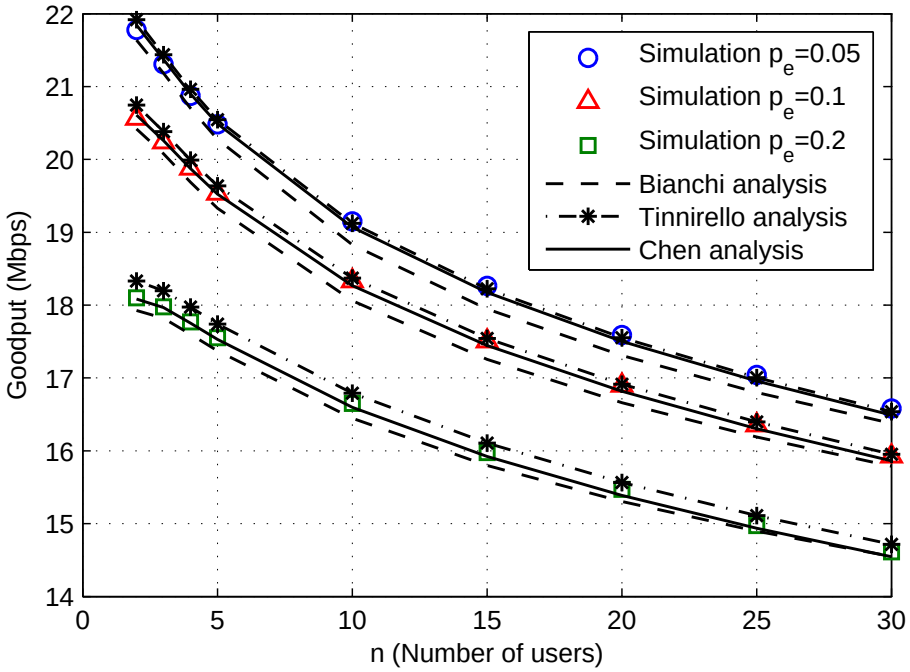


Figure 5.8: System goodput for MCS4 when using $R = 7$, $CW_{min} = 15$ and $a = 6$.

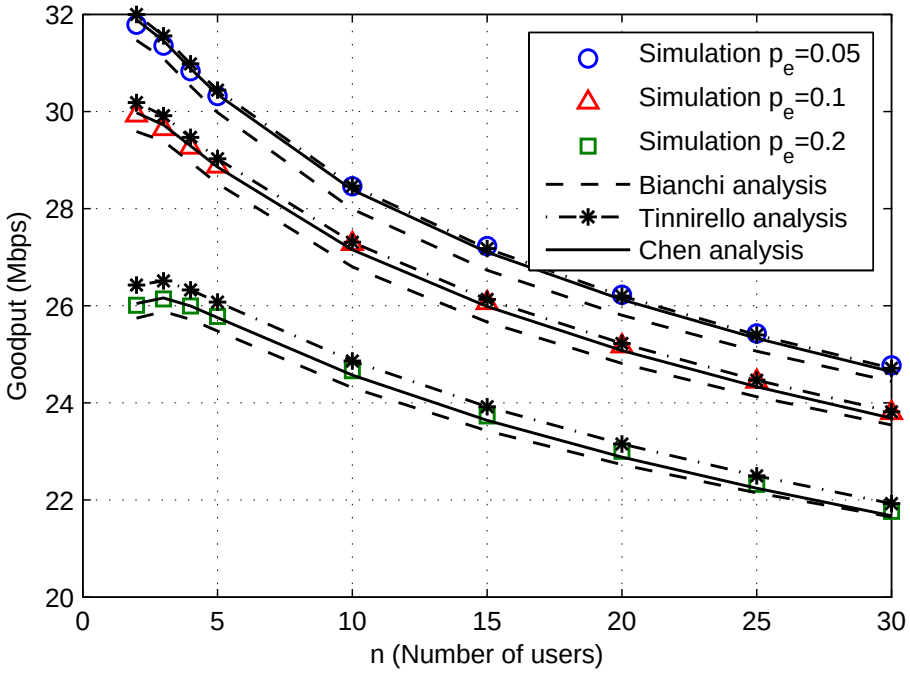


Figure 5.9: System goodput for MCS12 when using $R = 7$, $CW_{min} = 15$ and $a = 6$.

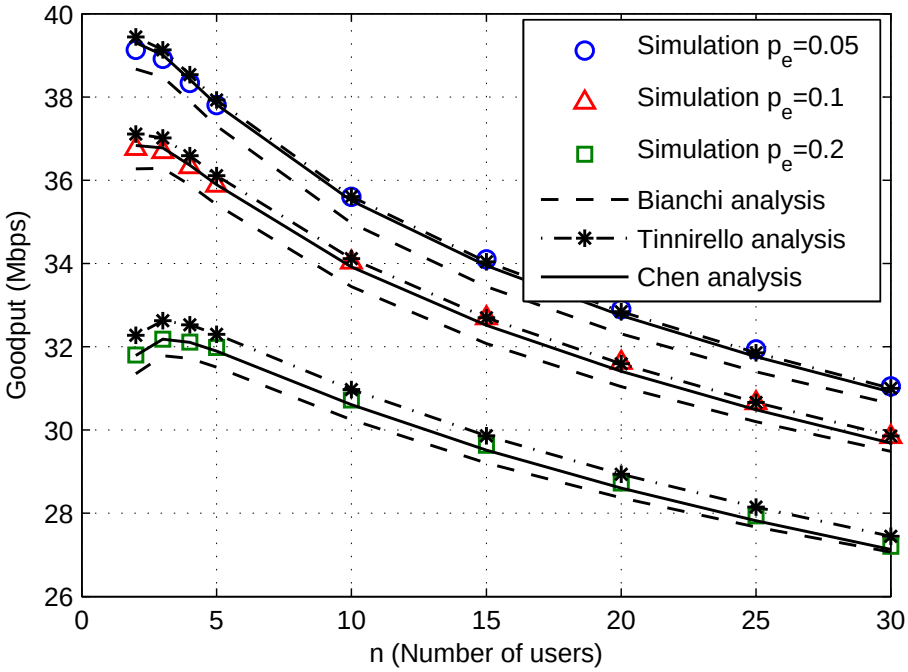


Figure 5.10: System goodput for MCS15 when using $R = 7$, $CW_{min} = 15$ and $a = 6$.

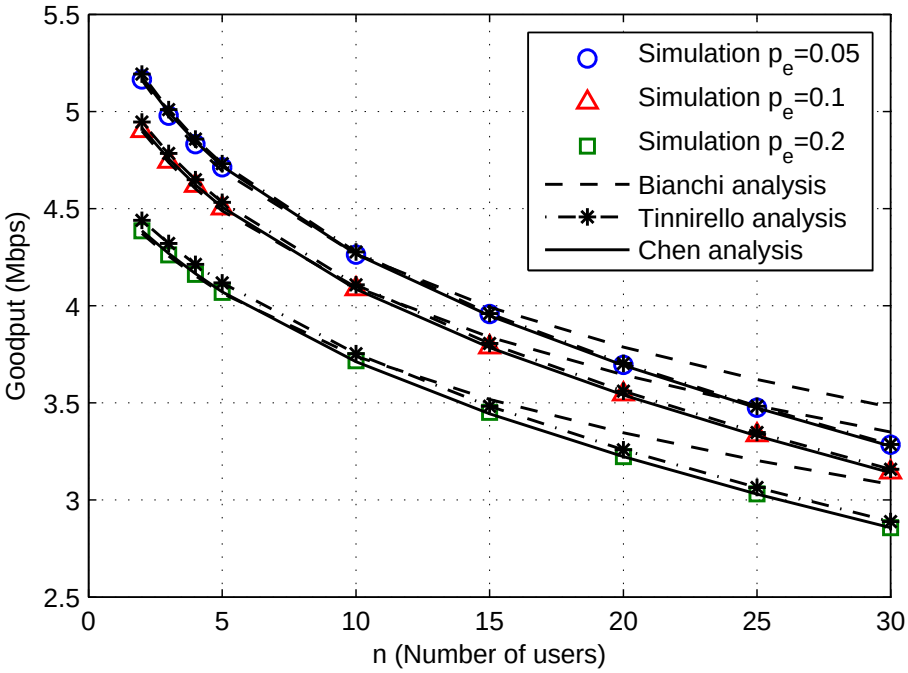


Figure 5.11: System goodput for MCS0 when using $R = 4$, $CW_{min} = 15$ and $a = 4$.

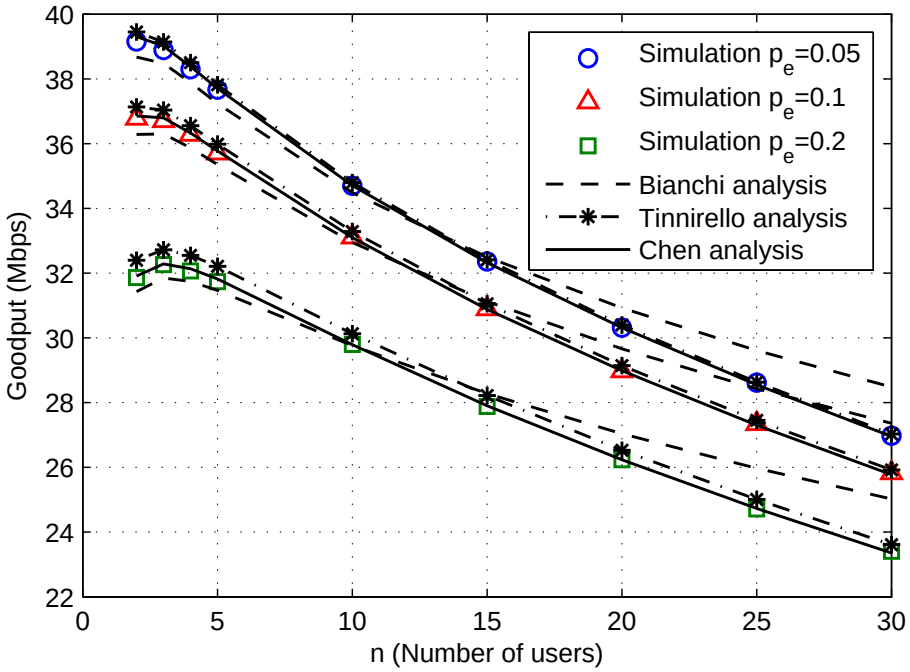


Figure 5.12: System goodput for MCS15 when using $R = 4$, $CW_{min} = 15$ and $a = 4$.

mance metrics such as goodput, probability of collision and STA probability of transmission. Lastly, the performance results of the analytic model have been validated by contrasting them to a custom simulator that considers all the DCF details.

CHAPTER 6

FAST LINK ADAPTATION PERFORMANCE UNDER SATURATED CONDITIONS

This chapter presents a comprehensive performance study of closed-loop FLA in the context of IEEE 802.11n, spanning the PHY and MAC layers. In particular, a semi-analytical model is derived for Basic and RTS/CTS access schemes of the DCF, that applies to both, open- and closed-loop strategies. Numerical results serve to demonstrate the accuracy of the proposed model and the superiority of FLA, in terms of MAC goodput, in comparison to open-loop policies. Realistic operating conditions such as outdated feedback information and the use of statistical packet length distributions, issues not treated in previous studies, have also been considered. Moreover, it is shown that incorporating a time-out mechanism in the FLA scheme, is a useful strategy to counteract the deleterious effects of outdated information. This strategy relies on a timeout-based mechanism that weights down the importance of the MCS feedback information as this becomes outdated.

Moreover, the performance of FLA in CSMA with enhanced collision avoidance (CSMA/E2CA) protocol is also presented. It is shown how the FLA algorithm benefits from the collision reduction experienced by CSMA/E2CA-based networks with respect to the IEEE 802.11 standard mechanism, CSMA/CA protocol. As in the DCF case, a model approach for the performance evaluation of CSMA/E2CA is also provided and validated. For completeness, the performance of FLA is compared with the ubiquitous ARF algorithm, showing that FLA clearly outperforms open-loop techniques such as ARF.

6.1 Motivation and Problem Statement

Adaptation plays a crucial role in dealing with the time varying nature of the wireless channel. Adaptive mechanisms allow the reconfiguration of system parameters, notably the transmission mode, in order to exploit the available instantaneous channel capacity while satisfying QoS constraints. One of the most widely used reconfigura-

tion techniques is AMC (also known as rate adaptation), which selects an appropriate MCS in response to changes in the environment or system behaviour. AMC algorithms can be broadly categorized as closed- or open-loop, depending on whether an explicit feedback channel between Rx and Tx is used or not. Open-loop setups operate in a heuristic manner and their rate of adaptation tends to be slow with respect to channel changes, thus risking the fulfilment of QoS constraints [19, 93]. In contrast, closed-loop mechanisms track more accurately the channel behavior and are more reactive to rapid channel variations [16].

Most IEEE 802.11-based systems employ the DCF with Basic Access scheme at the MAC sublayer and adopt open-loop AMC policies such as ARF [94] or one of its variants (e.g. collision-aware rate adaptation (CARA) [18], stochastic automata rate adaptation (SARA) [95]). Owing to its simplicity, ARF is by far the most popular algorithm in use. Unfortunately, the DCF with Basic Access scheme does not differentiate between collisions and transmission failures caused by poor channel conditions. Consequently, when the system experiences a high collision probability, ARF tends to use the lowest transmission rate even if the channel conditions would allow the use of higher transmission modes (see for example, [18, 96, 97, 98]). Many adaptive strategies have been proposed to solve this issue, but they may require either frame format changes [99], modifications to the medium access technique [18], or the use of channel quality indicators (e.g. signal strength indicator) [97, 99] and, in fact, none of them has achieved widespread use in current WLAN systems [100].

Alternatively, DCF can employ the RTS/CTS access scheme instead of the Basic Access scheme. The use of RTS/CTS shortens the collision duration and allows to differentiate between collisions and transmission failures at the expense of increased overhead. Consequently, RTS/CTS-based operation results in very different system characteristics whose performance is worth studying further.

A different approach to combat the deleterious effects collisions have in the AMC operation consists of modifying the backoff technique so that the probability of collision is drastically reduced. This is the rational behind a variant of the CSMA/CA protocol, called CSMA/E2CA, originally presented in [101] and further improved in [4]. The key idea is the use of a deterministic backoff after successful transmission. In particular, in CSMA/E2CA, a deterministic backoff is used for at least two consecutive transmissions after each successful transmission. The use of a deterministic backoff after successes substantially reduces the number of collisions and the system tends to converge to a collision-free operation, thus implying improvements for the goodput and AMC performance, specially to those algorithms whose performance is notably degraded by collisions, such as, ARF. Note that the performance of CSMA/E2CA under realistic channel conditions (AMC mechanisms, channel fading) has not yet been studied.

In Chapter 4, it was demonstrated that, in the context of IEEE 802.11n, the use of closed-loop techniques such as FLA offers important benefits in terms of physical layer throughput. However, current literature does not explore how this improvement reflects on the MAC goodput of a FLA-based system. This chapter presents a semi-analytical model that can be used to assess the goodput performance at the MAC layer of both, open- and closed-loop adaptive schemes targeting IEEE 802.11n. The proposed model expands the one presented in Chapter 5 by considering the use of AMC algorithms and modelling the error performance using realistic channel mod-

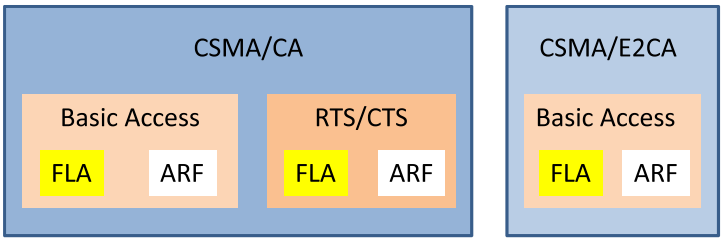


Figure 6.1: Scenario configurations.

els. Additionally, issues that may affect very significantly the practical implementation of closed-loop strategies such as having to cope with delayed (possibly outdated) feedback information and the possible utilization of different packet lengths are also considered in this study. For this reason, several novel strategies for the FLA-based scheme that minimise the effects of delayed feedback are presented and validated. Furthermore, a novel semi-analytical model approach to evaluate the performance of CSMA/E2CA networks using AMC algorithms is introduced, validating its performance using the same scenario conditions than CSMA/CA networks. Lastly, this chapter exhaustively evaluates the different combinations of AMC strategy (open- or closed-loop), network protocol (CSMA/CA or CSMA/E2CA) and access technique (Basic Access or RTS/CTS) with the aim of devising the optimum setup in terms of maximisation of MAC goodput while considering a target PER value and maintaining a high fairness index among users. In particular, the set of combinations studied in this chapter are schematized in Fig. 6.1.

The rest of the chapter is structured as follows. Section 6.2 briefly reviews the two adaptive schemes covered in this work, namely, ARF and FLA. For the FLA case, several timeout strategies are presented to cope with outdated MCS information. In Section 6.3, the analytical framework used to analyze the system goodput for DCF with CSMA/CA is presented. Section 6.4 describes the main characteristics of CSMA/E2CA and derives its goodput analysis model. Section 6.5 presents a physical layer abstraction model derived from results of Chapter 4. In Section 6.6, the numerical results comparing the performance of open and closed-loop AMC algorithms for DCF CSMA-/CA and CSMA/E2CA networks are presented assuming different configurations and access techniques. Finally, in Section 6.7, the main conclusions of this chapter are summarized.

6.2 Adaptive Modulation and Coding Strategies

6.2.1 AutoRate Fallback (ARF)

This algorithm adapts the transmission rate according to the number of consecutive transmission failures and successes, both reported by the ACK mechanism. The transmission rate is decreased after two consecutive transmission failures and increased after either ten consecutive successful packet transmissions or a timeout [94]. In order to improve the system adaptation during long intervals of inactivity, this timeout counter is reset after a transmission rate change or after a transmission failure. Ac-

ceptable timeout values lie in the range of 50-200 ms [99]. Note that, following a rate increase, the next data transmission is deemed as a probing transmission for the new mode. If an ACK is not received for this probing packet the system falls back to the previous data rate.

In order to implement ARF in IEEE 802.11n it is necessary to determine the available rates in the MCS set, denoted by $\mathscr{M}$. In contrast to previous IEEE 802.11 standards, in 802.11n different MCSs $\in \mathscr{M}$ can provide the same transmission rate [1], but only one of them can be used by the ARF algorithm. For this reason, the MCSs in $\mathscr{M}$ are reordered according to their transmission rate, and for those rates that can be attained using either SDM or STBC, only the STBC MCS is kept as it can be shown to be more robust against channel variations [80].

6.2.2 FLA and Extensions

Fast link adaptation is a closed-loop technique that relies on the availability of a feedback channel from the receiver to the transmitter. The main idea behind FLA is that the receiver, thanks to an accurate knowledge of the channel response, can compute a reliable prediction of the error rate for all available MCSs and choose the one maximising the instantaneous throughput while satisfying QoS constraints in the form of outage packet error rate probability. The selected MCS can then be communicated to the transmitter via the feedback channel. In this work we assume the use of the methodology presented in Chapter 4, where link performance prediction for each MCS is based on the EESM. Using this approach, the EESM for a given MCS can be easily associated to PER using look-up tables that have been previously computed during an off-line calibration phase.

Basic Access. Timeout-based strategies.

When using Basic Access with a large number of active users, a large delay between the MCS selection at the receiver and its use at the transmitter is to be expected. Often, this delay exceeds the channel's coherence time, significantly affecting the FLA operation. In DCF, all STAs have an equal long term probability of accessing the medium. Therefore, successive transmissions from a given STA are intertwined with transmissions from the other contending STAs and, consequently, the time between successive transmissions increases, thus affecting the MCS feedback delay. This delay becomes critical for FLA. When the delay exceeds the channel's coherence time, the provided MCS may have been determined for a channel response that is almost uncorrelated to the current channel response. This mismatch between current and prior channel states can increase the error probability due to a mistakenly selected (expired) MCS, causing several consecutive errors in the next retransmissions prior to packet discard.

In order to determine when the feedback MCS information is stale, a timeout mechanism has been incorporated at the transmitter side, measuring the time elapsed from the last successful transmission whose ACK provided feedback information to select the MCS. This timeout value should be chosen close to the channel's coherence time. In order to reduce the probability of error, the timeout expiration must trigger an MCS

[1] As stated in 3, the use of MIMO schemes in 802.11n enable that some transmission rates can be achieved by multiple MCSs.

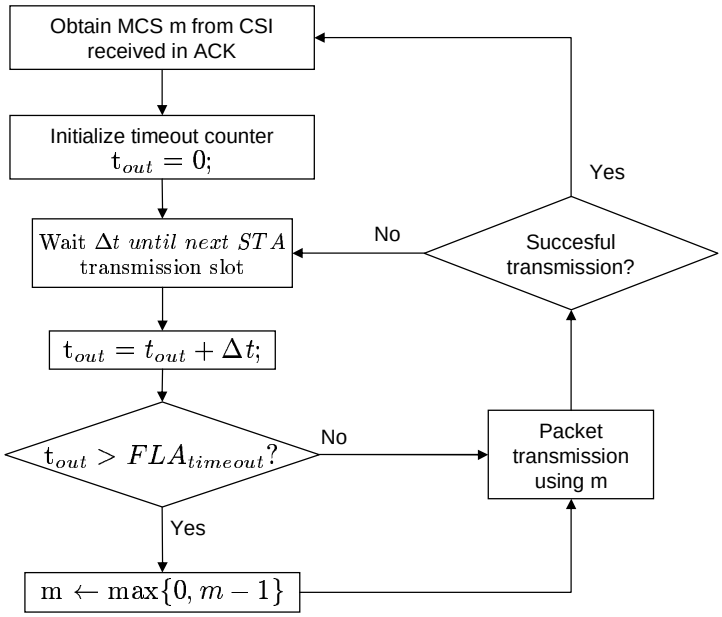


Figure 6.2: Block diagram of FLA *D*ownRate.

post-timeout update strategy. In this work, the following post-timeout strategies are proposed:

- *Rate decrease for each packet retransmission (DownRate)*

  Taking the last feedback MCS information as a reference, the STA decreases in one rate-level the MCS before packet transmission. This action is repeated for each subsequent retransmission until the packet is successfully received. The reception of the corresponding ACK triggers the timeout counter reset. This is a useful mechanism that effectively combats the effects delayed feedback information has on FLA-based CSMA/CA networks. Nevertheless, the selection of the timeout for each protocol type plays a crucial role when optimising system performance. The block diagram of this strategy is presented in Fig. 6.2.

- *RTS/CTS exchange prior to the packet transmission*

  Another post-timeout strategy is to activate the use of RTS/CTS frames until the packet is successfully transmitted.

  Unlike Basic Access, in the RTS/CTS scheme the MCS to be used in the next packet transmission is determined using the channel response affecting the RTS frame, and is returned to the sender in the CTS frame response, just prior to data transmission. Therefore, the MCS feedback delay is almost negligible compared to the channel's coherence time, and independent of the time elapsed since the last successful transmission.

  However, the use of the RTS/CTS mechanism increases the packet transmission overhead, thus lowering the goodput performance specially when using short packet lengths and/or when a small number of users are contending. As in the

previous strategy, the timeout counter is reset once the transmission is successfully completed.

RTS/CTS

When RTS/CTS is employed, as previously stated in the FLA post-timeout strategy with RTS/CTS, the FLA performance is not affected by the MCS feedback delay since this is negligible compared to the channel's coherence time. In contrast, the system performance may be notably degraded due to the protocol overhead introduced by the RTS/CTS frame exchange that precedes each packet transmission. This performance degradation increases whenever high transmission rates and/or short packet lengths are employed.

6.3 DCF Goodput Analysis

Following the analytical model for saturated conditions described in Chapter 5, the goodput analysis presented in this section is suitable for Basic and RTS/CTS access techniques when using DCF with CSMA/CA.

The system saturation goodput G can be defined as

$$G = \frac{E\{\text{payload information in a slot}\}}{E\{\text{duration of a slot}\}}. \tag{6.1}$$

As previously stated in Chapter 5, the duration of a slot refers to the time interval between two consecutive backoff counter decrements [23].

In any given slot, one out of four events can occur: a successful packet transmission (s), an error packet transmission (e), a collision (c) or an idle slot (i). From the point of view of the BEB algorithm, error transmissions and collisions are undistinguishable. Assuming a symmetric network with n active STAs, the conditional probability of the union of these events can be computed as

$$p = 1 - (1 - \overline{p_e})(1 - \tau)^{n-1}, \tag{6.2}$$

where $\overline{p_e}$ is the user error transmission probability obtained by simulation for the considered AMC algorithm, averaged across users, and τ is the stationary probability that a particular STA transmits in a given slot. Similarly to (5.17), this transmission probability can be obtained as

$$\tau = \frac{2(1 - p^{R+1})}{(1 - p)\Theta}, \tag{6.3}$$

with

$$\begin{aligned} \Theta = &\sum_{i=1}^{R} (W_i + 1) \left[\frac{\overline{p_e} p^{i-1} + (W_0 - 1) p^i}{W_0 + \overline{p_e} - 1} \right] + W_0 + 1 \\ &- \left(1 - \frac{1 - \overline{p_e}}{W_0 + \overline{p_e} - 1} \right) \left(1 - p^{R+1} \right), \end{aligned} \tag{6.4}$$

where $W_i = 2^{\min(i,a)}(CW_{min} + 1) - 1$, CW_{min} is the minimum contention window, $a = log_2\left(\frac{CW_{max}+1}{CW_{min}+1}\right)$ denotes the maximum backoff stage, CW_{max} is the maximum contention window size and $W_0 = CW_{min} + 1$. Notice that p and τ can be obtained by solving the nonlinear system of equations formed by (6.2) and (6.3).

Using τ, the probability that only one out of the n STAs transmits on a given slot achieving a successful transmission is

$$P_s^{(n)} = n\tau(1-\tau)^{n-1}(1-\overline{p_e}). \tag{6.5}$$

Similarly, the probability that only one STA transmits on a given slot resulting in an erroneous transmission is

$$P_e^{(n)} = n\tau(1-\tau)^{n-1}\overline{p_e}. \tag{6.6}$$

Furthermore, the probability that a given slot is idle is given by

$$P_i^{(n)} = (1-\tau)^n. \tag{6.7}$$

Among all possible events, only the successful packet transmission increases the payload information while any other event leads to a goodput degradation. Consequently, modifying the goodput expression (5.21) from Chapter 5, and taking into account the possible use of multiple transmission modes at each event, the system goodput can be expressed as

$$G = \frac{P_s^{(n)}\overline{L_s}}{P_i^{(n)}\sigma + P_s^{(n)}\overline{T}_s + P_e^{(n)}\overline{T}_{T_e} + (1-P_i^{(n)}-P_s^{(n)}-P_e^{(n)})\overline{T}_{T_c}}, \tag{6.8}$$

where $\overline{L_s} = E\{L_p\}\frac{W_0}{W_0+\overline{p_e}-1}$, with $L_p = L_M - L_h$ represents the packet payload length and L_h denoting the MAC sublayer overhead, and the time values $\overline{T}_s$, $\overline{T}_{T_c}$ and $\overline{T}_{T_e}$ correspond to the average elapsed time for successful, colliding and error transmissions, respectively. It should be pointed out that these average time values are determined by simulation and depend on the mean MPDU length (L_M), the number of STAs (n) and the probability of use of each MCS according to the AMC employed and the scenario configuration. Notice that $E\{L_p\}$ is multiplied by $\left\lfloor \frac{W_0}{W_0+\overline{p_e}-1} \right\rfloor$ in order to account for the additional information transmitted during anomalous slots [23] and

$$\overline{T}_s = \frac{W_0+\overline{p_e}}{W_0+\overline{p_e}-1}E\{T_{T_s}(m,L)\} \tag{6.9}$$

includes the anomalous slot duration.

The necessity of obtaining variables $\overline{p_e}$, $\overline{T}_s$, $\overline{T}_{T_c}$, $\overline{T}_{T_e}$ and $\overline{L_s}$ from numerical simulation is what renders our proposed model as semi-analytical. Note that this model is considerably more realistic than the one proposed by Tinnirello et al. in [23] and Chen et al. in [28], since it allows the treatment of AMC, statistical packet length distribution and non-ideal closed-loop FLA strategies, at the expense of relying on some semi-analytic parameters.

6.4 CSMA/E2CA

6.4.1 Protocol Description

The number of collisions in wireless networks can be notably reduced if a deterministic backoff is used after successful transmissions, as in the CSMA/E2CA protocol [101]. The deterministic backoff value after successful transmissions has to be the same for

all the STAs will be denoted by C. Two STAs that have successfully transmitted in their last transmission attempt have necessarily transmitted in different slots, otherwise their transmission would have resulted in a collision. Then, if those STAs backoff for exactly C slots before their next transmission attempt, they will also choose a different slot in their next transmission attempt. The fundamental principle is that two stations that have successfully transmitted in their last attempt and use a deterministic backoff will not collide among them in their next transmission slot.

If all the STAs have successfully transmitted in their last transmission attempt, this fundamental principle applies to each pair of stations. As a result, all the STAs will use a deterministic backoff C and they will not collide in their next transmission attempt. At this point, it can be stated that collision-free operation has been reached since all the STAs behave deterministically and successfully transmit in a round-robin fashion. Collision-free operation substantially improves the overall network performance in terms of delay, jitter, throughput and fairness [101].

Before reaching collision-free operation, the network goes through a transient state in which collisions occur. It is therefore highly desirable to reduce the length of this period. In [4], it is shown that setting $CW_{max} = CW_{min}$ reduces the length of the transient state and thus increases the network performance. If a value $CW_{max} = CW_{min}$ is chosen, it is mandatory to adjust the value of CW_{min} to the number of contenders.

In order to visualize the execution of the protocol, it is useful to imagine that the stations distributively choose a slot from a set of C slots. If two stations choose the same slot, they collide and then, they will use a random backoff value, thus probably choosing different slots in next transmission attempts. When a slot is chosen by a single station, this station succeeds (in case of no transmission errors) and will backoff for C slots, thus employing the same relative position slot in the set of C slots. In a sense, the station sticks to that relative position slot after a successful transmission.

This principle can be easily generalized to consider higher degrees of stickiness. If a stickiness degree of two is used, a STA will use a deterministic backoff for two consecutive transmissions after each successful transmission. When a higher degree of stickiness is used, the STAs have a stronger tendency to stick to a slot in which they have successfully transmitted. Stickiness shortens the duration of the transient state and has the additional benefit of providing extra robustness against channel errors. A STA with a stickiness degree of two will stick to its slot even in the occurrence of a channel error. This STA will only move back to the random behaviour if it suffers two consecutive channel errors in two consecutive transmission attempts, which is highly unlikely if the right MCS is used. Therefore, a stickiness degree of two offers an optimal protocol performance [4].

The CSMA/E2CA protocol considered in this chapter uses a degree of stickiness equal to two. This protocol represents only a subtle modification to the original CSMA-/CA and therefore both the new protocol and the legacy one can coexist in the same network. For the new STAs to be fair to the legacy ones, the deterministic backoff should be chosen as $C = \frac{CW_{min}+1}{2} = \frac{W_0}{2}$, which is the expected backoff after successful transmissions of the legacy STAs. In case of having to select a random backoff value due to two consecutive errors, then the backoff value is randomly selected in the range of $[0, CW_{min}]$. As previously stated, the value for CW_{max} is set to CW_{min} and CW_{min} is dynamically adjusted to keep the value of busy slots between $\frac{1}{8}$ and $\frac{1}{2}$ as detailed in [4].

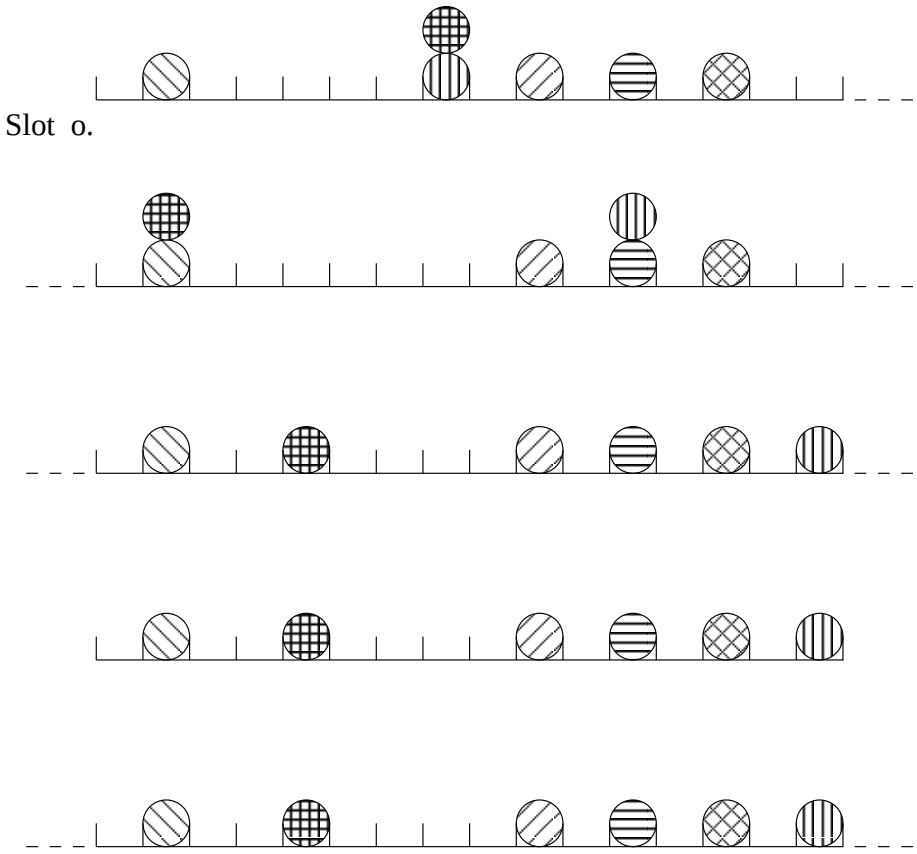


Figure 6.3: Example operation of CSMA/E2CA (extracted from [4]).

Example of operation

The operation of CSMA/E2CA is illustrated in Fig. 6.3. The figure shows a timeline which is divided in numbered slots. The transmissions are represented as balls in the slots, and the filling patterns of the balls are used to differentiate transmissions by different STAs. Note that in the figure all the slots are depicted as being equal for representation convenience, despite the fact that in reality busy slots are much longer than empty ones.

All the STAs transmit for the first time in the first row (slots from 0 to 15). Those stations that successfully transmit (in slots 1, 9, 11 and 13) use a deterministic backoff equal to 16, which in the figure is equivalent to sticking to the same column. Those STAs that suffer a collision (in slot 7) use a random backoff, which in this figure is equivalent to randomly choosing a new column for the next transmission attempt.

Moving to the second row of the figure, we can observe collisions in slots 17 and 27. Now, the station that successfully transmitted in slot 1 and then collided in slot 17 will use a deterministic backoff equal to 16, since a deterministic backoff is used for two consecutive times after each successful transmission. Contrastingly, the STA that collided in slot 7 and in slot 17 will choose a random backoff and will switch columns again. The situation is quite similar in slot 27. There is a collision but one of the STAs

will stick to the same column while the other one will switch columns.

In the third row of the figure all the STAs successfully transmit. At this point, all the STAs will use a deterministic backoff (which is equivalent to sticking to the same column) and will succeed in their next transmission attempt in the fourth row. It should be clear that collision-free operation has been reached, and therefore the behaviour of the system is completely deterministic and the participating STAs simply transmit in a round-robin fashion.

6.4.2 CSMA/E2CA Goodput Analysis

As in the DCF analysis, the CSMA/E2CA goodput analysis presented in this subsection has been performed assuming that all the STAs in the system are under saturation conditions, i.e., having a packet available to be transmitted at any moment. In this case, the goodput is also expressed as the average goodput per slot, thus using equation (6.1).

Considering ideal conditions, E2CA achieves collision-free operation and behaves as a round robin scheme where each user transmits once every C slots [4]. This leads to a constant probability of busy and idle slots with values $P_b^{(n)} = \frac{n}{C}$ and $P_i^{(n)} = 1 - P_b^{(n)} = \frac{C-n}{C}$, respectively, where n is the number of users (or contending STAs) and C is the value of the deterministic backoff. Under ideal conditions, the busy slots are uniquely filled by successful transmissions allowing the goodput to be expressed as [2]

$$G_{\text{E2CA}} = \frac{P_b^{(n)}\,\overline{L_p}}{P_b^{(n)}\,\overline{T}_{T_s} + (1 - P_b^{(n)})\sigma} \tag{6.10}$$

where $\overline{T}_{T_s}$ is the average time elapsed in a system successful transmission and σ is the duration of any idle slot.

Similarly, under non-ideal channel conditions with $n \leq \frac{C}{2}$ and $CW_{min} = CW_{max}$, $P_b^{(n)}$ and $P_i^{(n)}$ can be considered nearly constant when the error rate is not high [4]. The detailed analysis should also account for transmission errors and collisions. However, observing the weak performance degradation suffered in E2CA under non-ideal channel conditions when $n \leq \frac{C}{2}$, a tight upper-bound for the system goodput performance is given by

$$G_{\text{E2CA}} \leq \frac{P_b^{(n)}(1 - P_e^{(n)} - P_c^{(n)})\overline{L_p}}{P_b^{(n)}\left[(1 - P_e^{(n)} - P_c^{(n)})\overline{T}_{T_s} + P_c^{(n)}\overline{T}_{T_c} + P_e^{(n)}\overline{T}_{T_e}\right] + (1 - P_b^{(n)})\sigma}, \tag{6.11}$$

where $P_e^{(n)}$ and $P_c^{(n)}$ are the error and collision probabilities per transmission, respectively, and, $\overline{T}_{T_s}$, $\overline{T}_{T_e}$ and $\overline{T}_{T_c}$ are the average successful, collision and error duration, respectively.

As in the DCF case, the parameters $P_e^{(n)}$, $P_c^{(n)}$, $\overline{T}_{T_s}$, $\overline{T}_{T_e}$ and $\overline{T}_{T_c}$ are obtained by simulation, thus making our model semi-analytic.

[2] It is worth mentioning that this expression is only valid when having stable conditions. In this case, the saturation conditions are synonym of stability, since the transient state lapses for a brief period of time at the beginning of transmissions, as described in 6.4.1.

MCS	a_1	a_2
0	92,4138	1,1729
1	93,1221	0,5867
2	362,3555	0,5120
3	35,2028	0,1568
4	92,4112	0,1095
5	21,8023	0,0354
6	29,2704	0,0284
7	52,1024	0,0229

Table 6.1: BER curve parameters

6.5 Physical Layer modelling

Physical layer abstraction model

The PHY PER performance of the IEEE 802.11n for a given MCS and channel realization can be obtained by applying the physical layer abstraction model introduced in this subsection. This model relies on the BER prediction methodology presented in Sec. 4.3 and its calibration is derived from the physical layer results of Chapters 3 and 4. This model predicts the system PER performance for a given channel realization and MCS value.

This model works as follows. Firstly, the $SNR_{eff}^{(m)}$ of the current channel realization using MCS m is obtained by means of the EESM expression defined by eq. (4.4). This $SNR_{eff}^{(m)}$ is then used to find its corresponding BER value (P_{ϖ}) on the AWGN curve. To facilitate the calculation of P_{ϖ}, the AWGN BER curve achieved by simulation has been approximated by an exponential curve, that is,

$$P_{\varpi} = a_1 e^{-a_2 SNR_{eff}^{(m)}} \tag{6.12}$$

where a_1 and a_2 are the exponential curve fitting constants whose values for each one-stream MCS are presented in Table 6.1. The BER curves for two stream MCSs are derived from one stream curves, as stated in Chapter 4. For completeness, Fig. 6.4 shows the curve accuracy of the BER approximation for the set of one stream MCSs. Finally, P_{ϖ} is converted to PER using eqs. 4.7 and 4.8 parameterised by the current packet length and MCS value.

6.6 Results

In order to validate our semi-analytical model and compare the performance of FLA and ARF under different system configurations, such as, distinct number of users, diverse packet lengths, different access mechanisms (Basic Access, RTS/CTS) or different access protocols (CSMA/CA, CSMA/E2CA), an IEEE 802.11n system-level Matlab simulator has been implemented using the link-level parameters derived in section 6.5 and the procedures described in [5]. Simulator accuracy has been duly validated by reproducing results from [13, 23, 28] for single-mode transmission.

