Random Access on Multipacket Reception Channel with Applications to Satellite Communications

by

Majid Ghanbarinejad

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Communications

Department of Electrical and Computer Engineering
University of Alberta

Abstract

In the field of multiple-access systems, multipacket reception (MPR) is defined as the capability of the receiver to successfully receive/decode concurrent packets from multiple transmitters. While physical-layer capabilities and limitations of MPR-capable systems using technologies such as code-division multiple-access (CDMA) and multiple-antenna communications are well studied in the literature, an important question is how this capability affects the performance at the medium-access control (MAC) layer.

Random access on the MPR channel is studied in this thesis. First, we study throughput of ALOHA random access with Poisson arrivals and highlight the throughput advantage of the MPR channel compared to the conventional collision channel. The analysis shows that ALOHA performs more efficiently on the MPR channels where the maximum number of concurrent transmissions that can be received by the receiver is larger providing that the aggregate traffic on the medium is controlled and maintained below a threshold. If the average number of transmissions exceeds this threshold, however, the system throughput declines rapidly, which motivates efficient control mechanisms on top of random access. Bounds on the throughput performance of random access are examined and a closed-form approximation for throughput of genie-aided random access with Poisson arrivals is derived. A method of computing the optimal access probability is, then, presented for the cases where only partial information about contention on the medium is available.

Next, we examine a finite-size multiple-access system where the incoming traffic is the aggregate of interrupted on/off Markov sources. The resulting system is a

Markov decision process for which the optimal access algorithm is provided and compared to lower and upper bounds. A new rate/burstiness model of an on/off Markov source is proposed, which is used to present where the throughput curves lay compared to the computed bounds. Furthermore, it is shown that the gap between the upper and lower bounds reduces as the maximum number of packets that can be received increases. Queuing analysis is provided and closed-form equations for queuing delay as a function of throughput are derived.

While the above algorithm requires prior knowledge of traffic parameters, we also propose a framework for more practical random access systems where no prior knowledge of traffic and node population is available. The proposed scheme uses the theory of extended Kalman filters in order to track the contention on the medium and select the near-optimal access probability. The proposed scheme is shown to stabilize the operating point of the multiple-access system close to the point where throughput of ALOHA takes its maximum. This scheme is also examined for unideal cases such as delayed acknowledgement channel and shown to be a robust access scheme on the MPR channel.

Finally, we study two aspects of random access on the MPR channel that are related to satellite communications. First, a repetition random access scheme that employs transmission diversity and iterative collision resolution is analyzed for the case of MPR channels. Asymptotic analysis for this generalized model is provided and it is shown that larger loads can be supported in smaller frames. Next, throughput-delay tradeoff of random access over delayed links is analyzed and scaling laws are derived for the cases of the collision channel and the MPR channel as well as repetition random access. It is shown that multiuser detection and repetition schemes improve the multiple access performance in the sense that the inevitable compromise between throughput and delay is mitigated by joint detection capabilities and/or repetitions.

Preface

Chapter 3 of this thesis has been published as M. Ghanbarinejad and C. Schlegel, "Distributed Probabilistic Medium Access with Multipacket Reception and Markovian Traffic," *Telecommunication Systems*, vol. 56, no. 2, pp. 311-321, June 2014. I modeled, analyzed, and simulated the multiple-access system under the supervision of Prof. Christian Schlegel who contributed to the work through comments and suggestions about the system model, the analysis, and the presentation of the work.

To my dear family whose love knows no upper bound.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Professor Christian Schlegel and Professor Majid Khabbazian, for their patient guidance and very useful critiques of this research as well as all the life lessons that they have taught me.

I would also like to thank the committee of my candidacy examination and defense, Prof. Hai Jiang, Prof. Majid Ghaderi, Prof. Ivan Fair, and Prof. Janelle Harms, for many insightful comments and critiques of this work. Thanks to the Network Group at the Department of Computing Science, especially Prof. Ioanis Nikolaidis and Prof. Pawel Gburzynski, for their assistance and advice several times.

Special thanks to Ms. Pinder Bains, the department's graduate student advisor, who patiently answered my many questions many times and assisted me through my graduate program.

My colleagues at the High-Capacity Digital Communications Laboratory left me a lot of great memories that I gladly take as an invaluable souvenir from this journey. I learned a lot about research and industry from Sumeeth, Lukasz and Saeed. I wish to learn from Charmaine to be warmhearted, from Michiko to never run out of positive energy, and from Sheehan to have time for a coffee with friends in the busiest times. Thanks to Navid, Saina, Deyasini, David, Russell, Marcel, Parisa, Dmitry, Amirhossein, Wesam, Shuai, and Eric for all the fun in our gatherings.

And last but not least, I am sincerely grateful to my family and friends for all their support and kindness that cannot be described in words. Monir, Farhad, Hadi, Ali, Payam, and Amin are friends whose invaluable companionship will never be forgotten.

Table of Contents

List of Tables

List of Figures

List of Symbols

Symbol	Definition	First Use
λ	Offered Traffic Load (packets/slot)	2
R	Throughput (packets/slot)	2
N^{tx}	Number of Transmitted Packets	6
N^{rx}	Number of Received Packets	6
K	MPR Capability	6
M	Maximum Number of Nodes	14
$\mathsf{E}[X]$	Expected Value of X	17
$\mathsf{Var}(X)$	Variance of X	19
$f_{\text{bin}}(x; N, p)$	Binomial PMF of x with Parameters N and p	19
$f_{\text{pois}}(x; \lambda)$	Poisson PMF of x with Parameter λ	19
$F_{\text{bin}}(x; N, p)$	Binomial CDF of x with Parameters N and p	19
$F_{\text{pois}}(x; \lambda)$	Poisson CDF of x with Parameter λ	19
p	Access Probability	22
N	Number of Active Nodes	22
x^*	Throughput-Optimal x	72
$x^\dagger$	Delay-Optimal x	72

List of Abbreviations

Abbreviation	Description	First Use
MAC	Medium-Access Control	1
ACK	Acknowledgement (Packet)	1
IFT	Immediate First Transmission	2
DFT	Delayed First Transmission	2
TDMA	Time-Division Multiple-Access	2
CSMA	Carrier-Sense Multiple-Access	2
BEB	Binary Exponential Backoff	3
MPR	Multipacket Reception	4
MUD	Multiuser Detection	6
CDMA	Code-Division Multiple-Access	6
PHY	Physical (Layer)	6
SINR	Signal to Interference-plus-Noise Ratio	9
EB	Exponential Backoff	9
M2M	Machine-to-Machine	11
RRA	Repetition Random Access	13
CAC	Call Admission Control	14
GMAC	Gaussian Multiple-Access Channel	15
SNR	Signal to Noise Ratio	15
GM	Generalized Modulation	16
PDF	Probability Distribution Function	19
PMF	Probability Mass Function	19

CDF	Cumulative Distribution Function	19
PLR	Packet Loss Ratio	20
NACK	No-Acknowledgment (Packet)	35
FA	Forward Algorithm	36
CTP	Constant Transmission Probability	39
MILD	Multiplicative-Increase Linear-Decrease	47
EKF	Extended Kalman Filter	48
CRDSA	Collision Resolution Diversity Slotted Aloha	62
IRSA	Irregular Repetition Slotted Aloha	63
LDPC	Low-Density Parity-Check	63
AP	Access Probability	70

Chapter 1

Introduction

1.1 Preliminaries

In data communications networks, medium-access control (MAC) is a mechanism through which network devices (nodes) access a shared medium for transmission. Since each transmission contributes interference to other transmissions and degrades the quality of the received signal, MAC protocols are designed to limit the number of simultaneous transmissions in order to meet the receivers' requirement for signal quality.

Design and analysis of MAC protocols in wireless networks date back to Abramson's random access protocol called ALOHA [1]. ALOHA allows a node to send its packet – as soon as it is ready for transmission – with no consideration of the present interference and contention. This scheme is simple to implement and is suitable for multiple-access systems where data traffic is light and, therefore, simultaneous transmissions by more than one node are rare. If a transmitted packet is not received successfully, the node is informed immediately as it does not receive an acknowledgement (ACK) packet from its intended receiver. In this case, the node *backlogs* the packet and attempts a retransmission at a later time. Since the cause of the reception failure is possibly a *collision* with other simultaneously transmitted packets and since those packets are also possibly backlogged and will be retransmitted, each node attempts a retransmission after a random period of time called the *backoff* period in order to reduce the chance of further collisions with the other backlogged packets. For example, in *slotted* ALOHA, the retransmission probability p_r at each slot is set to a value smaller than 1.

The above variation of ALOHA was later called "immediate first transmission" (IFT) as the first transmission of a packet always occurs immediately after it is generated at the upper layers. The alternative variation is called "delayed first transmission" (DFT) [2] where the first transmission attempt also occurs with some probability $p < 1$ to reduce the chance of collisions in the presence of heavier traffic. In this thesis, we use the generic terms "random access" and "probabilistic access" interchangeably for any MAC scheme where nodes access the medium probabilistically, i.e., without deterministic coordinations with other nodes. In this sense, IFT ALOHA and DFT ALOHA are two special cases of a large family of protocols that will be of interest in this thesis.

In his analysis of the average throughput of ALOHA [3], Abramson assumed that concurrent transmissions fail to be received due to the signal quality being corrupted by interference. This model of channel/receiver, called the collision model, became the most commonly used model for design and analysis of MAC protocols. It was shown in [3] that, with Poisson-distributed incoming traffic, the throughput of ALOHA is obtained by

$$R = \lambda e^{-2\lambda} \ , \tag{1.1}$$

where λ (packets/T_{pkt}) is the average offered traffic, R (packets/T_{pkt}) is the average system throughput, and T_{pkt} denotes the duration of each packet. Figure 1.1 shows that ALOHA can utilize at most $\sim 18\%$ of the capacity of the collision channel where 100% corresponds to an ideal time-division multiple-access (TDMA) scheme.

In a slotted setup where packet transmissions are constrained to time slots, we have

$$R = \lambda e^{-\lambda} \ , \tag{1.2}$$

where the maximum throughput can be improved two-fold. But the resulting efficiency of $\sim 37\%$ is not yet impressive and, indeed, one can conclude that the simplicity of random access is obtained at the cost of low system throughput. Furthermore, one of the main challenges with random access schemes is to mitigate the likelihood of system instability; this subject attracted considerable attention since the advent of random access (e.g., see [4, 5, 6]).

Shortcomings of ALOHA led to other proposals for medium access such as the carrier-sense multiple-access (CSMA) [7] family of protocols, followed by proposals

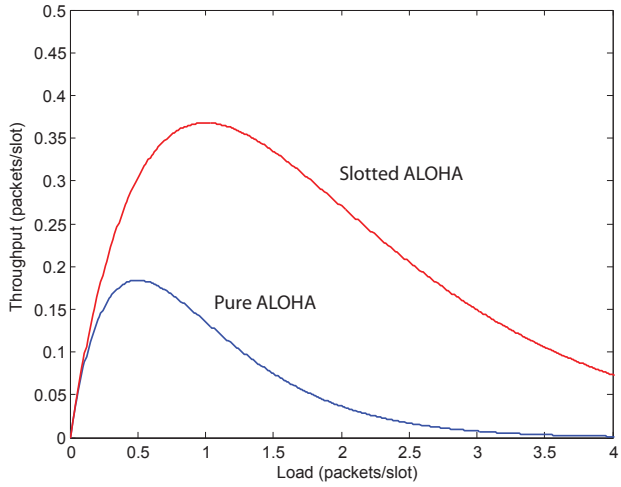

Figure 1.1: Average throughput of pure and slotted ALOHA with Poisson arrivals. Both axes are shown in units of packets/slot where a slot is defined as the duration of a packet T_{pkt}.

for collision detection and collision avoidance mechanisms. These protocols have served as the basis for the MAC protocol of several worldwide standards including the original IEEE 802.3 (Ethernet) and IEEE 802.11 (WiFi) family of standards.

In CSMA protocols, nodes sense the medium and transmit only if they detect no ongoing transmissions. The carrier sensing mechanism reduces the probability of collisions. In IEEE 802.11, for example, a mechanism called binary exponential backoff (BEB) is used where each node has to find the medium idle for a random period before any (re)transmission attempt. The random period is uniformly drawn from a window called the backoff window W. Any node with a new packet to transmit sets the window to the minimum $W \leftarrow W_{\min}$, which is a fixed value deteremined by the standard. Each time the (re)transmission of a packet fails, the node doubles its backoff window $W \leftarrow 2W$ unless it has reached the maximum value $W_{\max}$. Upon successful transmission, the node resets its window size to the minimum $W \leftarrow W_{\min}$.

Despite their differences, slotted ALOHA and IEEE 802.11 follow the same procedure for collision resolution: They reduce their access rate according to some predefined rule without learning about the overall traffic and the current number of nodes contending to access the medium. Although simple to implement, both of

these schemes suffer from a major drawback: The maximum load and/or the maximum number of nodes that can be supported are limited by the backoff parameters such as p_r, $W_{\min}$ and $W_{\max}$. For example, in slotted ALOHA, selecting a large value of p_r limits the maximum number of nodes contending at any moment while a small p_r limits the maximum load when only a few nodes are contending.

MAC efficiency, defined as the ratio of the system throughput to the shared bandwidth, is generally low in random access protocols. As mentioned, unslotted ALOHA (a.k.a. *pure* ALOHA) is known to reach a theoretical peak of 18% efficiency [3] when the aggregate of the new and backlogged traffic is adjusted to the optimal point at all times, a condition that is not generally easy to satisfy. Slotted ALOHA doubles this maximum at the cost of implementing packet synchronization. Practical values of MAC efficiency in IEEE 802.11 are below 15% [8]. Furthermore, legacy carrier-sense protocols suffer from fairness issues over conventional collision and capture channels, especially in ad hoc wireless networks (e.g., [9, 10, 11]).

One of the main reasons for the generally low throughput of random access is that the bandwidth cost of collisions is high while packets collide frequently. Therefore, in order to improve the MAC efficiency, we have to *i*) design MAC protocols that lower the probability of collisions, and *ii*) implement receivers that lower the probability of packet loss due to simultaneous transmissions. Conventionally, the former is the main focus of designing more efficient random access protocols, which has been shown to work only within inherent limits on the maximum achievable efficiency. The latter, which is the subject of signal processing for multiuser detection [12], is referred to as multipacket reception (MPR) in the context of MAC. MPR is the channel model used to capture the packet-level behavior of multiple-access systems where the receiver is capable of processing signals from multiple transmitters simultaneously. MPR channels have been shown to improve the MAC efficiency significantly while maintaining the desirably low complexity of random access [13]. In other words, the multiple-access bandwidth can be utilized more efficiently by moving a fraction of the MAC-layer complexity of the transmitters to the physical layer of the receiver(s).