In this work, focus is on the performance evaluation of the uplink scenario where, nevertheless, MAC control frame transmissions from AP to STA are also accounted for,

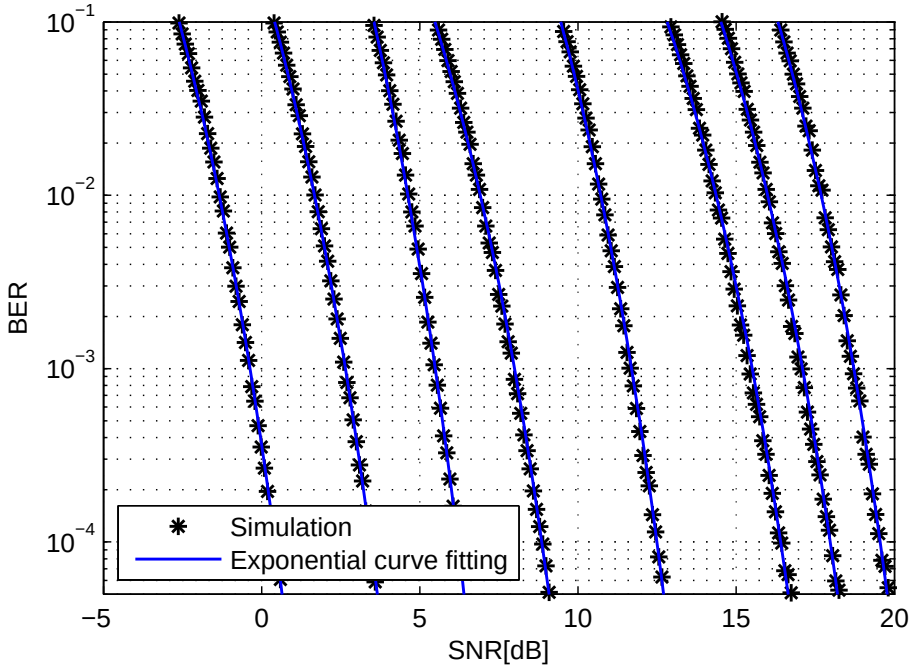


Figure 6.4: BER curve using exponential curve fitting for all the MCSs.

thus containing the corresponding CSI feedback overhead. All users in the system are assumed to use DCF with the same backoff algorithm (either BEB or E2CA) and access technique. Regarding the access scheme setup, CSMA/CA has been configured using the system parameters defined in the standard [5] (e.g., $CW_{min} = 15$, $CW_{max} = 1023$ and maximum number of packet retries ($R = 7$)) whereas CSMA/E2CA has been set up with a degree of stickiness of two, a maximum contention window $CW_{max} = CW_{min}$ and a fixed minimum contention window of $CW_{min} = 2^{\lceil log_2(4n) \rceil}$, thus satisfying $n \leq \frac{CW_{min}}{4}$.

The simulator generates a set of scenarios, where n static users are uniformly distributed in a circular area of radius 30 m around the AP. Note that this choice of radius assures a reduced distance between users, thus avoiding the appearance of the hidden terminal problem. Using the distance from each user to the AP, the simulator obtains the channel response from each user to the AP by employing the MIMO channel generation tool presented in [3]. Each user's channel response has been obtained assuming a 2×2 MIMO configuration, a carrier frequency $f_o = 2.4$ GHz and a bandwidth $B_w = 20$ MHz.

The simulator supports the first 16 MCS modes of IEEE 802.11n (MCS0-MCS15) employing the 20 MHz bandwidth mode and full GI of 0.8 μs, achieving data rates of up to 130 Mbps [5]. User transmit power and receiver noise power have been set to 20 dBm and -80 dBm, respectively. The simulator time constants have been configured according to the standard document as summarized in Table 6.2. Note that the High Throughput control field (HTC) required to feedback the MCS selection has been included in the frame size of ACK, RTS and CTS frames[3]. Regarding the AMC algorithm

[3]This is achieved by employing the Wrapper control frame that encapsules a control frame (i.e., ACK,

Table 6.2: System parameters used from the standard document [5, 6].

Name	Value	Name	Value
Idle slot duration	9μs	MAC header	320 bits
Short Interval Frame Space (SIFS)	16μs	ACK	160 bits
Distributed Interframe Space (DIFS)	34μs	RTS	208 bits
Propagation time	0μs	CTS	160 bits
PLCP preamble duration	40μs		
OFDM symbol period	4μs		

setup, the ARF timeout has been set to 60 ms and the FLA outage constraint for a PER objective of 10^{-1} (not including collisions) has been configured to 10%. In order to obtain an accurate estimate of the average system performance, $N_{sim} = 100$ realizations of duration $t_{sim} = 22$ seconds for each number of users (n) have been simulated.

In this model, and due to its negligible probability of occurrence, the possibility of error in the ACK, RTS and/or CTS transmissions has not been considered.

6.6.1 Basic Access

Figures 6.5a and 6.5b show the goodput performance and the conditional probabilities p and τ, respectively, as a function of the number of STAs. Results have been obtained for a fixed packet length of $L_p = 1500$ bytes and a maximum number of allowed retransmissions equal to either $R = 4$ or $R = 7$. A very accurate match between the semi-analytical and simulated system performance metrics for FLA- and ARF-based schemes can be appreciated. Figure 6.5a also reports the goodput performance obtained using the proposed semi-analytical model compared to a previous proposal[4], where the anomalous slot performance and the packet retry limit were not considered. Although the previous model provides valuable approximations to the simulation performance, the new semi-analytical framework results in improved modelling accuracy, especially when the system uses $R = 4$. Figure 6.5a also illustrates that, regardless of the retry limit, FLA-based schemes greatly outperform ARF-based strategies in terms of goodput performance. Note that FLA and ARF-based strategies with $R = 7$ outperform their counterparts with $R = 4$ in terms of goodput because a lower retry limit leads to a lower average backoff contention window, thus increasing the transmission probability (τ) and, consequently, the collision and error probability (p), see Fig. 6.5b.

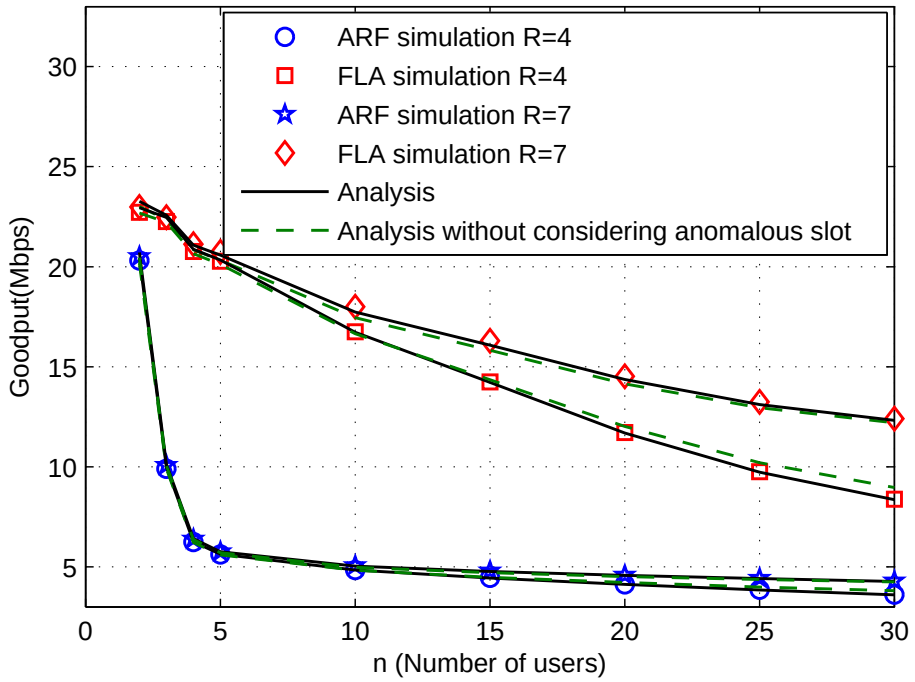
Another important system performance measure for WLAN scenarios is the transmission fairness among STAs. In this work, the fairness is measured by using the Jain's fairness index and is calculated as

$$I = \frac{\left(\sum_i^n \beta_i\right)^2}{n \sum_i^n \beta_i^2} \tag{6.13}$$

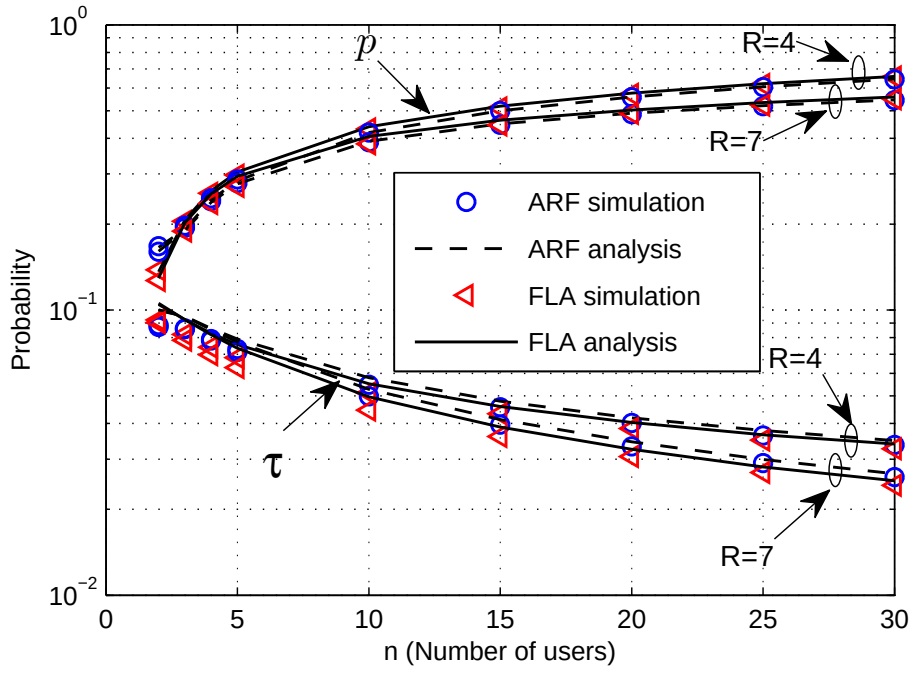
where β_i denotes the average number of transmissions for STA i. Note that $I = \frac{1}{n}$ implies an unfair system and $I = 1$ reflects a completely fair system.

RTS or CTS) and adds the HT control field required to feedback the MCS selection.

[4] Configured to $a = 4$ or $a = 6$ in order to be compared to the new model using $a = 6$ with $R = 4$ or $R = 7$, respectively.

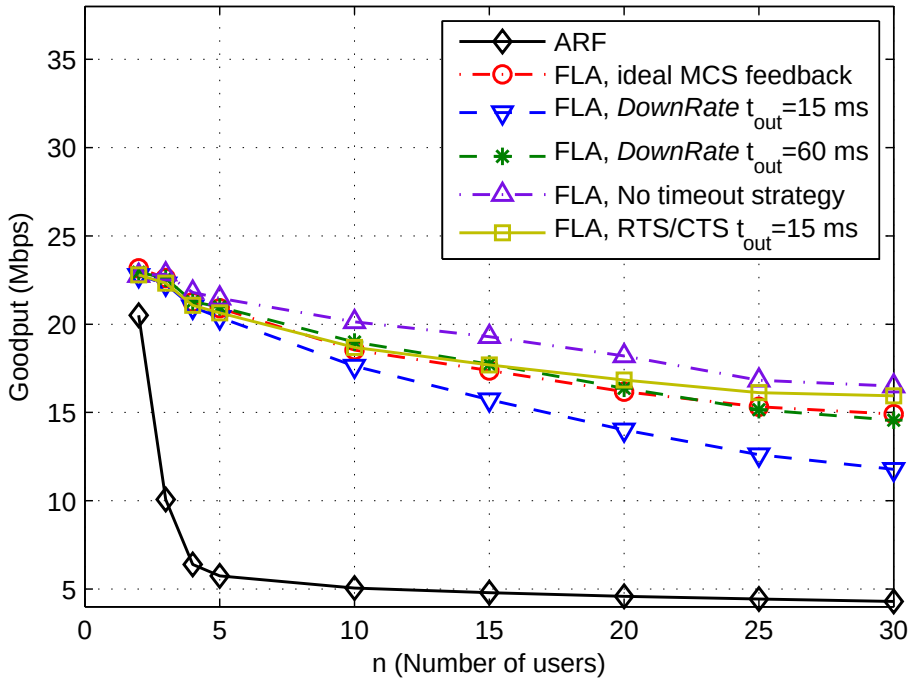


(a) System goodput.

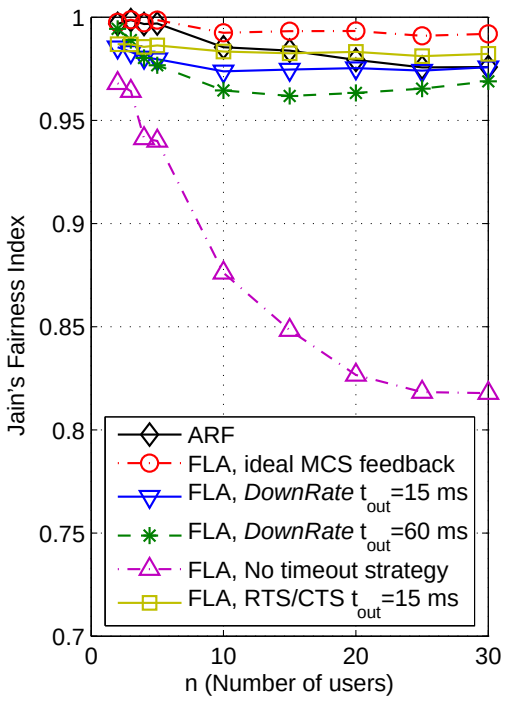


(b) τ and p.

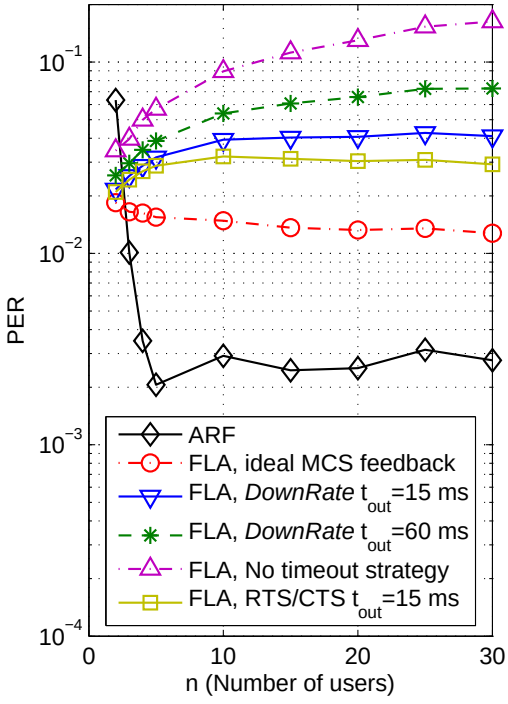
Figure 6.5: Semi-analytic and simulated system performance using $R = 4$ and $R = 7$ retransmissions.



(a) System goodput.



(b) System fairness.



(c) System PER.

Figure 6.6: Goodput, Jain's fairness index and PER of ARF and FLA strategies using $R = 7$.

Figures 6.6a, 6.6b and 6.6c, show the goodput, the fairness index and the PER[5] system performance, respectively, using the ARF and FLA algorithms described in 6.2. In addition to the proposed post-timeout strategies for FLA, an ideal mechanism that considers the existence of instantaneous feedback between receiver and sender is also presented, thus acting as a reference benchmark under the constraint of maximum system fairness. In real systems, however, it is not possible to implement such an ideal FLA-based scheme due to the 802.11n MCS feedback mechanism, which invariably introduces some delay in its transmission. Moreover, the performance of FLA without post-timeout strategy is also shown to evaluate the degradation derived from the use of outdated MCS feedback.

In Fig. 6.6a, a priori, ideal FLA could be expected to provide the maximum goodput, however it is outperformed by FLA without timeout strategy when $n > 5$. In order to understand this result it is crucial to observe (see Fig. 6.6b) that this goodput improvement is at the expense of a loss in the Jain's fairness index measured in terms of the per-STA transmission opportunity. This loss in fairness reflects that some stations are able to transmit more frequently than the others. These STAs are the ones whose channel conditions are so good (high SNR) that the MCS feedback delay has a negligible effect on their transmission error probability (see Fig. 6.7 where the probability of using each transmission rate in a successful transmission is depicted for different FLA configurations). Consequently, their contention window is mostly doubled due to collisions and rarely due to erroneous transmissions. In contrast, the other stations, which are affected by the MCS feedback delay, experience a considerably higher probability of error (see Fig. 6.6c). Accordingly, their DCF mechanism doubles their contention window after each erroneous transmission resulting in a lower transmission probability when compared with that of the STAs with good channel status. The combination of these two types of STAs leads to a system goodput higher than that of the ideal FLA because stations experiencing good channels, which use the highest transmission rates, have more chances of accessing the medium at the expense of the STAs with poor channel conditions.

Figure 6.6a shows that both FLA post-timeout strategies, *D*ownRate and RTS/CTS, achieve similar goodput performance than that of the ideal FLA. Different timeout values have been used, in particular, $t_{out} = 15\ ms$ and $t_{out} = 60\ ms$ for the RTS/CTS and the *D*ownRate strategies, respectively. In the *D*ownRate case and in accordance to the observations made in the no post-timeout strategy, the utilization of longer timeout values reduces the fairness performance, translating into a goodput improvement. Recalling that the *D*ownRate mechanism lowers the transmission rate irrespective of the channel state, it is reasonable to observe that the longer the timeout, the less rate decreases are experienced, thus maintaing the rate selection (see Fig: 6.7a and 6.7c) and increasing the PER (see Fig. 6.6c). In contrast, when small t_{out} values are employed, such as the *D*ownRate using $t_{out} = 15\ ms$ (see Fig. 6.6a), lower goodput results are obtained since the timeout expires more frequently, thus decreasing the average MCS rates (compare Fig. 6.7b and 6.7c rate selection performance) and reducing the PER probability since the transmission rates are more conservative (see Fig. 6.6c). Contrastingly, the frequent timeout expirations are adequately exploited by

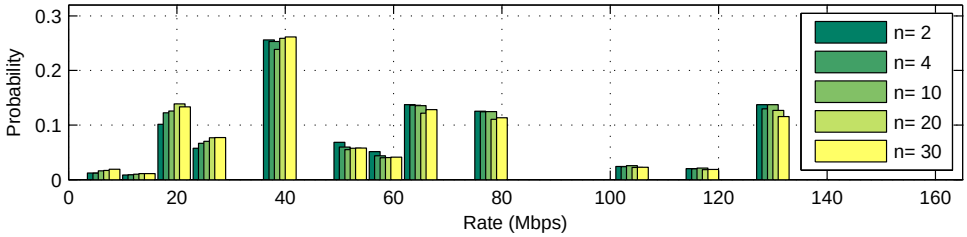
[5]The packet error rate (PER) presented in this chapter only refers to transmission errors caused by channel errors.

RTS/CTS post-timeout strategy to ameliorate the goodput performance. In this case, the utilization of RTS/CTS after the timeout expiration produces two positive effects. First, the average collision duration is reduced, since the collisions of packets using RTS/CTS mechanism only involve RTS frames. Second, the CTS frames provide the most suitable MCS to the current channel conditions, thereby maintaining an optimized transmission rate (see the similarities between Fig. 6.7a and 6.7d).

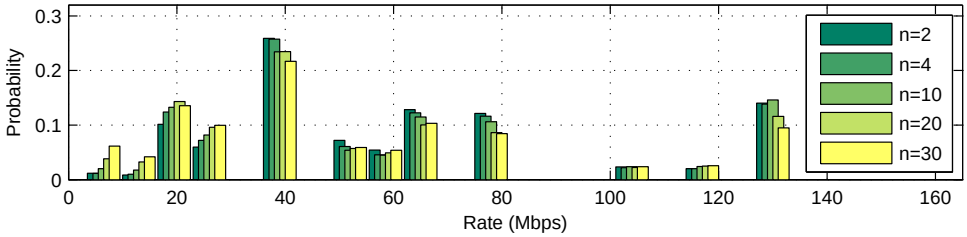
Note that the *D*ownRate mechanism using long timeout values (60 ms) achieves higher goodput performance than the one obtained using lower values than the channel coherence time (15 ms). To account for this behaviour, it should be noted that in CSMA/CA networks, the STAs involved in transmission errors (collisions or channel errors) wait on average more slots to transmit than STAs with no errors, because they are penalized by the BEB algorithm. As a consequence, the rest of STAs can take advantage of this situation by using the medium more intensively. For this reason, whenever STAs continue using stale MCSs, they leave the medium free for the rest of users until the *D*ownRate mechanism updates the MCS to an appropriate level. Note that the prize paid for this higher utilization of the medium is a huge fairness penalty as the users with a poorly selected timeout have less chances of accessing the channel. In case of setting the timeout value to the half of the channel coherence time, *D*ownRate MCS update mechanism appropriately lowers the transmission rate when required, thus reducing the error probability. Nevertheless, note that this comes at the cost of a throughput reduction as the MCS is selected in a more conservative manner. This problem is solved in the post-timeout RTS/CTS mechanism by updating the MCS feedback in the RTS frame.

Figure 6.8 shows the performance of FLA and ARF for 2 users scenario, employing a variable packet length (L_p) modelled as a doubly truncated exponential distribution between 40 and 10.000 bytes. When using FLA, it is assumed that the receiver knows the packet length to be used for the next transmission, hence permitting an accurate MCS estimation. This assumption is quite realistic since successive values of L_p (length of consecutive packets) are highly correlated between consecutive packets sent from the same STA in typical WLAN environments. Similarly to Fig. 6.8, the performance of FLA and ARF for 3, 5, 10 and 15 users is depicted in Figs. 6.9, 6.10, 6.11 and 6.12, respectively.

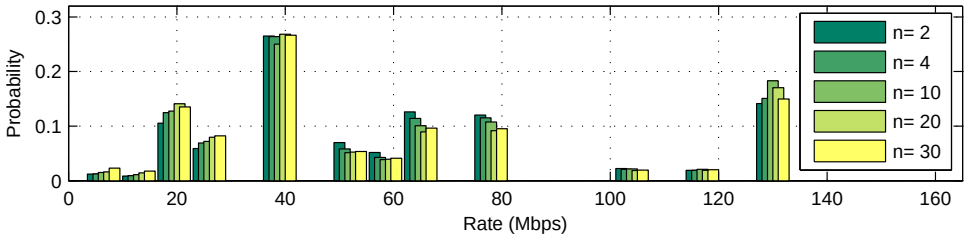
Due to the large overhead introduced by the DCF mechanism, the adoption of long $\overline{L_p}$ values improves the DCF protocol efficiency and consequently, the system goodput increases, especially for the FLA cases. Note that FLA is still outperforming ARF for any $\overline{L_p}$ and number of users, most notably for those cases where more than two users are contending for the medium. Notice that in comparison to results obtained using the other strategies, the goodput of the post-timeout RTS/CTS approach increases with $\overline{L_p}$ and n, due to the more frequent expiration of the timeout counter (see Figs. 6.11a and 6.12a). In contrast, no performance differences are observed when short $\overline{L_p}$ and/or low n values are used because of the low probability of a timeout counter expiration (see Figs. 6.8a, 6.9a and 6.10a). In other words, when the number of users is low and/or mostly short packets are transmitted, CSI information is rarely outdated, thus making the specific post-timeout strategy almost irrelevant. Notice that FLA with post-timeout RTS/CTS strategy outperforms the goodput of FLA with instantaneous feedback. This somewhat counterintuitive result is due to the use of the RTS/CTS mechanism prior to the packet transmission that leads to a collision dura-



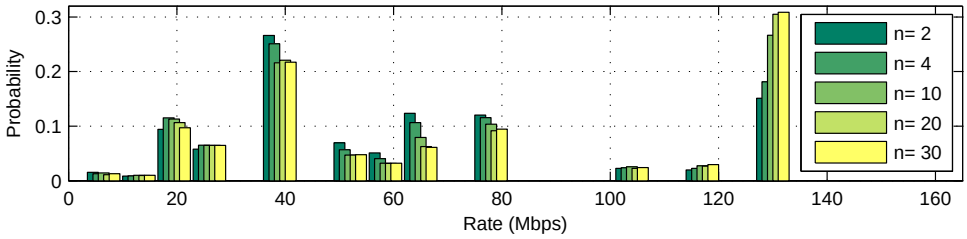
(a) FLA using non delayed MCS information.



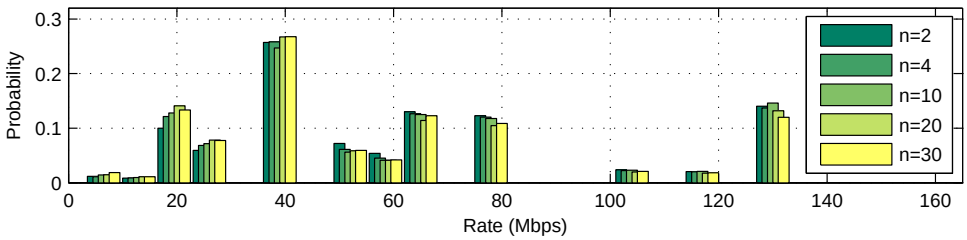
(b) FLA with DownRate strategy after t_{out} = 15 ms.



(c) FLA with DownRate strategy after t_{out} = 60 ms.

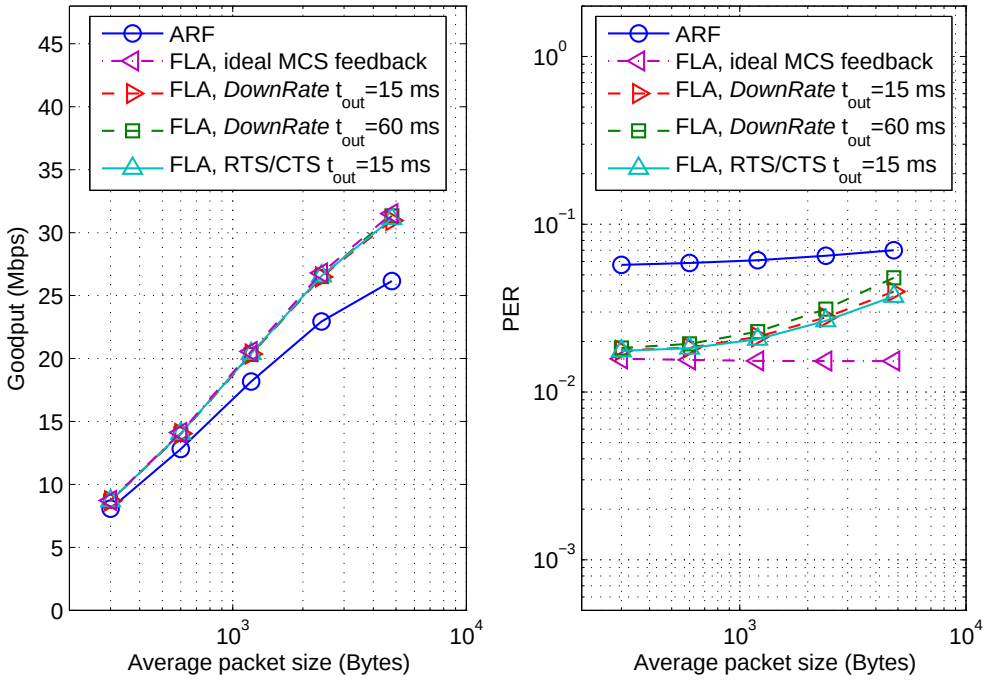


(d) FLA without post-timeout strategy.



(e) FLA with RTS/CTS after t_{out} = 15 ms

Figure 6.7: Probability of use of the different transmission rates given that the packet has been successfully received. Results are shown for different FLA settings using a system configuration with $R = 7$.



(a) System goodput.

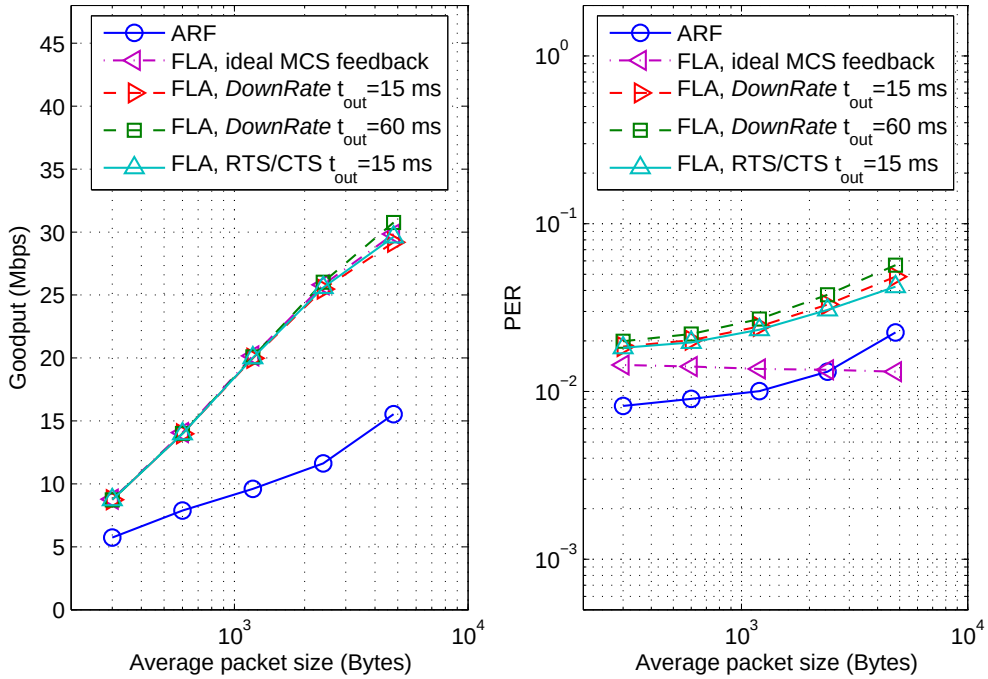
(b) System PER.

Figure 6.8: System goodput and PER performance as a function of $\overline{L_p}$ when $n = 2$.

tion reduction as some of the collisions involve RTS/CTS frames (considerably shorter than the data packets involved in the collisions under ideal feedback). Furthermore, it is worth mentioning that FLA with RTS/CTS post-timeout strategy fulfills the PER QoS constraint for all the considered configurations (see Figs. 6.8b, 6.9b, 6.10b, 6.11b and 6.12b). Moreover, although not shown in the figures, it is worth mentioning that all the studied strategies offer a high fairness index performance. Notice that the system PER performance of FLA with *D*ownRate strategy and ARF increases for long $\overline{L_p}$ as a consequence of the obvious increment of the average MCS feedback delay.

6.6.2 RTS/CTS

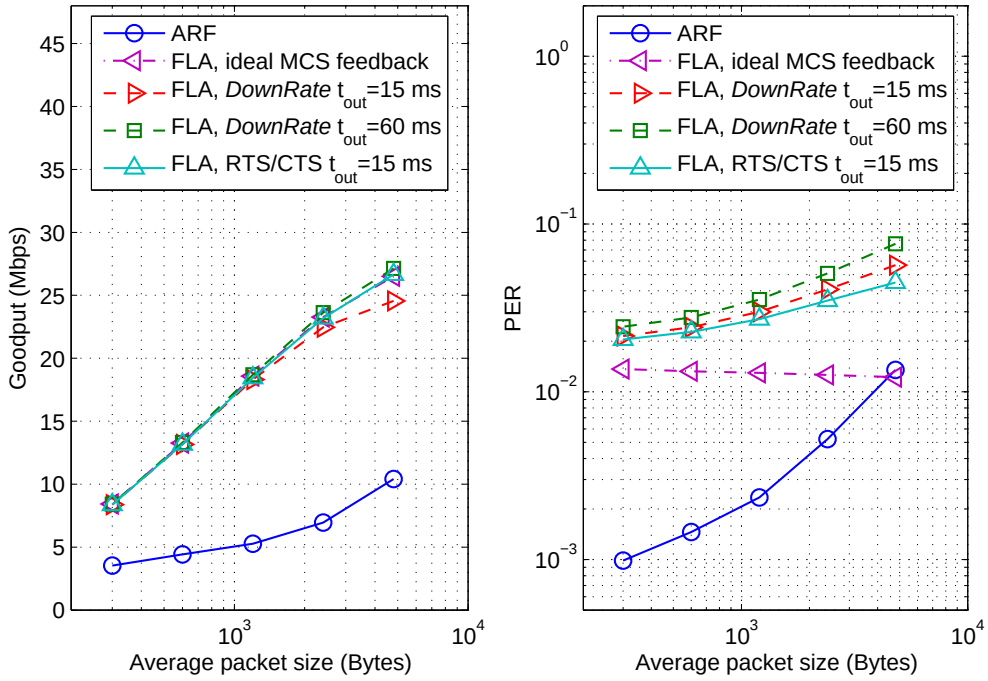
In Fig. 6.13, FLA and ARF system performance is presented for the Basic and RTS/CTS access techniques. ARF on RTS/CTS access is equivalent to the CARA-RTS algorithm defined in [18], with $P_{th} = 0$ (probing activation threshold) and $N_{th} = 1$ (consecutive failure threshold). Remarkably, the FLA algorithm used in Basic Access is the one that employs the RTS/CTS post-timeout strategy with $t_{out} = 15$ ms studied in subsection 6.6.1. Despite the use of RTS/CTS, FLA still outperforms ARF in terms of goodput, irrespective of the number of users in the system. However, the goodput improvement is significantly lower than that obtained in the Basic Access case (see Fig. 6.13a), as now ARF is capable of distinguishing the source of errors (collisions or channel errors). Note that, due to the reduced MCS feedback delay when employing RTS/CTS, FLA optimally selects the MCS and results in PER values similar to those obtained when using ideal FLA (see Figs. 6.13c and 6.6c). Remarkably, FLA with Basic Access is outper-



(a) System goodput.

(b) System PER.

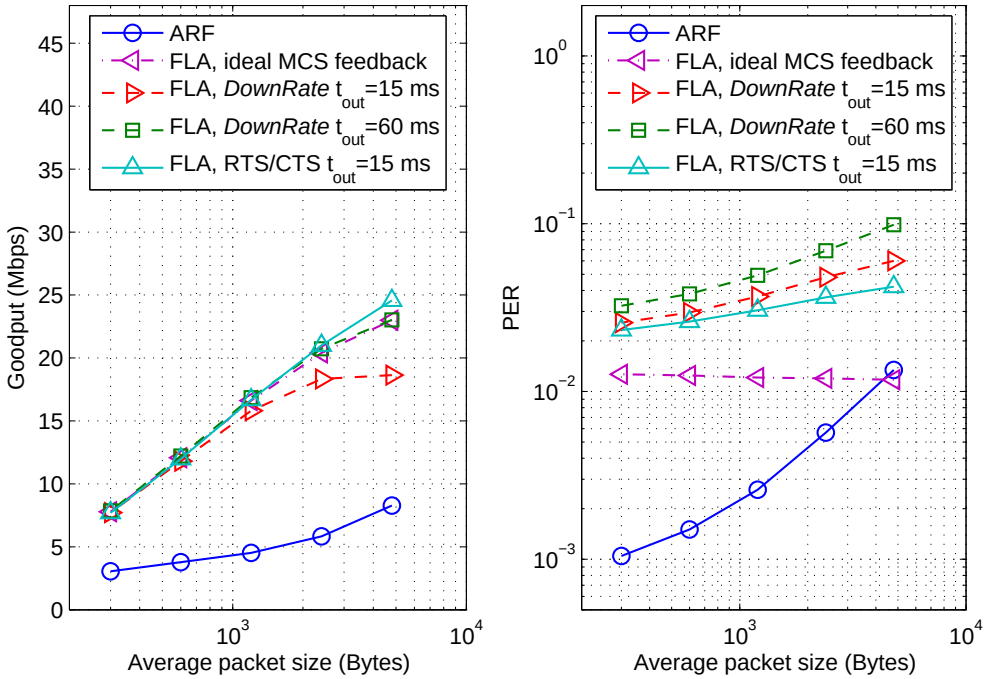
Figure 6.9: System goodput and PER performance as a function of $\overline{L_p}$ when $n = 3$.



(a) System goodput.

(b) System PER.

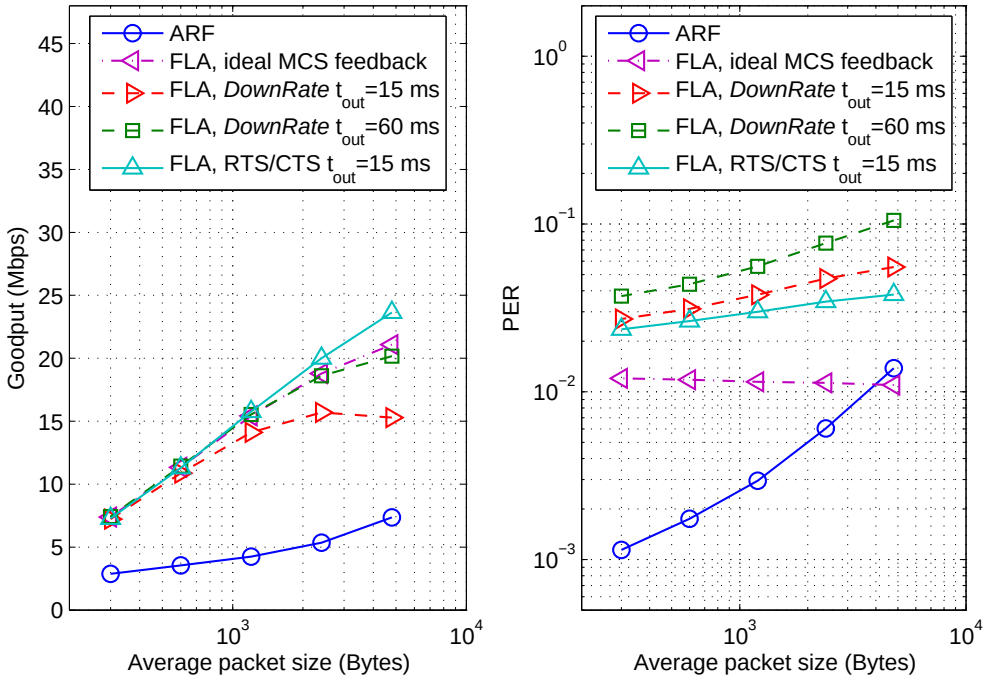
Figure 6.10: System goodput and PER performance as a function of $\overline{L_p}$ when $n = 5$.



(a) System goodput.

(b) System PER.

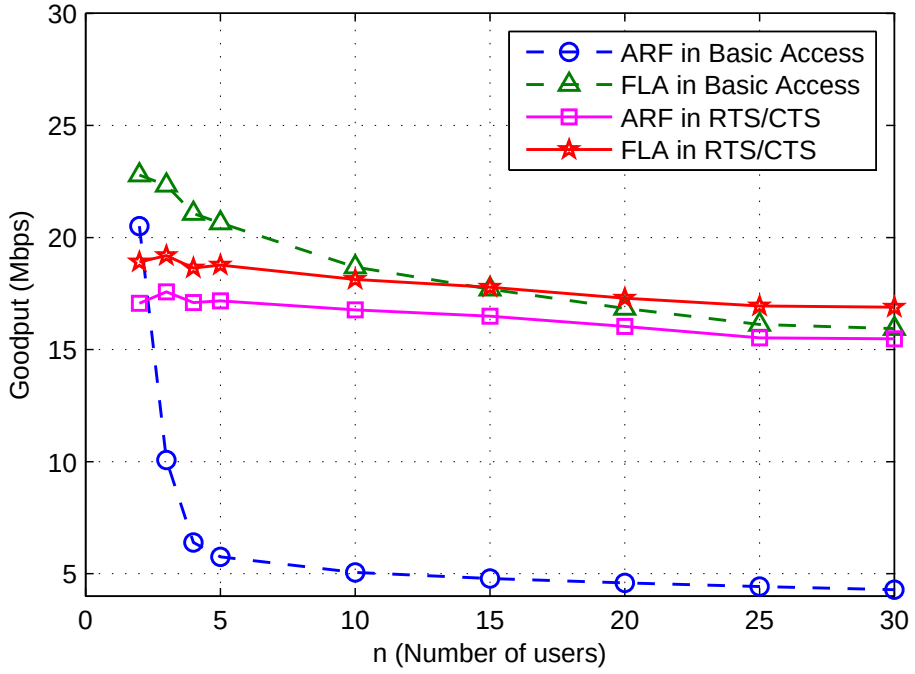
Figure 6.11: System goodput and PER performance as a function of $\overline{L_p}$ when $n = 10$.



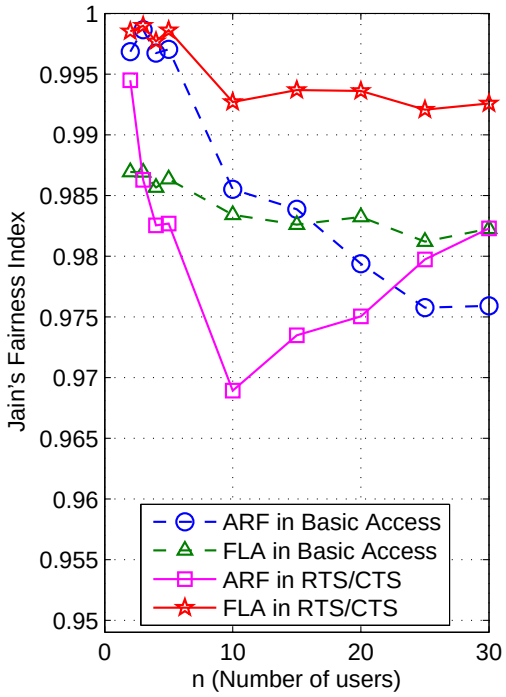
(a) System goodput.

(b) System PER.

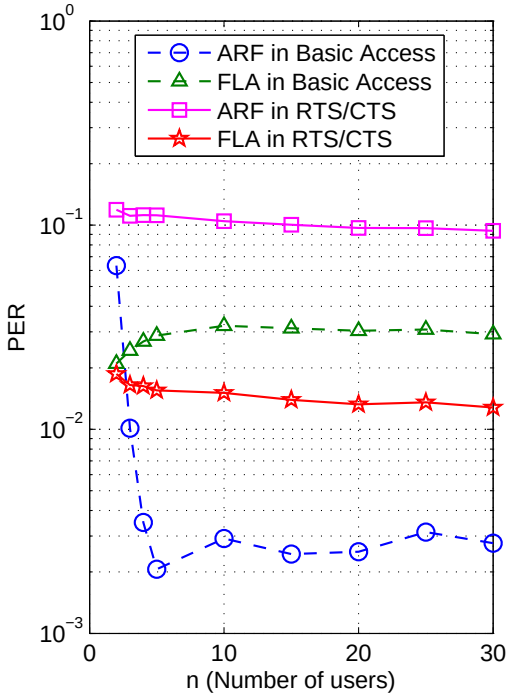
Figure 6.12: System goodput and PER performance as a function of $\overline{L_p}$ when $n = 15$.



(a) System goodput.



(b) System fairness.



(c) System PER.

Figure 6.13: FLA and ARF performance using RTS/CTS and Basic Access with fixed $L_p = 1500$ Bytes.

formed by FLA with RTS/CTS for $n > 15$, due to the better performance of RTS/CTS in dense environments with high collision probabilities. Finally, it should be mentioned that, as in Basic Access, both AMC strategies with RTS/CTS maintain a high degree of fairness irrespective of the number of users in the system (n) (see Fig. 6.13b).

Figure 6.14 shows analytical and simulation goodput performance for different n and $\overline{L_p}$ values, demonstrating the accuracy of the semi-analytical model and the clear superiority of FLA with respect to ARF irrespective of the access technique. Moreover, FLA on RTS/CTS, thanks to the reduction in the collision duration, outperforms FLA on Basic Access for long packet lengths and $n > 2$ (see Fig. 6.14a). Lastly, notice that ARF on RTS/CTS shows an improvement over ARF on Basic Access in terms of goodput for any $\overline{L_p}$ when $n > 2$, demonstrating that the overhead introduced by the RTS/CTS can be tolerated in order to provide reliable information to ARF.

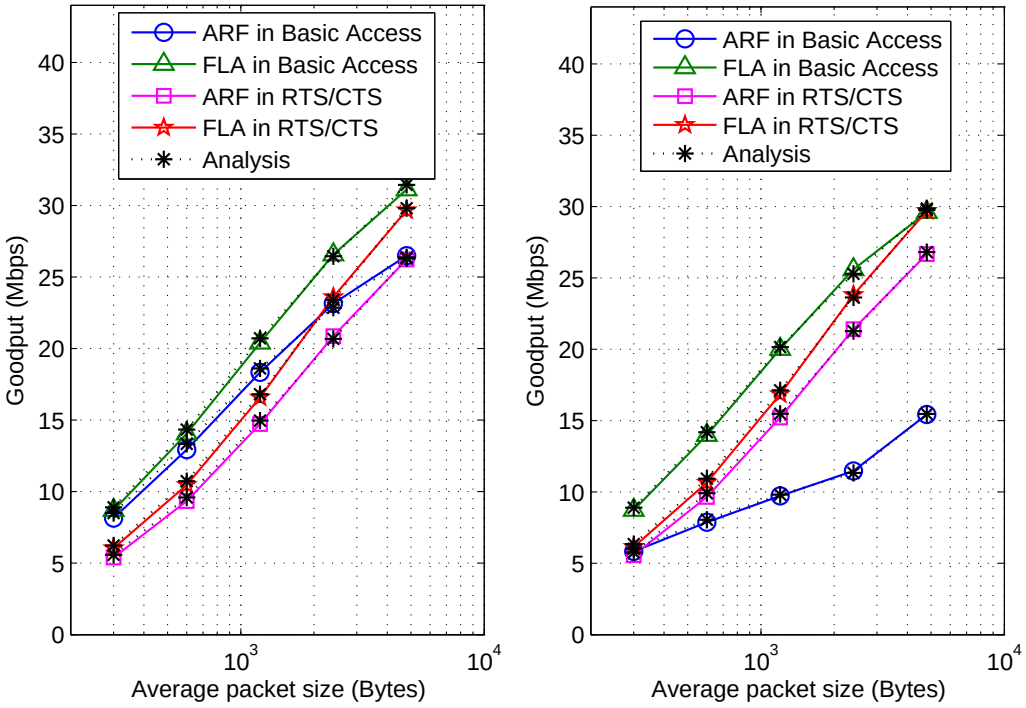
In case of a mixture of basic- and RTS/CTS-access schemes, and depending on the scenario configuration (number of users using each access scheme, packet size or user signal to noise ratio (SNR)), the throughput results will lie somewhere between the performance of Basic Access and RTS/CTS. In this case, and thanks to the DCF scheme, the system will preserve the medium access fairness between STAs. Remarkably, in this case, collisions involving users of Basic Access and RTS/CTS access scheme would appear and their duration would correspond to the longest transmission that would almost surely correspond to the Basic Access STAs.

6.6.3 FLA and ARF performance comparison in E2CA

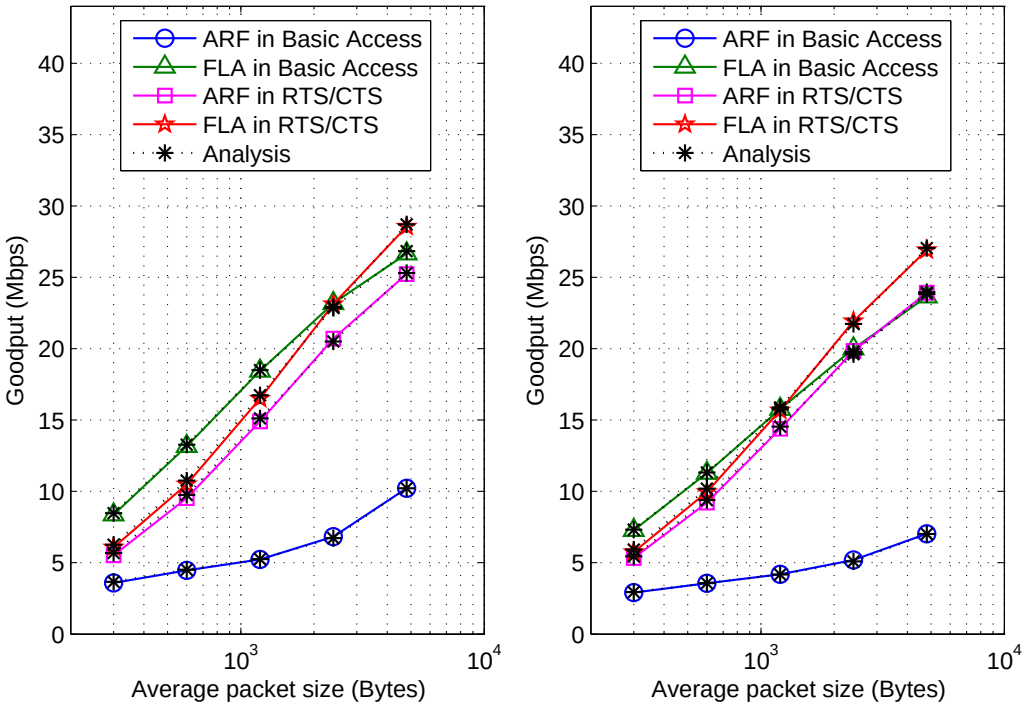
Figure 6.15 shows the probability of a busy slot $P_b^{(n)}$ for both the CSMA/CA- and CSMA-/E2CA-based networks. In CSMA/CA, $P_b^{(n)}$ increases gradually with n, outperforming CSMA/E2CA networks for $n > 8$. In contrast, the CSMA/E2CA $P_b^{(n)}$ performance is similar to that of a time division multiple access (TDMA) scheme whose frame length varies according to the number of STAs, providing n transmissions in a frame of $C = \frac{CW_{min}}{2}$ slots, where $CW_{min} = 2^{\lceil log_2(4n) \rceil}$. Consequently, $P_b^{(n)}$ improves as n increases while CW_{min} remains fixed, and decreases abruptly when the increment of n leads to the use of a higher CW_{min} value.

Looking at the probabilities of error and collision events, shown in Fig. 6.16, it can be observed that the higher the PER is in CSMA/CA networks, the larger the $P_b^{(n)}$ degradation becomes. This effect is caused by the BEB algorithm that increases the contention window of those STAs affected by erroneous transmissions, thus reducing $P_b^{(n)}$. In contrast, in CSMA/E2CA networks, similar $P_b^{(n)}$ values are obtained irrespective of the AMC PER performance (see ARF and FLA performance with CSMA/E2CA in Figs. 6.15 and 6.16b), because no access penalization is applied to those STAs with error transmissions.

Although CSMA/E2CA protocol does not always obtain the best performance in terms of $P_b^{(n)}$, it achieves an important reduction in the probability of collision $P_c^{(n)}$ when compared to the one obtained by the CSMA/CA protocol (see Fig. 6.16a and 6.16b). In fact, the collision probability is drastically reduced, approaching the collision-free operation, and the average wasted time in events that do not increase the information transfer is notably reduced. Therefore, a high performance improvement in terms of goodput should be expected in CSMA/E2CA-based networks. Furthermore,



(a) $n = 2$ (b) $n = 3$



(c) $n = 5$ (d) $n = 15$

Figure 6.14: System Goodput, as a function of n and $\overline{L_p}$, for FLA and ARF adaptation strategies when used with Basic and RTS/CTS access schemes.

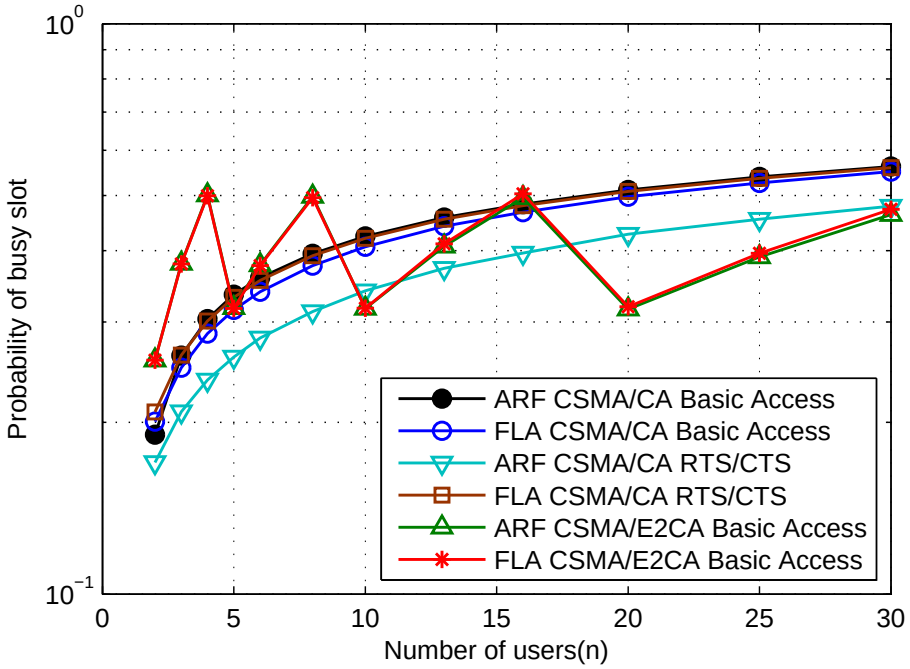
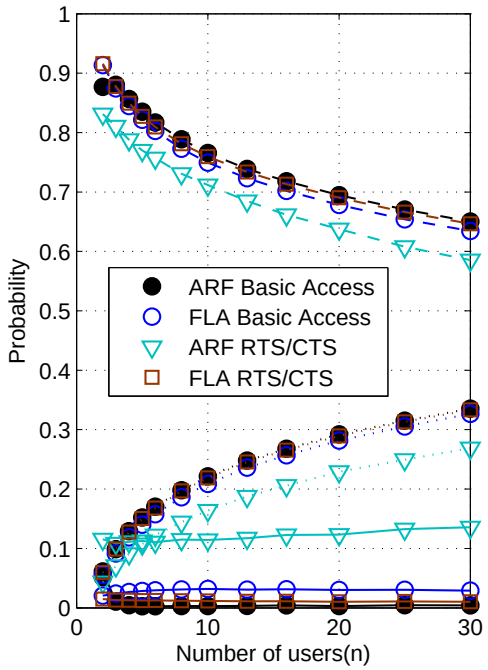


Figure 6.15: Probability of busy slot ($P_b^{(n)}$) in CSMA/CA and CSMA/E2CA networks.

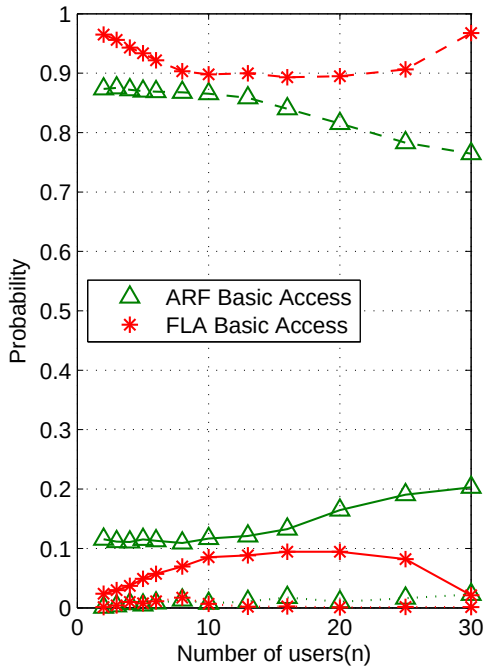
the collision avoidance allows to fix the malfunctioning problems derived from collisions, such as ARF and FLA MCS selection cases.

In Fig. 6.17, the goodput performance of both ARF and FLA AMC strategies over CSMA/CA and CSMA/E2CA protocols is presented for a fixed packet length of 1500 bytes. These results are also compared to those obtained by CSMA/E2CA semi-analytic model, thus validating the accuracy of the model for different packet lengths and AMC schemes. It is worth mentioning that the FLA algorithm used in Fig. 6.17 employs the RTS/CTS post-timeout strategy with t_{out} = 15 ms. Note that FLA on CSMA/E2CA networks obtains the best performance results, outperforming ARF under the same scenario conditions. The FLA algorithm satisfies the PER constraints because it selects the most appropriate MCS to be used according to the current channel characteristics. As it has been detailed before, the collision avoidance property of CSMA/E2CA protocol leads to a goodput improvement for both FLA and ARF strategies that outperforms the goodput obtained with any of the considered AMC schemes in CSMA/CA networks. Moreover, the goodput performance of FLA-based schemes is always better than that obtained with ARF-based strategies in all the considered scenarios (CSMA/CA with Basic Access, CSMA/CA with RTS/CTS and CSMA/E2CA with Basic Access).

Figures 6.18, 6.19 and 6.20, show the goodput and PER performance of ARF and FLA strategies when using variable packet lengths in scenarios with n = 5, 10 and 15 users, respectively. As it was explained in Subsection 6.6.1, FLA using the *D*ownRate mechanism, and assuming the use of an appropriate selection of the timeout value, satisfactorily solves the problem of delayed MCS feedback information in CSMA/CA networks. However, as it can be observed in Figs. 6.18a, 6.19a and 6.20a this timeout configuration (60 ms) causes the performance of CSMA/E2CA-based networks to



(a) CSMA/CA. Probability of success (dashes), error (lines) and collision (dots) events.



(b) CSMA/E2CA. Probability of success (dashes), error (lines) and collision (dots) events.

Figure 6.16: Event probabilities of busy slots.

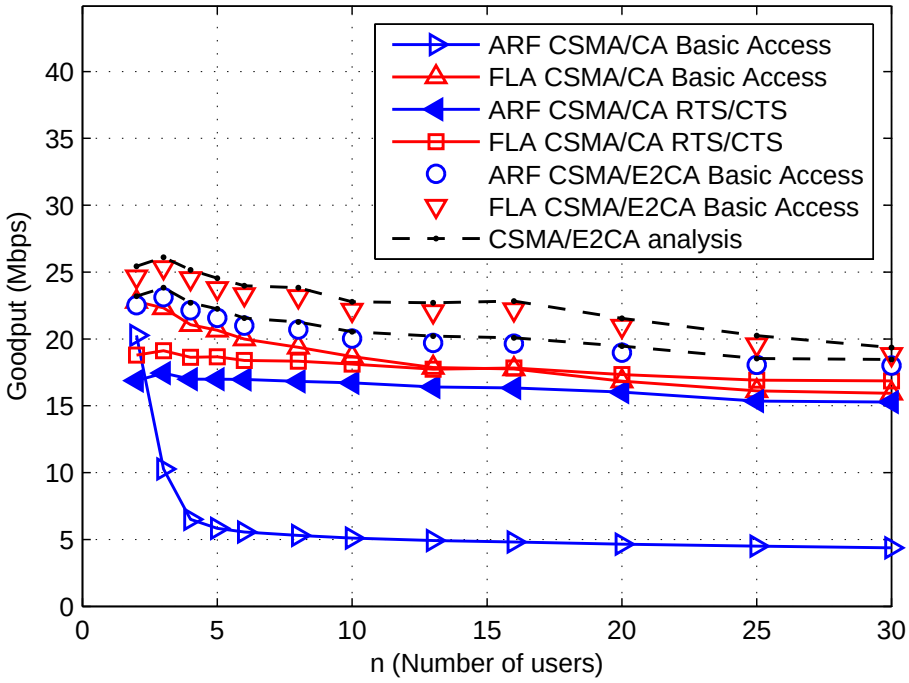
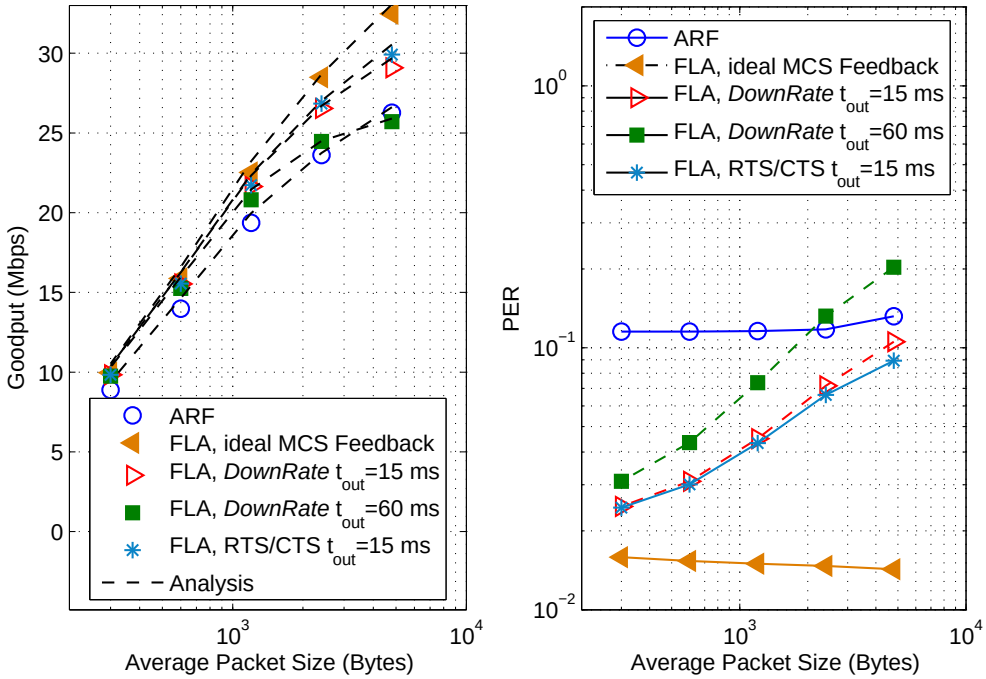


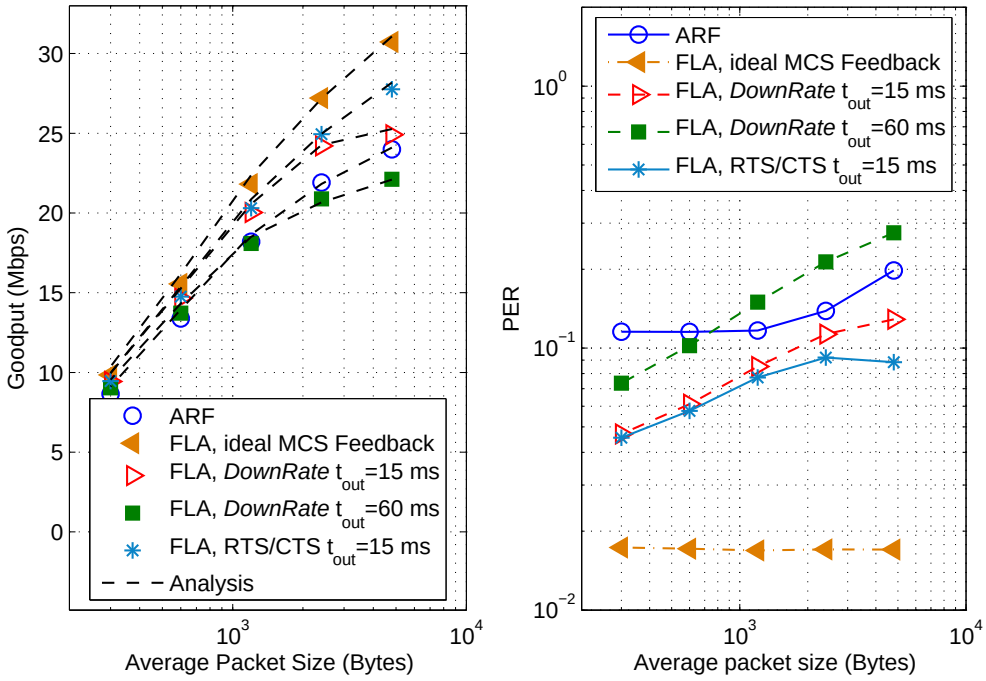
Figure 6.17: Goodput for ARF and FLA using CSMA/CA.



(a) System goodput.

(b) System PER.

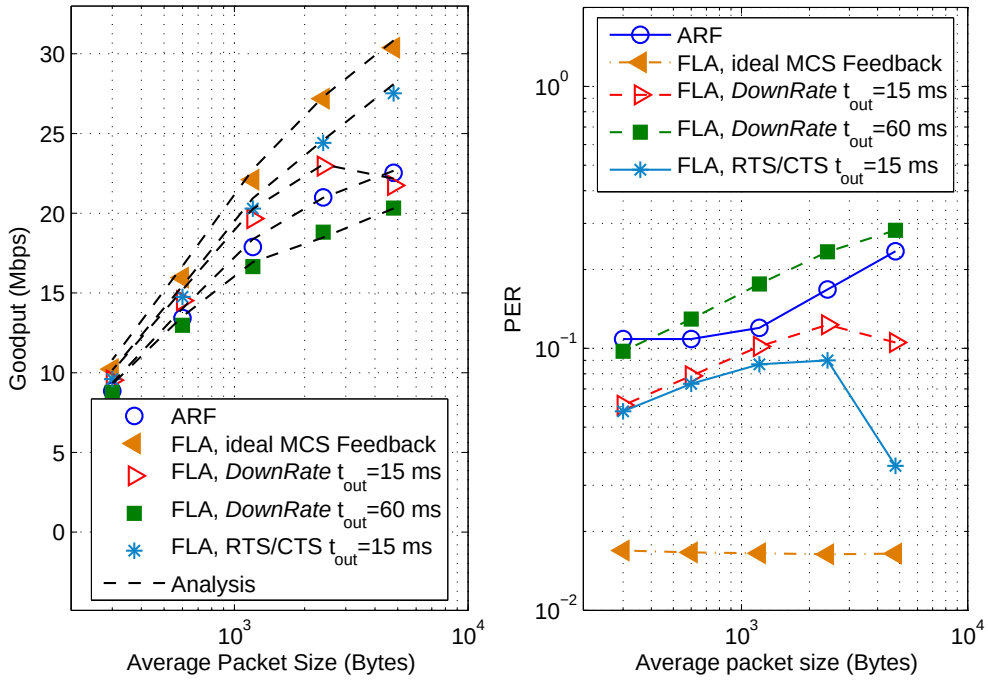
Figure 6.18: Goodput and PER system performance as a function of $\overline{L_p}$ and $n = 5$.



(a) System goodput.

(b) System PER.

Figure 6.19: Goodput and PER system performance as a function of $\overline{L_p}$ and $n = 10$.



(a) System goodput.

(b) System PER.

Figure 6.20: CSMA/E2CA Goodput and PER system performance as a function of $\overline{L_p}$ and $n = 15$.

be severely degraded. Note that in some cases FLA with *D*ownRate $t_{out} = 60$ ms is outperformed by the ARF algorithm. Although setting the timeout value to a lower time value (15 ms) partially solves this problem, the goodput is still lower than that obtained using FLA with RTS/CTS. In CSMA/E2CA networks, the error transmissions are not penalized, therefore, the transmitter must update the MCS as soon as it knows that the feedback MCS is stale. Even after applying the timeout reconfiguration, the performance of *D*ownRate mechanism does not fulfill the PER constraints and it is outperformed by FLA with RTS/CTS post-timeout strategy. Note that only FLA with ideal MCS feedback and FLA with RTS/CTS post-timeout strategies are able to fulfil the system PER constraint of 0.1.

6.7 Chapter Summary

This chapter has presented a semi-analytical framework for the performance modelling of MIMO-OFDM WLANs when using the Basic Access or RTS/CTS with CSMA/CA at the MAC layer. Unlike previous works, the proposed model is able to incorporate the effects of channel errors, the possibility of using open- or closed-loop transmission mode adaptation, the effect of the retry limit at the MAC layer and the use of outdated MCS feedback information. A complete study of FLA over 802.11n PHY/MAC in terms of goodput, fairness and system PER performance for a wide range of number of users and packet sizes has been presented and contrasted with those obtained using

ARF. Noteworthy, the influence of feedback delay on the performance of FLA has been assessed. In Basic Access scheme, the degradation caused by an outdated MCS information can be largely compensated with the use of a time-out strategy that weighs down the influence of the MCS feedback delay. This post-timeout strategy uses the RTS/CTS mechanism when the MCS feedback information is stale, thus updating the MCS prior to its use in the next packet transmission. Numerical results clearly show that as the number of users in the system grows, the FLA-based adaptation proves to be much more robust to collisions than ARF even when employing outdated MCS feedback information. This effect is clearly demonstrated by the fact that whereas ARF-based schemes suffer a dramatic reduction in goodput for more than two users, the FLA-based strategy exhibits a very graceful degradation thanks to a more accurate rate selection in the presence of collisions.

When using the CSMA/CA with RTS/CTS access scheme, FLA still beats ARF irrespective of the number of users and packet sizes, although the difference is not as significant as in Basic Access. Nevertheless, it should be remarked that the RTS/CTS frame exchange allows a delay-free selection of the MCS and results in a system PER similar to the one obtained in the ideal FLA case.

Lastly, the performance of FLA has been extensively evaluated on CSMA/E2CA networks observing significant goodput improvements over its CSMA/CA-based counterparts. It has been shown that this improvement is basically due to the collision reduction provided by the CSMA/E2CA protocol. Moreover, it has been confirmed that FLA outperforms the ARF algorithm even when operating with delayed MCS information. In parallel, a novel CSMA/E2CA semi-analytical modelling approach has been introduced and thoroughly validated. Overall, it can be concluded that FLA yields a goodput that significantly outperforms ARF for most system loads and access schemes, while keeping a large degree of fairness and satisfying prescribed PER-based QoS constraints.

CHAPTER 7

MAC SUBLAYER PERFORMANCE UNDER NON-SATURATED CONDITIONS

This chapter proposes a novel analytical model to evaluate the system performance of the IEEE 802.11 MAC sublayer under both light and heavy traffic loads. This approach, based on a three-dimensional Markov chain, generalizes the bi-dimensional Markov chain models introduced in previous chapters. Traditional bi-dimensional Markov chain models, while suitable for throughput analysis, are unable to capture QoS performance metrics such as the average packet delay, the average queue length and the blocking and discard probabilities due to the lack of a proper queueing model. We present a refined three-dimensional Markov chain approach that incorporates a queueing model and error prone channel effects. The Markov chain is efficiently solved by using our refined Collapsed Transition onto Basis approach, allowing the investigation of the whole set of QoS metrics for an IEEE 802.11 network under realistic conditions. The model is validated by thoroughly studying its performance for different system and scenario configurations and contrasting these results with extensive numerical simulations as well as with results obtained using other analytical methods that have been proposed in the literature.

7.1 Motivation and Problem Statement

WLANs based on the different flavors of the IEEE 802.11 series of standards have experienced a growing demand during recent years. Fueled by their widespread deployment, research efforts have been intensified in order to gain a better understanding of its MAC layer. Particularly, although the IEEE 802.11 MAC layer can also be based on the Point Coordination Function (PCF), a major focus has been put on the Distributed Coordination Function (DCF).

The basic access mechanism of DCF is based on the CSMA/CA protocol, which makes use of BEB to share the medium between competing stations. As a random access protocol, BEB-based CSMA/CA inherits the advantages of minimum coordi-

nation and distributed control; unfortunately, this random nature also complicates the derivation of suitable models and their corresponding performance evaluation. In fact, as shared collisions couple the service processes of stations competing for access, the performance evaluation of this multiple access protocol cannot be analytically investigated without resorting to judicious simplifying approximations.

As stated in Chapter 5, one of the most widely adopted modeling paradigms is the mean-field Markov model approach. It has its foundations on Bianchis's seminal papers [20, 13], where a two-dimensional Markov chain model was established to characterize the basic access mechanism of DCF. Bianchi's approach is based on the following simplifying assumptions:

1. The collision probability of a station attempting transmission is constant regardless of the network collision history or station considered (collision decoupling assumption).

2. Each station in the network always has packets ready for transmission (saturation assumption), thus enabling queueing dynamics to be ignored.

3. Packet errors caused by noise on the wireless medium can be ignored and, hence, transmission errors are only a result of collisions between packets sent by contending stations (clean channel assumption).

4. All stations sense the medium's state, busy or idle, correctly (ideal sensing assumption), thus getting around the difficulty of analyzing the so called hidden and exposed node problems.

Under these assumptions, this model is well supported by simulation results, and has proven to be a powerful, yet simple, analytical tool to evaluate the saturation throughput performance of DCF-based 802.11 networks.

Bianchi's basic approach has been adopted and refined by a series of follow-up models to include more practical assumptions and expand on its original range of applicability. A selection of these follow-up papers, relevant to our research, includes: [21, 22, 23], which present an improved analytical framework thoroughly accounting for the backoff freezing details of the DCF specifications; [24, 25, 26, 27, 28], which analyze the DCF for an error-prone channel; [29, 30, 31], which treat unsaturated bufferless stations; and [32, 33, 34], which consider the impact of unsaturated stations with infinite buffers.

Bufferless models [29, 30, 31] overestimate system throughput and underestimate collision probability. On the other hand, the use of infinite buffer models [20, 13, 32, 33, 34], forbids the analytical derivation of practical QoS performance measures such as the blocking probability or the queueing delay. In order to capture these QoS performance metrics, Liu *et al.* [35] proposed a finite buffer three-dimensional Markov chain model integrating contention resolution procedures and queueing processes. The extra third dimension, used to model the queue dynamics, allows the calculation of the throughput, transmission failure probability, collision probability, queue blocking probability, queue length, and system delay, without resorting to the obtention of an explicit service time distribution. A crucial aspect of Liu's three-dimensional model is the proposal of the so called Collapsed Transition onto Basis (CTB) approach to simplify the Markov chain analysis. The main drawback of this analytical model, however,

is that it is based on Bianchi's original approach and, thus, it does not take into account neither the refinements dealing with the freezing of backoff counter nor those considering the transmission over an error-prone channel.

Our goal in this chapter is to provide a substantial extension of the three-dimensional model proposed by Liu *et al.* in [35] by incorporating, on one hand, the freezing rules of backoff counters when the broadcast channel is sensed busy and, on the other hand, the error-prone channel conditions. As far as we know, our proposal is the first one dealing with the use of three-dimensional Markov models to face this complex mathematical problem. We address the issue of computational complexity by devising a computationally efficient method, the refined CTB approach, to solve for the collision and transmission probabilities, and the steady-state transition probabilities of the 3-D Markov chain. The complete Markov chain is collapsed onto an equivalent Markov chain basis, whose dimensions allow the application of the standard power method. The solution for the collapsed Markov chain can then be exploited to solve the original complete Markov chain, and to obtain analytical expressions for the whole set of QoS performance metrics of the system. In sharp contrast to analytical results obtained using 2-D Markov chain models [20, 13, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34], the analytical results obtained with the proposed refined 3-D Markov chain allow the evaluation/prediction of the impact of buffer size, block error probability and backoff freezing rules on all relevant QoS metrics of a DCF-based IEEE 802.11 network. Extensive simulations are used to validate the accuracy of the proposed approach.

The rest of this chapter is organized as follows. Sect. 7.2 presents the non-saturation analytical model based on a three-dimensional model. In Sect. 7.3, the analytical expressions to obtain the goodput, the average interface queue length, the blocking probability, the discard probability and the average packet delay at the MAC sublayer are provided. Analytical and simulation results are presented in Sect. 7.4 in order to validate the proposed refined three-dimensional model, and they are also compared with results obtained using previous analytical models. Finally, the conclusions of this chapter are presented in Sect. 7.5.

7.2 Analytical model

7.2.1 Model time scale

In his seminal work, Bianchi [13] proposed a Markov chain model to analyze the IEEE 802.11 DCF mechanism under saturated conditions. This analytical model implicitly relied on the assumption that the backoff counter should be decremented at the beginning of a slot time, regardless of whether the slot was idle or contained a frame transmission (either successful or erroneous). This was the same assumption made by Liu *et al.* [35] to extend Bianchi's model to incorporate non-saturated conditions. However, as was recognized by Tinnirello *et al.* in [23], the IEEE 802.11 standard specifies that the backoff procedure must be suspended if the medium is determined to be busy at any time during a backoff slot, revealing that a more conforming modeling assumption is to decrement the backoff counter at the end of a slot time. The immediate consequence of this interpretation is that the backoff counter is only decremented during idle slots. Hence, the slots right after a channel busy period are *anomalous* slots

since the corresponding probabilities of transmission and collision are different from those in other slots. As extensively explained in [23, Sect. III], two main slot anomalies may occur:

- A slot time immediately following a successful transmission can only be used by the STA that is currently transmitting if it selects a random backoff counter equal to 0. Obviously, such a transmission can only be corrupted by non-ideal channel conditions as no collision may occur.
- The extra slot after the end of an EIFS time following an erroneous transmission will not be used by any STA.

In order to accomodate the backoff freezing and the anomalous slots following a busy period, Tinnirello *et al.* [23] modified the model of Bianchi [13] by considering that the system time evolution is only triggered when a nontransmitting STA decrements its backoff counter. In this case, extending [13] and [23] with the assumption of an error-prone channel and a finite buffer and load, a *model slot time* can correspond to different periods of time:

- An idle backoff slot. This situation may arise if all active STAs in the system are either Idle (without any packet in the queue) or decrementing their backoff counters.
- A time interval including a failed transmission (either due to collision or a channel error), followed by an EIFS interval plus an extra backoff slot.
- A time interval including one or more consecutive successful transmissions and a final extra backoff slot. Notice that in the first successful transmission all active STAs in the system contend for the channel, while in the successive transmissions only the transmitting STA is granted access to the channel.
- A time interval including one or more consecutive successful transmissions. This case arises whenever the last transmission empties the queue and the backoff algorithm selects a random backoff counter equal to 0.
- A time interval including one successful transmission (contending with all other STAs in the system), zero or more successful transmissions (in an error-prone channel without contendings STAs), and an erroneous transmission (due to channel errors), followed by an EIFS interval plus an extra backoff slot.

The definition of the model slot time allows maintaining a synchronized backoff countdown among all contending STAs in the system. Nevertheless, as it will be shown in the next subsection, due to the assumption of a finite buffer and load, it will be necessary to dig into the anomalous slots, in which the STA state evolution is obviously triggered when the transmitting STA modifies its queue length, changes its backoff stage and asynchronously modifies its backoff counter with respect to other STAs in the system. The resulting Markov chain model will thus present a two-level hierarchical structure, with the low-level intra-anomalous slot states, characterized by asynchronous transitions, and the high-level states, characterized by their model slot time transition synchronicity.

7.2.2 Hierarchical three-dimensional Markov chain model

In the analytical model proposed by Bianchi [13], and further refined by Tinnirello *et al.* [23] and Chen [28], the contention process amongst all the STAs was described by considering as state parameters the backoff counter $b(t)$, representing the random delay applied to a packet transmission after gaining access to the channel, and the backoff stage $s(t)$, representing the number of retransmissions already suffered by the head-of-line (HOL) frame. The bi-dimensional process $\{s(t), b(t)\}$ resulted in a discrete-time bi-dimensional Markov chain, where the state of each STA was described by a 2-tuple (i,k). The index $i \in \{0,\ldots,R\}$, with R denoting the maximum number of allowed retransmissions, was used to model the backoff stage, and the index $k \in \{0,\ldots,W_i-1\}$ was used to model the backoff counter.

In contrast to the *traditional* bi-dimensional model, and based on the improvements proposed by Liu *et al.* in [35], a new random process is added to $s(t)$ and $b(t)$, denoted by $q(t)$, that represents the queue occupancy. In this way, the three-dimensional random process $\{q(t), s(t), b(t)\}$ results in a discrete-time three-dimensional Markov chain, shown in Fig. 7.1, where the state of each STA can be described by a 3-tuple (h,i,k). The index $h \in \{0,\ldots,L\}$, with L denoting the maximum queue length, is used to model the number of packets in the queue. The parameters p and p_e denote the probabilities of unsuccessful transmission in contending and non-contending time slots, respectively. Furthermore, q_{T_s}, q_{T_e}, q and q_δ are the probabilities that a packet arrives while in a successful transmission, a failed transmission, a backoff state or an *Idle* time slot, respectively.

The three-dimensional Markov chain has been divided into a set of L parallel planes. Each plane and its states have been represented using the same color, and plane $h \in \{1,\ldots,L\}$ contains all possible states corresponding to situations in which the number of packets in the queue is equal to h.

The top row of the three-dimensional Markov chain represents the post-backoff and idle states. The next L rows of states indexed $(h,0,0)$ to $(h,0,W_0-1)$ represent the stage-0 backoff when there are h packets in the queue, with $h \in \{1,\ldots,L\}$. Depending on the event that triggers the transition to any of these states, either a successful transmission or a packet drop due to exceeding the maximum number of allowed retransmissions, the backoff counter values can be drawn in the ranges $[0,W_0-2]$ or $[0,W_0-1]$, respectively. The transition to these states after a successful transmission or a packet drop will be indicated as a transition to stage 0^+ or 0^-, respectively. With a convenient slight abuse of notation, the transitions from the *Idle* state to states $(1,0,k)$, with $k \in \{0,\ldots,W_0-1\}$, will also be indicated as a transition to stage 0^-. The subsequent group of L rows of states indexed $(h,1,0)$ to $(h,1,W_1-1)$ correspond to the stage-1 backoff states when there are $h \in \{1,\ldots,L\}$ packets in the queue. The transition to any of these states can be triggered either by a collision when being in stage-0 backoff or by an erroneous transmission at the end of an anomalous slot. In both cases the backoff counter values can be drawn in the range $[0,W_1-1]$. The next $R-1$ groups of L rows of states indexed $(h,i,0)$ to (h,i,W_i-1), with $h \in \{1,\ldots,L\}$ and $i \in \{2,\ldots,R\}$, represent the stage-i backoff when there are h packets in the queue. In contrast to the other backoff stages, the transition to any state belonging to the stage-i backoff, $i \in \{2,\ldots,R\}$, can only be triggered by a collision when being in backoff stage-$(i-1)$, and the backoff counter values must be drawn in the range $[0,W_i-1]$.