1.2 Literature Review

1.2.1 MPR Models

There are several abstract models of multiple-access channels in the literature that deviate from the collision model. A popular example is the capture channel [9, 11] where, in the presence of multiple concurrent transmissions, one transmission has a nonzero chance of being received while the others will be corrupted and subject to retransmissions. Other models include probabilistic models that determine the probability of receiving a packet when a certain number of other packets are transmitted simultaneously. In the following, we will review some of the most well-known abstract models with emphasis on the MPR channel.

In [14, 15], the authors proposed a model for a symmetric MPR channel where a matrix of probabilities determine the chance of receiving a packet (in a slotted random access setup) where a certain number of concurrent packets are transmitted. Let ϵ_{nk} denote the probability of receiving k packets out of n packets transmitted simultaneously. Then, the following matrix uniquely defines a generic MPR channel:

$$\mathbf{E} = \begin{bmatrix} \epsilon_{10} & \epsilon_{11} & 0 & 0 & \cdots \\ \epsilon_{20} & \epsilon_{21} & \epsilon_{22} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \epsilon_{n0} & \epsilon_{n1} & \cdots & \epsilon_{nn} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} . \tag{1.3}$$

Obviously, we should have $\sum_k \epsilon_{nk} = 1$ for all n, i.e., rows of $\mathbf{E}$ should add up to 1. This model is generic and includes special cases of interest. For example, the collision channel corresponds to

$$\mathbf{E}_{\mathsf{col}} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} . \tag{1.4}$$

Another example is the capture channel that takes the form of

$$\mathbf{E}_{\mathsf{cap}} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 1-x_2 & x_2 & 0 & 0 & \cdots \\ 1-x_3 & x_3 & 0 & 0 & \cdots \\ 1-x_4 & x_4 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} , \tag{1.5}$$

where x_n denotes the probability of capture in the presence of n transmissions.

In the above model, $\mathbf{E}$ represents an MPR channel if at least for one $n \geq k > 1$, we have $\epsilon_{nk} > 0$. However generic and flexible, this model requires an analytical or experimental method to determine a potentially large number of parameters, namely ϵ_{nk}, for several values of n, k. Therefore, analytical usage of the model is mainly limited to more abstract parameters such as the "ergodic capacity" of the MPR channel $C_{\mathsf{erg}} := \lim_{n\to\infty} \sum_k k\epsilon_{nk}$.

The model of (1.3) is symmetric as it does not distinguish between different nodes. If the system is not symmetric and the probability of successful reception of a node's packet is different from other nodes', a more general model will be required where $\epsilon_{S,R}$ takes all the possible subsets S, R of the set of nodes instead of only their cardinalities $n = |S|$, $k = |R|$. Such a model was proposed in [16]. The model requires an exponentially growing number of parameters and has had limited use in the literature.

Multiuser detection (MUD) in code-division multiple-access (CDMA) systems [17, 12] is one of the most important examples of MPR-enabling technologies at the physical (PHY) layer. MUD techniques range from linear detection schemes, such as decorrelation and linear minimum mean-square error detection, to more practical schemes, such as nonlinear decision-driven and iterative interference cancelation methods. Multiuser multiple-antenna communications may also provide MAC-level MPR opportunities.

Since the emphasis in this thesis is on CDMA MUD systems for which, thanks to the exponentially growing advancements of the silicon technology, implementation of iterative receivers is practical, we use a simplified MPR model throughout this thesis that closely represents the behavior of these systems. Let N^{tx} and N^{rx} denote, respectively, the number of packets concurrently transmitted and successfully received. Then, the MPR channel used in this thesis is described as

$$N^{\mathsf{rx}} = \begin{cases} N^{\mathsf{tx}} & \text{if } N^{\mathsf{tx}} \leq K \\ 0 & \text{otherwise} \end{cases}, \tag{1.6}$$

where K is called the MPR capability. That is, simultaneous transmissions will be received successfully as long as the number of transmissions does not exceed the maximum allowed K. The first line of (1.6) represents the *success* of the transmis-

sions while the second line represents a *failure* or *outage*. This model is easy to use as it depends only on the single integer parameter K that is mainly a function of system parameters such as the spreading gain as well as the physical channels. Note that this model is a special case of (1.3); for example, $K = 4$ corresponds to

$$\mathbf{E}_{\mathsf{MPR}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} . \tag{1.7}$$

Also, note that $K = 1$ reduces the model to the collision channel.

Finally, aside from the signal processing approach to MPR in CDMA and multiuser multiple-antenna systems, a class of error control codes were specifically constructed in [18] for MPR on a noise-free additive multiple-access channel. This class of codes, known as T-out-of-N codes, allows a receiver to decode multiple superposed transmissions from (at most) T users out of N potential users with uniquely-assigned codebooks. Constraints apply to T and N; for example, N must be a prime number satisfying $N = m \cdot T + 1$ for some positive integer m. To the best of the author's knowledge, these codes are not used in any notable practical system, but they set an example of how coding theory rigorously supports the concept of MPR channels.

1.2.2 MAC Protocols over MPR Channels

MPR provides an opportunity to make significant improvements over random access systems. Nagaraj *et al.* [13] studied pure ALOHA on the MPR channel of (1.6) and showed through an approximate analysis that MPR potentially results in significant throughput gains when pure ALOHA is employed. Indeed, the authors showed that an MPR system could achieve the optimal throughput asymptotically as $K \to \infty$ providing that the offered traffic load does not exceed a threshold of order K. This is an important result as it shows that the throughput performance of MPR systems can surpass the inherent constraints imposed by the collision channel. What is also shown by that study is the significance of access control mechanisms to leverage the potential gains provided by MPR.

Centralized control of medium access on the MPR channel has been studied in the literature, e.g., [19, 20, 21, 22]. The protocol proposed in [19] uses bitmap-type explicit reservations before the transmission of data packets. In this scheme, time is divided into slots, each composed of two sub-slots: reservation and data transmission. Nodes with packets to transmit enter the reservation phase that specifies the winners deterministically. Hence, the proposed protocol should be categorized as a centralized access scheme. The performance of the protocol is compared to ALOHA-type protocols and is shown to outperform ALOHA for relatively light traffic loads. This performance gain tightly depends on the traffic load – the system is prone to instability for high traffic loads unless the reservation phase is expanded that will have a negative impact on the effective bandwidth efficiency.

Similarly, protocols proposed in [20] and [21] require central controllers. Reference [20] proposed a protocol where the central controller determines an access set based on quality-of-service requirements of data packets. The performance of this protocol seemingly depends on the ability of the users to saturate the medium, and the system efficiency supposedly declines with sporadic traffic, which can be addressed by random access schemes. A similar protocol was proposed in [21] where the optimal access set is determined by the central controller and broadcasted to the users prior to the data transmissions. Therefore, despite achieving near-optimal throughput performance in data transmission slots, the comparisons made in these papers to the throughput of ALOHA-type protocols are not fair in our view.

In [22], a multi-group priority queuing MAC protocol for cellular wireless networks was proposed. The protocol performs a user classification based on the activity history and applies a *deterministic* control policy over users' transmissions. The requirement of a central controlling entity, such as a base station or access point, puts the above protocols in a different category compared to random access methods that are of our interest in this thesis.

Random access protocols are generally prone to instability when the load is high or when the number of backlogged packets has become large. For example, it is known that ALOHA with Poisson arrivals on the collision channel is unstable when the number of transmissions per packet is not limited, because the number of backlogged packets cannot be bounded in the long run. An important question

is whether ALOHA is stable on the MPR channel. Stability of slotted ALOHA with unbounded user population was studied in [14, 15] for the probabilistic MPR model of (1.3). The first paper assumes a centralized control of the access probability in order to determine the stability region of a random access system with MPR. The second paper, then, proposes a probabilistic decentralized scheme that stabilizes the system within the derived stability region. What the papers ultimately conclude is that slotted ALOHA is stable for some nonzero system load (from an unbounded user population) if the probability of decoding a packet in the presence of infinitely many interferers is nonzero, a condition that does not seem realistic in practice.

Stability of finite-user slotted ALOHA was addressed in [16]. It was shown that slotted ALOHA can be more stable, in terms of the region of supportable loads, depending on the probability of decoding each packet in the presence of interference. That is, as one implements stronger MPR receivers, the resulting system can show more stability (and equivalently, support higher loads [23]) than ideal TDMA with the collision channel.

The majority of the papers reviewed so far use abstract multiple-parameter probabilistic MPR models. Spatial models of MPR based on received signal to interference-plus-noise ratio (SINR) were employed in [24, 25] to study the throughput performance of ad hoc wireless networks. A heuristic distributed backoff mechanism based on the history of decoding successes and failures was proposed in [24]. Each transmitting node maintains a status based on feedbacks from its intended receiver and selects the transmission probabilities accordingly. The paper studies the protocol under a spatially-distributed network and shows that it can address the problem of unfairness in multiple-access systems without power control. Reference [25] offers a computational analysis of the local and multihop throughput of slotted ALOHA. Multihop networks are out of the scope of this thesis.

Exponential backoff (EB) in slotted random access systems was analyzed in [26] and backoff parameters were optimized as a function of MPR channel parameters. The same authors also propose a cross-layer algorithm in [27] where the access point recognizes the identity of collided users and request retransmissions. Having obtained information about the number of contending nodes in the next slot, the nodes select the optimal access probability for retransmission of their packets. An

interesting observation in this paper is that the proposed scheme performs almost equally well for IFT and DFT transmissions because the number of concurrent retransmissions is softly controlled by the access point.

Reference [28] proposed a heuristic scheme to estimate the number of active nodes in a single-hop multiple-access system. The estimate was then used to select an access probability in a slotted random access setup. The estimation process proposed in that paper requires that the number of active nodes remains constant for a sufficiently long period of time. An interesting result presented in the paper is that the difference between the performance of ALOHA-type and CSMA-type random access is reduced by MPR. Nevertheless, we consider carrier sensing in MPR systems [28, 29] only hypothetical as it becomes increasingly difficult for a receiver to perform energy-based carrier sensing in multidimensional signal spaces that are essential to multiuser detection.

In this thesis, we are interested in slotted-ALOHA-type random access schemes as detailed in the next section. Particularly, we will propose schemes in Chapters 3 and 4 that estimate the number of active nodes efficiently, which, unlike the estimation scheme of [28], do not necessarily require the number of active nodes to remain constant. Due to the importance of EB mechanisms in practice, we will compare our proposed protocol in Chapter 4 to a number of commonly used random access protocols including EB where it will be shown that our proposed near-optimal access scheme outperforms an EB scheme with preset parameters for a large range of the number of users. Furthermore, we will show in Chapter 3 that, even if practical, implementation of receivers that are capable of recognizing collided receivers is not crucial for designing efficient random access schemes on the MPR channel.

1.3 Thesis Contributions and Organization

Basic concepts of medium access control, and random access in particular, were presented in this chapter. Throughput performance of random access over a shared collision channel was reviewed and was shown to be low compared to the capacity of the channel corresponding to a TDMA scheme. We also presented concepts and models of the MPR channel and briefly introduced the MPR model we will use in the rest of this thesis. The content of this chapter was mainly taken from the

existing literature.

In the next chapters, we will study the MPR channel and its potentials to support higher MAC efficiencies with random access, which is of significant importance and interest due to the inherently low complexity of random access schemes. We will aim at proposing a framework for very efficient random access schemes, mainly by employing the following elements:

- We focus on slotted random access. Slotted schemes are more efficient and more convenient to analyze. Furthermore, if time-division duplexing (TDD) is employed, a slotted setup is required for an MPR multiple-access system in order to ensure that nodes avoid transmissions to a receiver at the time it is busy transmitting ACK packets to other nodes.

- Since the focus of this thesis is on random access over MPR channels implemented through multiple-access techniques such as CDMA where energy-based carrier sensing is infeasible or impractical, we assume that nodes are informed about the traffic load only through ACK packets and not by carrier sensing or overhearing other transmitters. This makes the proposed framework highly advantageous for scenarios such as machine-to-machine (M2M) traffic [30] on the uplink of cellular and satellite communications systems where random access is practical due to the sporadic nature of the traffic while carrier sensing is infeasible.

- The backoff mechanism is implemented by selecting an access probability for each slot, which results in a *geometric* distribution of the backoff period that is different from the window-based *uniform* backoff distribution in IEEE 802.11. Nevertheless, similar results are applicable to IEEE 802.11 and similar protocols since uniformly distributed and geometrically distributed backoffs behave similarly at the large scale as long as the number of active nodes is not too small [31].

Chapter 2 presents analyses of the throughput efficiency of random access on the MPR channel. We will show that random access schemes tend to achieve larger throughput performance as the MPR capability of the channel improves. In particular, the average number of packets per time unit approaches the optimal throughput

that can be achieved through a perfect TDMA strategy with larger MPR capability K. This throughput can be obtained providing that contending nodes get timely updates of the contention level over the medium. We will introduce basic scenarios and derive equations and bounds on the throughput of random access on the MPR channel.

In Chapter 3, a theoretical model for the offered traffic is presented and bounds on the throughput are computed. In that chapter, the traffic is assumed to follow a decision Markov model. Two bounds are then computed based on the amount of nodes' instantaneous knowledge of the contention. Particularly, the upper and lower bounds correspond, respectively, to the two scenarios where *i*) the nodes are aware of the immediate number of contending nodes (the genie-aided scenario) and *ii*) the nodes are only aware of the long-term traffic parameters and transmit with a constant access probability. The modeled scenarios are simulated and shown to achieve throughput figures that are limited by the above bounds. We will show that measures related to burstiness of the traffic affect the actual throughput relative to these two bounds. Queuing analysis is also provided to compute the corresponding delay figures.

Chapter 4 presents a practical MAC framework for controlling the access probability for a generic traffic model. To this end, the system state, defined as the number of contending nodes, is modeled as a stationary process that, at each time instant, is the summation of the previous system state and a random component. Then, a state tracking mechanism based on the theory of extended Kalman filters is derived, which is shown through simulations to stabilize the operational point close to the optimum point. Since the scheme proposed in that chapter takes near-optimal action as a function of the information it receives from the feedback channel, it outperforms the random access schemes reviewed in the previous section such as exponential backoff.

The majority of the literature on random access schemes on the MPR channel, such as the papers reviewed in the previous section, are suitable only for terrestrial applications since they generally assume immediate feedback from the receiver and/or carrier sensing mechanisms, none of which are available in satellite multiple-access systems. Meanwhile, there is a recent research trend on a family of repetition

random access (RRA) schemes that aim at reducing packet error rate on the forward channel (rather than relying mainly on immediate ACK packets for collision resolution) through voluntary retransmissions by the users and iterative collision resolution at the receiver [32, 33, 34]. These papers consider single-user detection for design and analysis of their proposed protocols. In Chapter 5, we first present an analysis of an RRA scheme on the MPR channel and address the effect of multiuser detection on MAC parameters such as the required frame length for minimum probability of frame failure. Later in that chapter, we address the problem of selecting the delay-optimal access probability in the presence of non-negligible propagation delay. We will derive scaling laws on the ratio between throughput-optimal and delay-optimal access probabilities. We will conclude that MPR implementations allow us to smooth the essential compromise between throughput and delay in the presence of large propagation delays.