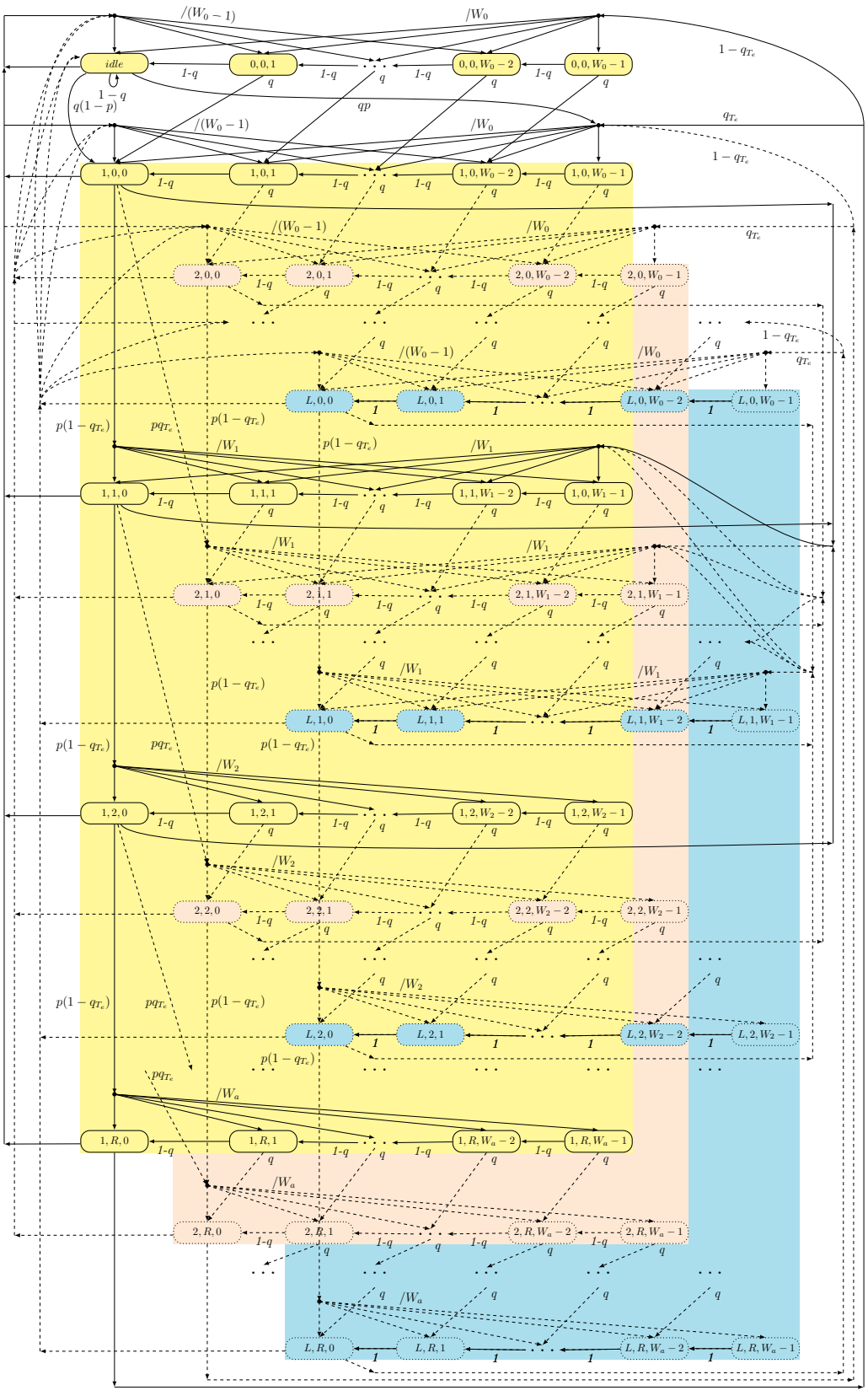


Figure 7.1: Three-dimensional Markov chain model.

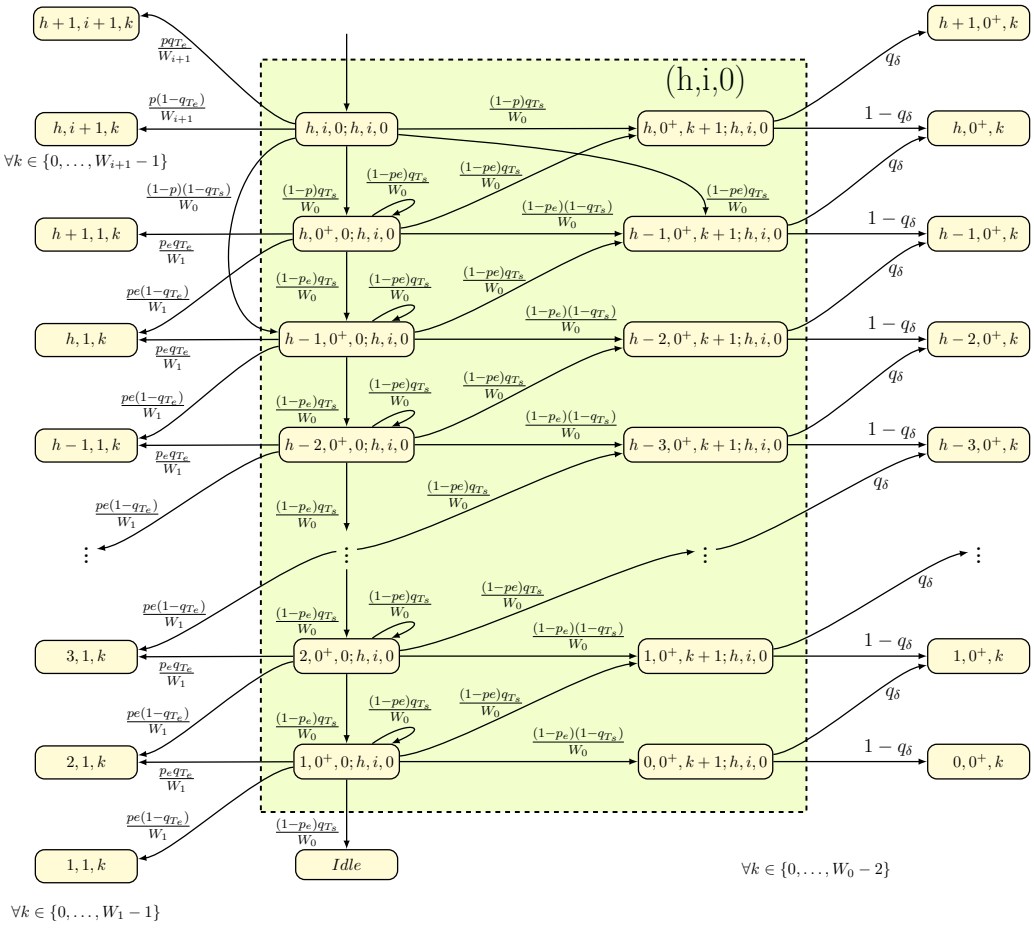


Figure 7.2: Markov chain model for a generic anomalous slot.

It is important to notice, however, that when entering a transmission model slot $(h,i,0)$, and depending on the behavior of the error-prone channel and the evolution of the packet arrival process, the STA can transition to special states denoted $(h',0,0;h,i,0)$, with $h' \in \{0,\ldots,h\}$, before leaving the anomalous slot. As shown in Fig. 7.2, states $(h',0,0;h,i,0)$ can only be accessed from state $(h,i,0;h,i,0)$ and thus, their steady-state probabilities can readily be obtained from that of state $(h,i,0)$, in a hierarchical fashion, after a proper normalization. As it will be shown in Section 7.3, there are important QoS performance measures like the average queue length or the average packet delay that can only be obtained from the steady-state probabilities of these intra-anomalous slot states.

7.2.3 Characterizing the low-level transition and steady-state probabilities

Transitions from transmission states $(h,i,0;h,i,0)$

Let us assume that the STA is in state $(h,i,0;h,i,0)$, with $h \in \{1,\ldots,L-1\}$ and $i \in \{0,\ldots,R-1\}$. As stated in Chapter 5, when transmitting the first MPDU after entering an anoma-

lous slot, all active STAs in the system are contending for the medium. In this case, if the transmission of the HOL packet fails, with probability p, the STA will move to state $(h, i+1, k)$, with $k \in \{0, \ldots, W_{i+1} - 1\}$, if the source does not generate a new packet during this time slot. If, instead, a packet arrival event occurs, the STA will move to state $(h+1, i+1, k)$. Hence, the corresponding one-step transition probabilities are

$$P\{h+1, i+1, k | h, i, 0; h, i, 0\} = \frac{p q_{T_e}}{W_{i+1}}$$
$$P\{h, i+1, k | h, i, 0; h, i, 0\} = \frac{p(1 - q_{T_e})}{W_{i+1}}.$$

For the special case in which $h = L$, if the transmission of the HOL packet fails, the STA will move to state $(L, i+1, k)$, with $k \in \{0, \ldots, W_{i+1} - 1\}$, either if a packet arrival event occurs or not. Thus,

$$P\{L, i+1, k | L, i, 0; L, i, 0\} = \frac{p}{W_{i+1}}.$$

Notice, however, that a packet arrival event will result in a buffer overflow. The queue blocking probability can be characterized from the steady-state probabilities of states with $h = L$ and, as stated by Liu *et al.* [35], this QoS performance measure could not be obtained with bufferless of infinite buffer models such as those proposed in [13, 102, 29, 33, 31, 103, 23, 28].

If the STA is in state $(h, R, 0; h, R, 0)$, with $h \in \{1, \ldots, L\}$ and the transmission of the HOL packet fails, with probability p, the packet will be dropped due to exceeding the maximum number of allowed retransmissions. Furthermore, if during this time slot the source does not generate a new packet, the STA will move to one of the states $(h-1, 0^-, k)$, with $k \in \{0, \ldots, W_0 - 1\}$, with uniform probability. If, instead, a packet arrival event occurs, the STA will move to one of the states $(h, 0^-, k)$. Hence, the corresponding one-step transition probabilities are

$$P\{h, 0^-, k | h, R, 0; h, R, 0\} = \frac{p q_{T_e}}{W_0}$$
$$P\{h-1, 0^-, k | h, R, 0; h, R, 0\} = \frac{p(1 - q_{T_e})}{W_0}.$$

If the transmission of the HOL packet succeeds, with probability $(1 - p)$, the STA will move to state $(h-1, 0, k; h, i, 0)$, irrespective of the value of i, if a random backoff counter $k \in \{0, \ldots, W_0 - 1\}$ is selected, with probability $1/W_0$, and the source does not generate a new packet during this time slot. If, instead, a packet arrival event occurs, the STA will move to state $(h, 0, k; h, i, 0)$. Hence, the corresponding one-step transition probabilities are

$$P\{h, 0^+, k; h, i, 0 | h, i, 0; h, i, 0\} = \frac{(1-p) q_{T_s}}{W_0}$$
$$P\{h-1, 0^+, k; h, i, 0 | h, i, 0; h, i, 0\} = \frac{(1-p)(1 - q_{T_s})}{W_0}.$$

Obviously, for the special case in which $h = 1$ the STA will move to the *Idle* state, irrespective of the value of i, if a random backoff counter $k = 0$ is selected and the source does not generate a new packet. That is,

$$P\{Idle | 1, i, 0; 1, i, 0\} = \frac{(1-p)(1 - q_{T_s})}{W_0}.$$

Transition probabilities from states $(h',0^+,0;h,i,0)$

After succeeding in the first transmission of an anomalous slot, the transmitting STA gains access to the channel and a non-contending period begins. If sometime in this period the STA enters the state $(h',0^+,0;h,i,0)$, with $h' \in \{1,\ldots,h\}$, the STA will transition to state $(h',0^+,k;h,i,0)$ if the packet transmission succeeds, with probability $(1-p_e)$, a packet arrival event occurs, and the random backoff counter generator is set to a value $k \in \{0,\ldots,W_0-1\}$, with probability $1/W_0$. On the contrary, if the source does not generate a new packet, then the STA will transition to state $(h'-1,0^+,k;h,i,0)$. Thus,

$$P\{h',0^+,k;h,i,0|h',0^+,0;h,i,0\} = \frac{(1-p_e)q_{T_s}}{W_0}$$
$$P\{h'-1,0^+,k;h,i,0|h',0^+,0;h,i,0\} = \frac{(1-p_e)(1-q_{T_s})}{W_0}.$$

For the special case in which $h'=1$ the STA will move to the *Idle* state, irrespective of the value of i, if a random backoff counter $k=0$ is selected and the source does not generate a new packet. That is,

$$P\{Idle|1,0^+,0;h,i,0\} = \frac{(1-p)(1-q_{T_s})}{W_0}.$$

If the packet transmission fails, with probability p_e, and a packet arrival event occurs, with probability q_{T_e}, the STA will transition to state $(h'+1,1,k)$, if the random backoff counter generator is set to a value $k \in \{0,\ldots,W_1-1\}$, with probability $1/W_1$. If, instead, the source does not generate a new packet, then the STA will transition to state $(h',1,k)$. Thus,

$$P\{h'+1,1,k|h',0^+,0;h,i,0\} = \frac{p_e q_{T_e}}{W_1}$$
$$P\{h',1,k|h',0^+,0;h,i,0\} = \frac{p_e(1-q_{T_e})}{W_1}.$$

For the special case in which $h=h'=L$, the STA will move to state $(L,1,k)$, with $k \in \{0,\ldots,W_1-1\}$, either if a packet arrival event occurs or not during this time slot. Hence,

$$P\{L,1,k|L,0^+,0;L,i,0\} = \frac{p_e}{W_1}.$$

A packet arrival event, however, will result in a buffer overflow, thus adding also to the queue blocking probability.

Transition probabilities from states $(h',0^+,k;h,i,0)$**, with** $k \in \{1,\ldots,W_0-1\}$

In all these cases, the anomalous slot ends with an empty time slot of duration δ. If a packet arrival event occurs, the STA transitions to state $(h'+1,0^+,k)$. Otherwise, it transitions to state $(h',0^+,k)$. That is,

$$P\{h'+1,0^+,k-1|h',0^+,k;h,i,0\} = q_\delta$$
$$P\{h',0^+,k-1|h',0^+,k;h,i,0\} = 1-q_\delta.$$

For the special case in which $h = h' = L$, the STA will move to state $(L, 0^+, k-1)$, either if a packet arrival event occurs or not during the empty time slot. Hence,

$$P\{L, 0^+, k-1 | L, 0^+, k; L, i, 0\} = 1.$$

Again, a packet arrival event will produce a buffer overflow.

Steady-state probabilities

For later convenience, let us now define the function $\mathscr{I}_\alpha(\beta_j)$ as

$$\mathscr{I}_\alpha(\beta_j) = \frac{(1-q_{T_s})^\alpha}{q_{T_s}^\alpha} \sum_{j=\alpha}^{\infty} \binom{j}{\alpha} \left(\frac{q_{T_s}(1-p_e)}{W_0} \right)^j \beta_j. \tag{7.1}$$

Using this function it is quite straightforward to show that the steady-state probabilities of states $(h', 0^+, 0; h, i, 0)$, for $i \in \{0, \ldots, R\}$, can be expressed as a function of the steady-state probability of $(h, i, 0;, h, i, 0)$ as

$$b(h, 0^+, 0; h, i, 0) = b(h, i, 0; h, i, 0) \frac{(1-p) q_{T_s}}{W_0} \mathscr{I}_0(1), \tag{7.2}$$

for $h \in \{1, \ldots, L\}$ and $h' = h$, and

$$\begin{aligned} b(h', 0^+, 0; h, i, 0) = {} & b(h, i, 0; h, i, 0)(1-p) \\ & \times \left[\frac{(1-q_{T_s})}{W_0} \mathscr{I}_{h-h'-1}(1) + \frac{q_{T_s}}{W_0} \mathscr{I}_{h-h'}(1) \right] \end{aligned} \tag{7.3}$$

for $h \in \{2, \ldots, L\}$ and $h' \in \{1, \ldots, h-1\}$.

The steady-state probabilities of states $(h', 0^+, k; h, i, 0)$, with $k \in \{1, \ldots, W_0 - 1\}$ and $i \in \{0, \ldots, R\}$, can now be written as

$$\begin{aligned} b(h, 0^+, k; h, i, 0) = {} & b(h, i, 0; h, i, 0) \frac{(1-p) q_{T_s}}{W_0} \\ & + b(h, 0^+, 0; h, i, 0) \frac{(1-p_e) q_{T_s}}{W_0}, \end{aligned} \tag{7.4}$$

for $h \in \{1, \ldots, L\}$ and $h' = h$,

$$\begin{aligned} & b(h-1, 0^+, k; h, i, 0) \\ & \quad = b(h, i, 0; h, i, 0) \frac{(1-p)(1-q_{T_s})}{W_0} \\ & \qquad + b(h, 0^+, 0; h, i, 0) \frac{(1-p_e)(1-q_{T_s})}{W_0} \\ & \qquad + b(h-1, 0^+, 0; h, i, 0) \frac{(1-p_e) q_{T_s}}{W_0} \end{aligned} \tag{7.5}$$

for $h \in \{2, \ldots, L\}$ and $h' = h-1$,

$$\begin{aligned} & b(0, 0^+, k; 1, i, 0) \\ & \quad = b(1, i, 0; 1, i, 0) \frac{(1-p)(1-q_{T_s})}{W_0} \\ & \qquad + b(1, 0^+, 0; 1, i, 0) \frac{(1-p_e)(1-q_{T_s})}{W_0} \end{aligned} \tag{7.6}$$

for the special case in which $h = 1$ and $h' = 0$,

$$\begin{aligned} & b(h', 0^+, k; h, i, 0) \\ & \quad = b(h'+1, 0^+, 0; h, i, 0)\frac{(1-p_e)(1-q_{T_s})}{W_0} \\ & \quad + b(h', 0^+, 0; h, i, 0)\frac{(1-p_e)q_{T_s}}{W_0} \end{aligned} \tag{7.7}$$

for $h \in \{3, \ldots, L\}$ and $h' \in \{1, \ldots, h-2\}$, and finally,

$$\begin{aligned} & b(0, 0^+, k; h, i, 0) \\ & \quad = b(1, 0^+, 0; h, i, 0)\frac{(1-p_e)(1-q_{T_s})}{W_0}, \end{aligned} \tag{7.8}$$

for $h \in \{2, \ldots, L\}$ and $h' = 0$.

7.2.4 Characterizing the high-level transition probabilities

Transitions from the *Idle* state

Let us assume that, when in the *Idle* state, a packet arrival event occurs. In this case, according to the DCF protocol, the STA enters the channel sensing mode and senses the channel for a DIFS time. If the channel is sensed free, the station moves to $(1,0,0)$ state to transmit the packet immediately. Otherwise, the station starts a period of backoff and moves to one of the $(1, 0^-, k)$ states, where $k \in \{0, \ldots, W_0 - 1\}$, with uniform probability. Therefore, the non null one-step transition probabilities from the *Idle* state result in[1]

$$\begin{aligned} P\{Idle|Idle\} &= 1-q \\ P\{1,0,0|Idle\} &= q(1-p) \\ P\{1, 0^-, k|Idle\} &= \frac{qp}{W_0}, \quad k \in \{0, \ldots, W_0 - 1\}. \end{aligned}$$

Transitions from non transmission states

If the station is in a backoff state (h, i, k), with a backoff counter value $k \in \{1, \ldots, W_i - 1\}$ (i.e., a non transmission state), and $h < L$ (i.e., the queue is not full), then a packet arrival event will trigger a transition to the next-plane state $(h+1, i, k-1)$. Otherwise, if the data source does not generate a new packet, the station will keep onto the same plane and will move to the state $(h, i, k-1)$. Hence, for $h \in \{0, \ldots, L-1\}$, $i \in \{0, \ldots, R\}$ and $k \in \{1, \ldots, W_i - 1\}$, it follows that

$$\begin{aligned} P\{h, i, k-1|h, i, k\} &= 1-q \\ P\{h+1, i, k-1|h, i, k\} &= q. \end{aligned}$$

When the queue is full (i.e., $k = L$), the station will keep onto the same plane whether a packet arrival event occurs or not. That is, for $i \in \{0, \ldots, R\}$ and $k \in \{1, \ldots, W_i - 1\}$,

$$P\{L, i, k-1|L, i, k\} = 1.$$

[1] As in [35], for the purpose of modeling, the channel sensing mode has been included in the *Idle* state.

Notice, however, that any packet arrival in one of these backoff states produces a buffer overflow.

Transitions from transmission states $(h,i,0)$

These high-level transition probabilities can be straightforwardly obtained using the low-level transition and steady-state probabilities.

Case I ($i \in \{1,\ldots,R-1\}$) In this case, the non null transition probabilities to states $(h',i+1,k)$ are

$$\begin{aligned} P\{h+1,i+1,k|h,i,0\} &= \frac{pq_{T_e}}{W_{i+1}}, \\ P\{h,i+1,k|h,i,0\} &= \frac{p(1-q_{T_e})}{W_{i+1}}, \end{aligned} \tag{7.9}$$

for $h \in \{1,\ldots,L-1\}$ and $k \in \{0,\ldots,W_{i+1}-1\}$, and

$$P\{L,i+1,k|L,i,0\} = \frac{p}{W_{i+1}}, \tag{7.10}$$

for $h = L$ and $k \in \{0,\ldots,W_{i+1}-1\}$.

The non null transition probabilities to states $(h',1,k)$, with $k \in \{0,\ldots,W_1-1\}$, can be obtained as

$$P\{1,1,k|h,i,0\} = \frac{p_e(1-q_{T_e})}{W_1}\frac{b(1,0^+,0;h,i,0)}{b(h,i,0;h,i,0)}, \tag{7.11}$$

for $h = \{1,\ldots,L\}$,

$$\begin{aligned} P\{h',1,k|h,i,0\} = &\frac{p_e(1-q_{T_e})}{W_1}\frac{b(h',0^+,0;h,i,0)}{b(h,i,0;h,i,0)} \\ &+ \frac{p_e q_{T_e}}{W_1}\frac{b(h'-1,0^+,0;h,i,0)}{b(h,i,0;h,i,0)}, \end{aligned} \tag{7.12}$$

for $h = \{1,\ldots,L-1\}$ and $h' = \{2,\ldots,h\}$,

$$P\{h+1,1,k|h,i,0\} = \frac{p_e q_{T_e}}{W_1}\frac{b(h,0^+,0;h,i,0)}{b(h,i,0;h,i,0)}, \tag{7.13}$$

for $h = \{1,\ldots,L-1\}$,

$$\begin{aligned} P\{h',1,k|L,i,0\} = &\frac{p_e(1-q_{T_e})}{W_1}\frac{b(h',0^+,0;L,i,0)}{b(L,i,0;L,i,0)} \\ &+ \frac{p_e q_{T_e}}{W_1}\frac{b(h'-1,0^+,0;L,i,0)}{b(L,i,0;L,i,0)}, \end{aligned} \tag{7.14}$$

for the special case in which $h = L$ and $h' = \{2,\ldots,L\}$, and finally

$$\begin{aligned} P\{L,1,k|L,i,0\} = &\frac{p_e}{W_1}\frac{b(L,0^+,0;L,i,0)}{b(L,i,0;L,i,0)} \\ &+ \frac{p_e q_{T_e}}{W_1}\frac{b(L-1,0^+,0;L,i,0)}{b(L,i,0;L,i,0)}, \end{aligned} \tag{7.15}$$

for $h' = L$.

The transition probabilities to the *Idle* state are

$$P\{Idle|h,i,0\} = \frac{(1-p_e)(1-q_{T_s})}{W_0}\frac{b(1,0^+,0;h,i,0)}{b(h,i,0;h,i,0)}, \tag{7.16}$$

for $h \in \{2,\ldots,L\}$, and

$$\begin{aligned} P\{Idle|1,i,0\} =& \frac{(1-p)(1-q_{T_s})}{W_0} \\ &+ \frac{(1-p_e)(1-q_{T_s})}{W_0}\frac{b(1,0^+,0;h,i,0)}{b(h,i,0;h,i,0)}, \end{aligned} \tag{7.17}$$

for $h = 1$.

Finally, the transition probabilities to states $(h',0^+,k)$, with $k \in \{0,\ldots,W_0-2\}$, can be obtained as

$$P\{0,0^+,k|h,i,0\} = (1-q_\delta)\frac{b(0,0^+,k+1;h,i,0)}{b(h,i,0;h,i,0)}, \tag{7.18}$$

for $h \in \{1,\ldots,L\}$,

$$\begin{aligned} P\{h',0^+,k|h,i,0\} =& (1-q_\delta)\frac{b(h',0^+,k+1;h,i,0)}{b(h,i,0;h,i,0)} \\ &+ q_\delta\frac{b(h'-1,0^+,k+1;h,i,0)}{b(h,i,0;h,i,0)}, \end{aligned} \tag{7.19}$$

for $h \in \{1,\ldots,L\}$ and $h' \in \{1,\ldots,h\}$, except for the special case in which $h = h' = L$ where

$$\begin{aligned} P\{L,0^+,k|L,i,0\} =& \frac{b(L,0^+,k+1;L,i,0)}{b(L,i,0;L,i,0)} \\ &+ q_\delta\frac{b(L-1,0^+,k+1;L,i,0)}{b(L,i,0;L,i,0)}, \end{aligned} \tag{7.20}$$

and

$$P\{h+1,0^+,k|h,i,0\} = q_\delta\frac{b(h,0^+,k+1;h,i,0)}{b(h,i,0;h,i,0)}, \tag{7.21}$$

for $h \in \{1,\ldots,L-1\}$.

Case II ($i = 0$) All the transition probabilities that have been obtained for Case I are still valid when the system is in backoff stage-0, except for the following ones:

$$\begin{aligned} P\{h+1,1,k|h,0,0\} =& \frac{pq_{T_e}}{W_1} \\ &+ \frac{p_e q_{T_e}}{W_1}\frac{b(h,0^+,0;h,0,0)}{b(h,0,0;h,0,0)}, \end{aligned} \tag{7.22}$$

for $h \in \{1,\ldots,L-1\}$ and $k \in \{0,\ldots,W_1-1\}$,

$$\begin{aligned} P\{h,1,k|h,0,0\} =& \frac{p(1-q_{T_e})}{W_1} \\ &+ \frac{p_e(1-q_{T_e})}{W_1}\frac{b(h,0^+,0;h,0,0)}{b(h,0,0;h,0,0)} \\ &+ \frac{p_e q_{T_e}}{W_1}\frac{b(h-1,0^+,0;h,0,0)}{b(h,0,0;h,0,0)}, \end{aligned} \tag{7.23}$$

for $h \in \{2,\ldots,L-1\}$ and $k \in \{0,\ldots,W_1-1\}$,

$$\begin{aligned} P\{1,1,k|1,0,0\} = &\frac{p(1-q_{T_e})}{W_1} \\ &+ \frac{p_e(1-q_{T_e})}{W_1}\frac{b(1,0^+,0;1,0,0)}{b(1,0,0;1,0,0)}, \end{aligned} \tag{7.24}$$

for $h = 1$ and $k \in \{0,\ldots,W_1-1\}$, and finally,

$$\begin{aligned} P\{L,1,k|L,0,0\} = &\frac{p}{W_1} \\ &+ \frac{p_e}{W_1}\frac{b(L,0^+,0;L,0,0)}{b(L,0,0;L,0,0)} \\ &+ \frac{p_e q_{T_e}}{W_1}\frac{b(L-1,0^+,0;L,0,0)}{b(L,0,0;L,0,0)}, \end{aligned} \tag{7.25}$$

Case III ($i = R$) If the HOL packet transmission fails when the STA is in backoff stage-R, the packet is dropped, the backoff counter is reset and the STA transits to one of the backoff stage-0^- states. Consequently, the transition probabilities obtained for Case I are again valid except for the following ones:

$$\begin{aligned} P\{h,0^-,k|h,R,0\} &= \frac{pq_{T_e}}{W_0}, \\ P\{h-1,0^-,k|h,R,0\} &= \frac{p(1-q_{T_e})}{W_0}, \end{aligned} \tag{7.26}$$

for $h \in \{1,\ldots,L\}$ and $k \in \{0,\ldots,W_0-1\}$.

7.2.5 Calculating the packet arrival probabilities

Assuming both a fixed packet error rate p_e and a given offered load, characterized by the number of active stations n and a per-STA packet arrival rate r_q, the packet arrival probabilities q_{T_s}, q_{T_e}, q and q_δ can be derived as

$$\begin{aligned} q_{T_s} &= r_q T_{T_s}, \\ q_{T_e} &= r_q T_{T_e}, \\ q &= r_q E_{s|n-1}, \\ q_\delta &= r_q \delta, \end{aligned} \tag{7.27}$$

where $E_{s|n-1}$ is the expected time spent per Markov state for states in which the STA is not transmitting.

According to the three-dimensional Markov chain model described in previous sections, although the transitions for the high-level Markov model are synchronized, the time spent in each model state may vary considerably from one state to another. For instance, when no station is transmitting and at least one of them is not *Idle* and, consequently, it is decrementing its backoff counter, the time between transitions is equal to δ. Furthermore, as an STA entering the *Idle* state loses synchronization with the other STAs, the concept of slot, for this particular STA, also loses its meaning. Nevertheless, as we have modeled the probability of transition from the *Idle* state back

onto itself as $(1-q)$, where q has been defined as the probability of generating a new packet during the average time spent in a transition when this particular STA does not transmit, it can be deducted that when all the STAs in the system are in the *Idle* state, the time spent in a transition also coincides with the empty slot time δ. Finally, when the backoff counter of an STA reaches zero and it has packets for transmission in the queue, it transmits and two events may happen:

- *Packet transmission failure (collision or packet error)*: in this case the time spent in the transition is equal to T_{T_e} and depends on both the packet size and the transmission mode.
- *Packet transmission success*: in this case the STA enters what we have denoted as an anomalous slot and the average time spent in the transition, which we denote as $\overline{T}_s$, depends on a list of parameters including, among others, the packet size, the transmission mode, the contention window size, the probability of error and the packet arrival rate.

Thus, putting all the pieces together, the expected time spent per Markov state, assuming a symmetric n-station network, can be obtained as

$$E_s^{(n)} = P_e^{(n)}\delta + P_s^{(n)}\overline{T}_s + P_c^{(n)}T_{T_e}, \tag{7.28}$$

where

$$\begin{aligned} P_e^{(n)} &= (1-\tau)^n, \\ P_s^{(n)} &= n\tau(1-p_e)(1-\tau)^{n-1}, \\ P_c^{(n)} &= 1 - P_e^{(n)} - P_s^{(n)}, \end{aligned} \tag{7.29}$$

are used to denote, respectively, the probability that no STA is transmitting (empty media), the probability that only one STA is successfully transmitting (anomalous slot), and the probability of suffering a packet transmission failure due to either a collision or a packet error. Assuming STA independence, $E_{s|n-1}$ is equivalent to $E_s^{(n-1)}$.

The time spent in an anomalous slot depends, amongst other parameters, on the instantaneous queue length Q at the beginning of this slot. Thus, by conditioning on the queue length Q, the average time spent in an anomalous slot can be computed as

$$\overline{T}_s = \frac{\sum_{h=1}^{L}\overline{T}_s^{(h)}\sum_{i=0}^{R}b(h,i,0)}{\sum_{h=1}^{L}\sum_{i=0}^{R}b(h,i,0)}, \tag{7.30}$$

where $\overline{T}_s^{(h)}$ denotes the average time spent in an anomalous slot if the STA accesses it with h packets in the queue and $b(h,i,k)$ denotes the probability of being in the (h,i,k) state. These conditional average times can be obtained by using the function $\mathscr{I}_\alpha(\beta_j)$, and letting the parameter β_j denote the duration of an anomalous slot containing $j+2$ transmissions. The following cases must be taken into account:

- Transition from $(h,i,0)$ to the *Idle* state: The STA leaves the anomalous slot after a first successful transmission in a contending state, j successful transmissions to reach state $(1,0,0;h,i,0)$, and a last successful transmission that empties the queue. The system makes a transition to the *Idle* state because during the last transmission the source does not generate a new packet and selects a random backoff counter $k=0$. Hence, $\beta_j = (j+2)T_{T_s}$.

- Transitions from $(h,i,0)$ to $(h',0^+,k)$: In these cases, the STA leaves the anomalous slot after $j+2$ successful transmissions and a backoff time slot of duration δ. Thus, $\beta_j = (j+2)T_{T_s} + \delta$.
- Transitions from $(h,i,0)$ to $(h',1,k)$: After entering the anomalous slot, the STA succeeds in the transmission of the first $j+1$ packets, but fails in the transmission of the last packet. Consequently, the STA transitions to one of the backoff stage-1 states after spending a time $\beta_j = (j+1)T_{T_s} + T_{T_e}$ in the anomalous slot.

The result is that

$$\begin{aligned}
\overline{T}_s^{(h)} =& \frac{p_e q_{T_s}}{W_0}\mathscr{I}_0\left((j+1)T_{T_s}+T_{T_e}\right) + \frac{W_0-1}{W_0}\left(T_{T_s}+\delta\right) \\
&+ \frac{p_e\left(1-q_{T_s}\right)}{W_0}\sum_{h'=1}^{h-1}\mathscr{I}_{h-h'-1}\left((j+1)T_{T_s}+T_{T_e}\right) \\
&+ \frac{p_e q_{T_s}}{W_0}\sum_{h'=1}^{h-1}\mathscr{I}_{h-h'}\left((j+1)T_{T_s}+T_{T_e}\right) \\
&+ \frac{(1-p_e)q_{T_s}(W_0-1)}{W_0^2}\mathscr{I}_0\left((j+2)T_{T_s}+\delta\right) \\
&+ \frac{(1-p_e)(W_0-1)(1-q_{T_s})}{W_0^2} \\
&\quad\times\sum_{h'=1}^{h-1}\mathscr{I}_{h-h'-1}\left((j+2)T_{T_s}+\delta\right) \\
&+ \frac{(1-p_e)(W_0-1)q_{T_s}}{W_0^2}\sum_{h'=1}^{h-1}\mathscr{I}_{h-h'}\left((j+2)T_{T_s}+\delta\right) \\
&+ \frac{(1-p_e)(1-q_{T_s})^2}{W_0^2}\mathscr{I}_{h-2}\left((j+2)T_{T_s}\right) \\
&+ \frac{(1-p_e)(1-q_{T_s})q_{T_s}}{W_0^2}\mathscr{I}_{h-1}\left((j+2)T_{T_s}\right),
\end{aligned} \tag{7.31}$$

for $h \in \{2,\ldots,L\}$, and for the special case in which $h=1$,

$$\begin{aligned}
\overline{T}_s^{(1)} =& \frac{p_e q_{T_s}}{W_0}\mathscr{I}_0\left((j+1)T_{T_s}+T_{T_e}\right) \\
&+ \frac{W_0-1}{W_0}\left(T_{T_s}+\delta\right) + \frac{(1-q_{T_s})}{W_0}T_{T_s} \\
&+ \frac{(1-p_e)q_{T_s}(W_0-1)}{W_0^2}\mathscr{I}_0\left((j+2)T_{T_s}+\delta\right) \\
&+ \frac{(1-p_e)(1-q_{T_s})q_{T_s}}{W_0^2}\mathscr{I}_0\left((j+2)T_{T_s}\right).
\end{aligned} \tag{7.32}$$

7.2.6 Solving the three-dimensional Markov chain model

The packet generation probabilities q_{T_s}, q_{T_e}, q and q_δ can be used to solve the three-dimensional Markov chain for the steady-state transition probabilities. The most direct approach to obtain the steady-state probabilities of the three-dimensional Markov

chain model would be to form the corresponding $N_s \times N_s$ transition matrix $\mathbf{\Pi}$, with $N_s = 1 + W_0 + LW_0(2^{a+1} + (R-a)2^a - 1)$ denoting the number of states of the Markov chain, and solve the system of equations

$$\begin{aligned} \mathbf{\Pi}\boldsymbol{b} &= \boldsymbol{b}, \\ \boldsymbol{b}^T \mathbf{1}_{N_s} &= 1, \end{aligned} \tag{7.33}$$

where $\boldsymbol{b}$ denotes the vector of steady-state probabilities. Using $\boldsymbol{b}$, the transmission probability τ could be obtained as

$$\tau = \sum_{h=1}^{L} \sum_{i=0}^{R} b(h,i,0) \tag{7.34}$$

in terms of p, p_e, q_{T_s}, q_{T_e}, q and q_δ. Of course, as the n contending stations are integrated into a network, the probability p depends itself on the values of packet error rate and probability of transmission experienced by all other stations in the network. That is, a given STA will succeed in transmitting the first packet (contention phase) if it is the only station transmitting in this time slot (absence of collision) and the receiver is able to detect the packet without errors. For a symmetric n-station network, this yields

$$p = 1 - (1-p_e)(1-\tau)^{n-1}. \tag{7.35}$$

Since p, τ and $E_s^{(n)}$ are all outputs of the three-dimensional Markov chain model, we could numerically obtain these parameters based on (7.28), (7.33) and (7.35). The problem with this approach is that for the typical 802.11 DCF system, with $W_0 = 16$, $L = 50$, $a = 6$ and $R = 7$, there are $N_s = 152,817$ states in the three-dimensional model and, although the transition matrix is sparse, it is a computationally inhibitive approach in practice.

7.2.7 Refined CTB approach to obtain $\boldsymbol{b}$, τ and $E_s^{(n)}$

Based on the CTB method proposed by Liu *et al.* [35], let us consider all the transitions between the set of transmission states for backoff stage-i to the set of transmission states for backoff stage-$(i+1)$. The sub-chain containing all these transitions can be isolated from the rest of the Markov chain and treated as a set of L *collapsed* transmission states with known multivariate input-output progression probabilities. In this way, as shown in Fig. 7.3, the three-dimensional Markov chain with N_s states is effectively translated onto a *collapsed* Markov chain with only $N_s^{\text{CTB}} = 1 + L(R+1)$ states, including the *Idle*/post-backoff state and backoff stage-i transmission states, for $i \in \{0,\ldots,R\}$.

By introducing the concept of *fan-out* backoff count selectors as the transition branching points that fan out just above each transmission state, the collapsed Markov chain can be solved by decomposing the calculations into sequential steps. Notice that the post-backoff state and backoff stage-0 transmission states can be accessed either from the fan-out state 0^+, after a successful transmission, or from the fan-out state 0^-, after a packet drop. Figure 7.4 shows the matrix representation of the collapsed Markov chain, where the multivariate state $(0,0)$ represents all the states in the post-backoff and backoff stage-0 sub-chains, and the multivariate state $(i,0)$ represents all

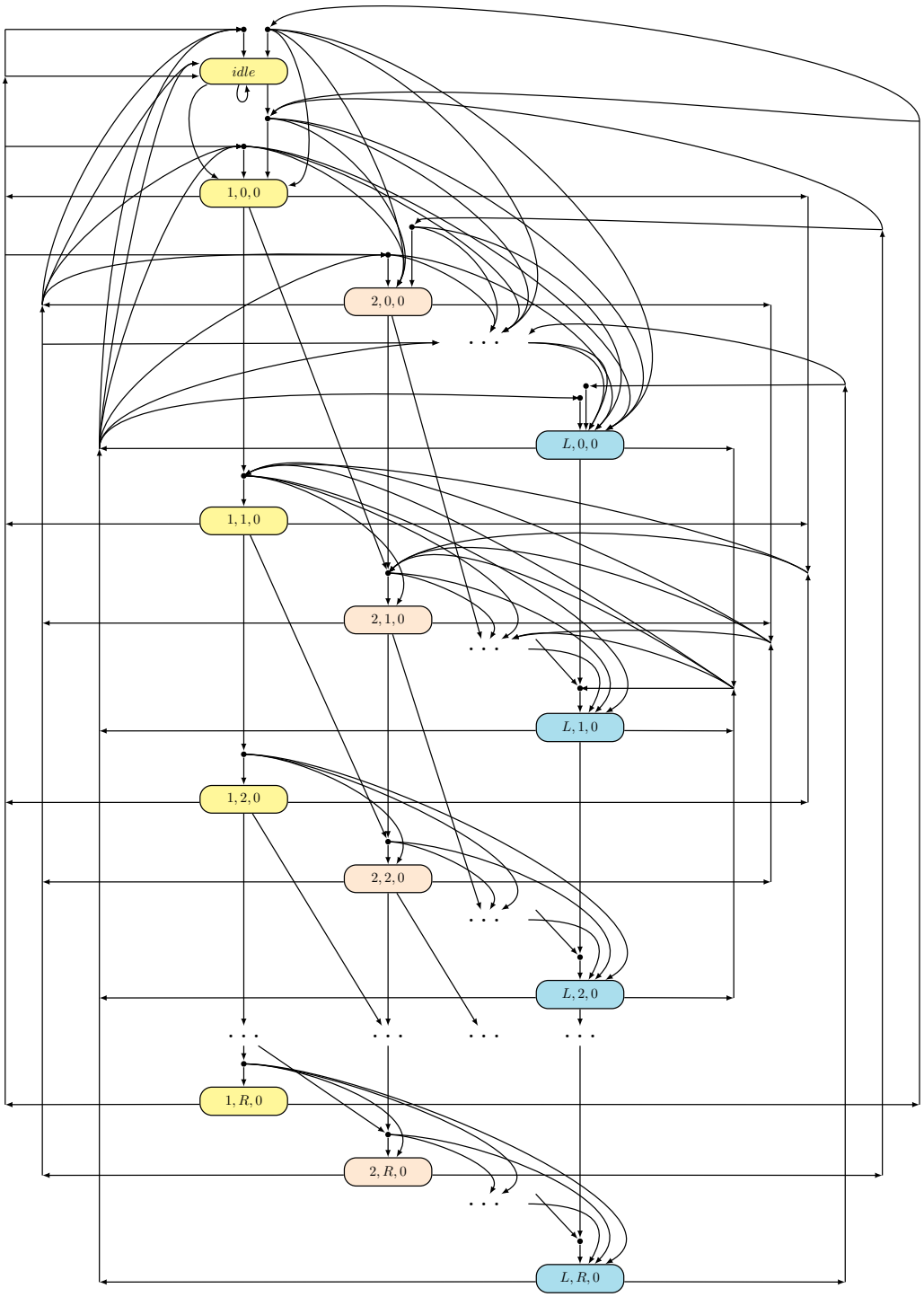


Figure 7.3: *Collapsed* Markov chain.

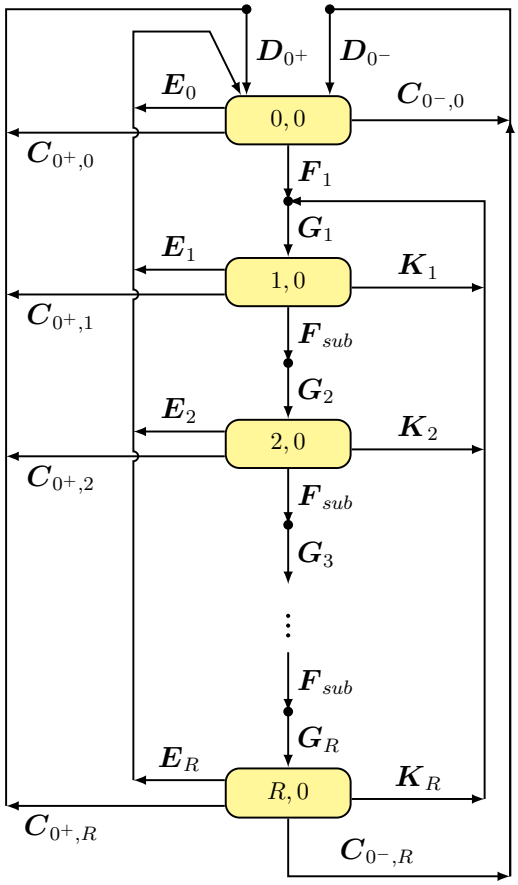


Figure 7.4: Matrix representation of the *collapsed* Markov chain.

the states in the backoff stage-i sub-chain, with $i \in \{1,\ldots,R\}$. In this matrix-based Markov chain model, the $L \times (L+1)$ matrix $\boldsymbol{F}_1$ collapses the progression probabilities corresponding to transitions from the set of post-backoff and backoff stage-0 transmission states in $(0,0)$ to the set of backoff count selectors for backoff stage-1, and the $L \times L$ matrices $\boldsymbol{F}_i = \boldsymbol{F}$, with $i \in \{2,\ldots,R\}$, collapse the progression probabilities from the backoff stage-$(i-1)$ transmission states to the set of backoff count selectors for backoff stage-i. That is,

$$\begin{aligned}\boldsymbol{F}_i[h,h'] &= \boldsymbol{F}[h,h'] \\ &= \begin{cases} p(1-q_{T_e}), & h' \in \{1,\ldots,L-1\},\ h = h', \\ pq_{T_e}, & h' \in \{1,\ldots,L-1\},\ h = h'+1, \\ p, & h = h' = L, \\ 0, & \text{otherwise}, \end{cases}\end{aligned} \tag{7.36}$$

for $i \in \{2,\ldots,R\}$, and

$$\boldsymbol{F}_1 = \begin{bmatrix} \boldsymbol{0}_L & \boldsymbol{K}_0 \end{bmatrix}, \tag{7.37}$$

with

$$\boldsymbol{K}_0[h,h'] = W_1 P\{h,1,k|h',0,0\}. \tag{7.38}$$

The $L \times L$ matrices $\boldsymbol{G}_i$, for $i \in \{1,\ldots,R\}$, collapse the progressing probabilities for the transitions from the set of *fan-out* backoff count selectors for backoff stage-i onto the transmission states for backoff stage-i. The probabilities in these matrices depend on the backoff contention window length W_i and can be obtained as

$$\boldsymbol{G}_i[h,h'] = \begin{cases} \frac{q^{h-h'}}{W_i} \sum_{j=h-h'}^{W_i-1} \binom{j}{h-h'} (1-q)^{j-h+h'}, & h' \in \{1,\ldots,L-1\} \\ & h \in \{h',\ldots,\min\{L-1,h'+W_i-2\}\}, \\ 1 - \sum_{j=h'}^{\min\{L-1,h'+W_i-2\}} \boldsymbol{G}_i[j,h'], & h' \in \{1,\ldots,L-1\}, \\ & h = \min\{L,h'+W_i-1\} \\ 1, & h = h' = L, \\ 0, & \text{otherwise.} \end{cases} \tag{7.39}$$

The progressing probabilities corresponding to the transitions from the backoff stage-i transmission states back to the set of *fan-out* backoff count selectors for backoff stage-1 are contained in the $L \times L$ matrices $\boldsymbol{K}_i$, for $i \in \{1,\ldots,R\}$. The elements of these matrices can be obtained directly out of the transition probabilities from transmission states $(h',i,0)$ to stage-1 backoff states $(h,1,k)$ as

$$\boldsymbol{K}_i[h,h'] = W_1 P\{h,1,k|h',i,0\}, \quad i \in \{1,\ldots,R\}. \tag{7.40}$$

The $(L+1) \times (L+1)$ matrix $\boldsymbol{C}_{0^+,0}$ collapses the progressing probabilities, after a successful transmission, for the transitions from the post-backoff and backoff stage-0 transmission states back to its set of *fan-out* backoff count selectors. Thus,

$$\boldsymbol{C}_{0^+,0}[h,h'] = (W_0-1)P\{h-1,0^+,k|h'-1,0,0\}. \tag{7.41}$$

With a slight abuse of notation, an $(L+1) \times (L+1)$ matrix $\boldsymbol{C}_{0^-,0}$ can also be used to collapse the progressing probabilities for the transitions from the *Idle* state back to the set of *fan-out* backoff count selectors for post-backoff and backoff stage-0. That is,

$$\boldsymbol{C}_{0^-,0}[h,h'] = \begin{cases} q\,p, & h = 2, h' = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{7.42}$$

Furthermore, the $(L+1) \times L$ matrices $\boldsymbol{C}_{0^+,i}$, with $i \in \{1,\ldots,R\}$, collapse the progressing probabilities, after a successful transmission, for the transitions from the backoff stage-i transmission states back to the set of *fan-out* backoff count selectors for post-backoff and backoff stage-0. After a packet drop, the $(L+1) \times L$ matrix $\boldsymbol{C}_{0^-,R}$ collapses the progressing probabilities for the transitions from the backoff stage-R transmission states back to the set of *fan-out* backoff count selectors for post-backoff and backoff stage-0. The elements of all these matrices can be obtained as

$$\boldsymbol{C}_{0^+,i}[h,h'] = (W_0-1)P\{h-1,0^+,k|h'-1,i,0\}, \tag{7.43}$$

for $i \in \{1,\ldots,R\}$, and

$$\boldsymbol{C}_{0^-,R}[h,h'] = W_0 P\{h-1,0^-,k|h'-1,R,0\}. \tag{7.44}$$

The loops are completed with the $(L+1)\times(L+1)$ transition matrices $\boldsymbol{D}_{0^+}$ and $\boldsymbol{D}_{0^-}$, collapsing the progressing probabilities for the transitions from the *fan-out* backoff count selectors for post-backoff and backoff stage-0 to the corresponding set of transmission states. These progressing probabilities can be obtained as

$$\boldsymbol{D}_{0^+}[h,h'] = \begin{cases} \frac{q^{h-h'}}{W_0-1}\sum_{j=h-h'}^{W_0-2}\binom{j}{h-h'}(1-q)^{j-h+h'}, & h' \in \{1,\ldots,L\}, \\ & h \in \{h',\ldots,\min\{L,h'+W_0-3\}\}, \\ 1-\sum_{j=h'}^{\min\{L,h'+W_0-3\}}\boldsymbol{D}_{0^+}[j,h'], & h' \in \{1,\ldots,L\}, \\ & h = \min\{L+1,h'+W_0-2\}, \\ 1, & h = h' = L+1, \\ 0, & \text{otherwise}, \end{cases} \tag{7.45}$$

and

$$\boldsymbol{D}_{0^-}[h,h'] = \begin{cases} \frac{q^{h-h'}}{W_0}\sum_{j=h-h'}^{W_0-1}\binom{j}{h-h'}(1-q)^{j-h+h'}, & h' \in \{1,\ldots,L\} \\ & h \in \{h',\ldots,\min\{L,h'+W_0-2\}\}, \\ 1-\sum_{j=h'}^{\min\{L,h'+W_0-2\}}\boldsymbol{D}_{0^-}[j,h'], & h' \in \{1,\ldots,L\}, \\ & h = \min\{L+1,h'+W_0-1\}, \\ 1, & h = h' = L+1, \\ 0, & \text{otherwise}. \end{cases} \tag{7.46}$$

The final transitions to account for are the transitions from the *Idle* state back onto itself and to state $(1,0,0)$, which can be collapsed onto on $(L+1)\times(L+1)$ matrix $\boldsymbol{E}_0$, and the transitions from the transmission states of backoff stage-i, for all $i \in \{1,\ldots,R\}$, back to the *Idle* state, without going through any *fan-out* backoff count selector, which can be collapsed onto the $(L+1)\times L$ matrices $\boldsymbol{E}_i$. The elements of these matrices are

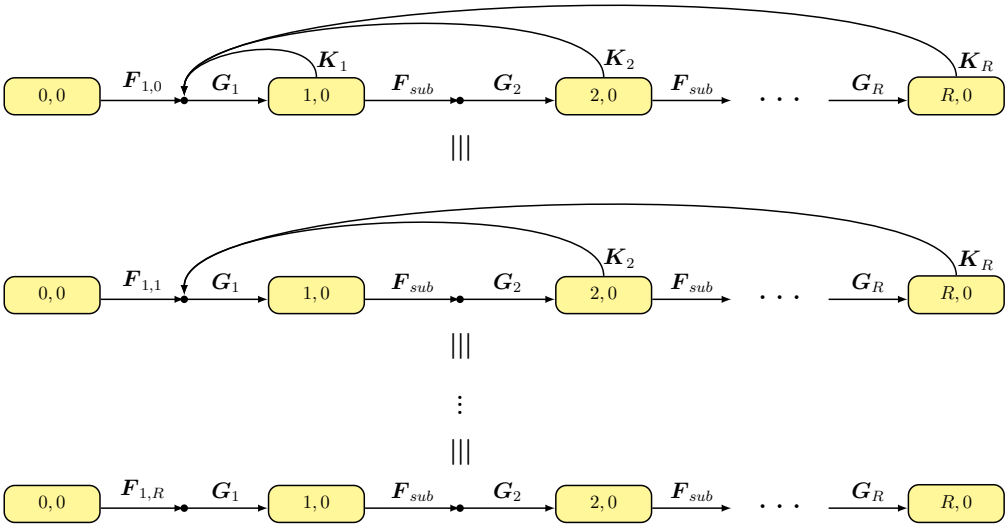
$$\boldsymbol{E}_0[h,h'] = \begin{cases} 1-q, & h = h' = 1, \\ q(1-p), & h = 2, h' = 1, \\ P\{Idle|h'-1,0,0\}, & h = 1, \\ & h' \in \{2,\ldots,L+1\}, \\ 0, & \text{otherwise}, \end{cases} \tag{7.47}$$

and

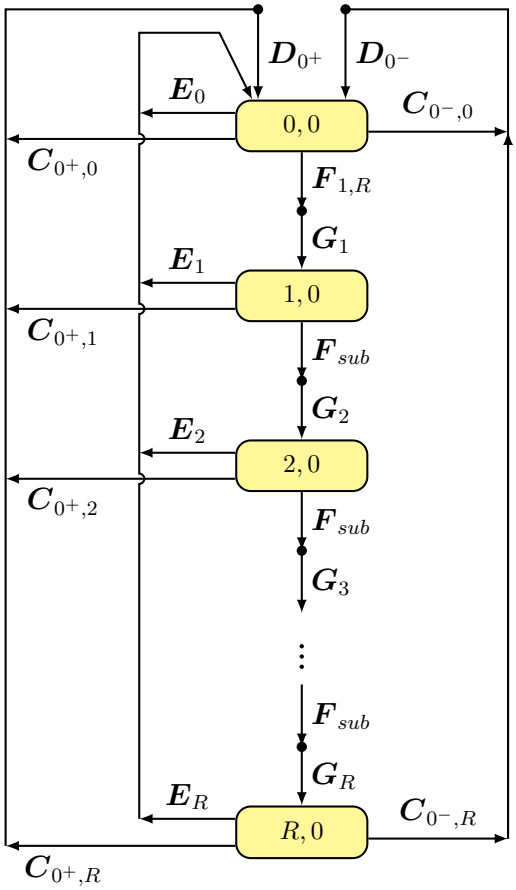
$$\boldsymbol{E}_i[h,h'] = \begin{cases} P\{Idle|h',i,0\}, & h = 1,\ h' \in \{1,\ldots,L\}, \\ 0, & \text{otherwise}. \end{cases} \tag{7.48}$$

Figure 7.5 shows an equivalent matrix representation of the collapsed Markov chain, where $\boldsymbol{F}_1^{\text{equiv}} = \boldsymbol{F}_{1,R}$, with

$$\begin{aligned} \boldsymbol{F}_{1,0} &= \boldsymbol{F}_1, \\ \boldsymbol{F}_{1,1} &= (\boldsymbol{I}-\boldsymbol{K}_1\boldsymbol{G}_1)^{-1}\boldsymbol{F}_{1,0}, \\ \boldsymbol{F}_{1,i} &= \left(\boldsymbol{I}-\boldsymbol{K}_i\left[\prod_{j=0}^{i-2}\boldsymbol{G}_{i-j}\boldsymbol{F}\right]\boldsymbol{G}_1\right)^{-1}\boldsymbol{F}_{1,i-1}, \end{aligned} \tag{7.49}$$



(a) Equivalent Markov chain.



(b) Complete matrix representation of the *collapsed* Markov chain.

Figure 7.5: Equivalent matrix representation of the *collapsed* Markov chain presented in Fig. 7.4.

for $i \in \{2,\ldots,R\}$. Defining the matrices

$$\begin{aligned}
\boldsymbol{B}_0 &= \boldsymbol{I}_{L+1},\\
\boldsymbol{B}_1 &= \boldsymbol{G}_1\boldsymbol{F}_{1,R}\boldsymbol{B}_0,\\
\boldsymbol{B}_i &= \boldsymbol{G}_i\boldsymbol{F}\boldsymbol{B}_{i-1}, \quad i \in \{2,\ldots,R\},
\end{aligned} \tag{7.50}$$

and using the equivalent Markov chain model, the overall transition probability matrix for the collapsed Markov chain can be obtained as

$$\begin{aligned}
\tilde{\boldsymbol{\Pi}} =& \boldsymbol{D}_{0^-}\left(\boldsymbol{C}_{0^-,0} + \boldsymbol{C}_{0^-,R}\boldsymbol{B}_R\right)\\
&+ \sum_{i=0}^{R} E_i\boldsymbol{B}_i + \boldsymbol{D}_{0^+}\sum_{i=0}^{R}\boldsymbol{C}_{0^+,i}\boldsymbol{B}_i,
\end{aligned} \tag{7.51}$$

such that,

$$\begin{aligned}
\tilde{\boldsymbol{\Pi}}\,\tilde{\boldsymbol{b}}_{0,0} &= \tilde{\boldsymbol{b}}_{0,0},\\
\tilde{\boldsymbol{b}}_{0,0}^T\mathbf{1}_{N_s^{\text{CTB}}} &= 1,
\end{aligned} \tag{7.52}$$

where $\tilde{\boldsymbol{b}}_{0,0}$ denotes the $(L+1)\times 1$ column vector containing the steady-state probabilities for the *Idle* and backoff stage-0 transmission states of the collapsed Markov chain. That is, the column vector $\tilde{\boldsymbol{b}}_{0,0}$ contains the steady-state probabilities for the collapsed Markov chain states with backoff stage index $i = 0$ and backoff counter $k = 0$. Notice that $\tilde{\boldsymbol{b}}_{0,0}$ is a scalar multiple of $\boldsymbol{b}_{0,0}$, i.e., $\boldsymbol{b}_{0,0} = \tilde{\boldsymbol{b}}_{0,0}/c$, where $\boldsymbol{b}_{0,0}$ denotes the vector of steady-state probabilities for the *Idle* and backoff stage-0 transmission states of the full Markov chain, and c is a normalizing constant ensuring that the steady-state probabilities of the full Markov chain sum to one.

To calculate the constant c it is necessary to evaluate and sum the probabilities, conditional on $\tilde{\boldsymbol{b}}_{0,0}$, of being in each state (h,i,k). That is,

$$\begin{aligned}
c =& \sum_{h=0}^{L}\sum_{k=0}^{W_0-2}\tilde{b}(h,0^+,k) + \sum_{h=0}^{L}\sum_{k=0}^{W_0-1}\tilde{b}(h,0^-,k)\\
&+ \sum_{i=1}^{R}\sum_{h=1}^{L}\sum_{k=0}^{W_0-1}\tilde{b}(h,i,k).
\end{aligned} \tag{7.53}$$

Except for the special cases of the *Idle* and $(1,0,0)$ states, the probability of being in state (h,i,k), conditional on $\tilde{\boldsymbol{b}}_{0,0}$, can be calculated as the probability of reaching the fan-out count selector (h',i) for backoff stage i and queue length $h' \leq h$ and multiplying it by the progressing probability to (h,i,k). This progressing probability can be

obtained as

$$P\{h,i,k|h',i\} = \begin{cases} & h' \in \{0,\ldots,L-1\} \\ \frac{q^{h-h'}}{\varpi_i}\sum_{j=0}^{\varpi_i-1-h+h'-k}\binom{h-h'+j}{j}(1-q)^j, & h \in \{h',\ldots,L-1\} \\ & k \in \{0,\ldots, \\ & \quad \max\{0,\varpi_i-1-h+h'\}\} \\ & h = L, \quad h' \in \{0,\ldots,L-1\} \\ q\sum_{k'=k+1}^{\varpi_i-L+h'} P\{L-1,i,k'|h',i\}, & k \in \{0,\ldots, \\ & \quad \max\{0,\varpi_i-1-L+h'\}\} \\ & h = h' = L \\ \frac{\varpi_i-k}{\varpi_i}, & k \in \{0,\ldots,\varpi_i-1\} \\ 0, & \text{otherwise,} \end{cases} \tag{7.54}$$

where

$$\varpi_i = \begin{cases} W_0 - 1, & i = 0^+, \\ W_i, & \text{otherwise.} \end{cases} \tag{7.55}$$

Using these expressions it can be shown that, irrespective of the fan-out count selector (h',i),

$$\sum_{h=h'}^{L}\sum_{k=0}^{\varpi_i-1+h-h'} P\{h,i,k|h',i\} = \frac{\bar{\varpi}_i+1}{2}. \tag{7.56}$$

Combining this result with the probability of reaching each backoff count selector, and adding the probability of reaching state $(1,0,0)$ directly from the *Idle* state and the probability of reaching the *Idle* state directly from itself and the transmission states, we obtain

$$\begin{aligned} c = & \left\| \left(\sum_{i=0}^{R} \boldsymbol{C}_{0^+,i}\boldsymbol{B}_i \right) \tilde{\boldsymbol{b}}_{0,0} \right\|_1 \frac{W_0}{2} \\ & + \left\| \left(\boldsymbol{C}_{0^-,0} + \boldsymbol{C}_{0^-,R}\boldsymbol{B}_R \right) \tilde{\boldsymbol{b}}_{0,0} \right\|_1 \frac{W_0+1}{2} \\ & + \left\| \boldsymbol{F}_{1,R}\tilde{\boldsymbol{b}}_{0,0} \right\|_1 \frac{W_1+1}{2} \\ & + \sum_{i=2}^{R} \left\| \boldsymbol{F}\boldsymbol{B}_{i-1}\tilde{\boldsymbol{b}}_{0,0} \right\|_1 \frac{W_i+1}{2} \\ & + \left\| \left(\sum_{i=0}^{R} \boldsymbol{E}_i\boldsymbol{B}_i \right) \tilde{\boldsymbol{b}}_{0,0} \right\|_1 . \end{aligned} \tag{7.57}$$

As a byproduct of the reasoning introduced in the last paragraph, the steady-state probabilities $b(h,i,k)$ of the full three-dimensional Markov chain can be straightforwardly obtained. Let us define the vectors $\boldsymbol{v}_i$ containing the probabilities of reaching the fan-out count selector i, conditional on the vector of normalized probabilities $\boldsymbol{b}_{0,0}$,

as

$$\begin{aligned}
\boldsymbol{v}_{0^+} &= \left(\sum_{i=0}^{R} \boldsymbol{C}_{0^+,i}\boldsymbol{B}_i\right)\boldsymbol{b}_{0,0},\\
\boldsymbol{v}_{0^-} &= \left(\boldsymbol{C}_{0^-,0} + \boldsymbol{C}_{0^-,R}\boldsymbol{B}_R\right)\boldsymbol{b}_{0,0},\\
\boldsymbol{v}_1 &= \boldsymbol{F}_{1,R}\boldsymbol{b}_{0,0},\\
\boldsymbol{v}_i &= \boldsymbol{F}\boldsymbol{B}_{i-1}\boldsymbol{b}_{0,0}, \quad i \in \{2,\ldots,R\}.
\end{aligned} \tag{7.58}$$

Now, using the progressing probabilities defined in (7.54) the steady-state probabilities can be obtained as

$$b(h,i,k) = \begin{cases}
\boldsymbol{b}_{0,0}[h+1], & i=k=0,\\ & h\in\{0,\ldots,L\},\\
\sum_{h'=0}^{L}\Big(P\{h,0^+,k|h',0^+\}\boldsymbol{v}_{0^+}[h'+1] & i=0,\ k\in\{1,\ldots,W_0-1\},\\
\quad +P\{h,0^-,k|h',0^-\}\boldsymbol{v}_{0^-}[h'+1]\Big), & h\in\{0,\ldots,L\},\\
\sum_{h'=1}^{L} P\{h,i,k|h',i\}\boldsymbol{v}_i[h'], & i\in\{1,\ldots,R\},\ k\in\{0,\ldots,W_i-1\},\\ & h\in\{1,\ldots,L\}.
\end{cases} \tag{7.59}$$

Thus, we can numerically obtain p, τ and $E_s^{(n)}$ based on (7.28), (7.59) and (7.35).

The steady-state probabilities of the whole Markov chain, including both high-level and low-level states, can be readily obtained as

$$\hat{b}(h,i,k) = \frac{b(h,i,k)}{\hat{c}} \tag{7.60}$$

and

$$\hat{b}(h',0^+,k;h,i,0) = \frac{b(h',0^+,k;h,i,0)\ b(h,i,0)}{\hat{c}\ b(h,i,0;h,i,0)}, \tag{7.61}$$

respectively, where the normalizing constant $\hat{c}$ is given by

$$\begin{aligned}
\hat{c} = &\sum_{h=0}^{L}\sum_{i=0}^{R}\sum_{k=0}^{W_i-1} b(h,i,k)\\
&+\sum_{h=1}^{L}\sum_{i=0}^{R}\sum_{h'=0}^{h}\sum_{k=0}^{W_0-1}\frac{b(h',0^+,k;h,i,0)\ b(h,i,0)}{b(h,i,0;h,i,0)}.
\end{aligned} \tag{7.62}$$

7.3 QoS performance analysis

7.3.1 MAC layer throughput

An interesting measure of system performance is the MAC layer throughput, also referred to as goodput, which can be defined as the number of useful information bits delivered by the network per unit of time. The amount of data considered excludes protocol overhead bits as well as retransmitted data packets, and the delivery time includes inter-packet time gaps caused by protocol timing, packet queuing delay, packet retransmission time, acknowledge time. In a symmetric network with n active stations, the average time spent in a model slot is given by $E_s^{(n)}$. Furthermore, a successful

transmission (i.e., transmission of useful information bits in an anomalous slot) may comprise multiple consecutive frames with a per-frame payload of P bits. Thus, to compute the MAC layer throughput it is necessary to obtain the average payload size accounting for multiple frames transmitted into the same model slot. The average payload size depends, amongst other parameters, on the instantaneous queue length Q at the beginning of an anomalous slot. Hence, by conditioning on the queue length Q, the average payload size in an anomalous slot can be computed as

$$\overline{P} = \frac{\sum_{h=1}^{L} \overline{P}^{(h)} \sum_{i=0}^{R} b(h,i,0)}{\sum_{h=1}^{L} \sum_{i=0}^{R} b(h,i,0)}, \tag{7.63}$$

where $\overline{P}^{(h)}$ denotes the average payload size of an anomalous slot if the STA accesses it with h packets in the queue, which can be obtained by using again the function $\mathscr{I}_{\alpha}(\beta_j)$, and letting the parameter β_j denote in this case the number of useful bits in an anomalous slot containing $j+2$ transmissions. The following cases must be taken into account:

- Transition from $(h,i,0)$ to the *Idle* state: The STA leaves the anomalous slot after a first successful transmission in a contending state, j successful transmissions to reach state $(1,0,0;h,i,0)$, and a last successful transmission that empties the queue. The system transitions to the *Idle* state because during the last transmission the source does not generate a new packet and selects a random backoff counter $k=0$. Hence, $\beta_j = (j+2)P$.

- Transitions from $(h,i,0)$ to $(h',0^+,k)$: In these cases, the STA leaves the anomalous slot after $j+2$ successful transmissions and a backoff time slot of duration δ. Thus, $\beta_j = (j+2)P$.

- Transitions from $(h,i,0)$ to $(h',1,k)$: After entering the anomalous slot, the STA succeeds in the transmission of the first $j+1$ packets, but fails in the transmission of the last packet. Consequently, the STA transitions to one of the backoff stage-1 states transmitting $\beta_j = (j+1)P$ useful bits in the anomalous slot.

The result is that

$$\begin{aligned}
\overline{P}^{(h)} =& \frac{p_e q_{T_s}}{W_0} \mathscr{I}_0\big((j+1)P\big) + \frac{W_0-1}{W_0} P + \frac{p_e\big(1-q_{T_s}\big)}{W_0} \sum_{h'=1}^{h-1} \mathscr{I}_{h-h'-1}\big((j+1)P\big) \\
&+ \frac{p_e q_{T_s}}{W_0} \sum_{h'=1}^{h-1} \mathscr{I}_{h-h'}\big((j+1)P\big) + \frac{(1-p_e) q_{T_s} (W_0-1)}{W_0^2} \mathscr{I}_0\big((j+2)P\big) \\
&+ \frac{(1-p_e)(W_0-1)(1-q_{T_s})}{W_0^2} \quad \times \sum_{h'=1}^{h-1} \mathscr{I}_{h-h'-1}\big((j+2)P\big) \\
&+ \frac{(1-p_e)(W_0-1) q_{T_s}}{W_0^2} \sum_{h'=1}^{h-1} \mathscr{I}_{h-h'}\big((j+2)P\big) + \frac{(1-p_e)(1-q_{T_s})^2}{W_0^2} \mathscr{I}_{h-2}\big((j+2)P\big) \\
&+ \frac{(1-p_e)(1-q_{T_s}) q_{T_s}}{W_0^2} \mathscr{I}_{h-1}\big((j+2)P\big),
\end{aligned} \tag{7.64}$$

for $h \in \{2,\dots,L\}$, and for the special case in which $h = 1$,

$$\begin{aligned}\overline{P}^{(1)} =& \frac{p_e q_{T_s}}{W_0}\mathscr{I}_0\big((j+1)P\big) + \frac{W_0-1}{W_0}P + \frac{(1-q_{T_s})}{W_0}P \\ &+ \frac{(1-p_e)q_{T_s}(W_0-1)}{W_0^2}\mathscr{I}_0\big((j+2)P\big) + \frac{(1-p_e)(1-q_{T_s})q_{T_s}}{W_0^2}\mathscr{I}_0\big((j+2)P\big).\end{aligned} \tag{7.65}$$

Consequently, the MAC layer throughput can be expressed as,

$$G = \frac{P_s^{(n)}\overline{P}}{P_e^{(n)}\delta + P_s^{(n)}\overline{T}_s + P_c^{(n)}T_{T_e}}. \tag{7.66}$$

7.3.2 Average interface queue length

Let us define $P_s[Q=h]$ as the probability of the interface queue staying in a state containing $Q = h$ packets. As we have obtained the steady-state probability vector $\hat{\boldsymbol{b}}$ of the hierarchical Markov chain, including low-level and high-level states, $P_s[Q=h]$ can be obtained as the sum of all the steady-state probabilities corresponding to states with $Q = h$, that is,

$$\begin{aligned}P_s[Q=h] =& \sum_{i=0}^{R}\sum_{k=1}^{W_i-1}\hat{b}(h,i,k) + \sum_{i=0}^{R}\hat{b}(h,i,0) \\ &+ \sum_{h'=h}^{L}\sum_{i=0}^{R}\hat{b}(h,0,0;h',i,0) \\ &+ \sum_{h'=h}^{L}\sum_{i=0}^{R}\sum_{k=0}^{W_0-2}\hat{b}(h,0^+,k+1;h',i,0),\end{aligned} \tag{7.67}$$

for $h \in \{1,\dots,L\}$. For the special case in which $Q = 0$,

$$\begin{aligned}P_s[Q=0] =& \sum_{k=0}^{W_i-1}\hat{b}(0,0,k) \\ &+ \sum_{h'=1}^{L}\sum_{i=0}^{R}\sum_{k=0}^{W_0-2}\hat{b}(0,0^+,k+1;h',i,0).\end{aligned} \tag{7.68}$$

Noting that the time spent in a transition between two states is highly dependent on the particular states this involves, it is of interest to find the probability of the interface queue length being $Q = h$ at any given time. Thus, to convert the state probabilities $P_s[Q=h]$ into time probabilities, denoted as $P_t[Q=h]$, the expected time with queue length equal to h must be normalized by the expected time per state. Hence,

$$\begin{aligned}P_t[Q=h] =& \frac{E_s^{(n-1)}}{\hat{E}_s^{(n)}}\sum_{i=0}^{R}\sum_{k=1}^{W_i-1}\hat{b}(h,i,k) \\ &+ \frac{pT_{T_e} + (1-p)T_{T_s}}{\hat{E}_s^{(n)}}\sum_{i=0}^{R}\hat{b}(h,i,0) \\ &+ \frac{p_e T_{T_e} + (1-p_e)T_{T_s}}{\hat{E}_s^{(n)}}\sum_{h'=h}^{L}\sum_{i=0}^{R}\hat{b}(h,0,0;h',i,0) \\ &+ \frac{\delta}{\hat{E}_s^{(n)}}\sum_{h'=h}^{L}\sum_{i=0}^{R}\sum_{k=0}^{W_0-2}\hat{b}(h,0^+,k+1;h',i,0),\end{aligned} \tag{7.69}$$

for $h \in \{1,\ldots,L\}$, and

$$\begin{aligned} P_t[Q=0] = & \frac{E_s^{(n-1)}}{\hat{E}_s^{(n)}} \sum_{k=0}^{W_i-1} \hat{b}(0,0,k) \\ & + \frac{\delta}{\hat{E}_s^{(n)}} \sum_{h'=1}^{L} \sum_{i=0}^{R} \sum_{k=0}^{W_0-2} \hat{b}(0,0^+,k+1;h',i,0), \end{aligned} \tag{7.70}$$

where the expected time per state, assuming a symmetric n-station network, can be obtained as

$$\begin{aligned} \hat{E}_s^{(n)} = & E_s^{(n-1)} \left[\sum_{k=0}^{W_0-1} \hat{b}(0,0,k) + \sum_{h=1}^{L} \sum_{i=0}^{R} \sum_{k=1}^{W_i-1} \hat{b}(h,i,k) \right] \\ & + \left[pT_{T_e} + (1-p)T_{T_s} \right] \sum_{h=1}^{L} \sum_{i=0}^{R} \hat{b}(h,i,0) \\ & + \left[p_e T_{T_e} + (1-p_e)T_{T_s} \right] \sum_{h'=1}^{L} \sum_{i=0}^{R} \sum_{h=1}^{h'} \hat{b}(h,0,0;h',i,0) \\ & + \delta \sum_{h'=1}^{L} \sum_{i=0}^{R} \sum_{h=0}^{h'} \sum_{k=0}^{W_0-2} \hat{b}(h,0^+,k+1;h',i,0). \end{aligned} \tag{7.71}$$

7.3.3 Blocking and dropping probabilities

The packet blocking probability P_B is the probability that an arriving packet will not join the input interface queue due to lack of space. If the queue is full ($Q = L$), the source generates a new packet, and the system is in a non-transmission state ($k \neq 0$), then the arriving packet will be blocked due to buffer overflow. If, instead, the system is in a transmission state ($k = 0$), the new arriving packet will only be lost if the transmission of the HOL packet fails. Notice, however, that if $i = R$ and the transmission fails, the HOL packet is dropped due to exceeding the maximum number of allowed retransmissions and, thus, it leaves room to accept the arriving packet. As such,

$$\begin{aligned} P_B = & \frac{E_s^{(n-1)}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \sum_{k=1}^{W_i-1} \hat{b}(L,i,k) \\ & + p \frac{T_{T_e}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R-1} \hat{b}(L,i,0) \\ & + p_e \frac{T_{T_e}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \hat{b}(L,0,0;L,i,0) \\ & + \frac{\delta}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \sum_{k=0}^{W_0-2} \hat{b}(L,0^+,k+1;L,i,0). \end{aligned} \tag{7.72}$$

A non-blocked packet is dropped due to exceeding the maximum number of allowed retransmissions if the STA is in a transmission state of backoff stage-R and the transmission of the HOL packet fails, either due to collision or channel error. Thus, the dropping probability of non-blocked packets can be expressed as

$$P_D = \frac{p}{r_q \hat{E}_s^{(n)}} \sum_{h=1}^{L} \hat{b}(h,R,0). \tag{7.73}$$

7.3.4 Average packet delay

The average packet delay D_t comprises the average time from a packet first joining the input interface queue until it reaches the HOL, plus the MAC delay, denoted as Δ and defined as the average time from a packet first reaching the HOL until it is successfully transmitted. Obviously, retransmission failures of packets other than that being traced do add to the delay. Hence, denoting with Δ_l the average time from a packet first reaching the HOL until it leaves the system, either by successful transmission or retransmission failure, the average packet delay D_t can be expressed as

$$\begin{aligned} D_t =& \Delta + \Delta_l \sum_{h=1}^{L-1} h \frac{P_t[Q=h]}{1-P_B} \\ &+ \left(\frac{T_{T_s}}{2} + (L-1)\Delta_l \right) \frac{P_t^{(T)}[Q=L]}{1-P_B}, \end{aligned} \tag{7.74}$$

where

$$\begin{aligned} P_t^{(T)}[Q=L] =& p \frac{T_{T_e}}{\hat{E}_s^{(n)}} \hat{b}(L,R,0) \\ &+ (1-p) \frac{T_{T_s}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \hat{b}(L,i,0) \\ &+ (1-p_e) \frac{T_{T_s}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \hat{b}(L,0,0;L,i,0), \end{aligned} \tag{7.75}$$

denotes the time probability of the system being in a transmission state with $Q = L$.

To calculate Δ_l, let us first obtain the average time per state, given a positive queue length, as

$$\hat{E}_{s|Q\geq 1}^{(n)} = \frac{\hat{E}_s^{(n)} \sum_{h=1}^{L} P_t[Q=h]}{\sum_{h=1}^{L} P_s[Q=h]}. \tag{7.76}$$

Now, let us find the state probability of a packet leaving the queue, either by retransmission failure or successful transmission, given a positive queue length, as

$$P_{l|Q\geq 1} = \frac{\sum_{h=1}^{L} P_{l\cap Q=h}}{\sum_{h=1}^{L} P_s[Q=h]}, \tag{7.77}$$

where the joint probability of a packet leaving the queue, either by retransmission failure or successful transmission, and the STA staying in a state corresponding to a positive queue length can be obtained as

$$\begin{aligned} P_{l\cap Q=h} =& (1-p) \sum_{i=0}^{R-1} \hat{b}(h,i,0) + \hat{b}(h,R,0) \\ &+ (1-p_e) \sum_{h'=h}^{L} \sum_{i=0}^{R} \hat{b}(h,0,0;h',i,0). \end{aligned} \tag{7.78}$$

Using $\hat{E}_{s|Q\geq 1}^{(n)}$ and $P_{l|Q\geq 1}$, Δ_l can be calculated as

$$\Delta_l = \frac{\hat{E}_{s|Q\geq 1}^{(n)}}{P_{l|Q\geq 1}} = \frac{\hat{E}_s^{(n)} \sum_{h=1}^{L} P_t[Q=h]}{\sum_{h=1}^{L} P_{l\cap Q=h}}. \tag{7.79}$$

Excluding the failed transmissions in backoff stage-R from the previous calculations, we can similarly obtain Δ as

$$\Delta = \frac{\hat{E}_s^{(n)} \sum_{h=1}^{L} \left(P_t[Q=h] - p \frac{T_{Te}}{\hat{E}_s^{(n)}} \hat{b}(h,R,0) \right)}{\sum_{h=1}^{L} \left(P_{l \cap Q=h} - p \hat{b}(h,R,0) \right)}. \tag{7.80}$$

7.4 Validation and comparison with previous models

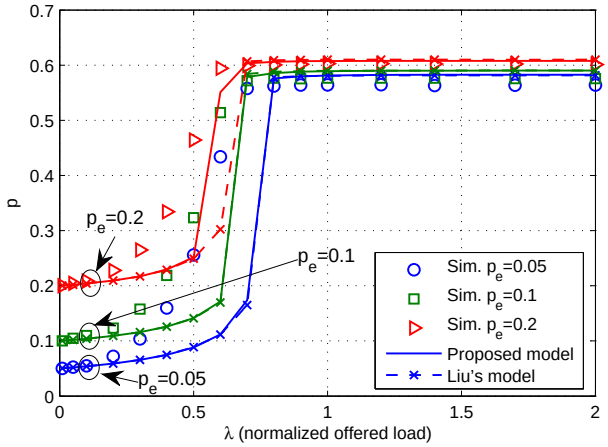
To validate the proposed 3D Markov model and refined CTB analytical approach, extensive simulations have been conducted using a Matlab-based simulator implementing the main tasks at both the MAC and PHY layers of a DCF-based IEEE 802.11n network. The simulator considers a network with n STAs within communication range of each other (absence of hidden terminal problem). Traffic is generated following a Poisson distribution for the packet interarrival times with mean value defined by $\lambda = r_q/(nT_{T_s})$. Moreover, a finite state machine has been implemented to manage a MAC layer conforming to the directives specified in the standard [6]. Except otherwise indicated, the simulation experiments for a particular network configuration have been run for 10^6 model slots using the default system parameters that have been summarized in Table 6.2, furthermore, an MPDU of duration $T_{MPDU} = 136\mu$s has been assumed.

Our refined 3D model analytical predictions, together with those provided by Liu's 3D model [35, 104], will be compared with simulation results as a function of the normalized arrival rate (offered load) and with the number of STAs, the packet error probability or the minimum contention window size as parameters. Theoretical and simulation results will be presented for a high-rate transmission mode, namely IEEE 802.11n MCS15 (see Table 3.3). Similar qualitative results have been obtained for all other adaptive modulation and coding formats.