Finally, Chapter 6 summarizes the main contributions of the thesis and points out future research directions on the subject.

Chapter 2

Random Access with Multipacket Reception

Studying a multiple access system with the conventional Poisson traffic provides insight on behaviors and limits of the system. In this chapter, we will review fundamentals of random access over the MPR channel modeled in Chapter 1 with emphasis on binomial and Poisson arrivals. First, we describe the model and introduce bounds on the throughput of random access over the MPR channel. Then, we introduce a method of maintaining belief vectors that assist us in tracking the system state. More practical scenarios will, then, be examined in the next chapters.[1]

2.1 System Model

We focus on slotted random access in the rest of this thesis; that is, time is divided into time slots of identical duration. Packets have identical length at the medium access level, each taking one time slot to be transmitted. Unless stated otherwise, we neglect the propagation time and assume the nodes to be slot-synchronized. These assumptions allow us to use a discrete time index $t \in \mathbb{N} \cup \{0\}$.

We study a *single-hop* multiple access system consisting of a receiver and $M \leq \infty$ transmitters. M may represent either the total number of nodes in the vicinity or the number of currently admitted nodes if connections follow a call-admission control (CAC) mechanism.[2]

[1] The content of this chapter was partially included in [35, 36, 37].

[2] The latter case offers a better representation of the system model in Chapter 3 as we will assume that the traffic rate and "burstiness" of the nodes are known. This assumption is well accommodated in a network with CAC.

Receivers are capable of multipacket reception with the model described by

$$N^{\mathsf{rx}} = \begin{cases} N^{\mathsf{tx}} & \text{if } N^{\mathsf{tx}} \leq K \\ 0 & \text{otherwise} \end{cases} , \tag{2.1}$$

where $K \geq 1$ is called the MPR capability, and N^{tx} and N^{rx} denote the number of concurrently transmitted and successfully received packets, respectively. The first line represents the *success* of the receiver while the second line represents a *failure* or *outage*.

Note that the MPR model of (2.1) is a special case of the following more general model briefly introduced in [38]: When n packets are transmitted simultaneously, the probability of reception failure per packet is 0 if $n \leq K$ and it is 1 otherwise. This "shifted" step function may be replaced by other nondecreasing functions that, for example, can be calculated based on received signal to interference-plus-noise ratio (SINR) and other system parameters. Let θ_n denote the probability of reception failure per packet in the presence of $n-1$ interferers. Then, this generalized MPR model can be represented by

$$\mathbf{E} = \begin{bmatrix} \theta_1 & 1-\theta_1 & 0 & 0 & 0 & \cdots \\ \theta_2^2 & \binom{2}{1}\theta_2(1-\theta_2) & (1-\theta_2)^2 & 0 & 0 & \cdots \\ \theta_3^3 & \binom{3}{1}\theta_3^2(1-\theta_3) & \binom{3}{2}\theta_3(1-\theta_3)^2 & (1-\theta_3)^3 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} . \tag{2.2}$$

This model can be simply described by the vector $(\theta_1, \theta_2, \theta_3, \cdots)$ rather than a matrix.

In the following, we briefly review three specific multiple-access scenarios and how they relate to the MPR models we have described.

Gaussian Multiple-Access Channel Here, we examine the case that any number of transmissions are received successfully as long as the total rate transmitted by all the users does not exceed the capacity of the Gaussian multiple-access channel (GMAC). Suppose all users share a common channel of bandwidth W. Then, if n nodes transmit with the common rate R, the receiver successfully decodes all the packets if

$$nR \leq W \log(1 + n\mathsf{SNR}) , \tag{2.3}$$

where SNR is the signal-to-noise ratio (SNR) per packet. In this model, K is the maximum number n of transmissions such that the above inequality holds.

CDMA with Matched-Filtering Consider a random-sequence CDMA system with power control at the transmitters and matched filtering (MF) at the receiver. When n nodes transmit simultaneously, the received signal to interference-plus-noise ratio (SINR) of the ith signal follows

$$\mathsf{SINR}_i = \frac{P_i}{\sigma^2 + \frac{1}{L}\sum_{j\neq i} P_j} \ , \tag{2.4}$$

where P_j is the post-MF received power of the jth user, σ^2 is the Gaussian noise power, and L is the spreading gain. Since we assumed power control at the users, we have

$$P_1 = P_2 = \cdots = P_n := P \ .$$

Hence, the received SINR of any of the users follows

$$\mathsf{SINR} = \frac{P}{\sigma^2 + \frac{n-1}{L}P} \ . \tag{2.5}$$

If the receiver has a threshold γ_{th} on the SINR for successful detection/decoding, then K is the maximum number of concurrent transmissions n such that the following inequality still holds

$$\frac{P}{\sigma^2 + \frac{n-1}{L}P} \geq \gamma_{\mathsf{th}} \ . \tag{2.6}$$

Therefore,

$$K := \left\lfloor L\left(\frac{1}{\gamma_{\mathsf{th}}} - \frac{1}{\gamma_{\mathsf{u}}}\right)\right\rfloor + 1 \ , \tag{2.7}$$

where $\gamma_{\mathsf{u}} := P/\sigma^2$ is the post-MF received SNR of each of the users. More complex receivers may demand more complex calculations of the effective interference to obtain K (e.g., see [39]).

Generalized Modulation on Satellite Uplink Generalized modulation (GM) [40, 41] enables multiuser detection with relatively low complexity and high spectral efficiency theoretically approaching the GMAC capacity (2.3). We are interested to examine the MPR performance of GM in a subscriber-to-satellite (uplink) multiple-access system.

Here, we examine a satellite multiple-access system where a number of users transmit at the same time on the same frequency band. The users employ GM

with antipodal modulation, $M = 4$ repetitions per symbol and a spreading gain of $L = 4$ (see [41] for details). Three different environments (rural, suburban, urban) are considered and the corresponding channel parameters, i.e., the mean and the variance of the log-normal path-loss, are obtained from [42, Table VII]. Iterative demodulation is employed at the receiver. We use the following equation to compute SINR evolutions in the iterative demodulator [41]:

$$\sigma_{k,i}^2 = \sum_{k' \neq k} \frac{P_{k'}}{L} g\left(\frac{M-1}{M} \frac{P_k}{\sigma_{k,i-1}^2}\right) + \sigma^2 \quad . \tag{2.8}$$

Here, $\sigma_{k,i}^2$ denotes the power of the residual interference-plus-noise on the kth signal at the ith iteration, P_k is the received power of the kth signal, σ^2 is the power of the Gaussian noise, and $g(\cdot)$ is defined as [41]

$$g(x) \triangleq \left(\mathsf{E}\left[1 - \tanh(x + \sqrt{x}\xi)\right]\right)^2 \quad ,$$

where $\mathsf{E}[\cdot]$ denotes the expected value, and $\xi \sim \mathcal{N}(0,1)$. The signal from user k is assumed detected if the corresponding SINR grows above a certain threshold.

We simulated each scenario several times for different number of users n and computed the SINR evolutions for each case from which the probability of success per user $1-\theta_n$ was obtained. The result is shown in Figure 2.1. It can be seen that for rural environments where channel variations are very limited, the MPR model of (2.1) closely approximates the behavior of the multiuser detection with GM. For suburban and urban areas, let

$$K := \left\lfloor \sum_{n=1}^{\infty} (1 - \theta_n) \right\rceil \quad , \tag{2.9}$$

where $\lfloor \cdot \rceil$ denotes the rounding function. Note that this definition of K is consistent with the special case of (2.1) where $1 - \theta_n$ equals 1 for $1 \leq n \leq K$ and equals 0 for $n > K$. We will briefly show in Section 2.4 that the MPR model of (2.1) where K is obtained by (2.9) is a good model for practical cases such as the examples examined here.

A key point in employing joint detection techniques is that when a detector is focussing only on a subset of the concurrent transmissions, the receiver may still need

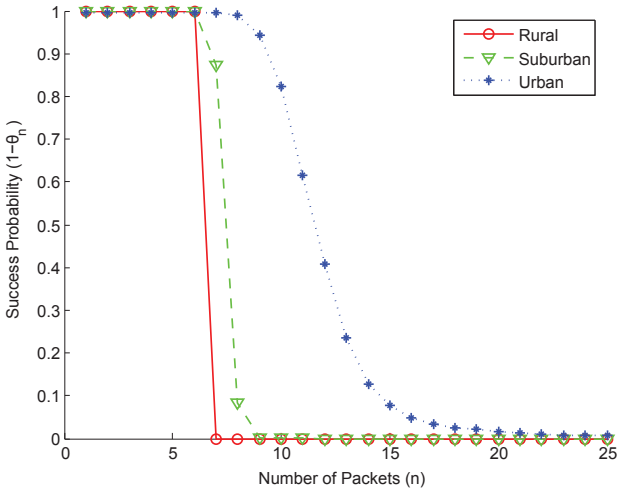

Figure 2.1: Success probability per packet as a function of the number of transmissions.

to decode all transmissions for the purpose of canceling interference. This means that assumptions about which receiver is decoding which packets are generally irrelevant at the medium access level. Therefore, in the rest of this thesis, we assume that the system has one receiver willing to decode all packets.

2.1.1 MAC Efficiency versus Spectral Efficiency

In this thesis, we mainly focus on system throughput and average packet delay as the main performance measures. The system throughput is defined as the average number of packets successfully received and decoded at the receiver per time unit. Since the focus is on slotted random access, the duration of a time slot is considered the time unit for simplicity. A time slot equals the transmission time of a packet plus guard intervals. In this framework, a question is how the MAC-level system throughput is related to the spectral efficiency, maximizing which is one of the main goals of designing more efficient multiple-access systems.

Conventionally, the MAC efficiency of a protocol is considered a value in $[0, 1]$ where 1 corresponds to the TDMA-equivalent system; that is, the goal of a MAC protocol is usually to achieve as bigger portion as possible from the TDMA throughput. This roots in the use of the collision model where concurrent transmissions can only be destructive and, therefore, the $\sim \log n$ capacity gain of the GMAC capac-

ity with respect to the number of concurrent transmissions n is usually ignored. As mentioned before, modern MUD schemes such as GM [40, 41] are capable of approaching the GMAC capacity, hence using simultaneous transmissions to their advantage in terms of the achievable total rate. In this thesis, we normalize the system throughput by K to take into consideration the extra bandwidth resources that are used to implement the MPR channel. In this sense, the MAC efficiency is equivalent to the multiple-access spectral efficiency for any multiuser detector that is capable of achieving a linearly-growing capability K as a function of bandwidth (or any other resources that provide extra signal dimensions). Therefore, the aforementioned $\sim \log n$ advantage of capacity-achieving techniques such as GM is not reflected in our throughput diagrams.

2.1.2 Notations

We let $P_X(x) := \Pr\{X = x\}$ denote the probability distribution function (PDF) of random variable X. Accordingly, we let $P_{X,Y}(x,y) := \Pr\{X = x, Y = y\}$ and $P_{X|Y}(x|y) := \Pr\{X = x \mid Y = y\}$. Expected value and variance of random variable X are denoted by $\mathsf{E}[X]$ and $\mathsf{Var}(X)$, respectively. We use $X \sim \mathcal{B}(N,p)$ and $Y \sim \mathcal{P}(\lambda)$ to state that X and Y follow the binomial distribution and Poisson distribution, respectively; that is,

$$P_X(x) = f_{\text{bin}}(x; N, p) \triangleq \binom{N}{x} p^x (1-p)^{N-x} \ ,$$

$$P_Y(y) = f_{\text{pois}}(y; \lambda) \triangleq \exp(-\lambda)\frac{\lambda^y}{y!} \ ,$$

where $f_{\text{bin}}(x; N, p)$ and $f_{\text{pois}}(y; \lambda)$ denote probability mass functions (PMF) of binomial and Poisson distribution, respectively. Similarly, cumulative distribution functions (CDF) of these distributions are denoted by $F_{\text{bin}}(x; N, p)$ and $F_{\text{pois}}(y; \lambda)$, i.e.,

$$F_{\text{bin}}(x; N, p) \triangleq \sum_{i=0}^{x} f_{\text{bin}}(x; N, p) \quad , \quad F_{\text{pois}}(y; \lambda) \triangleq \sum_{i=0}^{y} f_{\text{pois}}(y; \lambda) \ .$$

These notations will be consistently used throughout the thesis.

2.2 Lower Bound: No State Information

In this section, we examine a system with the conventional Poisson arrival model. In this model, the aggregate offered traffic in terms of the number of packets N_t at any time slot t follows a Poisson distribution with mean λ; that is, $N_t \sim \mathcal{P}(\lambda)$ for all t independent of the history of the system. Note that this system is equivalent to $M \to \infty$ where each of the M nodes offers a traffic of λ/M packets per slot in average. We sometimes omit the time index for simplicity and call N the system *state*. This is illustrated in Figure 2.2.

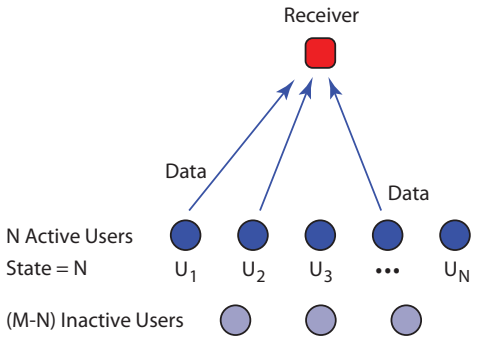

Figure 2.2: A snapshot of a system of M users and a single receiver.

In the case where no state information is available to the N transmitters, we may assume that all the N packets will be transmitted simultaneously;[3] that is, $N^{\mathsf{tx}} := N$. The resulting system throughput follows:

$$\begin{aligned} S_K(\lambda) &= \mathsf{E}_N\left[N^{\mathsf{rx}}\right] && (2.10)\\ &= \sum_{n=1}^{K} n \Pr\left\{N^{\mathsf{tx}} = n\right\} && (2.11)\\ &= \sum_{n=1}^{K} n \exp(\lambda)\frac{\lambda^n}{n!} \quad \text{packets/slot} \ . && (2.12) \end{aligned}$$

The above throughput can also be written in terms of the packet loss probability (PLR). A packet fails to be received successfully if more than $K-1$ *other* packets are transmitted simultaneously at the same slot. This probability follows:

$$\mathsf{PLR}(\lambda) = \Pr\left\{\text{fail} \mid \text{transmission}\right\} = \sum_{n=K}^{\infty} \exp(\lambda)\frac{\lambda^n}{n!} = P(K,\lambda) \ , \qquad (2.13)$$

[3]This is assumed with no loss of generality; if, for example, each packet is transmitted with a preset probability $p < 1$, the outcome will be similar to a system with $N \sim \mathcal{P}(p\lambda)$.

where $P(\cdot,\cdot)$ denotes the regularized lower incomplete gamma function defined as

$$P(s,x) = \frac{1}{\Gamma(s)} \int_0^x t^{s-1}\, e^{-t}\, \mathrm{d}t \ .$$

Then, the throughput can be obtained by:

$$S_K(\lambda) = \lambda\big(1 - \mathsf{PLR}(\lambda)\big) = \lambda Q(K,\lambda) \quad \text{packets/slot} \ , \tag{2.14}$$

where $Q(x,y) = 1 - P(x,y)$ is the regularized upper incomplete gamma function.