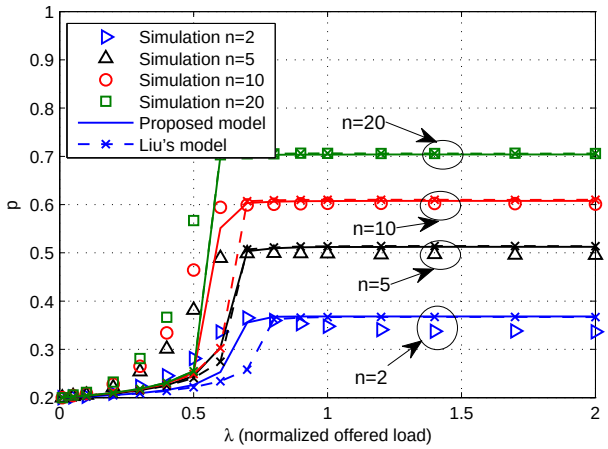
7.4.1 p, τ and $E_s^{(n)}$

The transmission probability, the collision (plus error) probability, and the expected time spent per Markov state, are the outputs of the refined 3D Markov chain model and have been numerically obtained by solving (7.28), (7.33) and (7.35). The behavior of these parameters is shown in Figs. 7.6-7.8 as a function of the normalized offered load, and for various network configurations. A close match between analytical results provided by our proposed model (in solid lines) and the simulation results (in markers) can be observed, irrespective of the number of STAs in the network, the packet error probability and/or the minimum contention window size. Remarkably, our analytical model predicts the normalized offered load values that set off the network congestion, evidenced by a sudden increase of collision probability from a low level towards a saturated value. Collisions produce a subsequent escalade of retransmissions and hence, a sharp increase of transmission probability and expected time spent per Markov state, the later due to the decrease of the probability of the network being in an Idle state.

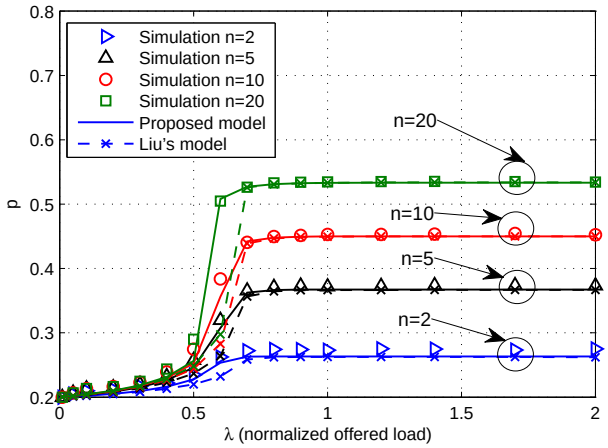
As clearly observed in Fig. 7.6a, under non-saturated conditions the value of p is dominated by the packet error probability and thus, an increase of p_e translates to an almost equivalent increase of p. Under network congestion conditions, however,



(a) $p_e \in \{0.05, 0.1, 0.2\}$, $CW_{\min} = 3$, $n = 10$

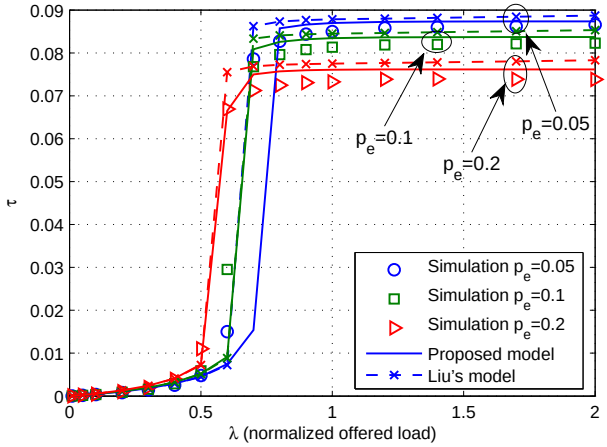


(b) $n \in \{2, 5, 10, 20\}$, $CW_{\min} = 3$, $p_e = 0.2$

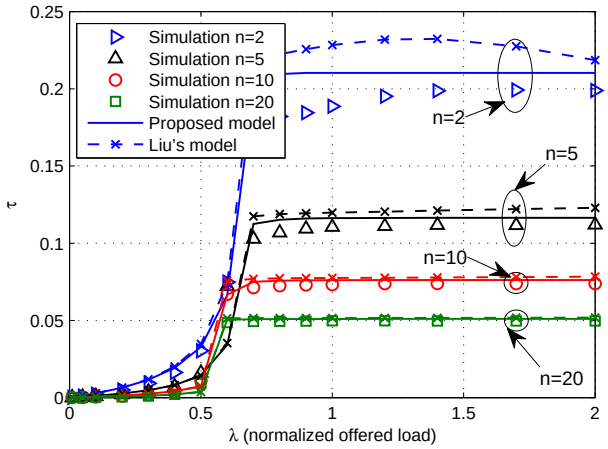


(c) $n \in \{2, 5, 10, 20\}$, $CW_{\min} = 15$, $p_e = 0.2$

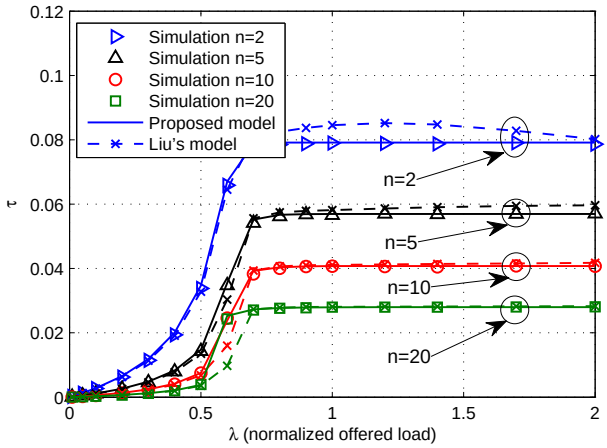
Figure 7.6: Collision probability vs normalized offered load



(a) $p_e \in \{0.05, 0.1, 0.2\}$, $CW_{\min} = 3$, $n = 10$

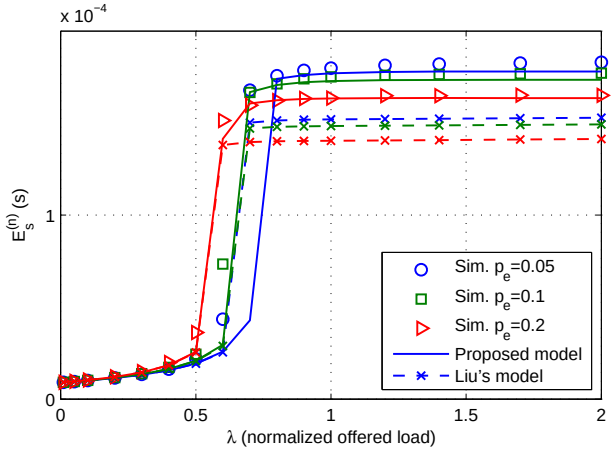


(b) $n \in \{2, 5, 10, 20\}$, $CW_{\min} = 3$, $p_e = 0.2$

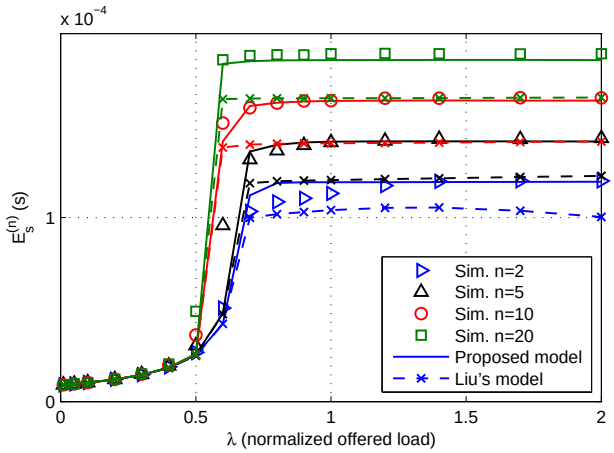


(c) $n \in \{2, 5, 10, 20\}$, $CW_{\min} = 15$, $p_e = 0.2$

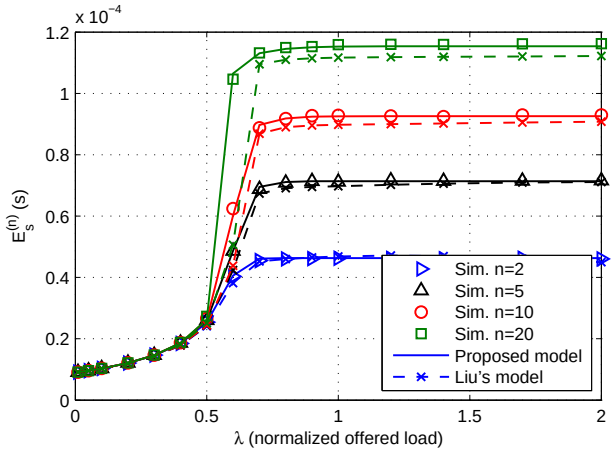
Figure 7.7: Transmission probability vs normalized offered load



(a) $p_e \in \{0.05, 0.1, 0.2\}$, $CW_{\min} = 3$, $n = 10$



(b) $n \in \{2, 5, 10, 20\}$, $CW_{\min} = 3$, $p_e = 0.2$



(c) $n \in \{2, 5, 10, 20\}$, $CW_{\min} = 15$, $p_e = 0.2$

Figure 7.8: Time spent per Markov state vs normalized offered load

where the value of p is dominated by collisions, the increase of p_e only produces a slight effect on p. Nonetheless, this small increase in p produces a noticeable decrease in both the probability of transmission and the expected time spent per Markov state (see Figs. 7.7a and 7.8a.

As shown in Figs. 7.6b and 7.6c, increasing the number of STAs in the network produces an increase in the collision probability. However, thanks to the operation of the BEB algorithm, the number of collisions under network congestion can be kept to a saturation value by reducing the transmission probability (see Figs. 7.7b and 7.7c). Furthermore, notice that the larger the number of STAs the higher is the expected time spent per Markov state (see Figs. 7.8b-7.8c) due to the combined effects of increased collision and decreased transmission probabilities.

As expected, comparing Fig. 7.6b with Fig. 7.6c it can be deducted that configuring the BEB algorithm with a minimum contention window size $CW_{\min} = 3$ provides a poorer protection against collisions than that obtained with $CW_{\min} = 15$. In fact, under network congestion conditions, using $CW_{\min} = 15$ instead of $CW_{\min} = 3$ translates to a higher transmission probability and, consequently, to a higher collision probability. Comparing the expected time spent per Markov state does not make sense because the number of states of the Markov chain in both cases is very different.

Let us now compare our refined 3D model with the original 3D model proposed by Liu *et al.* [35, 104], that can be considered as one of the best unsaturated models so far. Clearly, although our refined 3D model produces slightly more accurate predictions for the transmission and collision probabilities, both analytical methods could be considered as very accurate. The reason is that the mere fact of considering the anomalous slot, one of the main differences between the two analytical methods, does not have a significant effect on the calculation of these probabilities. However, turning our attention to the expected time spent per Markov state, while our refined approach provides highly accurate predictions, the method proposed by Liu *et al.* clearly underestimates its duration and, as discussed later, this will result in a poor prediction of QoS performance metrics like, for instance, the goodput, the average queue length, the blocking and dropping probabilities or the average packet delay. It is worth pointing out that as $CW_{\min}$ decreases, and therefore, the probability of an STA transmitting more than a single packet in an anomalous slot increases, the differences between the predictions provided by both theoretical models become increasingly larger.

7.4.2 Goodput performance

Figures 7.9-7.12 show predicted and simulated goodputs against overall normalized offered load, under a wide variety of network configurations. In all cases, there is a very good agreement between results obtained using our refined 3D model and those obtained by simulation. In particular, our model captures the linear relation between the offered load and goodput when the network operates in light traffic conditions. It also captures the presence of an overshoot region characterizing the transition from light to heavy traffic regimes[2]. Interestingly, this overshoot becomes more pronounced as the number of STAs in the network increases and as the buffer length and the packet error probability decrease. Furthermore, although the lower the packet error rate and

[2] As stated by Liu *et al.* in [35], the fact that there exists a value of λ optimizing the goodput could be exploited by using some form of traffic control or IEEE 802.11e QoS mechanisms.

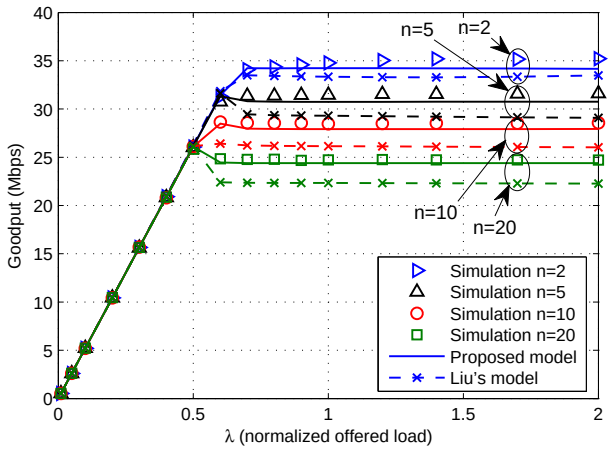


Figure 7.9: Goodput vs normalized offered load ($L = 30$, $p_e = 0.2$, $CW_{\min} = 3$).

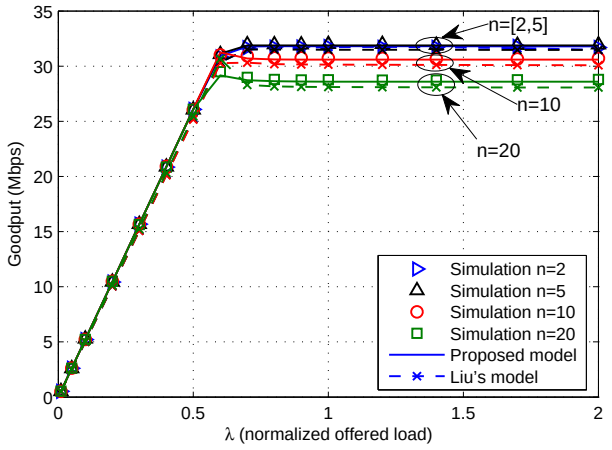


Figure 7.10: Goodput vs normalized offered load ($L = 30$, $p_e = 0.2$, $CW_{\min} = 15$).

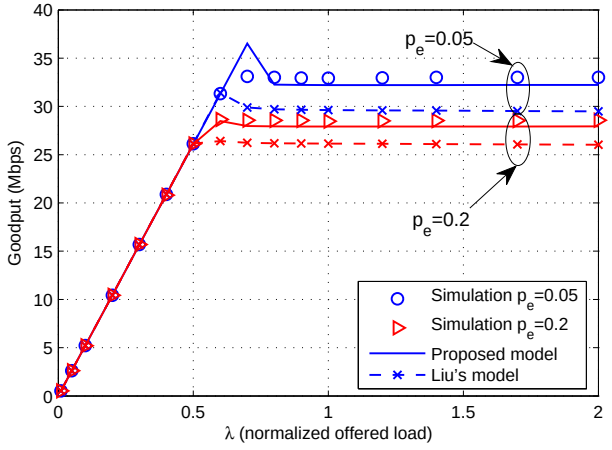


Figure 7.11: Goodput vs normalized offered load ($n = 10$, $L = 30$, $CW_{\min} = 3$).

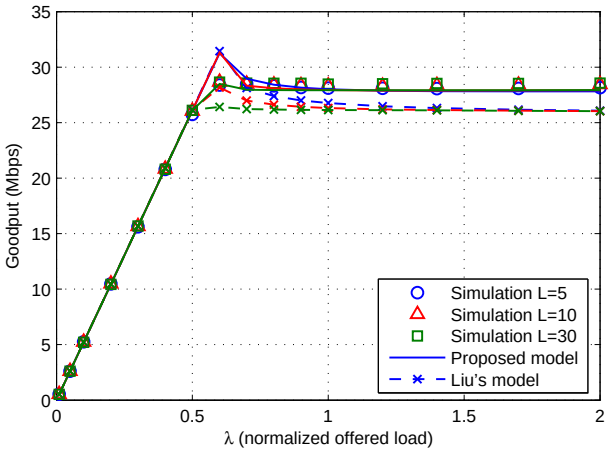


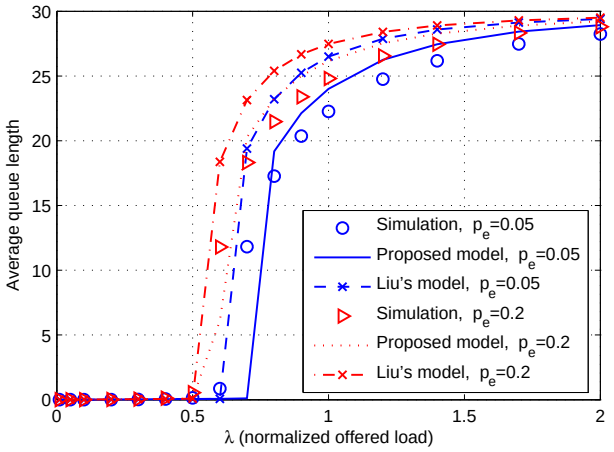
Figure 7.12: Goodput vs normalized offered load ($n = 10$, $p_e = 0.2$, $CW_{\min} = 3$).

the lower the number of STAs the higher is the value of λ at which the maximum of the goodput occurs, this value is quite insensitive to the buffer size. Our model also provides accurate estimates of the saturation goodput as predicted by Chen's two-dimensional model [28] and simulations. Liu's model, however, as it does not take into account the occurrence of anomalous slots, is unable to provide accurate predictions of this saturation goodput values. These shortcomings become more apparent as the minimum contention window decreases, as it can be observed when comparing Figs. 7.9 and 7.10.

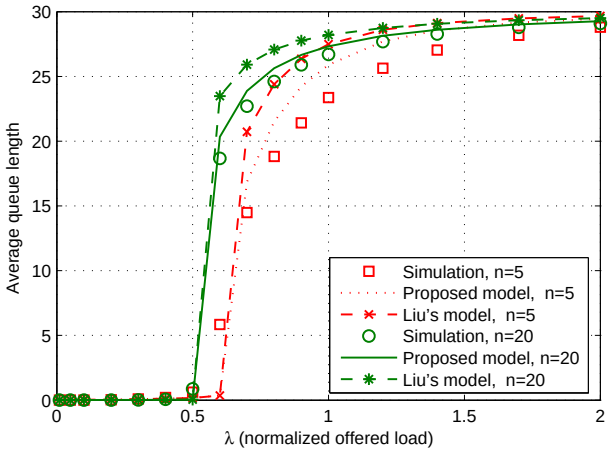
The rise in the number od STAs in the network leads to an increase in the number of collisions as it can be observed in Figs. 7.6b and 7.6c, thus resulting in a clear goodput degradation (see Figs. 7.9 and 7.10). The shorter the minimum congestion window is the more evident this degradation becomes. The increase in the probability of error, as shown in Fig. 7.11, has a similar effect. In contrast, goodput is rather insensitive to the queue length (see Fig. 7.12). In fact, once the buffers are full, whatever their length, there will almost always be a packet ready for transmission and consequently the overall goodput will not change; however, as it will be shown over the next subsections, the change of the buffer length has a dramatic impact on QoS metrics such as the blocking probability of the average packet delay. Again, our refined 3D model is shown to accurately match the simulated results, clearly showing the effects of the use of finite buffers and outperforming predicted results obtained using Liu's model.

7.4.3 Average queue length and packet loss analysis

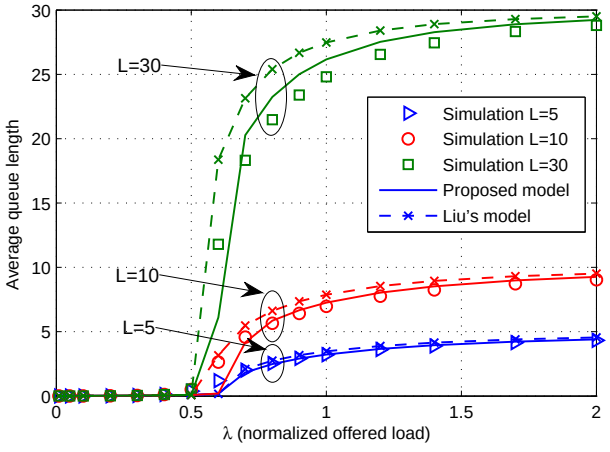
In IEEE 802.11 networks, a packet is lost when either it is discarded or blocked, conditions arising after R unsuccessful transmissions or after a buffer overflow due to the lack of space in the STA queue, respectively. Note that a packet that has been lost at the MAC sublayer can be recovered using upper layers mechanisms such as ARQ, which retransmit the packets that have been dropped or blocked at the cost of higher delays. The performance evaluation of these upper layer mechanisms is out of the scope of this paper, however, their application within the context of the IEEE 802.11 MAC



(a) $p_e \in \{0.05, 0.2\}$, $n = 10$, $L = 30$, $CW_{\min} = 3$

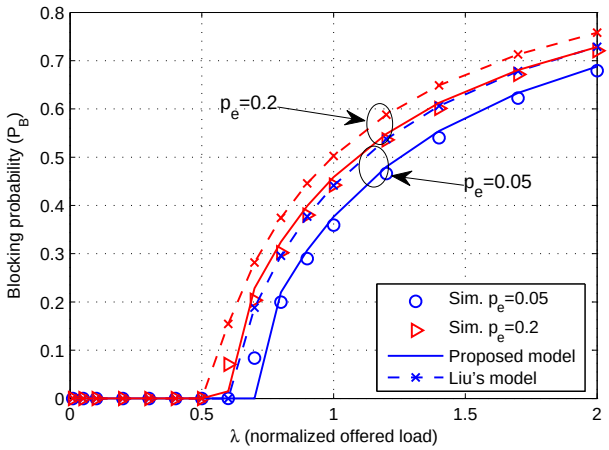


(b) $n \in \{5, 20\}$, $p_e = 0.2$, $L = 30$, $CW_{\min} = 3$

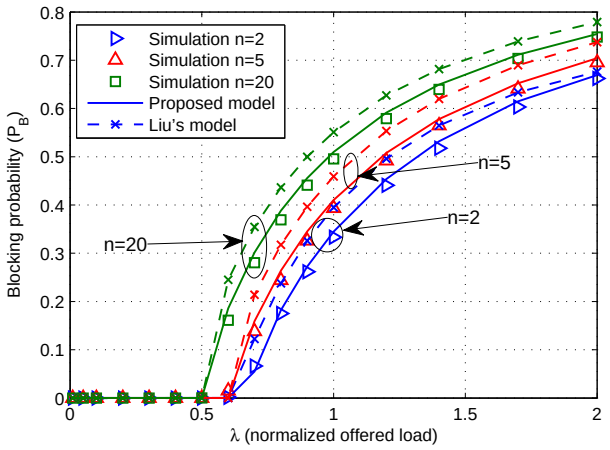


(c) $L \in \{5, 10, 30\}$, $n = 10$, $p_e = 0.2$, $CW_{\min} = 3$

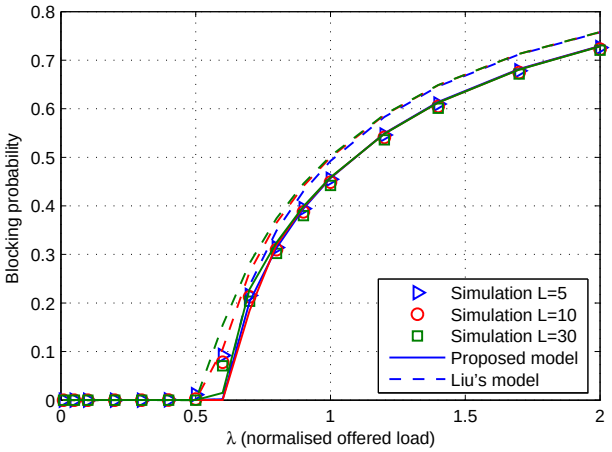
Figure 7.13: Average queue length vs normalized offered load



(a) $p_e \in \{0.05, 0.2\}$, $n = 10$, $L = 30$, $CW_{\min} = 3$

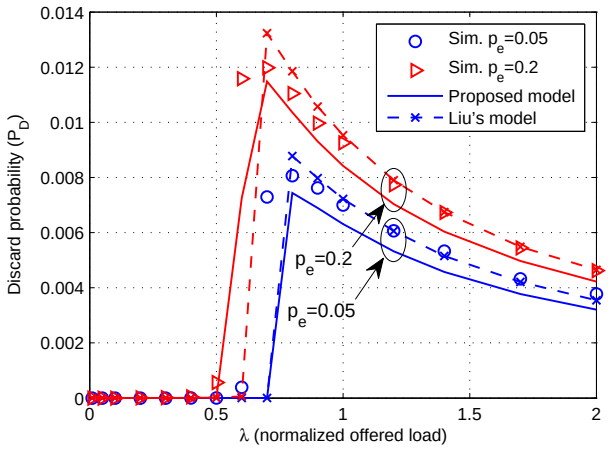


(b) $n \in \{2, 5, 20\}$, $p_e = 0.2$, $L = 30$, $CW_{\min} = 3$

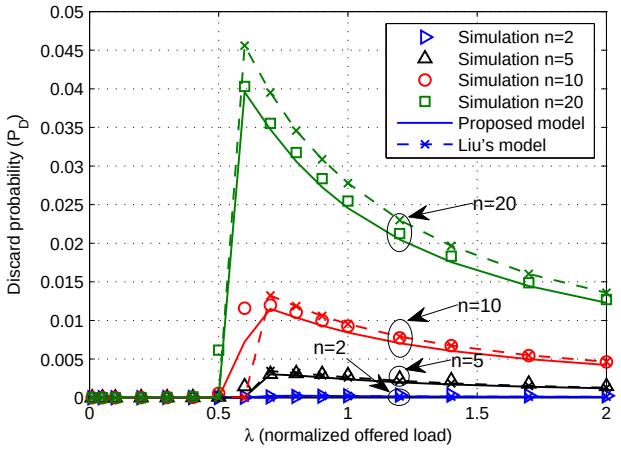


(c) $L \in \{5, 10, 30\}$, $n = 10$, $p_e = 0.2$, $CW_{\min} = 3$

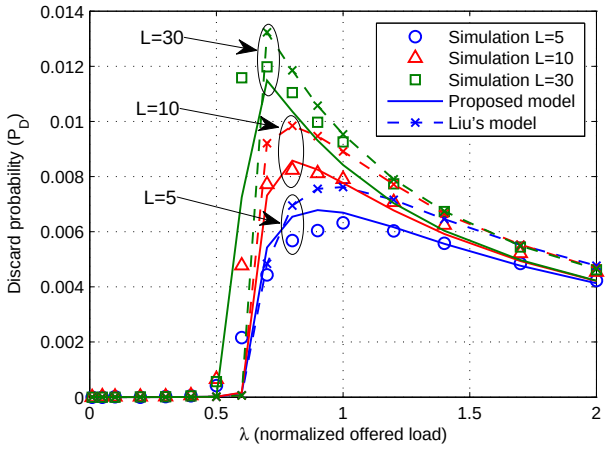
Figure 7.14: Blocking probability vs normalized offered load



(a) $p_e \in \{0.05, 0.2\}$, $n = 10$, $L = 30$, $CW_{\min} = 3$



(b) $n \in \{2, 5, 10, 20\}$, $p_e = 0.2$, $L = 30$, $CW_{\min} = 3$



(c) $L \in \{5, 10, 30\}$, $n = 10$, $p_e = 0.2$, $CW_{\min} = 3$

Figure 7.15: Discard probability vs normalized offered load

sublayer can constitute an interesting avenue for future research.

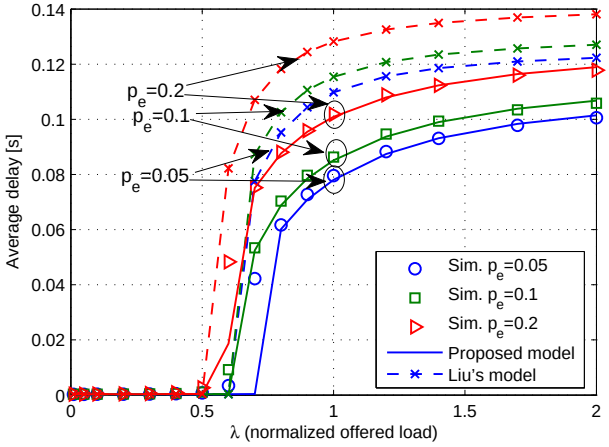
Figures 7.13-7.15 show predicted and simulated results for the average queue length, as well as the blocking and dropping probabilities, measured against normalized offered load for network configurations with a wide variety of packet error rates, number of STAs in the network and buffer lengths. It can be observed that these QoS metrics are all near zero for very light traffic conditions ($\lambda \leq 0.5$), even for very small buffer sizes. This is because for low offered loads, arriving packets are almost instantaneously transmitted and do not accumulate in the queues. For normalized offered loads characterizing the transition regime from light to heavy traffic loads ($0.5 < \lambda \leq 0.8$), these QoS performance metrics undergo an abrupt increase in response to the sudden increase in the probability of collision experienced in the network. For heavy traffic conditions ($\lambda > 0.8$), the average queue length tends to the maximum queue size L, the blocking probability keeps raising since most arriving packets find full queues, and the dropping probability smoothly decays towards zero because the total number of discarded packets remains virtually fixed while the number of arriving packets keeps growing. This means that, under network congestion, the packet losses at the MAC/PHY layers are mainly due to buffer overflows rather than to channel errors and collisions.

The average queue length, as well as the blocking and dropping probabilities, increase with n and p_e. When increasing L, however, although the average queue length and the dropping probability also increase, the blocking probability remains virtually invariable, since larger buffer sizes do not translate to an increase in the number of packets served by the network. It is worth pointing out that our analytical model captures the behavior of these QoS performance metrics and accurately predicts the network performance in the transition regime from light to heavy traffic loads. It also helps predicting the saturation tendency for these QoS metrics and remarkably, provided results are clearly more accurate than those obtained using previous analytical approaches [31, 35, 104].

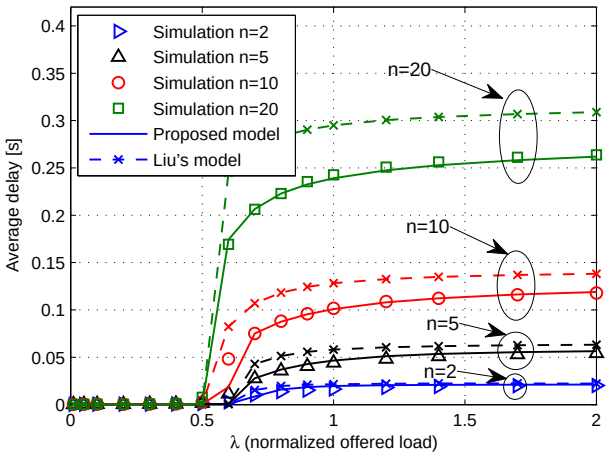
7.4.4 Delay performance

Predicted and simulated average delay experienced by a packet after joining the STA queue is shown in Fig. 7.16 as a function of the overall normalized offered load, under a variety of network configurations. Results obtained with our refined 3D model exhibit a close match with those obtained by simulation, contrasting with those obtained using Liu's model that clearly overestimate the measured average packet delay. In fact, it is interesting to remark that successful packet transmissions taking place in an anomalous slot, not considered in Liu's approach, result in a clear reduction of the average packet delay, since all these packets avoid the time spent in channel access contention periods.

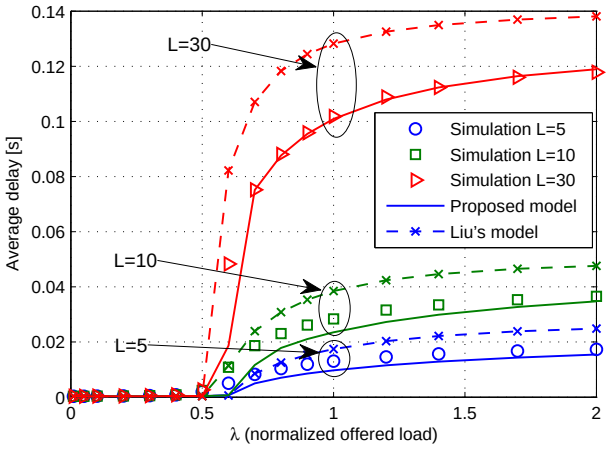
Results presented in Fig. 7.16 show that increasing the buffer size dramatically increases the average delay under network congestion. Moreover, observing results presented in Figs. 7.12 and 7.14, this increment in buffer size does not reduce the packet blocking probability nor improves the overall system goodput. Therefore, it seems that there may be little benefit in designing IEEE 802.11 networks with very large buffers.



(a) $p_e \in \{0.05, 0.1, 0.2\}$, $n = 10$, $L = 30$, $CW_{\min} = 3$



(b) $n \in \{2, 5, 10, 20\}$, $p_e = 0.2$, $L = 30$, $CW_{\min} = 3$



(c) $L \in \{5, 10, 30\}$, $n = 10$, $p_e = 0.2$, $CW_{\min} = 3$

Figure 7.16: Average delay vs normalized offered load

7.5 Chapter Summary

An analytic model of the IEEE 802.11 MAC sublayer that evaluates the system performance under saturated and non-saturated conditions has been presented. In contrast to previous work on this topic, our new proposed model considers the effects of anomalous slots and channel errors, according to the IEEE 802.11 standard. The resulting Markov chain model presents a two-level hierarchical structure, with the low-level intra-anomalous slot states, characterized by asynchronous transitions, and the high-level states, characterized by their model slot time transition synchronicity. Furthermore, the so-called refined Collapsed Transition onto Basis (CTB) approach has been proposed in order to simplify the analytical complexity of the method used to obtain the steady-state probabilities of the Markov chain.

Making use of the steady-state probabilities of the Markov chain, mathematical expressions to estimate QoS performance metrics such as the Goodput, the average queue length, the blocking and discard probabilities, and the average packet delay have been derived. Finally, the model has been validated by comparing analytical results for different network configurations with those obtained by simulation and those obtained using previous analytical approaches (i.e., Liu's *et al.* 3D model). Our refined 3D Markov chain model has been shown to accurately match the simulated results and remarkably, it provides results clearly more accurate than those obtained using previous analytical approaches.

CHAPTER 8

FAST LINK ADAPTATION UNDER NON-SATURATED CONDITIONS

This chapter introduces a semi-analytical approach to evaluate how the use of various AMC strategies affects the system performance under both saturated and non-saturated conditions. In particular, ARF- and FLA- based AMC algorithms are thoroughly studied, hence validating the accuracy of the model. Furthermore, it is shown that, when using appropriate mechanisms to deal with delayed channel state information, and irrespective of the traffic load, FLA-based strategies outperform ARF-based ones.

8.1 Motivation and problem statement

WLAN networks often do not work under saturated conditions, since real traffic flows are nonpersistent and do not lead to true saturation [105]. Therefore, it is important to derive a model accounting for practical network operations.

Recalling Markov chain introduced in Chapter 7, the assumption of saturated stations notably simplifies the analysis avoiding the queue buffer analysis, resulting in the removal of the the third dimension of the Markovian Chain model corresponding to the queue buffer evaluation and leaving only backoff stages and backoff counter dimensions. However, as stated in [35], [104], [106] and Chapter 7, their application to non-saturated conditions does not provide accurate estimation results for the metrics of interest.

Our aim in this chapter is to provide a semi-analytic model to evaluate the performance of WLAN networks under more realistic conditions, in particular, considering the application of AMC algorithms under saturated and non-saturated conditions. This model is an extension of the analysis provided in Chapter 7, that incorporates the restriction of finite buffers hence allowing the measurement of other performance metrics, such as blocking probability or packet delay. Furthermore, the new adaptive strategy previously detailed in Chapters 4 and 6 to cope with outdated CSI is now

evaluated under unsaturated conditions in order to validate its affectivity within this context.

This chapter is structured as follows. Firstly, Section 8.2 briefly reviews the evaluated AMC algorithms, whose operation was thoroughly detailed in Chapter 6. In Section 8.3, a semi-analytical model that estimates the system performance when using AMC algorithms under non-saturated conditions. Section 8.4 presents the performance results of AMC algorithms under non-saturated conditions, that serves to evaluate their performing under such conditions and, moreover, validate the proposed semi-analytical method. Finally, the conclusions of this chapter are given in Section 8.5.

8.2 Adaptive modulation and coding algorithms

ARF and FLA are the AMC mechanisms extensively evaluated in this chapter. In particular, FLA has been investigated with the 2 post-timeout strategies introduced in Chapter 6, namely, the *DownRate* mechanism and the RTS/CTS mechanism. For completeness, and to allow comparisons to be drawn, the ideal MCS feedback scenario has also been examined.

8.3 Semi-analytical modelling approach

The availability of multiple AMC transmission modes modifies the conditions used in the analytical model previously presented in Sections 7.2 and 7.3.

In the single-MCS model, errors and collisions were jointly treated as a single event, since they have the same duration. However, with the use of AMC algorithms each STA can now employ distinct MCSs based on the AMC selections, hence varying the duration of error and collision events. Consequently, T_{T_s}, T_{T_c} and T_{T_e}, defined as the duration of successful, collision and, instead, error transmissions, are no longer fixed and their average value can be obtained as:

$$\overline{T}_{T_x} = \sum_{m=0}^{N_{MCS}-1} P_{MCS}^{x}(m)\, T_{T_x}(m) \tag{8.1}$$

where $x \in \{s, c, e\}$ denotes the event type (success (s), error (e) or collision (c)), N_{MCS} is the number of available MCSs, $P_{MCS}^{x}(m)$ is the probability of using MCS m when evaluating event x and $T_{T_x}(m)$ corresponds to the time lapsed by event x when using MCS m.

In this chapter, note that $P_{MCS}^{s}(m)$, $P_{MCS}^{c}(m)$ and $P_{MCS}^{e}(m)$ can be analytically evaluated, whereas $\overline{T}_{T_s}$, $\overline{T}_{T_c}$ and $\overline{T}_{T_e}$ values are obtained by simulation, thus making this model semi-analytical.

The corresponding packet arrival probabilities during these events can be expressed as:

$$\begin{aligned} q_{T_s} &= r_q \overline{T}_{T_s}, \\ q_{T_e} &= r_q \overline{T}_{T_e}, \\ q_{T_c} &= r_q \overline{T}_{T_c}. \end{aligned} \tag{8.2}$$

The fact that T_{T_c} and T_{T_e} can be different leads to the necessity of distinguishing when each of these events is triggered. To this end, let us define the probability p of collision or error as,

$$p = p_c + p_e - p_c p_e. \tag{8.3}$$

where p_e is obtained by simulation and p_c is defined by

$$p_c = 1 - (1 - \tau)^{n-1}. \tag{8.4}$$

Due to these new time definitions, the *model slot time* presented in section 7.2 needs to be refined, by splitting errors and collisions into two separated events. Notwithstanding these modifications, the new model does not suffer important variations with respect to the original approach thoroughly described and evaluated in Chapter 7. Nonetheless, this division also affects, albeit slightly, the 3D Markov Chain transition probabilities presented in Fig. 7.1. In particular, the modifications of the 3D Markov chain are:

- State transitions caused by errors and collisions (that happen with p probability) are now divided into two different state transitions, error and collision transitions. The transiting state probability caused by a collision is defined by p_c while the probability of transiting due to an error is $p_e(1 - p_c)$.
- Packet arrival probabilities referred to collision and error events (defined by q_{T_e} in the previous chapter) are also now differentiated into two different packet arrival rates, q_{T_c} and q_{T_e}, denoting the probability of packet arrival during a collision and error, respectively.
-

state transitions caused by collisions and/or errors with probability p must be replaced by different events with probabilities p_c and $p_e(1 - p_c)$, respectively, and the corresponding transitions should each be divided into two transitions, p_c and p_e. Lastly, the packet arrival probability should be set accordingly to collision and error events, respectively. As it will be later presented, this changes define new transition probabilities inside the Markov chain states.

8.3.1 Refinements on the model slot time

The *model slot time* of this chapter is a refined version of the one presented in Sec. 7.2, splitting the error and collision event used in previous model into two different events. Therefore, the refined *model slot time* now consists of:

- An idle backoff slot. This situation arises when all active STAs in the system are either Idle (without any packet in the queue) or decrementing their backoff counters.
- A time interval including a failed transmission due to collision, followed by an EIFS interval plus an extra backoff slot.
- A time interval including a failed transmission due to a channel error, followed by an EIFS interval plus an extra backoff slot.