Figure 2.3 shows the PLR as a function of the average offered traffic. The average packet arrival λ is normalized by K for ease of illustration. Similarly, both packet arrival rate and expected throughput are normalized by K in Figure 2.4. We use this convention in the rest of this thesis. [4]

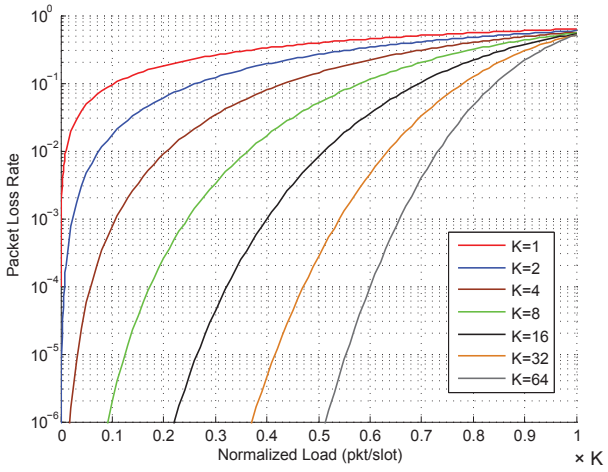

Figure 2.3: Packet loss rate vs. average load for different values of K.

Figure 2.5 shows the maximum throughput of ALOHA[5] for different values of K. The average load at which the maximum throughput occurs is also shown. This figure suggests that $\lambda_K^*/K \to 1$ and $S_K^*/K \to 1$ as $K \to \infty$. This is indeed the case since we have

$$\lim_{K\to\infty} Q(K,1) = 1 \ . \tag{2.15}$$

This means that ALOHA-type random access is asymptotically optimal on the MPR channel as long as $\lambda < K$ (see the dashed curve in Figure 2.4).

[4] We assume that the resources allocated for guard intervals, transmitting ACK packets, etc. are also proportional to K so that the comparisons between normalized measures are fair.

[5] Slotted ALOHA is simply called ALOHA for brevity in the rest of this thesis.

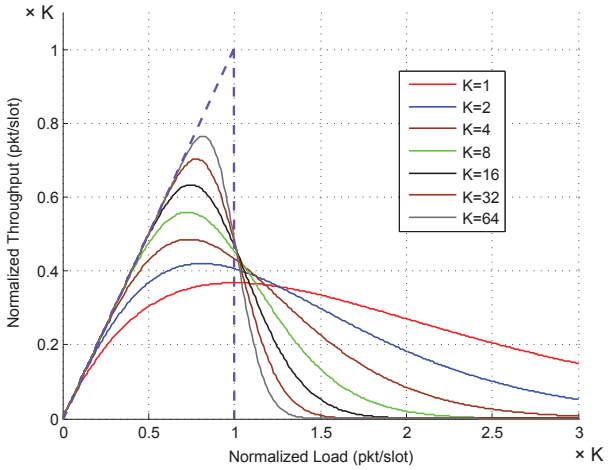

Figure 2.4: Normalized system throughput vs. average load for different values of K. The dashed curve correspond to the asymptotic case $K \to \infty$.

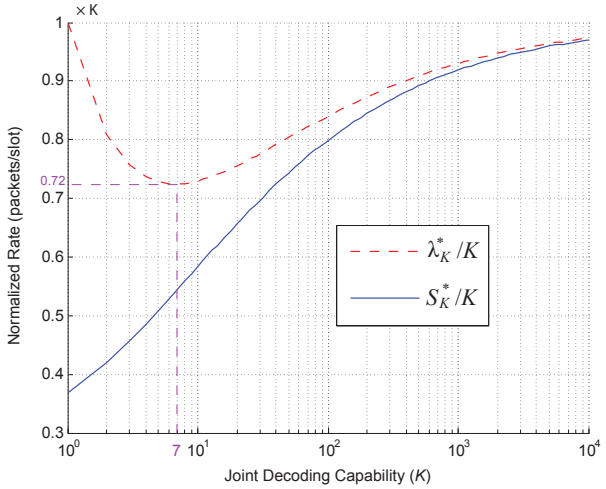

Figure 2.5: Maximum throughput of ALOHA S_K^* for different values of K as well as the corresponding average load: $S_K^* = S_K(\lambda_K^*)$.

2.3 Upper Bound: Perfect State Information

In this part, we examine "genie-aided" random access, i.e., when each node attempts to access the shared medium with a probability p based on perfect state information available to all the nodes. We assume the symmetric scenario where all the traffic has the same priority and, therefore, all nodes select the same probability p.

When N nodes attempt to access the medium, each with probability p, the

expected throughput follows

$$\begin{aligned} R_K(N,p) &\triangleq \mathsf{E}\left[N^{\mathsf{rx}} \mid N,p\right] \\ &= \sum_{n=0}^{K} n\binom{N}{n}p^n(1-p)^{N-n}. \end{aligned} \tag{2.16}$$

We define the throughput-optimal transmission probability by

$$p_K^*(N) \triangleq \underset{0<p\leq 1}{\arg\max}\ R_K(N,p)\ . \tag{2.17}$$

We have $p_K^*(N) = 1$ for $N \leq K$. For $N \geq K$, however, this probability is a strictly decreasing function of N and $p_K^*(N) \propto 1/N$ for large N in order to yield a bounded nonzero expected number of transmissions $\mathsf{E}[n] = Np_K^*(N)$.

Figure 2.6 compares the throughput of genie-aided probabilistic access (perfect state information) to ALOHA (no state information). It can be observed that, for each value of K, the genie-aided scheme ensures an asymptotic throughput that is at least as large as the peak of ALOHA throughput.

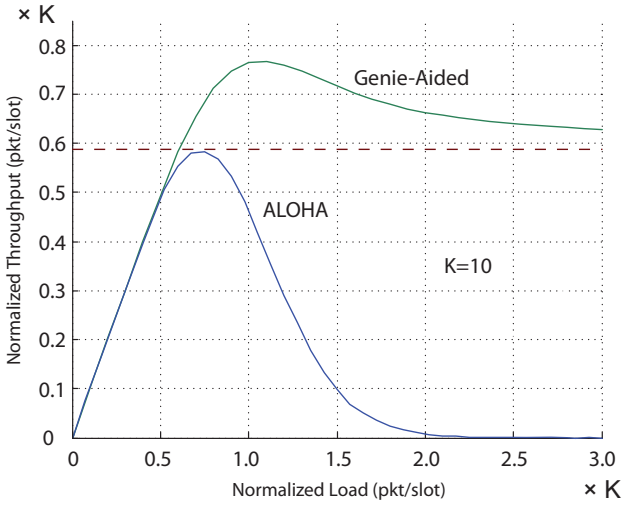

Figure 2.6: Throughput of genie-aided random access compared to ALOHA. Curves are obtained by both simulations and numerical calculations.

2.3.1 An Approximation to the Genie-Aided Rate

First, we show that $p_K^*(N)$ decreases at a rate of approximately λ_K^*/N for large N.

Proposition 1. *For a finite constant K, we have*

$$\lim_{N\to\infty} Np_K^*(N) = \lambda_K^*\ . \tag{2.18}$$

Proof As $N \to \infty$, we should have $p_K^*(N) \to 0$ in order to yield a bounded expected number of transmissions $\mathsf{E}[n] = Np_K^*(N)$. Denoting $p_K^* = p_K^*(N)$ for brevity, we have

$$\begin{aligned} \lim_{N\to\infty} \Pr\left\{N^{\mathsf{tx}} = n\right\} &= \lim_{N\to\infty} \binom{N}{n} {p_K^*}^n \left(1-p_K^*\right)^{N-n} \\ &= \lim_{N\to\infty} \frac{(Np_K^*)^n}{n!} \left(1-p_K^*\right)^N \ , \end{aligned} \tag{2.19}$$

for any finite n. We have $Np_K^*(N) = N$ for $N \leq K$. But the subsequence $\left[Np_K^*(N)\right]_{N=K}^{\infty}$ is decreasing and lower-bounded and, therefore, it has a limit L_K (see Figure 2.7). Replacing p_K^* by L_K/N in the above equation, we conclude that the asymptotic distribution of N^{tx} is Poisson with mean L_K. Therefore, $L_K = \lambda_K^*$ as it would otherwise imply a contradiction with the optimality of $p_K^*(N)$. □

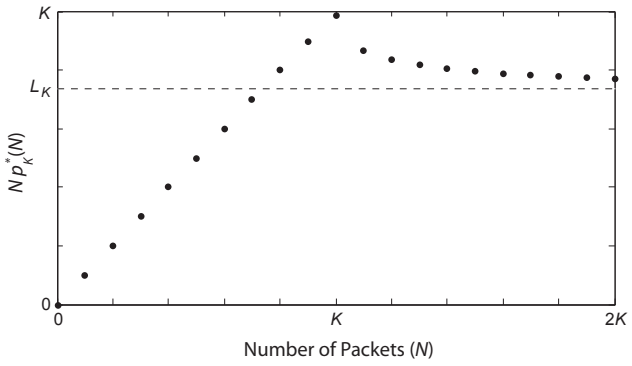

Figure 2.7: Convergence of the sequence $Np_K^*(N)$.

Then, we approximate the genie-aided throughput with Poisson traffic as follows. We have $N^{\mathsf{tx}} \leq K$ when either the number of active nodes is $N = N^{\mathsf{tx}}$ and all of them transmit with probability 1, or $N > K$ active nodes transmit with some probability

$p_K^*(N) < 1$. It then follows that

$$\begin{aligned}
\Pr\left\{N^{\mathsf{tx}} = n\right\} &= \sum_{N=n}^{\infty} \exp(-\lambda)\frac{\lambda^N}{N!}\binom{N}{n} {p_K^*}^n \left(1-p_K^*\right)^{N-n} \\
&= \exp(-\lambda)\frac{\lambda^n}{n!} + \sum_{N=K+1}^{\infty} e^{-\lambda}\frac{\lambda^N}{N!}\binom{N}{n} {p_K^*}^n \left(1-p_K^*\right)^{N-n} \\
&\simeq \exp(-\lambda)\frac{\lambda^n}{n!} + \sum_{N=K+1}^{\infty} e^{-\lambda}\frac{\lambda^N}{N!}\binom{N}{n} \left(\frac{\lambda_K^*}{N}\right)^n \left(1-\frac{\lambda_K^*}{N}\right)^{N-n} \quad (2.20) \\
&\simeq \exp(-\lambda)\frac{\lambda^n}{n!} + \sum_{N=K+1}^{\infty} \exp(-\lambda)\frac{\lambda^N}{N!}\exp(-\lambda_K^*)\frac{{\lambda_K^*}^n}{n!} \quad (2.21) \\
&= \exp(-\lambda)\frac{\lambda^n}{n!} + \exp(-\lambda_K^*)\frac{{\lambda_K^*}^n}{n!}P(K+1,\lambda) \ , \quad (2.22)
\end{aligned}$$

where (2.20) is by Proposition 1 for all $N > K$, (2.21) is the Poisson approximation to the binomial distribution, and (2.22) holds due to the Poisson-gamma relation. Finally, the expected throughput of genie-aided probabilistic access is approximated by

$$\begin{aligned}
R_K^*(\lambda) &\triangleq \mathsf{E}_N\left[R_K\left(N, p_K^*(N)\right)\right] \\
&\simeq \sum_{n=0}^{K} n\left(e^{-\lambda}\frac{\lambda^n}{n!} + e^{-\lambda_K^*}\frac{{\lambda_K^*}^n}{n!}P(K+1,\lambda)\right) \\
&= \lambda Q(K,\lambda) + \lambda_K^* Q(K,\lambda_K^*)P(K+1,\lambda), \quad (2.23)
\end{aligned}$$

where $P(\cdot,\cdot)$ and $Q(\cdot,\cdot)$ are incomplete gamma functions, and $N \sim \mathcal{P}(\lambda)$.

Note that the approximation obtained from Proposition 1 gives a strictly sub-optimal transmission probability while errors due the Poisson approximation are essentially negligible. Therefore, the approximation of (2.23) behaves as a lower bound for most values of interest as depicted in Figure 2.8.

2.4 An Alternative MPR Model

We showed through an example in Section 2.1 that the actual behavior of the multiuser detectors may be slightly different from the MPR model we use in this thesis. Here, we examine ALOHA and genie-aided random access for the case of urban environment in Figure 2.1 and compare the results with the MPR model suggested by (2.1) and (2.9).

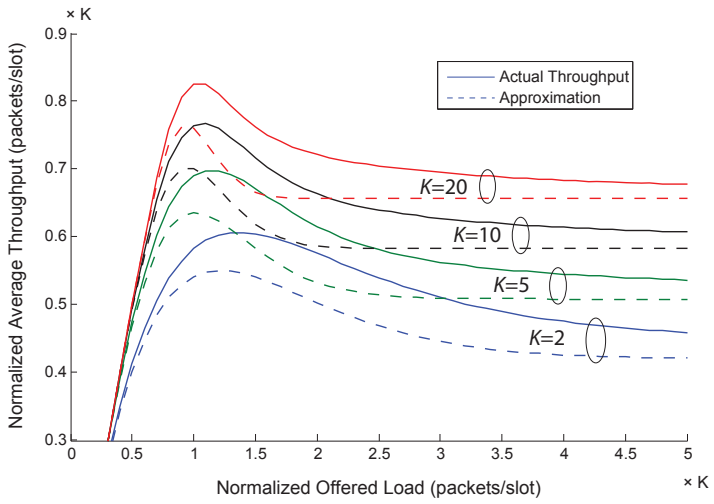

Figure 2.8: Throughput of genie-aided random access compared to the derived approximation.

For the example of urban environment in Figure 2.1, we have

$$K := \left\lfloor \sum_{n=1}^{\infty} (1 - \theta_n) \right\rceil = \lfloor 11.4 \rceil = 11 \ .$$

Figure 2.9 shows the results of simulations comparing the actual system with the equivalent model $K = 11$. For the case of genie-aided random access, the optimal access probability is computed for $K = 11$. It can be seen that the suggested approximate MPR model closely models the behavior of the system. Note that for larger K, the gap between the lower and upper bounds shrinks, which results in increasing accuracy of the approximate model.

2.5 Random Access by Using Belief Vectors

In this section, we continue to use the slotted MAC model and MPR model from the previous section. An *active node* is defined as a node that has a packet ready to be transmitted in the next time slot. It is assumed that an active node acquires an estimate of the number of active nodes N in the vicinity of its intended receiver. The estimation mechanism, in general, may include PHY/MAC cross-layering, e.g. monitoring and tracking the medium traffic by overhearing other connections and transmissions, and may be aided by explicit announcements.

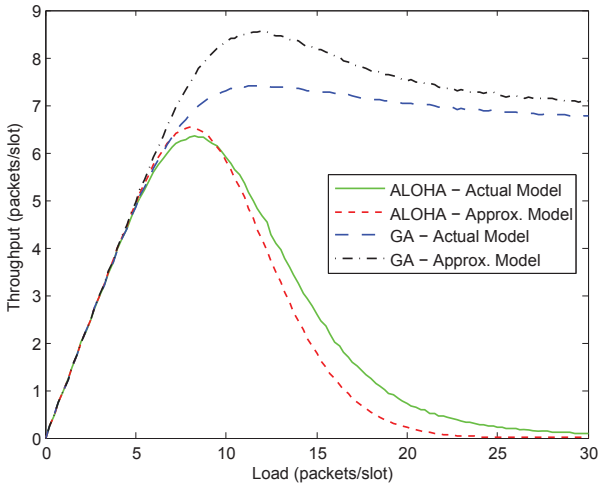

Figure 2.9: Throughput of the actual MPR model obtained in Section 2.1 compared to the suggested approximate model.

Based on the acquired statistics of the current traffic, an active node creates a belief vector of the form

$$\mathbf{b} = (b_1, b_2, \ldots, b_M) \ ,$$

where $M \leq \infty$ is the total number of nodes and b_i is the node's belief that there are currently $N = i$ active nodes, including itself, in the vicinity. The belief distribution is essentially a function of the estimation process, network parameters, and the instantaneous offered traffic.

2.5.1 Analysis of the Uniform Belief Distribution

Since all active nodes observe almost the same environment, in the following analysis we consider the case that an active node assumes that all the other active nodes select the same transmission probability as the one it selects. Hence, the expected throughput (according to the node's current belief) and the corresponding optimal transmission probability respectively follow as

$$R_K^*(\mathbf{b}) \triangleq \max_p \sum_{i=1}^{M} b_i R_K(i,p) \ , \tag{2.24a}$$

$$p_K^*(\mathbf{b}) \triangleq \arg\max_p \sum_{i=1}^{M} b_i R_K(i,p) \ , \tag{2.24b}$$

where $R_K(i,p)$ is defined by (2.16).

The following example discusses the above formulation for a simple scenario. Suppose that an active node has estimates of the minimum and maximum number of currently active nodes, and assumes a belief vector $\mathbf{b}$ of the following form

$$b_i = \begin{cases} (N_2 - N_1)^{-1} & N_1 < i \leq N_2 \\ 0 & \text{otherwise} \end{cases} ,$$

where $0 < N_1 < N_2 < \infty$ are given integer constants. We use the following approximations in order to calculate $p_K^*(\mathbf{b})$:

- N_1 and N_2 are large enough so that $R_K(N,p) \simeq S_K(Np)$, where $S_K(\lambda)$ denotes the throughput of ALOHA for average traffic λ (see Figure 2.6); in other words, the binomial distribution can be approximated by the Poisson distribution.

- K is large enough so that we can approximate $S_K(\lambda)$ by a piecewise linear function $T(\lambda)$ of the form

$$T(\lambda) = \begin{cases} \lambda & 0 \leq \lambda \leq \lambda_1 \\ -\alpha_1 \lambda + \alpha_2 & \lambda_1 < \lambda \leq \lambda_2 \\ 0 & \lambda_2 < \lambda \end{cases} , \tag{2.25}$$

  as illustrated in Figure 2.10.

- We ignore the edge effects of rounding large non-integer quantities, and approximate sums with integrals.

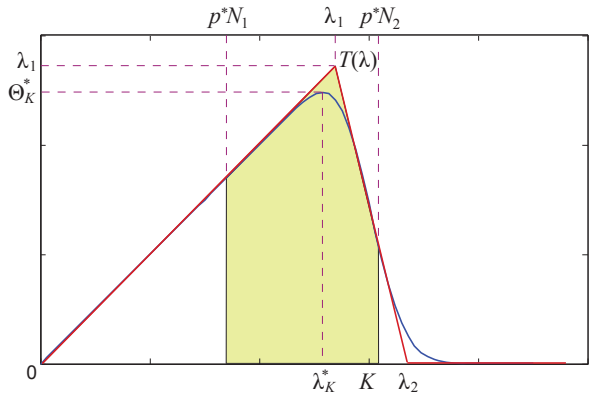

Figure 2.10: Illustration of the uniform belief scenario. $S_K(\lambda)$ is approximated by the piecewise linear function $T(\lambda)$.

Let $p_K^*(\mathbf{b})$ denote the optimal transmission probability given K and $\mathbf{b}$. From Figure 2.10, it is clear that $p^* := p_K^*(\mathbf{b})$ satisfies $p^* N_1 \leq \lambda_1 \leq p^* N_2$. Therefore, it

follows

$$\begin{aligned} p^* &\simeq \arg\max_p \frac{1}{N_2 - N_1} \sum_{i=N_1}^{N_2} T(pi) \\ &\simeq \arg\max_p \frac{1}{p(N_2 - N_1)} \int_{pN_1}^{pN_2} T(\lambda) d\lambda \\ &= \sqrt{\frac{\alpha_2 \lambda_1}{{N_1}^2 + \alpha_1 {N_2}^2}} \\ &= \frac{1}{\mu_{\mathbf{b}}} \sqrt{\frac{\alpha_2 \lambda_1}{(3{d_{\mathbf{b}}}^2 + 1)(\alpha_1 + 1) + 2\sqrt{3} d_{\mathbf{b}} (\alpha_1 - 1)}} \ , \end{aligned}$$

where $\mu_{\mathbf{b}}$ and $d_{\mathbf{b}} = \sigma_{\mathbf{b}}/\mu_{\mathbf{b}}$ are the mean and relative standard deviation of $\mathbf{b}$, respectively.

From (2.25), we obtain

$$\lambda_K^* \simeq \lambda_1 = \frac{\alpha_2}{\alpha_1 + 1} \ .$$

We observe that $p^* \simeq \lambda_K^*/\mu_{\mathbf{b}}$ when $d_{\mathbf{b}} = 0$. However, as the node becomes uncertain, i.e. $d_{\mathbf{b}} > 0$, its selection of p^* decreases by the following factor

$$\frac{p_K^*(\mathbf{b})}{p_K^*(\mathbf{b}^0)} = \sqrt{\frac{\alpha_1 + 1}{(3{d_{\mathbf{b}}}^2 + 1)(\alpha_1 + 1) + 2\sqrt{3} d_{\mathbf{b}} (\alpha_1 - 1)}} \ , \tag{2.26}$$

where $\mathbf{b}^0 := (b_0^0, b_1^0, \cdots)$ with

$$b_i^0 = \begin{cases} 1 & i = \lfloor \mu_{\mathbf{b}} \rceil \\ 0 & \text{other} \end{cases} \ ,$$

where $\lfloor \cdot \rceil$ denotes the rounding function.

Figure 2.11 shows the ratio (2.26) as well as the resulting normalized expected throughput $R_K^*(\mathbf{b})/K$ for different uncertainty levels $d_{\mathbf{b}}$. The above analysis shows how a node becomes conservative (dashed lines in Figure 2.11), when its certainty about the offered traffic decreases, in order to optimize the expected throughput. It can be seen that although uncertainty has an adverse effect on the expected throughput, increasing K results in good performance even for larger uncertainty levels. For example, an uncertainty of $d_{\mathbf{b}} = 10\%$ with $K = 20$ results in almost the same expected throughput as $d_{\mathbf{b}} = 25\%$ with $K = 100$.

Finally, as $K \to \infty$ we have $\alpha_1 \to \infty$ and therefore

$$\lim_{K \to \infty} \frac{R_K^*(\mathbf{b})}{S_K^*} = \lim_{K \to \infty} \frac{p_K^*(\mathbf{b})}{p_K^*(\mathbf{b}^0)} = \frac{1}{\sqrt{3}\, d_{\mathbf{b}} + 1} \ ,$$

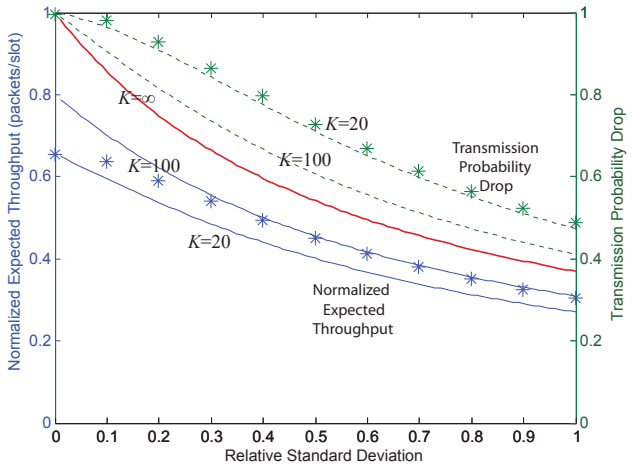

Figure 2.11: Transmission probability and normalized throughput with a uniform belief distribution. Results of the approximate analysis are shown for $K \in \{20, 100, \infty\}$. Results obtained through numerical optimizations are also presented for $K = 20$ (asterisks).

which is marked in Figure 2.11 by $K = \infty$.

Figure 2.11 also shows the actual optimal transmission probabilities as well as the corresponding expected throughput values when only the first approximation $R_K(N, p) \simeq S_K(Np)$ is applied. It can be seen that the above analysis is acceptable for a wide range of uncertainty levels for $K = 20$. Note that the analysis becomes more accurate for larger K as $T(\lambda)$ represents $S_K(\lambda)$ more precisely (see Figure 2.4).

The closed-form result (2.26) is significant especially from the viewpoint of implementation complexity. First, it reduces the optimization (2.24) of a double summation to one of a single summation, followed by a multiplicative factor that solely depends on $d_{\mathbf{b}}$. Second, $p_K^*(\mathbf{b}^0)$ depends only on an integer value, $\lfloor \mu_{\mathbf{b}} \rceil$, and can be tabulated.

2.5.2 Simulation Results

We simulate a system where at each time slot a random number of nodes uniformly drawn from the integer interval $[N_1, N_2]$ become active and send their packets with p^* calculated in the previous section. The simulations are run for different values of N_1 and N_2, but the numbers are fixed in each run and, therefore, a fixed p^* is used at all the nodes in each run. Obviously, a real-world scenario following the above

uniform statistics would have the complexity advantage that nodes are not required to calculate transmission probabilities continuously for all time slots.

Figure 2.12 compares the simulation and analytical results for $K = 20$.

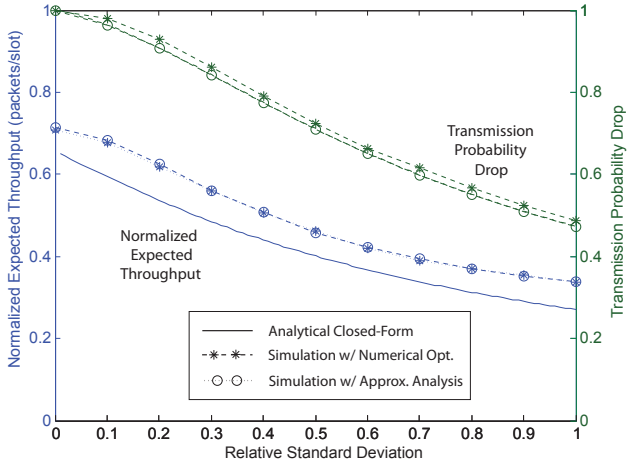

Figure 2.12: Simulation results of a uniform distribution scenario for $K = 20$ and $\mu_{\mathbf{b}} = 40$.

2.6 Summary

In this chapter, we studied random access on the MPR channel with the conventional Poisson traffic. The study provided insight on benefits and limits of employing MUD at the PHY layer in order to gain throughput at the MAC layer. We introduced two bounds on the throughput of random access: *i*) the genie-aided scenario where nodes are informed about the number of contenders before transmission, and *ii*) the scenario in which no information about the system state is available to the nodes. It was shown that, in the genie-aided scenario, the asymptotic throughput is guaranteed to be at least as large as the peak of the ALOHA curve for the corresponding value of the MPR capability K.

Then, a method of maintaining belief vectors was introduced through which nodes can select the best access probability given imperfect information about the system state. We will make further use of this method in the next chapter where a practical algorithm is proposed for a Markovian class of offered traffic.

Chapter 3

Controlled Random Access with Markovian Traffic

We briefly reviewed in Chapter 2 how a node can select the access probability when perfect state information is not available. In this chapter, we use a Markovian traffic model resulting from aggregation of interrupted on/off packet sources as detailed in the following section. We will examine lower and upper bounds and how effectively nodes can optimize the system throughput by tracking the system state through the inherent feedback of acknowledgement packets.[1]

3.1 System Model

In the system modeled in this chapter, packets are generated by on/off Markov sources [45] as depicted in Figure 3.1. An on/off Markov source S is in the *active* state when it generates a new packet, and is in the *idle* state otherwise. The probabilities α and δ together with the slot duration determine/model the packet generation rate of the source as well as its level of "burstiness" – when α and δ are small, the source shows more *inertia* to remain in a given state. Smaller transition probabilities increase the predictability of the source and, as we will see, can be utilized in estimating the offered traffic. This will be explained in more details in Section 3.4.

We then propose a model for the traffic offered by each user. The model is illustrated in Figure 3.2. Each transmitting node has an on/off Markov source

[1] This study was published in [43, 44].

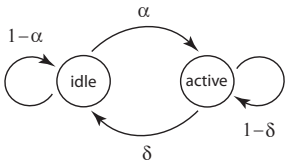

Figure 3.1: On/off Markov source used to model packet generation in this chapter.

with parameters α and δ generating packets that enter a queue and wait to be serviced. At each time slot, if the queue is occupied, TX transmits a queued packet with probability p, and immediately receives feedback informing the node if the transmission was successful. In case of success, the packet is removed from the queue. Otherwise, the packet is backlogged. Similarly to the definitions of active and idle for a *source*, we call a *node* (transmitter) active if its queue is occupied, and we call it idle otherwise.

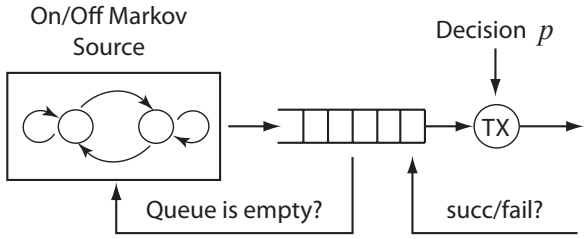

Figure 3.2: Abstract model of a transmitter.