- A time interval including one or more consecutive successful transmissions and a final extra backoff slot. Notice that in the first successful transmission all active STAs in the system contend for the channel, while in the successive transmissions only the transmitting STA is granted access to the channel.

- A time interval including one or more consecutive successful transmissions. This case arises whenever the last transmission empties the queue and the backoff algorithm selects a random backoff counter equal to 0.

- A time interval including one successful transmission (contending with all other STAs in the system), zero or more successful transmissions (in an error-prone channel without contendings STAs), and an erroneous transmission (due to channel errors), followed by an EIFS interval plus an extra backoff slot.

8.3.2 Refinements on the low-level transition and steady-state probabilities

This subsection only presents the transition probabilities of those cases which have been modified with respect to those presented in Sect. 7.2.

Transitions from transmission states $(h,i,0;h,i,0)$

Let us assume that the STA is in state $(h,i,0;h,i,0)$, with $h \in \{1,\ldots,L-1\}$ denoting the number of packets in the STA queue with L corresponding to its maximum packet capacity, and $i \in \{0,\ldots,R-1\}$ defining the current backoff index stage with R denoting the maximum number of retransmissions. As stated in Chapter 5, when transmitting the first MPDU after entering an anomalous slot, all active STAs in the system are contending for the medium. In this case, if the transmission of the HOL packet fails, with probability $p = p_c + p_e - p_c p_e$, the STA will move to state $(h,i+1,k)$, with $k \in \{0,\ldots,W_{i+1}-1\}$ denoting the current backoff index, if the source does not generate a new packet during this time slot. Alternatively, if a packet arrival event occurs, the STA will move to state $(h+1,i+1,k)$. Hence, the corresponding non null one-step transition probabilities are

$$P\{h+1,i+1,k|h,i,0;h,i,0\} = \frac{p_c q_{T_c} + p_e(1-p_c)q_{T_e}}{W_{i+1}}$$
$$P\{h,i+1,k|h,i,0;h,i,0\} = \frac{p_c(1-q_{T_c}) + p_e(1-p_c)(1-q_{T_e})}{W_{i+1}}.$$

8.3.3 Refinements on the high-level transition probabilities

Due to the changes experienced on the transition state probabilities explained in the previous subsection, some high-level transition probabilities are modified. These modified high-level transition probabilities can be straightforwardly obtained using the low-level transition and steady-state probabilities.

Transitions from transmission states $(h,i,0)$

Case I ($i \in \{1,\ldots,R-1\}$)

$$\begin{aligned} P\{h+1,i+1,k|h,i,0\} &= \frac{p_c q_{T_c} + p_e(1-p_c)q_{T_e}}{W_{i+1}}, \\ P\{h,i+1,k|h,i,0\} &= \frac{p_c(1-q_{T_c}) + p_e(1-p_c)(1-q_{T_e})}{W_{i+1}}, \end{aligned} \tag{8.5}$$

for $h \in \{1,\ldots,L-1\}$ and $k \in \{0,\ldots,W_{i+1}-1\}$, and

$$P\{L,i+1,k|L,i,0\} = \frac{p}{W_{i+1}}, \tag{8.6}$$

for $h = L$ and $k \in \{0,\ldots,W_{i+1}-1\}$.

Case II ($i = 0$)

$$P\{h+1,1,k|h,0,0\} = \frac{p_c q_{T_c} + p_e(1-p_c)q_{T_e}}{W_1} + \frac{p_e q_{T_e}}{W_1}\frac{b(h,0^+,0;h,0,0)}{b(h,0,0;h,0,0)}, \tag{8.7}$$

for $h \in \{1,\ldots,L-1\}$ and $k \in \{0,\ldots,W_1-1\}$,

$$\begin{aligned} P\{h,1,k|h,0,0\} = &\frac{p_c(1-q_{T_c}) + p_e(1-p_c)(1-q_{T_e})}{W_1} + \frac{p_e(1-q_{T_e})}{W_1}\frac{b(h,0^+,0;h,0,0)}{b(h,0,0;h,0,0)} \\ &+ \frac{p_e q_{T_e}}{W_1}\frac{b(h-1,0^+,0;h,0,0)}{b(h,0,0;h,0,0)}, \end{aligned} \tag{8.8}$$

for $h \in \{2,\ldots,L-1\}$ and $k \in \{0,\ldots,W_1-1\}$,

$$P\{1,1,k|1,0,0\} = \frac{p_c(1-q_{T_c}) + p_e(1-p_c)(1-q_{T_e})}{W_1} + \frac{p_e(1-q_{T_e})}{W_1}\frac{b(1,0^+,0;1,0,0)}{b(1,0,0;1,0,0)}, \tag{8.9}$$

for $h = 1$ and $k \in \{0,\ldots,W_1-1\}$, and finally,

$$P\{L,1,k|L,0,0\} = \frac{p}{W_1} + \frac{p_e}{W_1}\frac{b(L,0^+,0;L,0,0)}{b(L,0,0;L,0,0)} + \frac{p_e q_{T_e}}{W_1}\frac{b(L-1,0^+,0;L,0,0)}{b(L,0,0;L,0,0)}, \tag{8.10}$$

Case III ($i = R$)

$$\begin{aligned} P\{h,0^-,k|h,R,0\} &= \frac{p_c q_{T_c} + p_e(1-p_c)q_{T_e}}{W_0}, \\ P\{h-1,0^-,k|h,R,0\} &= \frac{p_c(1-q_{T_c}) + p_e(1-p_c)(1-q_{T_e})}{W_0}, \end{aligned} \tag{8.11}$$

for $h \in \{1,\ldots,L\}$ and $k \in \{0,\ldots,W_0-1\}$.

8.3.4 Modifications to the refined CTB approach

In this model, the expected time spent per Markov state, assuming a symmetric n-station network and dividing error and collision events, can be obtained as

$$E_s^{(n)} = P_i^{(n)}\delta + P_s^{(n)}\overline{T}_s + P_c^{(n)}\overline{T}_{T_c} + P_e^{(n)}\overline{T}_{T_e}, \tag{8.12}$$

where

$$\begin{aligned} P_i^{(n)} &= (1-\tau)^n, \\ P_s^{(n)} &= n\tau(1-p_e)(1-\tau)^{n-1}, \\ P_e^{(n)} &= n\tau p_e(1-\tau)^{n-1}, \\ P_c^{(n)} &= 1 - P_i^{(n)} - P_e^{(n)} - P_s^{(n)}, \end{aligned} \tag{8.13}$$

are used to denote, respectively, the probability that no STA is transmitting (empty media), the probability that only one STA is successfully transmitting (anomalous slot), the probability of suffering a packet transmission failure due to a packet error, and the probability of suffering a packet transmission failure due to a collision. Note that, due to STA independence, $E_{s|n-1}$ is equivalent to $E_s^{(n-1)}$.

Regarding the matrix definitions, only $\boldsymbol{F}_i$ presents some slight differences with respect to that presented in (7.36). In particular,

$$\begin{aligned} \boldsymbol{F}_i[h,h'] &= \boldsymbol{F}[h,h'] \\ &= \begin{cases} p_c(1-q_{T_c}) + p_e(1-p_c)(1-q_{T_e}), & h' \in \{1,\ldots,L-1\},\ h = h', \\ p_c q_{T_c} + p_e(1-p_c)q_{T_e}, & h' \in \{1,\ldots,L-1\},\ h = h'+1, \\ p, & h = h' = L, \\ 0, & \text{otherwise}, \end{cases} \end{aligned} \tag{8.14}$$

for $i \in \{2,\ldots,R\}$, and

$$\boldsymbol{F}_1 = \begin{bmatrix} \boldsymbol{0}_L & \boldsymbol{K}_0 \end{bmatrix}, \tag{8.15}$$

with

$$\boldsymbol{K}_0[h,h'] = W_1 P\{h,1,k|h',0,0\}. \tag{8.16}$$

8.3.5 Refinements on QoS performance analysis

MAC layer throughput expression

The MAC sublayer throughput can be now expressed as,

$$G = \frac{P_s^{(n)}\overline{P}}{P_i^{(n)}\delta + P_s^{(n)}\overline{T}_s + P_e^{(n)}\overline{T}_{T_e} + P_c^{(n)}\overline{T}_{T_c}}. \tag{8.17}$$

Blocking probability

The blocking probability P_B considering error and collision events separately is now defined as,

$$\begin{aligned} P_B = &\frac{E_s^{(n-1)}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \sum_{k=1}^{W_i-1} \hat{b}(L,i,k) \\ &+ \frac{p_c\overline{T}_{T_c} + p_e(1-p_c)\overline{T}_{T_e}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R-1} \hat{b}(L,i,0) \\ &+ p_e \frac{\overline{T}_{T_e}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \hat{b}(L,0,0;L,i,0) \\ &+ \frac{\delta}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \sum_{k=0}^{W_0-2} \hat{b}(L,0^+,k+1;L,i,0). \end{aligned} \tag{8.18}$$

Average queue length

The probability of the interface queue length being $Q = h$ at any given time is now defined as

$$\begin{aligned} P_t[Q=h] = {} & \frac{E_s^{(n-1)}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \sum_{k=1}^{W_i-1} \hat{b}(h,i,k) \\ & + \frac{p_c \overline{T}_{T_c} + p_e(1-p_c)\overline{T}_{T_e} + (1-p)\overline{T}_{T_s}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \hat{b}(h,i,0) \\ & + \frac{p_e \overline{T}_{T_e} + (1-p_e)\overline{T}_{T_s}}{\hat{E}_s^{(n)}} \sum_{h'=h}^{L} \sum_{i=0}^{R} \hat{b}(h,0,0;h',i,0) \\ & + \frac{\delta}{\hat{E}_s^{(n)}} \sum_{h'=h}^{L} \sum_{i=0}^{R} \sum_{k=0}^{W_0-2} \hat{b}(h,0^+,k+1;h',i,0), \end{aligned} \tag{8.19}$$

for $h \in \{1,\ldots,L\}$, and

$$\begin{aligned} P_t[Q=0] = {} & \frac{E_s^{(n-1)}}{\hat{E}_s^{(n)}} \sum_{k=0}^{W_i-1} \hat{b}(0,0,k) \\ & + \frac{\delta}{\hat{E}_s^{(n)}} \sum_{h'=1}^{L} \sum_{i=0}^{R} \sum_{k=0}^{W_0-2} \hat{b}(0,0^+,k+1;h',i,0), \end{aligned} \tag{8.20}$$

where the expected time per state, assuming a symmetric n-station network, can be obtained as

$$\begin{aligned} \hat{E}_s^{(n)} = {} & E_s^{(n-1)} \left[\sum_{k=0}^{W_0-1} \hat{b}(0,0,k) + \sum_{h=1}^{L} \sum_{i=0}^{R} \sum_{k=1}^{W_i-1} \hat{b}(h,i,k) \right] \\ & + \left[p_c \overline{T}_{T_c} + p_e(1-p_c)\overline{T}_{T_e} + (1-p)\overline{T}_{T_s} \right] \sum_{h=1}^{L} \sum_{i=0}^{R} \hat{b}(h,i,0) \\ & + \left[p_e \overline{T}_{T_e} + (1-p_e)\overline{T}_{T_s} \right] \sum_{h'=1}^{L} \sum_{i=0}^{R} \sum_{h=1}^{h'} \hat{b}(h,0,0;h',i,0) \\ & + \delta \sum_{h'=1}^{L} \sum_{i=0}^{R} \sum_{h=0}^{h'} \sum_{k=0}^{W_0-2} \hat{b}(h,0^+,k+1;h',i,0). \end{aligned} \tag{8.21}$$

Average packet delay

Finally, the latest changes on the analysis correspond to the modelling of the average packet delay. In particular, $P_t^{(T)}[Q=L]$ and Δ are now defined as

$$\begin{aligned} P_t^{(T)}[Q=L] = {} & \frac{p_c \overline{T}_{T_c} + p_e(1-p_c)\overline{T}_{T_e}}{\hat{E}_s^{(n)}} \hat{b}(L,R,0) \\ & + (1-p)\frac{\overline{T}_{T_s}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \hat{b}(L,i,0) \\ & + (1-p_e)\frac{\overline{T}_{T_s}}{\hat{E}_s^{(n)}} \sum_{i=0}^{R} \hat{b}(L,0,0;L,i,0), \end{aligned} \tag{8.22}$$

and

$$\Delta = \frac{\hat{E}_s^{(n)} \sum_{h=1}^{L} \left(P_t[Q=h] - \frac{p_c \overline{T}_{Tc} + p_e(1-p_c)\overline{T}_{Te}}{\hat{E}_s^{(n)}} \hat{b}(h,R,0) \right)}{\sum_{h=1}^{L} \left(P_{l \cap Q=h} - p\hat{b}(h,R,0) \right)}, \tag{8.23}$$

respectively.

8.4 Performance results

8.4.1 System configuration

In order to validate the semi-analytical model of this chapter, the WLAN Matlab simulator described in Chapter 6 has been extended to model the system performance under non-saturated conditions. This simulator allows us to compare the performance of FLA and ARF under circumstances not previously considered such as, different STA queue lengths or variable offered loads.

To this end, it has been configured a simulation tool employing the same system conditions presented in Sect. 6.6.

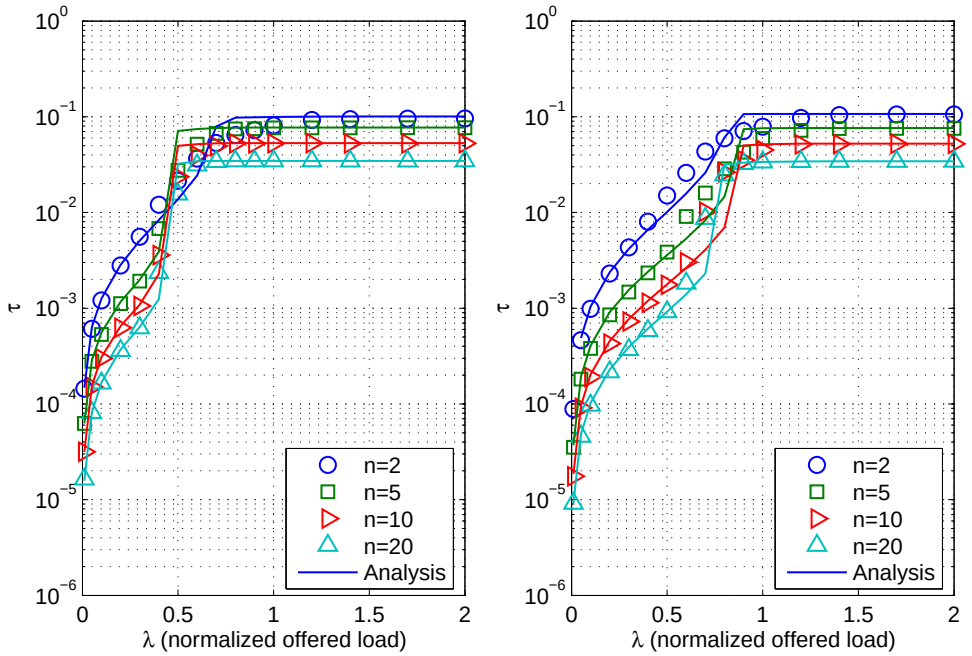
In this model, and due to its negligible probability of occurrence, the possibility of error in the ACK, RTS and/or CTS transmissions has not been considered. Finally, it should be stated that the packet arrival rate denoted by λ is defined as, $\lambda = r_q/(nT_{base})$, where $T_{base} = 500\mu s$ is a time value that fixes the arrival rate per second . In this case it no longer depends on T_{T_s}, since each AMC employs different T_{T_s} values, therefore, different λ values would be obtained if λ from Chapter 7 was used.

8.4.2 performance of p, τ and PER

Probabilities τ and p for each AMC algorithm are shown in Figs. 8.1 and 8.2, respectively, considering different number of users.

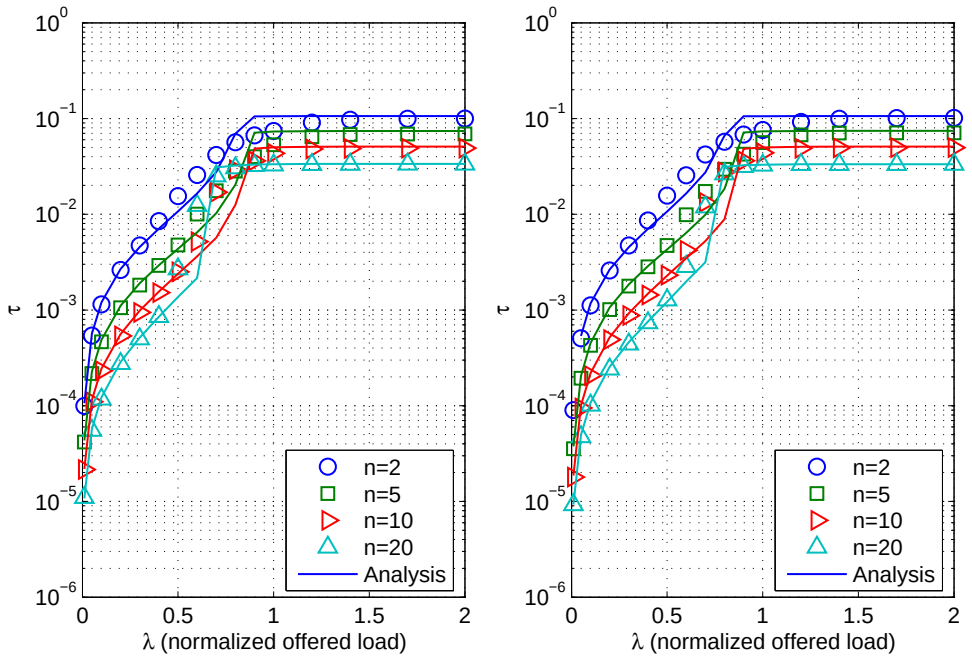
In these figures, note that the accuracy of τ and p estimations varies according to the current traffic conditions. For saturated conditions, p and τ are precisely estimated, providing the same performance values obtained by the 2-D Markov model in Chapter 6. For unsaturated traffic conditions, however, while the semi-analytical model precisely estimates τ, it slightly underestimates p. This slight loss of accuracy was already noticed in [106], where it was stated that analytical models for non-saturated conditions rooted on the Bianchi approach, do not provide precise results since slot events under non-saturated conditions are not independent. Nonetheless, they offer a good performance approximation, as it can also be observed in the aforementioned figures. Under unsaturated conditions, where most STAs have empty buffers, the collision probability varies according to the events of preceding slots. In particular, if one of these slots was occupied by one transmission, the probability of packet arrival on the other STAs is notably increased, hence increasing also the possibility of collision. Otherwise, in case of empty slots, the probability of packet arrival is reduced, thereby lowering the collision probability.

Additionally, Fig. 8.3 shows the packet error rate (PER) only considering channel errors (removing collisions) for each AMC technique and different number of users. Under non-saturated conditions, note that all realistic AMC algorithms experience elevated PER values, specially the ARF algorithm. Only the ideal FLA algorithm with



(a) ARF.

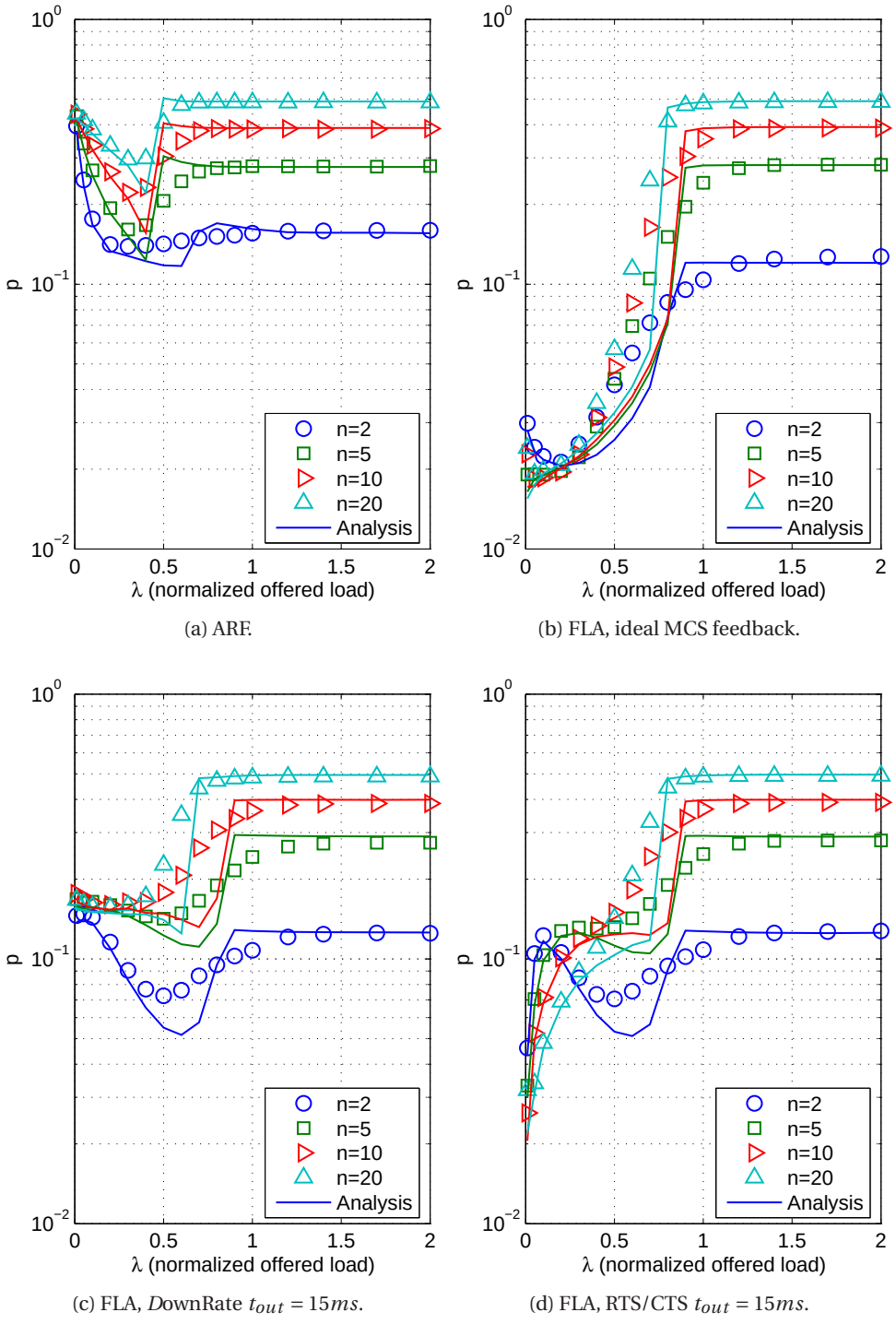
(b) FLA, ideal MCS feedback.



(c) FLA, *D*ownRate $t_{out} = 15ms$.

(d) FLA, RTS/CTS $t_{out} = 15ms$.

Figure 8.1: τ probability using $L = 30$ for different AMCs.



(a) ARF.

(b) FLA, ideal MCS feedback.

(c) FLA, *D*ownRate $t_{out} = 15ms$.

(d) FLA, RTS/CTS $t_{out} = 15ms$.

Figure 8.2: p probability using $L = 30$ for different AMCs.

non delayed feedback information achieves a low PER value irrespective of the offered load, since it selects the best MCS according to the current channel conditions with reduced PER value. At the same time, this elevate PER values in realistic AMCs influences the performance of p under non-saturated conditions, hence notably increasing its value, as it can be appreciated in Fig. 8.2. Note that for the ideal FLA, the p probability is still lower than any of the other AMCs, since the collision probability (p_c) under non-saturated conditions is notably lower than in saturated conditions. Consequently, most of lost packets under non-saturated conditions are caused by transmissions during unfavorable channel conditions for the selected MCS.

Under non-saturated conditions, the ARF algorithm suffers from the largest PER values (see Fig. 8.3a). ARF schemes adapts their transmission rate slowly to the channel conditions. Under non-saturated conditions, the time lapsed between arriving packets might notably overpass the channel coherence time, and then, current MCS does not match current channel conditions. Furthermore, the ARF algorithm after long periods of inactivity decides to increase the transmission rate without considering the current channel conditions, hence increasing the probability of error. As traffic load increases, and the system approaches saturation, the PER results converge to those presented in Chapter 6.

The FLA strategy with no MCS feedback delay performs as an ideal AMC with fixed p_e probability. In this case, PER probability remains fixed irrespective of the number of users (see Fig. 8.3), fulfilling the FLA algorithm PER constraint. For non-saturated conditions, p is only affected by channel errors, therefore, p is equal to the p_e value (see Chapter 4 for more details). For higher λ values, collisions increase the probability of p while the corresponding PER remains fixed.

FLA with *DownRate* and RTS/CTS strategies attain similar performance in terms of p because their post-timeout strategies are triggered after the same timeout value (15 ms). Nonetheless, RTS/CTS results in lower p and PER performance results since it selects the most suitable MCS for current conditions once the timeout has expired. In contrast, the *DownRate* strategy lowers the transmission rate irrespective of the channel conditions, hence not guaranteeing a successful transmission and requiring another retransmission in case of error. As previously stated in Chapter 6, the *DownRate* mechanism has difficulties in accurately tracking the channel variations.

Under non-saturated conditions, the retransmissions of a packet usually take place in a short time span, thus using the same MCS (since the timeout has not yet expired) and resulting in similar probability of error. In case of larger delays, the timeout will expire and perform one post-timeout strategy. For the RTS/CTS case, the most appropriate MCS with reduced PER value is selected. Otherwise, for the *DownRate* strategy, the STA will use another MCS with presumable lower probability of error than the previous, however, without ensuring a reduced error probability. Therefore, it is clear that the usage of non ideal MCS feedback information seriously degrade the MCS selection of FLA.

It should be clear at this point that FLA algorithms appropriately exploit the system capacity, hence outperforming the ARF algorithm. Furthermore, note that FLA with *DownRate* and RTS/CTS strategies do not fulfill the PER contraint at saturation when using t_{out} a value of 15 ms. A shorter timeout is required to fulfil the prescribed PER constraint. To this end, Fig. 8.3 introduces the performance of FLA with RTS/CTS strategy using a shorter timeout value of 13 ms. This shorter timeout value leads to the

fulfilling of the PER constraint.

8.4.3 Goodput performance

Figure 8.4 presents the system goodput of each aforementioned AMC technique for different number of users. This figure shows that the semi-analytical model accurately estimates the system performance, specially in denser users scenarios (n = 10 and n = 20). Note also that their goodput under saturation conditions matches the results presented in Chapter 6 for each AMC algorithm. Tiny performance differences are observed between FLA strategies.

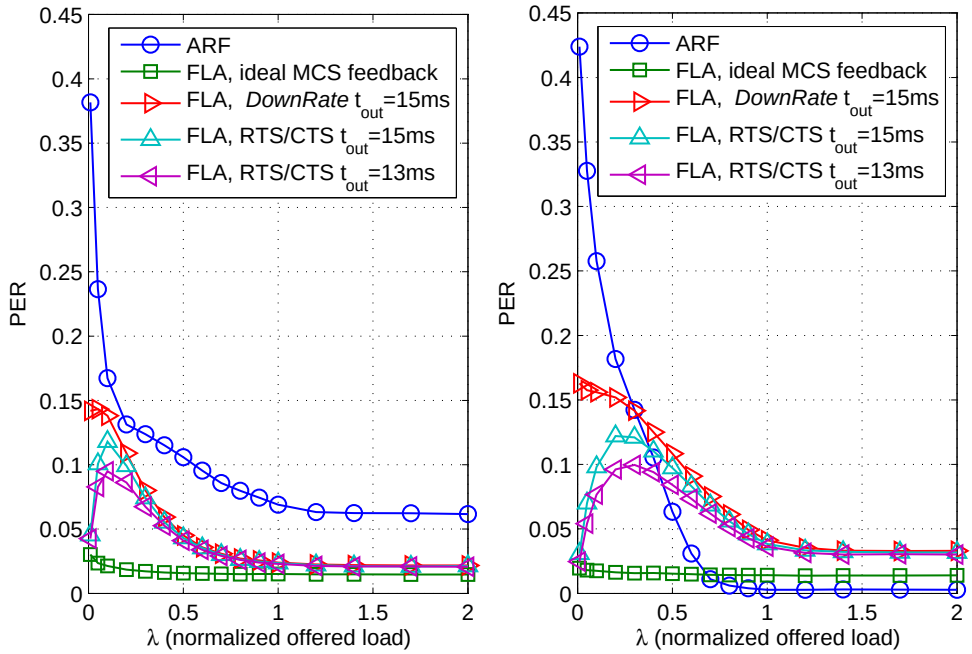
Under non-saturated conditions, all the AMC algorithms achieve identical goodput performance since they are able to serve the totality of incoming packets. For the FLA case, the goodput increases in accordance with the offered load, until it reaches the saturation goodput. In contrast, in ARF, the system goodput initially exceeds its saturation value, since the presence of collisions does not seriously affects the MCS selection and then higher goodputs are achieved. However, this decreases as λ increases, since more collisions are obtained hence degrading the MCS selections and lowering the goodput to the saturation stable value presented in Chapter 6.

8.4.4 Blocking probability, discard probability and queue occupancy

The average queue length (Q) for all AMCs considering scenarios with 10 and 20 users is depicted in Figs. 8.5a and 8.5b, respectively. In these figures, the semi-analytical approach accurately estimates the average queue length, hence validating the proposed model. Note that the utilization of ARF notably increases the average queue lengths with respect to the FLA algorithm. For the ARF case, the maximum queue length is reached at λ = 0.6 whereas for the FLA algorithm the maximum is near λ = 1.6. Furthermore, note that all evaluated post-timeout strategies of FLA achieve similar performance, as it was already described when evaluating τ or the system goodput. Although not presented in these figures, the use of other queue sizes do not bring along important differences on the performance of the various AMC algorithms.

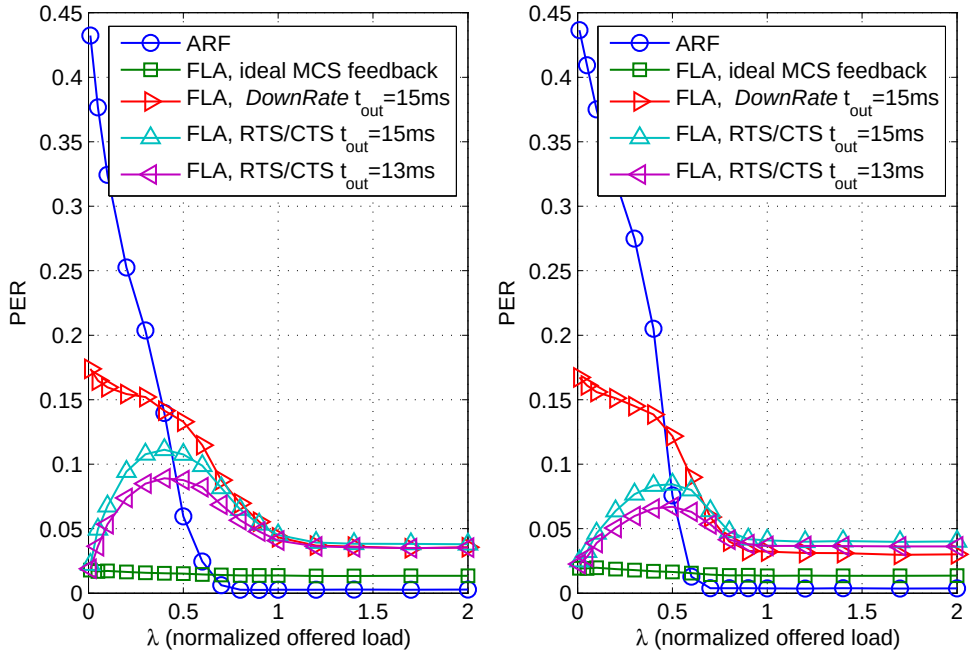
The blocking probability (P_B) of each AMC for different number of users is presented in Fig. 8.6. Similarly to previous metrics, P_B is accurately estimated by the semi-analytical model. In Figs. 8.6b, 8.6c and 8.6d, similar P_B performance is observed independently of the employed post-timeout strategies. Note that these P_B probabilities smoothly increases with the number of users, thus reducing the number of served packets per λ value, since more collisions are experienced. Notice how, except for n = 2, the ARF malfunction seriously affects the serving rate performance. Similarly to FLA, P_B is affected by the collisions, hence increasing with the number of users.

Figure 8.7 presents the discard probability (P_D) of each AMC technique using different number of users. In this figure, P_D is only precisely estimated at saturated conditions. Under non-saturated conditions, which is based on p estimation, P_D is underestimated. The discard probability over non-saturated conditions decreases as the offered load increases. Their values can be easily associated to the simulation p values. The higher the p values, the greater the probability of discard.



(a) $n = 2$.

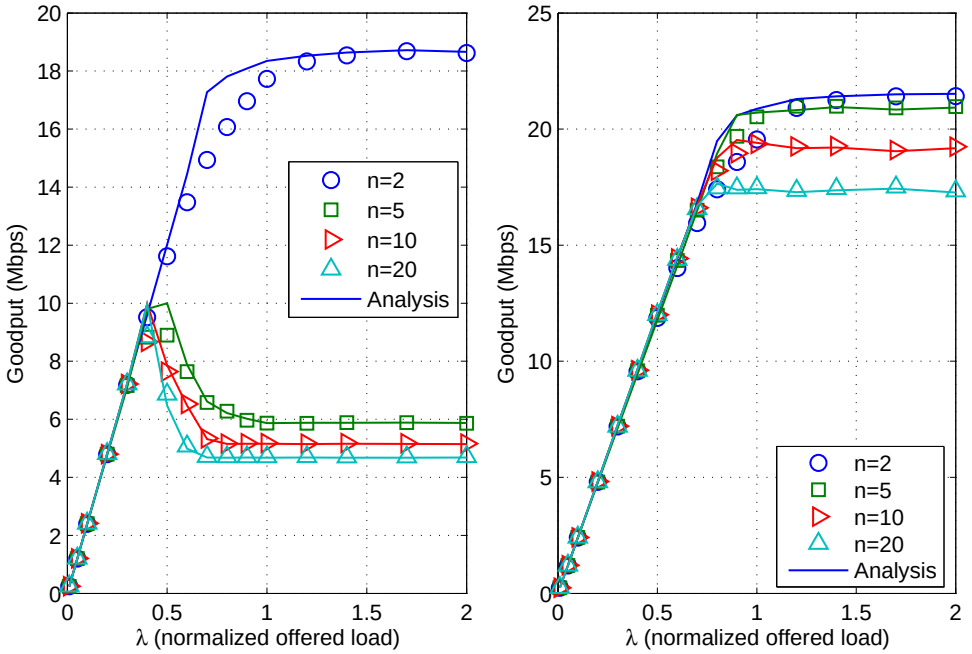
(b) $n = 5$.



(c) $n = 10$.

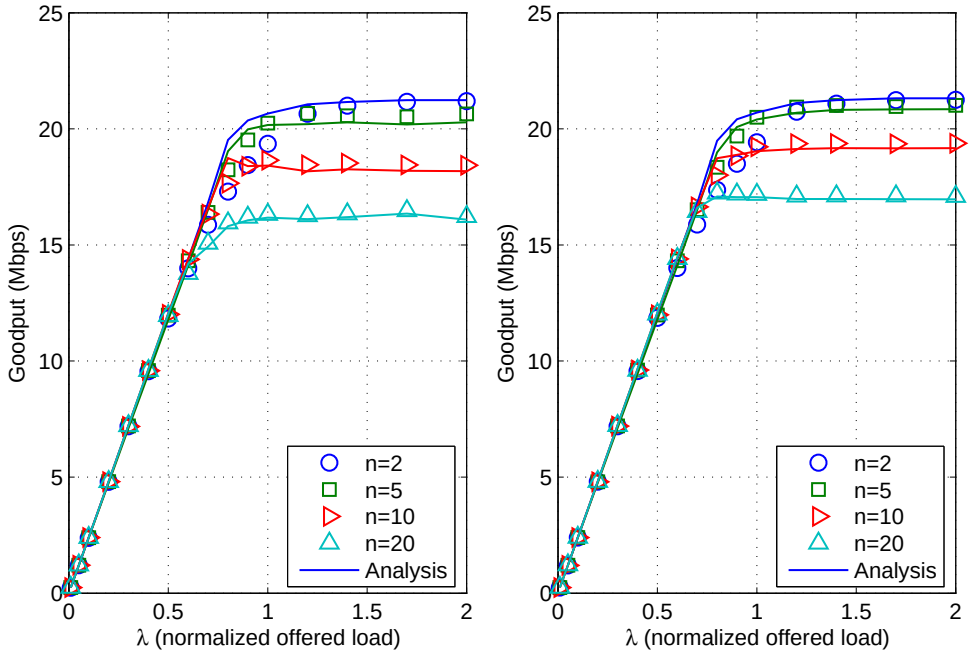
(d) $n = 20$.

Figure 8.3: PER using $L = 30$ for different AMCs.



(a) ARF.

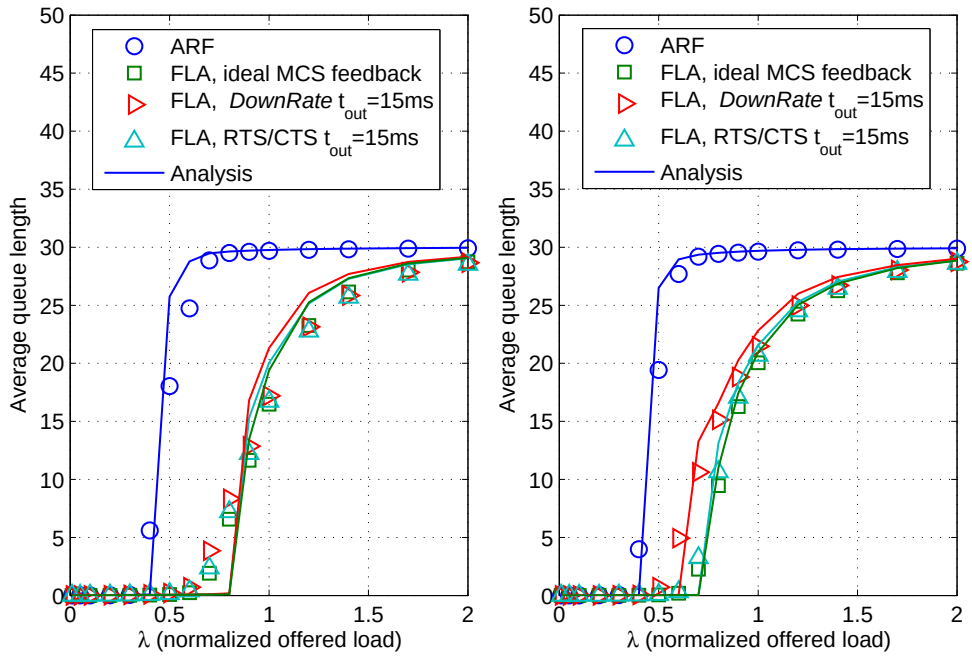
(b) FLA, ideal MCS feedback.



(c) FLA, *D*ownRate $t_{out} = 15ms$.

(d) FLA, RTS/CTS $t_{out} = 15ms$.

Figure 8.4: Goodput using $L = 30$ for different AMCs.



(a) n=10 (b) n=20

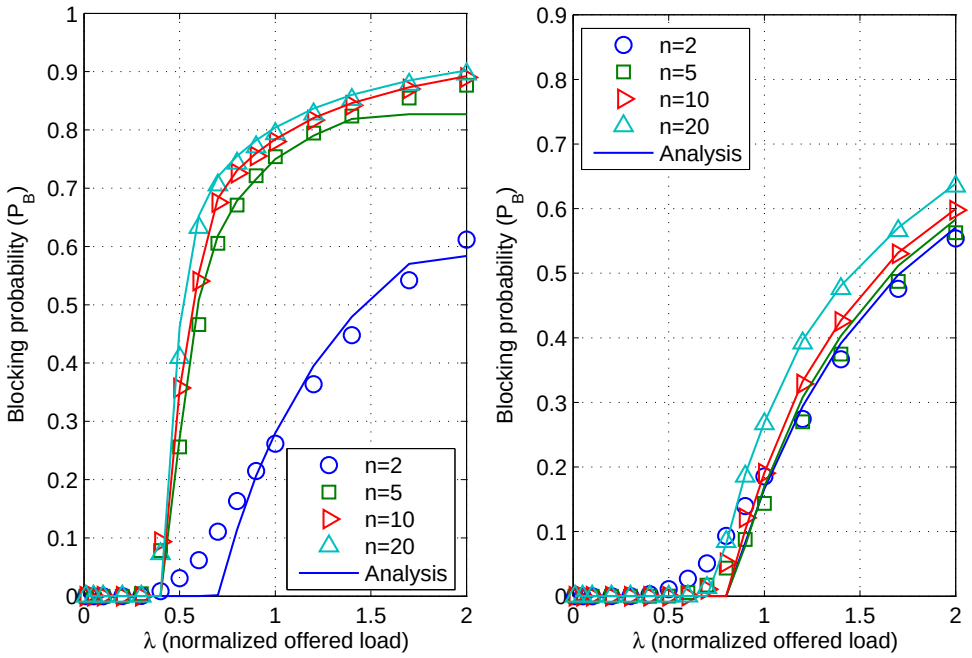
Figure 8.5: Average queue size using $L = 30$ for different number of users.