For simplicity of analysis, we suppose that the source's transitions between the idle and active states are synchronized with the time slots. To avoid the problem of unbounded queuing delay, the source *pauses* when it is in the idle state and the queue is occupied. This, in part, guarantees that the source does not start a new burst of packets before the previously queued burst is completely serviced. Moreover, it transfers the Markovian behavior of the source to the output of TX for the sake of analytical tractability. We will address this in more details later.

3.2 System Analysis

3.2.1 A Decision-Theoretic Description

Let us look at the stochastic process of the traffic generated by the M nodes. If $M \leq K$, the MPR capability is sufficient to receive packets no matter how many nodes are active. That is to say, the persistent-transmission strategy ($p = 1$) is optimal. In this case, the number of active nodes N forms a *stationary* Markov chain with the state space $\{0, 1, \cdots, M\}$.

If $M > K$, however, the probability of N exceeding K is nonzero, packets become backlogged eventually, and the optimal decision is some $0 < p \leq 1$ in general. As a result, the system state follows a Markov process with the transition probability matrix

$$\mathbf{T} = \mathcal{T}\big(\alpha, \delta, M, K, p\big) \triangleq [\tau_{ij}]_{ij}$$

described by

$$\tau_{ij} = \tau_{ij}^{\mathsf{succ}} + \tau_{ij}^{\mathsf{fail}} \quad , \tag{3.1}$$

$$\tau_{ij}^{\mathsf{succ}} = \begin{cases} 1 & i \leq K \\ \displaystyle\sum_{(m,n,k)\in\mathcal{Z}_{ij}} f_{\mathrm{bin}}\left(m; i, p\right) f_{\mathrm{bin}}\left(n; m, \delta\right) f_{\mathrm{bin}}\left(k; M-i, \alpha\right) & 0 < i - K \leq j \\ 0 & \text{otherwise} \end{cases} , \tag{3.2a}$$

$$\tau_{ij}^{\mathsf{fail}} = \begin{cases} f_{\mathrm{bin}}\left(j-i; M-i, \alpha\right) \displaystyle\sum_{m=K+1}^{i} f_{\mathrm{bin}}\left(m; i, p\right) & K < i \leq j \\ 0 & \text{otherwise} \end{cases} , \tag{3.2b}$$

$$\mathcal{Z}_{ij} \triangleq \Big\{ (m,n,k) \quad | \quad \begin{aligned} &\max(0, i-j) \leq m \leq \min(i, K), \\ &\max(0, m+j-M) \leq n \leq \min(m, m+j-i), \\ &k = m - n + j - i \Big\} \ . \end{aligned}$$

where $\tau_{ij}^{\mathsf{succ}}$ and $\tau_{ij}^{\mathsf{fail}}$ are the transition probabilities following the success or failure of the receiver, respectively. This process is obviously non-stationary since it is a function of the nodes' *decision* on p.

The resulting model is called a partially-observable Markov decision process [46], i.e., a Markov decision process with *hidden* states. State transitions of an individual

transmitter in this system are illustrated in Figure 3.3. Note that P_s is, in general, a function of the system state (i.e., the states of all the nodes) as well as p.

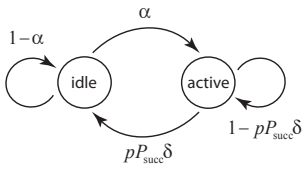

Figure 3.3: State transitions of an arbitrary individual transmitter.

Similarly to the argument for the genie-aided case in Chapter 2, we focus on optimization of the expected throughput of the *current* time slot obtained by (2.24b) due to complexity considerations, and neglect the slight improvement that a long-term-optimal decision on p may achieve through a computationally complex dynamic-programming procedure [47].

3.3 Belief Update and Access Control

In this section, we assume that CAC mechanisms are devised that allow users to know the number of admitted users and the traffic they offer; that is, M, α and δ are known to the users. In the proposed MAC scheme, nodes receive feedback $\mathcal{F}_t$ at the end of every time slot t informing them of the outcome of the transmissions in that slot.[2] A typical example is the binary success/failure feedback through ACK/NACK messages from the receiver as depicted in Figure 3.4. The type of feedback depends, in general, on the application, signal processing capability of the receiver, bandwidth of the feedback channel, etc. Examples of feedback types are introduced and studied later. Having received $\mathcal{F}_t$, nodes compute the updated belief vector $\mathbf{b}_t$.

Given a hidden Markov process and a sequence of observations (i.e., limited feedback), what is the probability of the process being in a given state at a given time instant? The forward-backward algorithm [48] solves this problem. The algorithm consists of two stages: the forward pass which is executed in the forward time direction and is used to update posterior marginal beliefs based on previous observations, and the backward pass to smooth the past beliefs. In our application,

[2]Generalization of the algorithm to cases of less frequent feedback is straight-forward.

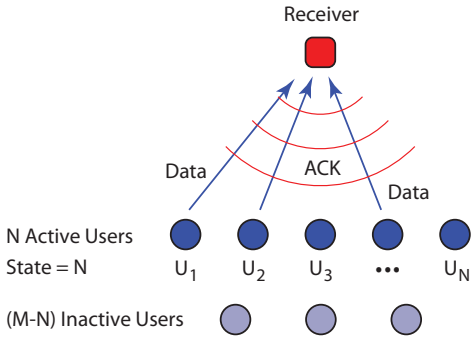

Figure 3.4: A snapshot of a system of M users and a single receiver. Each of the N active users transmit with a common access probability p and, then, receive ACK/NACK messages. All the M nodes receive this feedback and update their estimates of the system state N.

we only implement the forward pass, a.k.a. the forward algorithm, since smoothing beliefs of the past time slots does not affect a node's current decision.

Forward Algorithm (FA) Suppose that the hidden Markov process is stationary, and let state space $\mathcal{S}$ and event space $\mathcal{E}$ denote, respectively, the set of all possible hidden states and the set of all possible events observed from the underlying process. Let $\mathbf{T} = [\tau_{ij}]_{ij}$ be the $|\mathcal{S}| \times |\mathcal{S}|$ transition probability matrix given by (3.1) and (3.2). Also, let the entry ϵ_{ij} of the $|\mathcal{S}| \times |\mathcal{E}|$ event matrix $\mathbf{E}$ be the *posterior* probability of observing event j when the process is in state i. We use the discrete time index t to distinguish quantities in different time slots. For a stationary process, the belief-update output of the FA can be summarized as

$$\mathbf{b}_t = \kappa_t^{-1} \mathbf{b}_{t-1} \mathbf{T} \operatorname{diag}(\boldsymbol{\epsilon}_{\mathcal{F}_t}) \quad , \tag{3.3}$$

where $\operatorname{diag}(\mathbf{x})$ denotes the diagonal matrix taking vector $\mathbf{x}$ on its diagonal, and $\boldsymbol{\epsilon}_{\mathcal{F}_t}$ is the column of $\mathbf{E}$ indexed by the observation $\mathcal{F}_t \in \mathcal{E}$ at time t. The normalization factor κ_t^{-1} is to ensure $\sum_i \mathbf{b}_{t,i} = 1$.

We use the FA and adapt it to accommodate the non-stationarity of the underlying decision process. Suppose a node is active at the beginning of time slot t. The node's currently computed belief vector is $\mathbf{b}_{t-1}$, which was last updated after receiving $\mathcal{F}_{t-1}$ and is, therefore, one time slot old. To take into account possible changes of the system state from $t-1$ to t (before transmissions at slot t occur),

the node adjusts its belief by using the *constant* matrix $\mathbf{A}$ defined by

$$\mathbf{A} \triangleq \mathcal{T}\big(\alpha, \delta, M, K, 0\big) \ , \tag{3.4}$$

that is, the transition probability matrix with no transmission ($p = 0$). The optimal decision is therefore obtained as

$$p_t := p_K^*(\mathbf{b}_{t-1}\mathbf{A}) \ . \tag{3.5}$$

After transmissions, nodes receive $\mathcal{F}_t$ from which $(\boldsymbol{\epsilon}_t)_{\mathcal{F}_t}$ is computed as a function of p_t. Note that in order to reduce computational effort, it is sufficient to compute only the $\mathcal{F}_t$th column of $\mathbf{E}_t$ since the matrix changes over time. Finally, beliefs are updated as

$$\mathbf{b}_t = \kappa_t^{-1}\mathbf{b}_{t-1}\mathbf{T}_t \ \mathrm{diag}\left((\boldsymbol{\epsilon}_t)_{\mathcal{F}_t}\right) \ , \tag{3.6}$$

where $\mathbf{T}_t = \mathcal{T}(\alpha, \delta, M, K, p_t)$. Nodes may initialize $\mathbf{b}_0$, for example, by the uniform probability vector or the steady state probability vector corresponding to

$$\mathbf{T} = \mathcal{T}\left(\alpha, \delta, M, K, p_K^*\left(\frac{M\alpha}{\alpha+\delta}\right)\right) \ .$$

3.4 Performance Analysis and Results

3.4.1 Rate/Burstiness Description of a Source

We examine the throughput performance of the MAC scheme proposed in the previous section. The goal is to study the performance of the proposed controlled probabilistic access with limited feedback when nodes have a known degree of predictability in generation of the traffic. In the following, we translate the transition probabilities α and δ into parameters that distinguish packet generation rate and burstiness level. We define the packet generation rate of a source S as the steady state probability of S being active, i.e.,

$$\lambda_S = \frac{\alpha}{\alpha+\delta} \ . \tag{3.7}$$

It is sensible to assume that $\alpha \leq 1 - \delta$, which gives $\lambda_S \geq \alpha$.

We need a second parameter, in addition to λ_S, in order to describe an on/off Markov source. Here, we introduce a quantity called the burstiness overhead γ of the source S defined as

$$\gamma \triangleq \frac{\log(\alpha+\delta)}{\log\alpha} \ , \tag{3.8}$$

which gives $\lambda_S = \alpha^{1-\gamma}$.

Note that the pair (λ_S, γ) uniquely describes an on/off Markov source. For $\gamma = 0$ we have $\lambda_S = \alpha$, which corresponds to the case where the generation of packets are i.i.d. Bernoulli experiments with parameter $\alpha = 1 - \delta$. Note that the Poisson traffic is the special case where $M \to \infty$ and $\alpha = 1 - \delta = \lambda/M$.

3.4.2 Limited Feedback

Let $f_{\text{bin}}(\cdot;\cdot,\cdot)$ and $F_{\text{bin}}(\cdot;\cdot,\cdot)$ denote the probability mass function and cumulative distribution function of the binomial distribution. We simulate systems with the following types of feedback.

- With binary feedback $\mathcal{E} = \{\mathsf{fail}, \mathsf{succ}\}$ from the receiver informing the nodes of decoding failure or success, we have

$$\begin{cases} \Pr\{\mathcal{F}_t = \mathsf{fail} \mid N_t = i\} = 1 - F_{\text{bin}}(K; i, p_t) \\ \Pr\{\mathcal{F}_t = \mathsf{succ} \mid N_t = i\} = F_{\text{bin}}(K; i, p_t) \end{cases} .$$

- When, in the case of successful decoding, the receiver informs the users of the number of *received* packets, i.e. $\mathcal{E} = \{0, 1, \cdots, K, \mathsf{fail}\}$, we have

$$\begin{cases} \Pr\{\mathcal{F}_t = n \mid N_t = i\} = f_{\text{bin}}(n; i, p_t) \\ \Pr\{\mathcal{F}_t = \mathsf{fail} \mid N_t = i\} = 1 - F_{\text{bin}}(K; i, p_t) \end{cases} .$$

- Finally, if it is feasible to devise a feedback mechanism that reports the number of *transmitted* packets even in the case of decoding failure, we have $\mathcal{E} = \{0, 1, \cdots, M\}$, which gives

$$\Pr\{\mathcal{F}_t = n \mid N_t = i\} = f_{\text{bin}}(n; i, p_t) \quad .$$

The conditional probabilities above are the entries of $(\boldsymbol{\epsilon}_t)_{\mathcal{F}_t}$ computed upon receiving $\mathcal{F}_t$.

3.4.3 Simulation and Numerical Results

We simulate a system with $K = 5$ and $M = 10$. The average rate of the aggregate traffic offered to the medium $\lambda = M\lambda_S$ is fixed at $\lambda^* = 4.4$ for the best achievable throughput (see Figure 3.6). We examine this system for $\gamma \in [0, 0.9]$. The resulting traffic is characterized by the average burst and idle durations obtained by

$$\mathsf{E}[T_{\mathsf{burst}}] = 1/\delta, \quad \mathsf{E}[T_{\mathsf{idle}}] = 1/\alpha \quad .$$

Table 3.1 summarizes these averages for some of the values in the simulated interval.

Table 3.1: Average Burst and Idle Durations for $\lambda_S = 1/16$.

γ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\mathsf{E}[T_{\mathsf{burst}}]$	1.79	1.96	2.19	2.54	3.09	4.06	6.12	12.13	47.64	2889
$\mathsf{E}[T_{\mathsf{idle}}]$	2.27	2.49	2.79	3.23	3.93	5.17	7.79	15.43	60.64	3677

Figure 3.5 shows the throughput performance of the proposed scheme with the three aforementioned types of feedback. Two additional cases are examined as comparison baselines:

- The genie-aided case that can hypothetically achieve the best throughput among probabilistic access schemes.
- The case of constant transmission probability (CTP) where no feedback is available and nodes use only the *a priori* information to optimize the steady-state throughput. The optimal constant probability is obtained by

$$p^*_{\text{ctp}} \triangleq \arg\max_{0<p\leq 1} \sum_{i=0}^{M} \pi_i(p)\ R_K(i,p),$$

  where $\boldsymbol{\pi}(p) = \big(\pi_0(p), \pi_1(p), \cdots, \pi_M(p)\big)$ denotes the probability distribution of the steady-state of the *stationary* Markov process $\mathcal{T}(\alpha, \delta, M, K, p)$ when p is fixed.

It can be seen that for the system with the given parameters: i) different feedback types result in almost the same performance, and ii) increasing γ improves the throughput performance, starting from the CTP performance for $\gamma = 0$ and increasing to the genie-aided performance. These values, which turn out to remain almost constant for fixed λ_S and variable γ, can be used as approximate[3] bounds for the throughput of probabilistic access with limited feedback. The above results make sense because, with non-bursty traffic, the information conveyed by feedback about the system state in the subsequent time slots is low and the resulting performance is approximately equal to the case where no instantaneous information about the state N_t is available, while this information increases as γ approaches 1. Indeed, for

[3]The reason why these values do not act as exact bounds is the slight difference between short-term-optimal and long-term-optimal selections of p explained earlier.

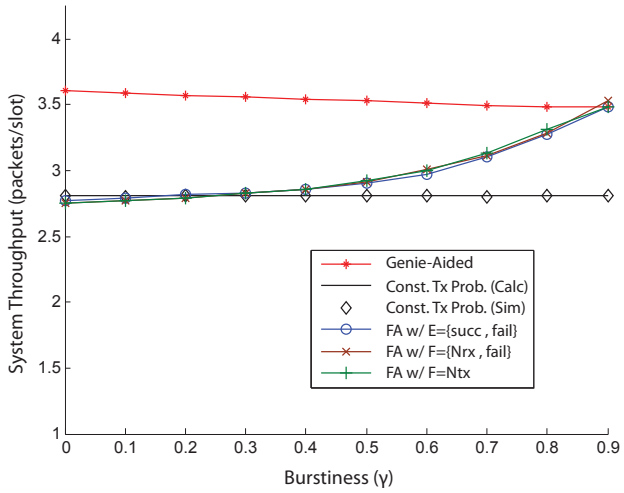

Figure 3.5: Throughput of FA with different types of feedback versus the throughput by the state-aware genie and constant transmission probability.

γ very close to 1, nodes rarely change their states and, hence, N_t remains almost constant from slot to slot, leaving sufficient time for the nodes to learn N.

Consider the extreme case of $\gamma = 1$, which corresponds to $\delta = 0$. Suppose each node chooses to become active or idle randomly and remains at the same state for all time. That is, N_t equals a constant N for all t. Let us consider the case of $\mathcal{E} = \{0, 1, \cdots, M\}$. Suppose that as the observation history $\mathcal{H} = \{\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_t\}$ grows, a maximum likelihood (ML) estimator is being used, for example, to estimate N. Then, since the ML estimator is asymptotically consistent and efficient, the estimation error tends to zero as $t \to \infty$. The FA does perform as well as any estimator, in the sense of the expected error, and therefore the belief vector tends to put asymptotically all the probability mass on N; that is, $\lim_{t\to\infty}(b_i)_t$ is 1 for $i = N$ and 0 otherwise. This argument applies to any feedback type – what is possibly affected by the type of feedback is only the rate of convergence.

It is observed that the bounds remain almost constant as γ varies. Let us fix γ to an arbitrary value to evaluate the bounds for different values of λ_S. The result is illustrated in Figure 3.6. It can be seen that the lower bound monotonically increases while the upper bound increases up to a maximum and then converges with the lower bound. The two bounds meet at $\lambda_S = 0$ and $\lambda_S = 1$, i.e., where the traffic behavior of the nodes are perfectly known *a priori*.

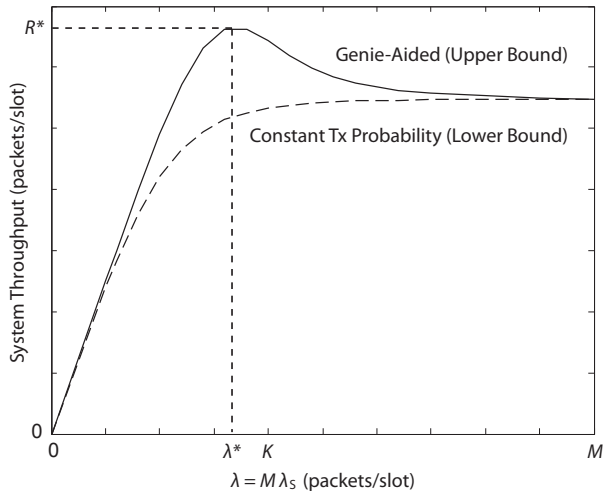

Figure 3.6: Behavior of the system throughput bounds (evaluated for $K = 4$, $M = 10$, $\gamma = 0.2$ in this figure).

Figure 3.7 shows the bounds on the throughput for different values of K. It is clear from this figure that devising a larger MPR capability increases the maximum achievable system throughput.

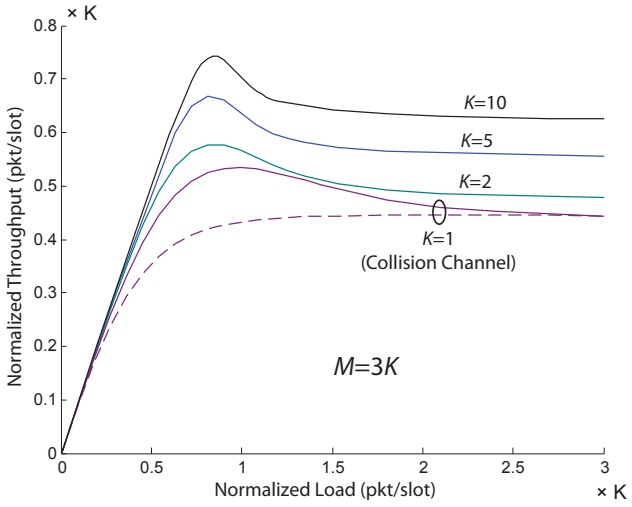

Figure 3.7: Bounds on the MAC efficiency (R/K) for $K = 1, 2, 5, 10$, $M = 3K$, $\gamma = 0.2$. The total offered rate $\lambda = M\lambda_S$ is normalized by a factor of $1/K$. Some lower bounds are omitted for clarity.

Figure 3.8 illustrates the effect of M on the maximum achievable throughput. It can be seen that smaller numbers of admitted nodes are generally better in terms of

the total system throughput providing that this number can offer sufficient amount of traffic load to avoid bandwidth under-utilization. The figure illustrates how a cross-layer admission/access control strategy may play a positive role in maximizing the MAC efficiency. Furthermore, it is evident that the transport layer can also shape the offered traffic to induce more burstiness and approach the upper bound.

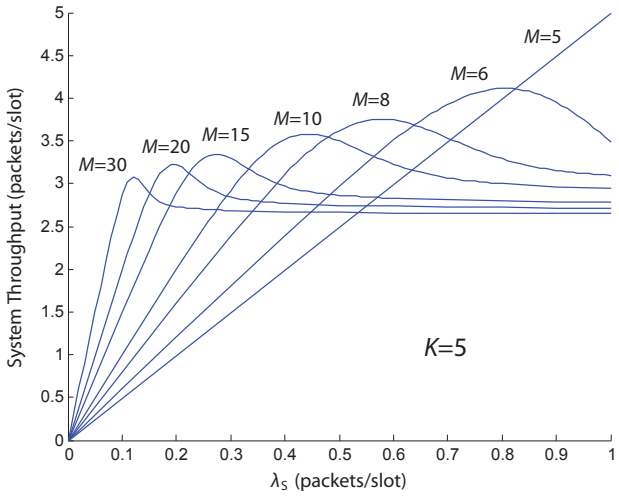

Figure 3.8: Upper bound on the throughput for $K = 5$ and different values of M. Smaller M allows to achieve higher throughput providing that each transmitter offers sufficient traffic.

3.4.4 Discussion

Figure 3.7 illustrates the performance improvement afforded by larger K. Indeed, as was also shown in Figure 2.4 for the case of Poisson traffic, one can theoretically approach 100% MAC efficiency with a sufficiently large K. It also shows how flow control allows to achieve higher system throughput. By adjusting the offered packet rate and burstiness, nodes can boost the system throughput by a factor of approximately 25% in the examined scenarios. Note, however, that the need for strict flow control to achieve the best possible MAC efficiency is gradually mitigated as K grows larger.

It can be observed from Figure 3.8 that controlling the number of admitted nodes through CAC mechanisms can further improve performance – smaller M allows larger achievable throughput as long as the aggregate offered packet rate of the admitted nodes saturates the medium. The maximum throughput promised by

the upper bound, however, is not achieved unless the traffic is sufficiently bursty.

As an illustrative example, let $K = 5$ and suppose there are no delay constraints. As a result, nodes can shape their traffic arbitrarily resulting in a desired value of γ. From Figure 3.8, the optimal choice for the number of admitted nodes is $M = 5$ if $\lambda_S > 0.82$. For a smaller value of λ_S, however, this number results in bandwidth under-utilization and a larger number of nodes should be admitted.

In the case there are delay constraints, throughput bounds should be considered as functions of both λ_S and γ.

3.4.5 Queuing Delay

Let $P_{\mathsf{srv}} = pP_s$ denote the service probability of a packet in a time slot, i.e., the probability that a packet is transmitted *and* successfully received in a time slot when the queue is occupied. We approximate this quantity by its average obtained by

$$(1 - P_e)P_{\mathsf{srv}} = \frac{R}{M} \ , \tag{3.9}$$

where P_e is the probability of the queue being empty, and R/M is a node's share of the total system throughput R. The average service time of TX is $1/P_{\mathsf{srv}}$. In order to calculate P_e, note that

$$\mathsf{E}\left[T_q\right] = \frac{1}{\delta P_{\mathsf{srv}}}, \quad \mathsf{E}\left[T_e\right] = \frac{1}{\alpha} \ , \tag{3.10}$$

where T_q (resp. T_e) denotes the average time that a queue is occupied (resp. empty); note that the second equality is ensured by the flow control mechanism. Using a renewal theory argument, we have

$$P_e = \frac{T_e}{T_e + T_q} = \frac{\delta P_{\mathsf{srv}}}{\delta P_{\mathsf{srv}} + \alpha} \ . \tag{3.11}$$

From (3.9) and (3.11), we obtain

$$P_{\mathsf{srv}} = \frac{R\alpha}{M\alpha - R\delta} \ . \tag{3.12}$$

Let $P_B(m)$ be the probability that a generated packet is the mth in the current burst, which is equal to the probability that the "reverse" chain of an on/off Markov source (Figure 3.1), starting at the active state, generates a burst of exactly m packets. Since the on/off Markov chain is equivalent to its reverse, we have

$$P_B(m) = (1 - \delta)^{m-1}\delta \ .$$

Also, define $P_Q(n)$ as the probability that a packet is the nth in the queue upon its entrance to the queue. For this to occur, if the packet is the mth one generated in the current burst, TX must have already serviced exactly $m-n$ packets out of the previous $m-1$ packets, and must have delayed the queue for exactly $n-1$ slots. Therefore,

$$\begin{aligned} P_Q(n) &= \sum_{m=n}^{\infty} P_B(m) \binom{m-1}{m-n} P_{\mathsf{srv}}^{m-n} (1-P_{\mathsf{srv}})^{n-1} \\ &= \frac{(1-\delta)^{n-1}\delta(1-P_{\mathsf{srv}})^{n-1}}{(1-P_{\mathsf{srv}}+P_{\mathsf{srv}}\delta)^n} \quad . \end{aligned}$$

It can be easily verified that $\sum_n P_Q(n) = 1$. Finally, the average queuing delay (in units of slots) is obtained by

$$\bar{D} = \sum_{n=1}^{\infty} n \left(\frac{1}{P_{\mathsf{srv}}}\right) P_Q(n) = \frac{1-P_{\mathsf{srv}}}{P_{\mathsf{srv}}\delta} + 1 = \frac{M\alpha - R(\alpha+\delta)}{R\alpha\delta} + 1 \ , \tag{3.13}$$

where (3.12) is used in the last step. Note that we have $\bar{D} = 1$ with $P_{\mathsf{srv}} = 1$.

Figure 3.9 shows the average queuing delay that a packet experiences. It is observed that the delay increases dramatically as γ approaches one.

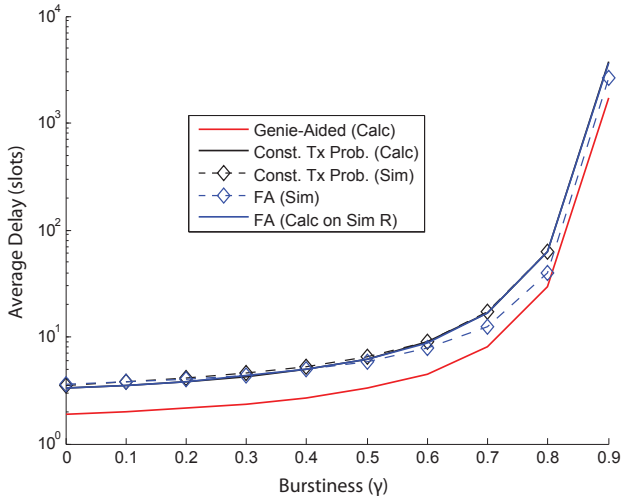

Figure 3.9: Average queuing delay versus burstiness. The results are obtained for $K = 5$, $M = 10$, $\lambda_S = 0.44$, and $\mathcal{E} = \{0, 1, \cdots, K, \mathsf{fail}\}$.

3.5 Summary

Throughput performance of probabilistic medium access in MPR-capable systems was examined in this chapter. It was shown that an increased MPR capability K results, in general, in better system throughput. In a system of nodes offering Markovian traffic, burstiness of the packetized data traffic improves the predictability of the system state and increases the throughput when only limited feedback is available. A Markov decision model was described for the system and a throughput-optimal MAC scheme was examined, which updates the nodes' belief about the system state and optimizes the access probability. Approximate lower and upper bounds on the throughput were introduced that correspond, respectively, to no information and perfect information about the instantaneous system state. It was observed that the throughput with fixed traffic load may vary between the introduced bounds as burstiness varies, starting from the lower bound for $\gamma = 0$ and approaching the upper bound with $\gamma \to 1$. This may mean a 25% improvement of the MAC efficiency, depending on K and the number of nodes, by controlling the traffic burstiness. For $M \gg K$, however, the gap between the two bounds vanishes and the feedback becomes unnecessary as long as the average traffic parameters are known. Finally, a closed-form equation for the average delay of the system was derived as a function of the throughput.

Chapter 4

Controlled Random Access Using Adaptive Filtering

In the previous chapter, we computed bounds on the throughput performance for cases where some information about the offered traffic was available. In this chapter, we assume no prior information about the traffic and the number of potential transmitters.

In this chapter, we use adaptive filtering to design a MAC protocol that tracks the number of active nodes and, in essence, keeps the nodes "conscious" about the contention level on the medium. This contrasts with most protocols used in practice such as slotted Aloha and IEEE 802.11 where nodes blindly change their backoff parameters according to simple predefined rules. The design of the EKF is inspired by the work in [49] where adaptive filtering was used to adjust the backoff window of the IEEE 802.11 MAC according to feedback from the medium.[1]

4.1 Related Work

Early after the advent of random access schemes in 1970s, researchers showed that random access is prone to instability depending on the number of nodes and backoff parameters. For example, it was shown in [51] that Aloha may show a bistability behavior – a bistable system switches between two stable "equilibrium points" where one operating point corresponds to high throughput and low delay while nodes experience frequent collision and long delays at the other point. Whether the system

[1]This study was partially published in [50].

is bistable depends on the number of nodes in the system and the retransmission probability p_r. Jenq analyzed the bistability behavior of slotted ALOHA through drift analysis [4] and provided a method for optimizing p_r when the number of nodes and the offered traffic per node is known *a priori* [52].

Dynamic adjustment of backoff parameters in random access protocols over the collision channel has been an active field of research for years. In an early effort, Kleinrock and Lam analyzed the dynamics of slotted ALOHA systems with Markovian traffic [6] and proposed a Markov-decision method for dynamic adjustment of the retransmission probability. Note that the resulting throughput of employing the above schemes is still limited by the theoretical peaks of 18% and 37% for unslotted and slotted protocols.