8.4.5 Delay

The packet delay of each AMC obtained by simulation and via the semi-analytical model is presented in Fig. 8.8 for different number of users and offered load. As expected, the delay increases with the number of users regardless of the considered AMC. Remarkably, for ARF, this is notably greater than the rest of FLA algorithms (by nearly a factor of 3). Similar delays are experienced by the rest of FLA strategies. Notice that, as explained previously, the usage of shorter queue lengths could reduce these packet delays, at the cost of higher blocking probabilities.

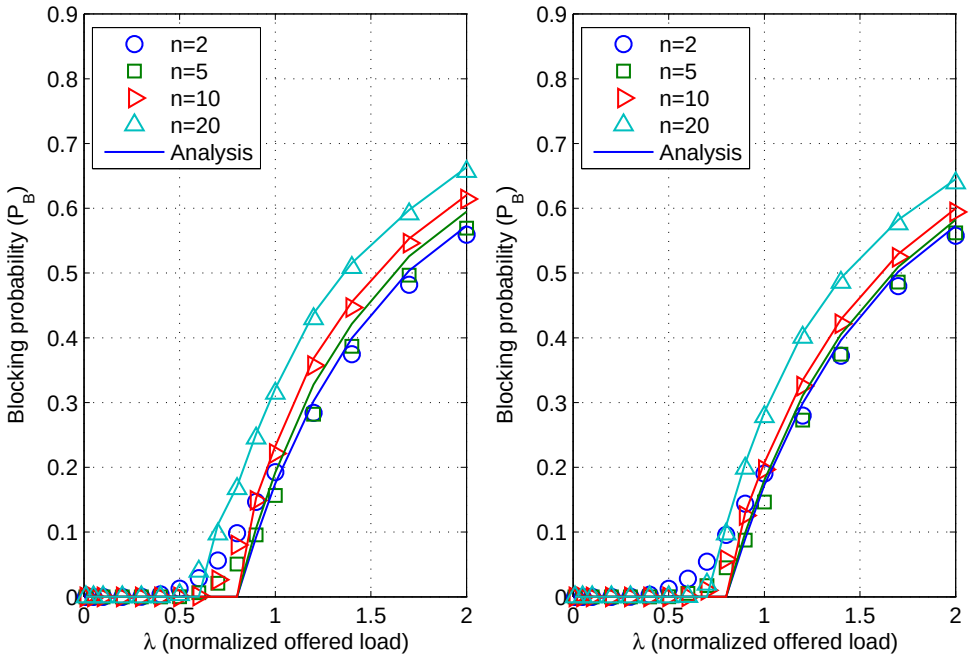
8.5 Conclusions

This chapter has expanded the semianalytical model of Chapter 6, derived for saturated performance, to be able to also handle the operation under no saturation. In particular, ARF and FLA AMC algorithms have been thoroughly studied, hence validating the accuracy of the model. Additionally, it has been shown that FLA improves the ARF system performance under no saturation irrespective of the traffic load, when using appropriate mechanisms to deal with delayed channel information, highlighting the importance of the timeout strategy to reduce the error rate.



(a) ARF.

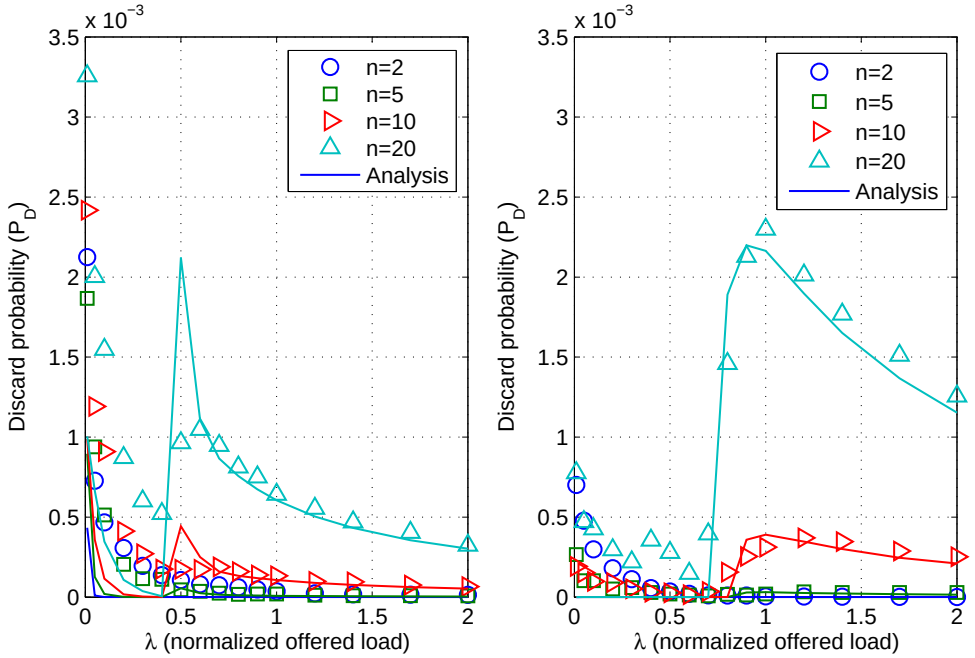
(b) FLA, ideal MCS feedback.



(c) FLA, *D*ownRate $t_{out} = 15ms$.

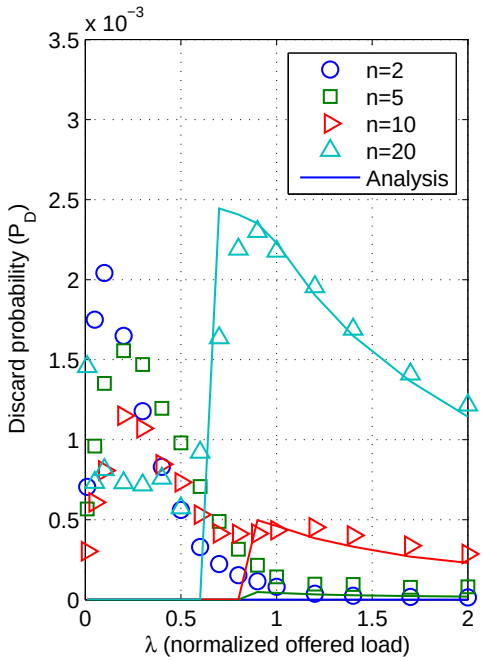
(d) FLA, RTS/CTS $t_{out} = 15ms$.

Figure 8.6: Blocking probability using $L = 30$ for different AMCs.

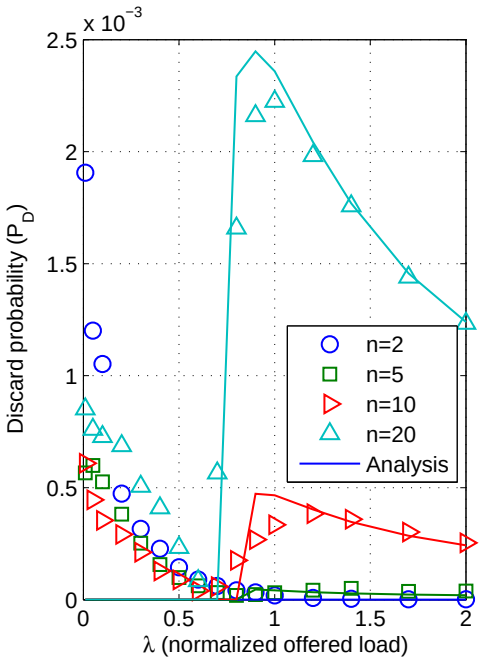


(a) ARF.

(b) FLA, ideal MCS feedback.

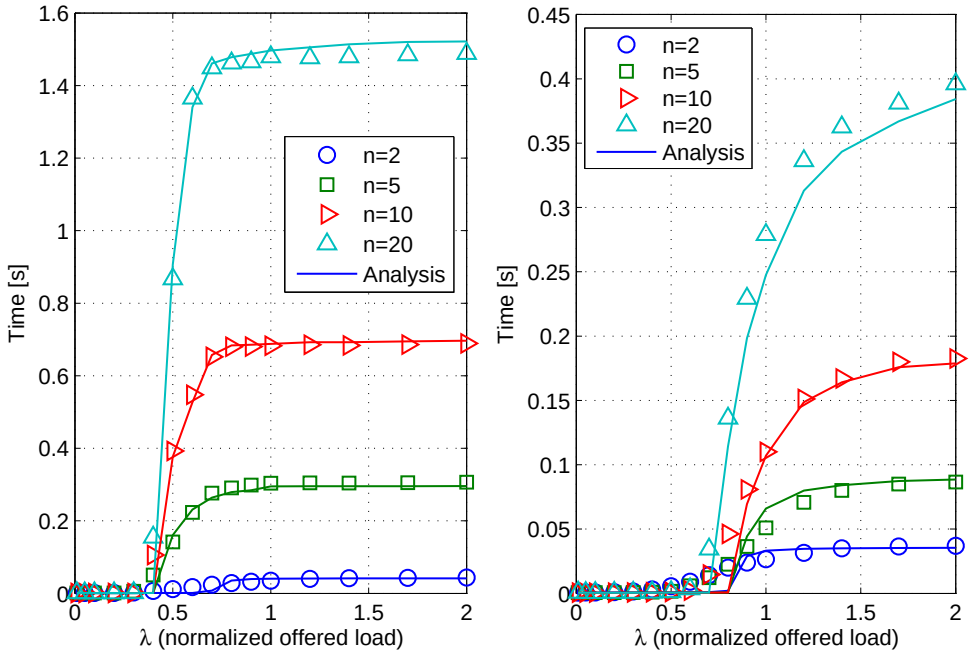


(c) FLA, *D*ownRate $t_{out} = 15ms$.



(d) FLA, RTS/CTS $t_{out} = 15ms$.

Figure 8.7: Discard probability using $L = 30$ for different AMCs.



(a) ARF.

(b) FLA, ideal MCS feedback.

(c) FLA, *D*ownRate $t_{out} = 15ms$.

(d) FLA, RTS/CTS $t_{out} = 15ms$.

Figure 8.8: Delay using $L = 30$ for different AMCs.

Part III

Concluding Remarks

CHAPTER 9

CONCLUSIONS AND FUTURE DIRECTIONS

The single main conclusion to be drawn from this thesis is that closed-loop AMC techniques are able to optimize the performance of WLANs, specially when these techniques are designed to take into account the fact that channel-related information available at the transmitter side might be outdated.

The next sections describe in detail the main outcomes achieved on each of the problems (P1, P2 and P3) posed in the introductory chapter of this thesis.

9.1 MCS selection in the physical layer of MIMO-OFDM systems

After extensive simulations using our purpose-built IEEE 802.11n PHY simulator, the proposed FLA algorithm has been confirmed as an AMC solution to provide optimum MCS selection according to the instantaneous channel conditions.

Even with the availability of perfect channel state information, the use of the average SNR across subcarriers and/or streams is not a good performance metric in the context of MIMO-OFDM systems, hence requiring of other link quality metrics, such as EESM. Results demonstrate that EESM-based FLA with ideal channel estimation meets the prescribed quality of service constraints while performing very close to the PBA (ideal MCS selection) in terms of throughput. Furthermore, FLA notably outperforms fixed MCSs or rate adaptive schemes based on an average SNR metric. Moreover, besides the transmission mode selection, this algorithm is able to provide a solution to the MIMO mode selection issue. Note that MIMO selection within the context of WLANs has only become important since the appearance of IEEE 802.11n, and the performance impact it has when combined with AMC has hardly received any attention.

When imperfect channel estimation is considered, it has been demonstrated that the PER prediction methodology of FLA is still capable of effectively selecting MCS values. Only small degradations in terms of throughput and PER have been observed

in highly-frequency selective channels, whose estimations are notably poorer. Consequently, these effects should be taken into consideration during the system design phase. In case of not meeting the prescribed QoS constraints, more conservative MCS selection thresholds can be employed to fulfil the QoS constraints at the cost of slightly affecting the throughput performance.

In order to improve the applicability of the FLA strategy, a new adaptive algorithm based on BER estimation instead of PER has been proposed. Results demonstrate its suitability for the IEEE 802.11n physical layer as it performs almost identically than its PER-based counterpart, while due to its independence from the packet length, it simplifies the costly calibration/prediction procedure.

9.1.1 Future work

Future efforts in the context of MCS selection should be devoted to the evaluation of other transmission techniques, such as the incorporation of the multi-user (MU)-MIMO technology or the effects last generation coding schemes such as LDPC may have in the MCS selection. Both techniques are envisaged to become technological pillars of the next generation of WLANs, currently under standardisation within IEEE 802.11ac. Research towards simplifying the off-line calibration phase required in FLA would constitute another important thread for future work. Finally, another interesting topic should be to study the performance of FLA adapting the packet size to the current channel conditions, hence making use of the packet aggregation mechanism provided by IEEE 802.11n.

9.2 AMC performance in the MAC sublayer of IEEE 802.11n

In this dissertation, the performance study of AMC algorithms over the MAC sublayer has been addressed by extending the PHY-layer simulator to also encompass the MAC operation. This platform was developed considering all the system performance details of the IEEE 802.11n MAC sublayer (see Chapter 5), and specifically allowing the configuration of different traffic loads, number of users, access schemes, access protocols and packet sizes.

Under saturated conditions (see Chapter 6), a complete study of FLA over 802.11n PHY/MAC in terms of goodput, fairness and system PER performance for a wide range of users and packet sizes has been presented and contrasted with those obtained using ARF. Noteworthy, the influence of feedback delay on the performance of FLA has been assessed. In Basic Access scheme, the degradation caused by an outdated MCS information can be largely compensated with the use of a time-out strategy that weighs down the influence of the MCS feedback delay. Numerical results clearly show that as the number of users in the system grows, the FLA-based adaptation proves to be much more robust to collisions than ARF, even when relying on outdated MCS information. This effect is clearly demonstrated by the fact that whereas ARF-based schemes suffer a dramatic reduction in goodput when more than two users are in the system, the FLA-based strategy exhibits a very graceful degradation thanks to a more accurate rate selection that takes into account the presence of collisions.

When considering other performance metrics, such as, delay or blocking probability, the FLA algorithm has also shown its superiority in comparison to ARF-based

methods (see chapter 8). In particular, FLA achieves nearly three times lower delay than ARF, which in turn implies lower blocking probability for a given traffic load.

FLA still beats ARF, although by a modest margin, irrespective of the number of users and packet sizes, when using the CSMA/CA with RTS/CTS access scheme. Remarkably, the RTS/CTS frame exchange allows a delay-free selection of the MCS and results in a system PER similar to the one obtained with the ideal FLA.

Lastly, the performance of FLA has been extensively evaluated on CSMA/E2CA networks observing significant goodput improvements over its CSMA/CA-based counterparts. It has been shown that this improvement is basically due to the collision reduction provided by the CSMA/E2CA protocol.

Overall, it can be concluded that FLA yields a goodput that significantly outperforms ARF for most system loads and access schemes, while keeping a large degree of fairness and satisfying prescribed PER-based QoS constraints.

Under non-saturated conditions (see Chapter 8), the performance study has been limited to Basic access scheme in CSMA/CA networks. It has been observed how different AMC schemes obtained similar results in terms of goodput, delay or blocking probability. Only some performance differences have been appreciated when evaluating the system PER of each AMC, where FLA consistently outperforms ARF. Furthermore, it has also been observed how collisions dramatically affect the ARF MCS selection mechanism as traffic load increases, hence reaching the saturation point with traffic loads lighter than when using FLA.

9.2.1 Future work

Future research efforts in this line can be directed towards the evaluation of the impact that the advanced PHY features envisaged in the IEEE 802.11ac may have at the MAC sublayer when combined with AMC. Moreover, the AMC performance study can be extended to consider other standard MAC sublayers, such as PCF, EDCA or HCCA. Lastly, and as a rather ambitious goal, the influence that the AMC algorithm may have in upper layers, such as TCP, could constitute an interesting research avenue.

9.3 IEEE 802.11n MAC sublayer analysis

This dissertation has presented analytical and semi-analytical models to evaluate the performance of IEEE 802.11 MAC sublayer systems when using fixed MCS and AMC algorithms, respectively. The newly proposed model has been shown to be a generalization of previous proposals that serve to encapsulate in one single model many features that had only been treated individually. In particular, by taking into account channel errors and the presence of anomalous slots, the proposed models have been shown to be significantly more accurate than previous proposals.

Furthermore, a semi-analytical model to evaluate the MAC sublayer performance under saturated conditions when considering different AMC algorithms has been proposed.

Additionally, an analytic approach has been introduced that estimates the system performance under both, saturated and non-saturated conditions. This model has expanded the previous saturation model presented in Chapter 6 by incorporating an extra dimension to the Markov chain that serves to model the queue buffer state of a

given STA. It has been shown that this analysis permits the performance evaluation of additional important metrics such as, delay, blocking probability, average queue size or discard probability. Performance metrics already presented under saturated conditions have now also been extended to non-saturated conditions.

Lastly, the analytical approach for non-saturated and saturated conditions has been extended to a semi-analytical model in order to consider the performance influence of AMC algorithms. To this end, the analytical model has been slightly changed to consider the different time definitions of error and collision events resulting from using different transmission modes. Results demonstrate the accuracy of the proposed framework, hence validating its suitability irrespective of the system configuration under study (number of users, AMC, packet length or access schemes).

9.3.1 Future work

As future work, the analysis of WLAN networks could be extended to consider the MU-MIMO technology that will notably enlarge the MAC sublayer capacity. Another interesting thread could be to consider the analysis of other MAC layer functions, such as, HCF or PCF mechanisms. Finally, the analysis could also be extended to consider non-symmetric traffic at each stations, or different traffic categories.

BIBLIOGRAPHY

[1] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005. (document), 1.2, 2, 2.1.1, 2.1.2, 2.3, 2.4, 2.5, 3.1.1, 3.1.2, 4.1

[2] V. Erceg, L. Schumacher, P. Kyritsi, A. Molisch, D. Baum, A. Gorokhov, C. Oestges, Q. Li, K. Yu, N. Tal *et al.*, "TGn channel models," *IEEE 802.11-03/940r4*, 2004. (document), 2.1.1, 2.2, 2.2.1, 2.2.1, 2.2.1, 2.1, 2.2.2, 2.6, 2.2, 2.2.2, 4.5

[3] J. Kermoal, L. Schumacher, K. Pedersen, P. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, 2002. (document), 2.2.1, 2.2.1, 2.2.1, 2.3, 3.4, 4.3, 4.5, 6.6

[4] J. Barcelo, B. Bellalta, C. Cano, A. Sfairopoulou, M. Oliver, and K. Verma, "Towards a Collision-Free WLAN: Dynamic Parameter Adjustment in CSMA/E2CA," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 708617, p. 11, 2011. (document), 6.1, 6.4.1, 6.3, 6.4.2, 6.4.2

[5] IEEE, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput," *IEEE Std 802.11n-2009*, 2009. (document), 3.3.1, 4.5, 5.1, 5.1.1, 5.1.2, 6.6, 6.2

[6] IEEE, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, Dec. 2007. (document), 5.2, 5.2.1, 6.2, 7.4

[7] B. Crow, I. Widjaja, J. G. Kim, and P. Sakai, "IEEE 802.11 Wireless Local Area Networks," *IEEE Communications Magazine*, vol. 35, no. 9, pp. 116–126, 1997. 1.1

[8] J. Gilb, "The wireless dictionary," *IEEE Wireless Dictionary [IEEE Standards Wireless Series]*, pp. c1–113, 2005. 1.1

[9] R. Gass and C. Diot, "An Experimental Performance Comparison of 3G and Wi-Fi," in *Passive and Active Measurement*, ser. Lecture Notes in Computer Science (LNCS). Springer Berlin Heidelberg, 2010, vol. 6032, pp. 71–80. 1.1

[10] E. H. Ong and J. Khan, "A Unified QoS-Inspired Load Optimization Framework for Multiple Access Points Based Wireless LANs," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2009, pp. 1–6. 1.1

[11] T. Joshi, D. Ahuja, D. Singh, and D. Agrawal, "SARA: Stochastic Automata Rate Adaptation for IEEE 802.11 Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 11, pp. 1579–1590, Nov. 2008. 1.2.2, 4.1

[12] K. Brueninghaus, D. Astely, T. Salzer, S. Visuri, A. Alexiou, S. Karger, and G.-A. Seraji, "Link performance models for system level simulations of broadband radio access systems," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 2306–2311, Sept. 2005. 1.2.2, 1.5, 4.2.1, 4.2.2

[13] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535 –547, March 2000. 1.4, 1.5, 5.1.1, 5.2, 5.3.2, 5.4, 6.6, 7.1, 7.1, 7.2.1, 7.2.2, 7.2.3

[14] S. Simoens, S. Rouquette-Léveil, P. Sartori, Y. Blankenship, and B. Classon, "Error prediction for adaptive modulation and coding in multiple-antenna OFDM systems," *Elsevier Signal Processing*, vol. 86, no. 8, pp. 1911–1919, 2006. 1.5, 4.1, 4.2.1

[15] P. Tan, Y. Wu, and S. Sun, "Link adaptation based on adaptive modulation and coding for multiple-antenna OFDM system," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1599–1606, 2008. 1.5, 4.1

[16] T. Jensen, S. Kant, J. Wehinger, and B. Fleury, "Mutual information metrics for fast link adaptation in IEEE 802.11n," in *IEEE International Conference on Communications (ICC)*, May 2008, pp. 4910–4915. 1.5, 4.1, 4.5, 6.1

[17] J. Choi, K. Park, and C.-K. Kim, "Analysis of Cross-Layer Interaction in Multirate 802.11 WLANs," *IEEE Transactions on Mobile Computing*, vol. 8, no. 5, pp. 682–693, 2009. 1.5

[18] S. Kim, L. Verma, S. Choi, and D. Qiao, "Collision-Aware Rate Adaptation in multi-rate WLANs: Design and implementation," *Computer Networks*, vol. 54, no. 17, pp. 3011 – 3030, 2010. 1.5, 5.1.1, 6.1, 6.6.2

[19] Y. Song, X. Zhu, Y. Fang, and H. Zhang, "Threshold optimization for rate adaptation algorithms in IEEE 802.11 WLANs," *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 318–327, 2010. 1.5, 6.1

[20] G. Bianchi, "IEEE 802.11-saturation throughput analysis," *Communications Letters, IEEE*, vol. 2, no. 12, pp. 318 –320, dec. 1998. 1.5, 5.2, 7.1, 7.1

[21] E. Ziouva and T. Antonakopoulos, "CSMA/CA performance under high traffic conditions: throughput and delay analysis," *Computer Communications*, vol. 25, no. 3, pp. 313 – 321, 2002. 1.5, 5.2, 7.1

[22] C. H. Foh and J. Tantra, "Comments on IEEE 802.11 saturation throughput analysis with freezing of backoff counters," *IEEE Communications Letters*, vol. 9, no. 2, pp. 130 – 132, feb. 2005. 1.5, 5.2, 7.1

[23] I. Tinnirello, G. Bianchi, and Y. Xiao, "Refinements on IEEE 802.11 Distributed Coordination Function Modeling Approaches," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 3, pp. 1055 –1067, 2010. 1.5, 5.2, 5.2.1, 5.2.1, 5.4, 6.3, 6.3, 6.3, 6.6, 7.1, 7.2.1, 7.2.2, 7.2.3

[24] P. Chatzimisios, A. Boucouvalas, and V. Vitsas, "Influence of channel BER on IEEE 802.11 DCF," *IET Electronics Letters*, vol. 39, no. 23, pp. 1687–1689, 2003. 1.5, 7.1

[25] Q. Ni, T. Li, T. Turletti, and Y. Xiao, "Saturation throughput analysis of error-prone 802.11 wireless networks," *Wireless Communications and Mobile Computing*, vol. 5, no. 8, pp. 945–956, 2005. 1.5, 7.1

[26] X. Dong and P. Varaiya, "Saturation throughput analysis of IEEE 802.11 wireless LANs for a lossy channel," *IEEE Communications Letters*, vol. 9, no. 2, pp. 100–102, 2005. 1.5, 7.1

[27] F. Daneshgaran, M. Laddomada, F. Mesiti, M. Mondin, and M. Zanolo, "Saturation throughput analysis of IEEE 802.11 in the presence of non ideal transmission channel and capture effects," *IEEE Transactions on Communications*, vol. 56, no. 7, pp. 1178–1188, 2008. 1.5, 7.1

[28] H. Chen, "Revisit of the Markov Model of IEEE 802.11 DCF for an Error-Prone Channel," *IEEE Communications Letters*, vol. 15, no. 12, pp. 1278 –1280, december 2011. 1.5, 5.2, 5.2.2, 5.3, 5.4, 6.3, 6.6, 7.1, 7.2.2, 7.2.3, 7.4.2

[29] M. Ergen and P. Varaiya, "Throughput Analysis and Admission Control for IEEE 802.11a," *Mobile Networks and Applications*, vol. 10, pp. 705–716, 2005. 1.5, 5.2, 7.1, 7.2.3

[30] K. Duffy, D. Malone, and D. Leith, "Modeling the 802.11 distributed coordination function in non-saturated conditions," *IEEE Communications Letters*, vol. 9, no. 8, pp. 715–717, 2005. 1.5, 7.1

[31] D. Malone, K. Duffy, and D. Leith, "Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions," *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 159–172, 2007. 1.5, 7.1, 7.2.3, 7.4.3

[32] H. Zhai, Y. Kwon, and Y. Fang, "Performance analysis of IEEE 802.11 MAC protocols in wireless LANs," *Wireless communications and mobile computing*, vol. 4, no. 8, pp. 917–931, 2004. 1.5, 7.1

[33] G. Cantieni, Q. Ni, C. Barakat, and T. Turletti, "Performance analysis under finite load and improvements for multirate 802.11," *Computer Communications*, vol. 28, no. 10, pp. 1095–1109, 2005. 1.5, 7.1, 7.2.3

[34] K. Duffy and A. Ganesh, "Modeling the impact of buffering on 802.11," *IEEE Communications Letters*, vol. 11, no. 2, pp. 219–221, 2007. 1.5, 7.1

[35] R. Liu, G. Sutton, and I. Collings, "A new queueing model for QoS analysis of IEEE 802.11 DCF with finite buffer and load," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2664–2675, 2010. 1.5, 7.1, 7.2.1, 7.2.2, 7.2.3, 1, 7.2.7, 7.4, 7.4.1, 2, 7.4.3, 8.1

[36] B. Sklar, "Rayleigh fading channels in mobile digital communication systems .I. Characterization," *IEEE Communications Magazine*, vol. 35, no. 7, pp. 90 –100, jul 1997. 2.1.1, 2.1.2

[37] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge Univ Pr, 2005. 2.1.1

[38] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," *Review of the Electrical Communnication Laboratory*, vol. 16, no. 9, pp. 825–73, 1968. 2.1.1

[39] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, 1980. 2.1.1

[40] E. C. in the Field of Science and T. R. E.-C. 231, "Urban transmission loss models for mobile radio in the 900 and 1800 Mhz bands," vol. Revision 2, The Hague, the Netherlands, September 1991. 2.1.1

[41] A. Hatami and K. Pahlavan, "Comparative statistical analysis of indoor positioning using empirical data and indoor radio channel models," in *IEEE Consumer Communications and Networking Conference (CCNC)*, vol. 2, 2006, pp. 1018–1022. 2.1.1

[42] V. Erceg, L. Greenstein, S. Tjandra, S. Parkoff, A. Gupta, B. Kulic, A. Julius, and R. Bianchi, "An empirically based path loss model for wireless channels in suburban environments," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 7, pp. 1205–1211, 1999. 2.1.1

[43] S. Ghassemzadeh, L. Greenstein, A. Kavcic, T. Sveinsson, and V. Tarokh, "UWB indoor path loss model for residential and commercial buildings," in *IEEE Vehicular Technology Conference (VTC)*, vol. 5, oct. 2003, pp. 3115 – 3119 Vol.5. 2.1.1

[44] A. Saleh and R. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128 –137, february 1987. 2.1.2, 2.2

[45] R. Clarke, "A statistical theory of mobile-radio reception," *Bell System Technical Journal*, vol. 47, no. 6, pp. 957–1000, 1968. 2.1.2

[46] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of communication systems: modeling, methodology, and techniques*. Kluwer Academic/Plenum publishers, 2000. 2.1.3

[47] Q. Spencer, B. Jeffs, M. Jensen, and A. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 347 –360, mar 2000. 2.2

[48] R.-M. Cramer, R. Scholtz, and M. Win, "Evaluation of an ultra-wide-band propagation channel," *Antennas and Propagation, IEEE Transactions on*, vol. 50, no. 5, pp. 561 –570, may 2002. 2.2

[49] A. Poon and M. Ho, "Indoor multiple-antenna channel characterization from 2 to 8 Ghz," in *IEEE International Conference on Communications (ICC)*, vol. 5, may 2003, pp. 3519 – 3523 vol.5. 2.2

[50] J.-G. Wang, A. Mohan, and T. Aubrey, "Angles-of-arrival of multipath signals in indoor environments," in *IEEE Vehicular Technology Conference (VTC)*, vol. 1, apr-1 may 1996, pp. 155 –159. 2.2

[51] C.-C. Chong, D. Laurenson, and S. McLaughlin, "Statistical characterization of the 5.2 GHz wideband directional indoor propagation channels with clustering and correlation properties," in *IEEE Vehicular Technology Conference (VTC)*, vol. 1, 2002, pp. 629 – 633 vol.1. 2.2

[52] G. German, Q. Spencer, L. Swindlehurst, and R. Valenzuela, "Wireless indoor channel modeling: statistical agreement of ray tracing simulations and channel sounding measurements," in *IEEE International Conference Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, 2001, pp. 2501 –2504. 2.2

[53] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyösti, D. Laurenson, G. Matz, A. F. Molisch, C. Oestges, and H. Özcelik, "Survey of channel and radio propagation models for wireless MIMO systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, pp. 56–56, Jan. 2007. 2.2

[54] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996. 2.2.1, 3.1.2, 3.3.1

[55] J. Kermoal, L. Schumacher, P. Mogensen, and K. Pedersen, "Experimental investigation of correlation properties of MIMO radio channels for indoor picocell scenarios," in *IEEE Vehicular Technology Conference (VTC)*, vol. 1, 2000, pp. 14 –21 vol.1. 2.2.1, 2.2.1

[56] L. Schumacher, K. Pedersen, and P. Mogensen, "From antenna spacings to theoretical capacities - guidelines for simulating MIMO systems," in *Personal, Indoor and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on*, vol. 2, sept. 2002, pp. 587 – 592 vol.2. 2.2.1

[57] J. Medbo and P. Schramm, "Channel models for HIPERLAN/2 in different indoor scenarios," *ETSI/BRAN 3ERI085B*, 1998. 2.2.2

[58] J. Medbo, J. Berg, and H. Andersson, "Measured radiowave propagation characteristics at 5 GHz for typical HIPERLAN/2 scenarios," *ETSI/BRAN document*, no. 3ERI084A, 1998. 2.2.2

[59] M. S. Kuran and T. Tugcu, "A survey on emerging broadband wireless access technologies," *Elsevier Computer Networks*, vol. 51, no. 11, Aug. 2007. 3.1.1

[60] V. Chakravarthy, A. Nunez, J. Stephens, A. Shaw, and M. Temple, "TDCS, OFDM, and MC-CDMA: a brief tutorial," *IEEE Communications Magazine*, vol. 43, no. 9, pp. S11 – S16, sept. 2005. 3.1.1

[61] J. Winters, "On the Capacity of Radio Communication Systems with Diversity in a Rayleigh Fading Environment," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 5, pp. 871 – 878, jun 1987. 3.1.2

[62] G. Foschini and M. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998. 3.1.2

[63] E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999. 3.1.2

[64] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744 –765, mar 1998. 3.1.2

[65] A. van Zelst, "Space division multiplexing algorithms," in *IEEE Mediterranean Electrotechnical Conference (MELECON)*, vol. 3, may 2000, pp. 1218 – 1221 vol.3. 3.1.2

[66] ——, "MIMO OFDM for wireless LANs," Ph.D. dissertation, Technische Universiteit Eindhoven, 2004. 3.1.2, 3.1.2, 3.1.2, 3.1.2

[67] X. Li, H. Huang, G. Foschini, and R. Valenzuela, "Effects of iterative detection and decoding on the performance of BLAST," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 2, 2000, pp. 1061 –1066. 3.1.2

[68] A. van Zelst, "Per-antenna-coded schemes for MIMO OFDM," in *IEEE International Conference on Communications (ICC)*, vol. 4, may 2003, pp. 2832 – 2836. 3.1.2

[69] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct 1998. 3.1.2, 3.1.2

[70] J. Barry, E. Lee, and D. Messerschmitt, *Digital communication*. Springer, 2004. 3.1.2, 3.1.2

[71] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 68 – 73, oct. 2004. 3.1.2

[72] H. Zhang, A. Molisch, and J. Zhang, "Applying Antenna Selection in WLANs for Achieving Broadband Multimedia Communications," *IEEE Transactions on Broadcasting*, vol. 52, no. 4, pp. 475 –482, dec. 2006. 3.1.2

[73] J. Ha, A. N. Mody, J. H. Sung, J. R. Barry, S. W. McLaughlin, and G. L. Stüber, "LDPC Coded OFDM with Alamouti/SVD Diversity Technique," *Wireless Personal Communications*, vol. 23, pp. 183–194, 2002. 3.1.2

[74] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *IEEE International Conference on Communications (ICC)*, vol. 3, may 1993, pp. 1630 –1634 vol.3. 3.1.2

[75] G. Bauch and J. Malik, "Cyclic delay diversity with bit-interleaved coded modulation in orthogonal frequency division multiple access," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2092 –2100, aug. 2006. 3.1.2, 1

[76] A. Dammann, R. Raulefs, and S. Kaiser, "Beamforming in combination with space-time diversity for broadband OFDM systems," in *IEEE International Conference on Communications (ICC)*, vol. 1, 2002, pp. 165 –171. 3.1.2

[77] D. Gore, S. Sandhu, and A. Paulraj, "Delay diversity codes for frequency selective channels," in *IEEE International Conference on Communications (ICC)*, vol. 3, 2002, pp. 1949 – 1953 vol.3. 3.1.2

[78] S. Kaiser, "Performance of multi-carrier CDM and COFDM in fading channels," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1B, 1999, pp. 847 –851 vol. 1b. 3.1.3

[79] R. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21 –28, january 1962. 3.2

[80] Y.-S. Choi and S. Alamouti, "A pragmatic PHY abstraction technique for link adaptation and MIMO switching," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 960–971, August 2008. 3.3.1, 4.2.3, 6.2.1

[81] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge university press, 2003. 3.3.2

[82] F. Tosato and P. Bisaglia, "Simplified soft-output demapper for binary interleaved COFDM with application to HIPERLAN/2," in *IEEE International Conference on Communications (ICC)*, vol. 2, 2002, pp. 664–668. 3.3.2

[83] T. Jensen, S. Kant, J. Wehinger, and B. Fleury, "Fast Link Adaptation for MIMO OFDM," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 8, pp. 3766–3778, 2010. 3.4

[84] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 67–75, 2002. 3.4.1

[85] M. Lampe, T. Giebel, H. Rohling, and W. Zirwas, "Per-prediction for PHY mode selection in OFDM communication systems," in *IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2003, pp. 25–29. 4.2

[86] F. Zheng and J. Nelson, "Adaptive Design for the Packet Length of IEEE 802.11n Networks," in *IEEE International Conference on Communications (ICC)*, 2008, pp. 2490–2495. 4.3

[87] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement," in *IEEE International Conference on Computer Communications (INFOCOM)*, vol. 2, 2002, pp. 599 – 607 vol.2. 5.2

[88] Y. Xiao, "A simple and effective priority scheme for IEEE 802.11," *IEEE Communications Letters*, vol. 7, no. 2, pp. 70 –72, feb. 2003. 5.2

[89] V. Vishnevsky and A. Lyakhov, "802.11 LANs: Saturation Throughput in the Presence of Noise," in *IFIP*, ser. NETWORKING '02. Springer-Verlag, 2002, pp. 1008–1019. 5.2

[90] Z. Hadzi-Velkov and B. Spasenovski, "An analysis of CSMA/CA protocol with capture in wireless LANs," in *IEEE Wireless Communications and Networking (WCNC)*, vol. 2, march 2003, pp. 1303 –1307 vol.2. 5.2

[91] T.-C. Hou, L.-F. Tsao, and H.-C. Liu, "Throughput analysis of the IEEE 802.11 DCF scheme in multi-hop ad hoc networks," in *International Conference on Wireless Networks (ICWN)*, june 2003, pp. 653–659. 5.2

[92] P. Chatzimisios, A. Boucouvalas, and V. Vitsas, "IEEE 802.11 packet delay-a finite retry limit analysis," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 2, dec. 2003, pp. 950 – 954 Vol.2. 5.2

[93] S. H. Y. Wong, H. Yang, S. Lu, and V. Bharghavan, "Robust rate adaptation for 802.11 wireless networks," in *ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2006, pp. 146–157. 6.1

[94] A. Kamerman and L. Monteban, "WaveLAN-II: a high-performance wireless LAN for the unlicensed band," *Bell Labs Technical Journal*, vol. 2, no. 3, pp. 118–133, 1997. 6.1, 6.2.1

[95] T. Joshi, D. Ahuja, D. Singh, and D. Agrawal, "SARA: stochastic automata rate adaptation for IEEE 802.11 networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 11, pp. 1579–1590, 2008. 6.1

[96] J. He, D. Kaleshi, A. Munro, and J. McGeehan, "Modeling Link Adaptation Algorithm for IEEE 802.11 Wireless LAN Networks," in *IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Valencia, Spain, Sept. 2006. 6.1

[97] J. Zhang, K. Tan, J. Zhao, H. Wu, and Y. Zhang, "A Practical SNR-Guided Rate Adaptation," in *IEEE International Conference on Computer Communications (INFOCOM)*, Phoenix, AZ, April 2008. 6.1

[98] H. Jung, T. Kwon, Y. Choi, and Y. Seok, "A Scalable Rate Adaptation Mechanism for IEEE 802.11e Wireless," in *IEEE Future generation Communication and Networking (FGCN)*, vol. 1, Jeju-Island, Korea, Dec. 2007. 6.1

[99] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive MAC protocol for multi-Hop wireless networks," in *ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2001, pp. 236–251. 6.1, 6.2.1

[100] J. Choi, J. Na, Y. sup Lim, K. Park, and C. kwon Kim, "Collision-aware design of rate adaptation for multi-rate 802.11 WLANs," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1366 –1375, 2008. 6.1

[101] J. Barcelo, A. Toledo, C. Cano, and M. Oliver, "Fairness and Convergence of CSMA with Enhanced Collision Avoidance (ECA)," in *IEEE International Conference on Communications (ICC)*, may 2010, pp. 1 –6. 6.1, 6.4.1

[102] G. Ahn, A. Campbell, A. Veres, and L. Sun, "Supporting service differentiation for real-time and best-effort traffic in stateless wireless ad hoc networks (SWAN)," *IEEE Transactions on Mobile Computing*, vol. 1, no. 3, pp. 192–207, 2002. 7.2.3

[103] O. Tickoo and B. Sikdar, "Modeling queueing and channel access delay in unsaturated IEEE 802.11 random access MAC based wireless networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 4, pp. 878–891, 2008. 7.2.3

[104] R. P. Liu, G. Sutton, X. Yang, and I. Collings, "Modelling QoS Performance of IEEE 802.11 DCF under Practical Channel Fading Conditions," in *IEEE International Conference on Communications (ICC)*, 2011, pp. 1–6. 7.4, 7.4.1, 7.4.3, 8.1

[105] Q. Zhao, D. H. K. Tsang, and T. Sakurai, "Modeling Nonsaturated IEEE 802.11 DCF Networks Utilizing an Arbitrary Buffer Size," *IEEE Transactions on Mobile Computing*, vol. 10, no. 9, pp. 1248–1263, 2011. 8.1

[106] K. Huang and K. Duffy, "On a buffering hypothesis in 802.11 analytic models," *IEEE Communications Letters*, vol. 13, no. 5, pp. 312–314, 2009. 8.1, 8.4.2