Several papers have proposed improvements over the simplistic BEB employed by IEEE 802.11. Haas and Deng [8] proposed to employ a backoff adjustment algorithm similar to the multiplicative-increase / linear-decrease (MILD) algorithm [53] as an alternative to BEB, which increased the MAC efficiency from $\sim 13\%$ to about $\sim 19\%$. In [49], it was shown that employing adaptive filtering to estimate the number of contending nodes achieves similar results.

As mentioned before, MPR channels increase the maximum efficiency that is achievable through random access. Nagaraj *et al.* [13] showed that ALOHA-type protocols can in principle achieve 100% efficiency asymptotically as $K \to \infty$, where K denotes the number of packets that the receiver can process jointly. For example, with slotted ALOHA, devising $K = 100$ can theoretically result in up to 80% MAC efficiency compared to a mere maximum of 37% with the collision channel [3]. However, it was also shown that the penalty of overloading the system is more severe for larger values of K.

We studied in Chapter 3 how (prior and learned) knowledge of the parameters of the offered traffic can be utilized to optimize the access probability. In this chapter, we address the more general problem where the prior knowledge of traffic parameters may not be available to the nodes. This is a valid assumption in dynamic mobile wireless networks with sporadic communications where CAC mechanisms are inefficient or infeasible and adaptive control mechanisms are favorable.

4.2 Extended-Kalman-Filter Estimation of N

Similarly to the previous chapters, a slotted MAC protocol is considered where all packets are of the same transmission duration. M nodes generate packets to be transmitted to a common receiver. Let $\lambda_n \leq 1$ (packets/slot) denote the average traffic offered by node n. Then, the aggregate load of the system is obtained by $\lambda = \sum_{n=1}^{M} \lambda_n$ packets/slot. Packets generated at each node enter the queue of that node and wait to be transmitted. The number of nodes with nonempty queues, called the active nodes, is denoted by N and is called the system *state* in that slot. In the following, we propose a random access scheme by using the theory of extended Kalman filters (EKF) where nodes track the system state based on statistics of the ACK messages received by the nodes.

Suppose the system state N follows

$$N_t = N_{t-1} + w_{t-1} \ , \tag{4.1}$$

where $t = 1, 2, 3, \cdots$ is the slot number. Here, $w_t = N_{t+1} - N_t$ models the dynamics of the generated traffic. In the context of adaptive filters, w_t is called the process noise and is assumed to follow a Gaussian distribution. We make a similar assumption as follows. Suppose that the MAC protocol is stable. Then, the average of w_t must be zero as $t \to \infty$ in order to yield a finite, nonzero average number of active nodes. For simplicity, let us assume that w_t is stationary and therefore $\mathsf{E}[w_t] = 0$.

Let us consider the symmetric scenario where each of the N active nodes transmit with probability p. Then, the probability of success in that slot follows the binomial CDF at $N^{\mathsf{tx}} = K$, i.e.,

$$p_s = \Pr\left\{N^{\mathsf{tx}} \leq K\right\} = F_{\text{bin}}(K; N, p) \ . \tag{4.2}$$

If a node measures p_s, it can calculate an estimate of N given K and p using the inverse function of the binomial CDF. To this end, each node monitors and stores the number of ACK messages in a window of B time slots and calculates the *success rate*

$$\xi \triangleq \left(\sum_{\tau=t-B}^{t-1} a_\tau \right) / B \ , \tag{4.3}$$

where

$$a_t := \begin{cases} 1 & \text{if ACK received in slot } t \\ 0 & \text{otherwise} \end{cases} \ . \tag{4.4}$$

Assuming p_s is almost constant for the duration of B slots, the numerator of (4.3) is binomially distributed with parameters B and p_s. Hence,

$$\begin{aligned} \mathsf{E}[\xi] &= p_s \ , \\ \mathsf{Var}(\xi) &= \frac{p_s(1-p_s)}{B} \ . \end{aligned}$$

Let $h(N) := p_s = F_{\text{bin}}(K; N, p)$ given K and p. We can rewrite the *observation process* as

$$\xi_t = h(N_t) + v_t \ , \tag{4.5}$$

where v_t is a zero-mean binomial random variable with variance $h(N)(1-h(N))/B$.

Equations (4.1) and (4.5) resemble continuous-time formulations of the system state and observation process used in the context of extended Kalman filters. We can now use the following to track the system state N_t [54, Ch. 8]:

Prediction:

$$\begin{aligned} \hat{N}_{t|t-1} &= \hat{N}_{t-1} \ , \\ V_{t|t-1} &= F_{t-1}^2 V_{t-1} + V_{t-1}^{(w)} \ , \end{aligned}$$

Update:

$$\begin{aligned} y_t &= \xi_t - h(\hat{N}_{t|t-1}) \ , \\ V_t^{(y)} &= H_t^2 V_{t|t-1} + V_t^{(v)} \ , \\ \kappa_t &= \frac{V_{t|t-1} H_t}{V_t^{(y)}} \ , \\ \hat{N}_t &= \hat{N}_{t|t-1} + \lfloor \kappa_t y_t \rceil \ , \\ V_t &= (1 - \kappa_t H_t) V_{t|t-1} \ , \end{aligned}$$

where $\lfloor \cdot \rceil$ denotes the round function and:

- V_t is the variance of the estimate $\hat{N}_t$,
- $V_t^{(x)}$ denotes the variance of the random variable x,
- y_t is the *innovation residual*,
- κ_t is the *near-optimal* Kalman gain, and

- F_t and H_t are the following derivatives:

$$F_t = \frac{\partial \hat{N}_{t|t-1}}{\partial \hat{N}_{t-1}} = 1 \quad , \quad H_t = \frac{\partial h}{\partial N} \mid_{N=\hat{N}_{t|t-1}} \quad .$$

We relax the constraint of N being an integer in order to be able to derive H_t. We have

$$h(N) = F_{\text{bin}}\left(K; N, p\right) = I_{1-p}((N-K)^+, K+1) \ , \tag{4.6}$$

where $I_x(\cdot,\cdot)$ denotes the regularized incomplete beta function, and $x^+ \triangleq \max(x, 0)$. Assuming N is sufficiently large so that the binomial distribution approaches the Poisson distribution, we can simplify $h(N)$ to

$$h(N) \simeq F_{\text{pois}}\left(K; Np\right) = \frac{\Gamma(K+1, Np)}{K!} \ , \tag{4.7}$$

where $\Gamma(K+1, Np)$ denotes the upper incomplete gamma function. Equations (4.6) and (4.7) are compared in Figure 4.1. It can be seen that the approximation holds when $N \gg K$. When N is close to K, however, this approximation may introduce inaccuracies in the calculation of the Kalman coefficient. Nevertheless, we use this approximation for simplicity of the equations derived in the following.

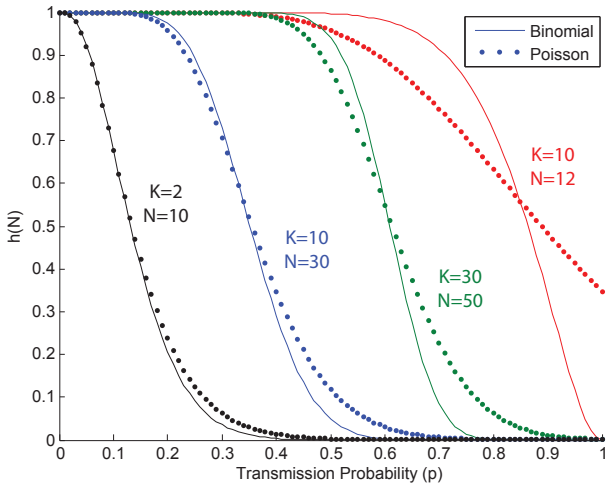

Figure 4.1: $h(N)$ calculated based on Poisson approximation to binomial distribution. It can be seen that the approximation holds when $N \gg K$.

It follows that

$$H_t \simeq \frac{-p\, e^{-\hat{N}_{t-1}p}(\hat{N}_{t-1}p)^K}{K!} \ , \tag{4.8}$$

for large N. Also, following up on the earlier discussion, we choose

$$V_t^{(v)} = \frac{h(\hat{N}_{t-1})\big(1 - h(\hat{N}_{t-1})\big)}{B} \ .$$

Finally, the EKF can be summarized as

$$\hat{N}_t = \hat{N}_{t-1} + \lfloor \kappa_t y_t \rceil \ , \tag{4.9}$$

where

$$\begin{aligned} y_t &= \xi_t - h(\hat{N}_{t-1}) \ , \\ \kappa_t &= \frac{BH_t \left(V_{t-1} + V_{t-1}^{(w)}\right)}{BH_t^2 \left(V_{t-1} + V_{t-1}^{(w)}\right) + h(\hat{N}_{t-1})\big(1 - h(\hat{N}_{t-1})\big)} \ . \end{aligned}$$

V_t is updated as

$$V_t = (1 - \kappa_t H_t) \left(V_{t-1} + V_{t-1}^{(w)}\right) \ ,$$

and H_t is given by (4.8).

Once a node updates its estimate $\hat{N}_t$, it sets its access probability to $p := p^*(N)$. Since there is no known closed form for computation of $p^*(N)$ other than approximations of the form $p^*(N) \simeq \lambda^*/N + c/N^2$, a lookup table can be precomputed and stored for implementation purposes.

4.2.1 An Alternative Observation Statistics

Instead of averaging over a window of the last B time slots, one may use the following alternative observation statistics:

$$\zeta_t = \gamma \zeta_{t-1} + (1 - \gamma) a_{t-1} \ , \tag{4.10}$$

where a_t is the last received binary feedback. Similar to the previously defined sliding window statistics, the probability distribution of this statistics tends to a Gaussian distribution. The relation between the length of the sliding window B and the *forgetting factor* γ can be obtained by equalizing the variances of the two statistics. Assuming a constant success rate p_s and, hence, stationary ξ_t and ζ_t, we have

$$\mathsf{Var}\,(\zeta) = \gamma^2 \mathsf{Var}\,(\zeta) + (1 - \gamma)^2 p_s (1 - p_s) \ . \tag{4.11}$$

Therefore,

$$\begin{aligned} \mathsf{Var}\left(\zeta\right) &= \frac{(1-\gamma)^2}{(1-\gamma^2)} p_s(1-p_s) \\ &= \frac{1-\gamma}{1+\gamma} p_s(1-p_s) \ . \end{aligned} \tag{4.12}$$

Equalizing this variance with $\mathsf{Var}\left(\xi\right) = p_s(1-p_s)/B$, we obtain $B = (1+\gamma)/(1-\gamma)$ and $\gamma = (B-1)/(B+1)$. The equations derived in the previous section can then be used by replacing $B := (1+\gamma)/(1-\gamma)$.

4.2.2 Other Parameters

To complete the design of the proposed random access scheme, two parameters remain to be set:

Variance of the process noise $V_t^{(w)}$ This quantity may be preset if prior knowledge of the traffic is available. Dynamic solutions may also be possible. Figure 4.2 compares the tracking ability of the EKF simulated with two preset values of the process noise. In both scenarios, $K = 10$ and $\gamma = 0.5$. The artificially preset system state N_t is 50 from slot 0 to slot 100 when it suddenly doubles. It can be seen that the EKF with preset $V_t^{(w)} = 10$ is faster in tracking N_t compared to the EKF with preset $V_t^{(w)} = 1$, but the estimate is noisier. We preset this parameter to $V_t^{(w)} = \lambda^*$ in our simulations.

The window size B or γ We will see in Section 4.3 that the value of B or γ that minimizes $|\mathsf{Var}\left(\hat{N}_t\right) - \mathsf{E}\left[V_t\right]|$ for a fixed N_t is an appropriate value and provides good stability behavior.

4.3 Simulation Results

In this section, we present results of simulating the multiple access system described earlier where nodes use the derivations of Section 4.2 at the end of each slot in order to calculate the access probability, common among all the transmitting nodes, for the next slot. In an actual system, each node performs the calculations independently of other nodes, but since all the nodes observe the same environment, they are expected to reach essentially similar results.

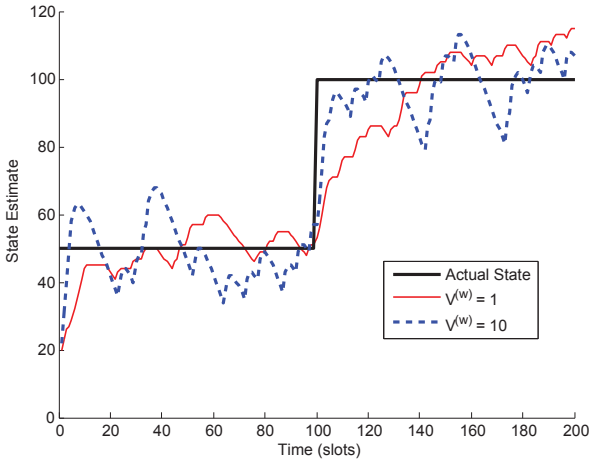

Figure 4.2: Comparison of the tracking ability with different values of the process noise $V_t^{(w)}$.

As mentioned before, a parameter to set at the design phase is the window size B or, equivalently, the forgetting factor γ. This parameter may be obtained by attempting to minimize $|\mathsf{Var}\left(\hat{N}_t\right) - \mathsf{E}\left[V_t\right]|$ for different values of (fixed) system state N as mentioned in Section 4.2. Extensive simulations show that $\gamma = 1/3$ is close to the optimal value for different values of K and N. The equivalent $B = 2$ is also shown by simulations to be the optimal window size for the observation statistics. Table 4.1 lists the best values of γ obtained for a few values of K while $N \leq 100$. We preset $\gamma = 1/3$ for the simulations in this section.

Table 4.1: Optimal γ for different K.

K	1	2	5	10	20
γ	0.34	0.35	0.37	0.30	0.36

First, we examine the throughput performance of the proposed scheme and compare it to the following baselines:

GA In the genie-aided scenario, nodes are informed of the exact value of the system state N and select the optimal access probability $p^*(N)$.

SA With slotted ALOHA, each node in the transmission mode transmits a packet with probability 1. If the transmission is not successful, the node switches to the backlog mode and retransmits the same packet with probability $p_r =$

0.1 until the packet is received successfully. Then, it switches back to the transmission mode.

BEB Nodes start transmitting new packets with probability $p_{\max} = 1/16$. Upon failure, the node halves its access probability unless it has reached the minimum value $p_{\min} = 1/1024$. Nodes switch to $p_{\max} = 1/16$ upon success.

MILD Nodes start from $p_{\max} = 1/16$. Upon failure, the node multiplies its access probability by $2/3$ unless it has reached the minimum $p_{\min}$. Upon success, the node reduces its average backoff $1/p$ by one slot unless the access probability p has reached the maximum $p_{\max}$.

Figures 4.3 and 4.4 compare the throughput of the proposed EKF-based scheme to the above baselines for $K = 4$ and $K = 8$, respectively. It can be seen that, compared to SA, BEB and MILD algorithms, our proposed scheme achieves the closest throughput to the GA bound for a large range of the number of nodes M.

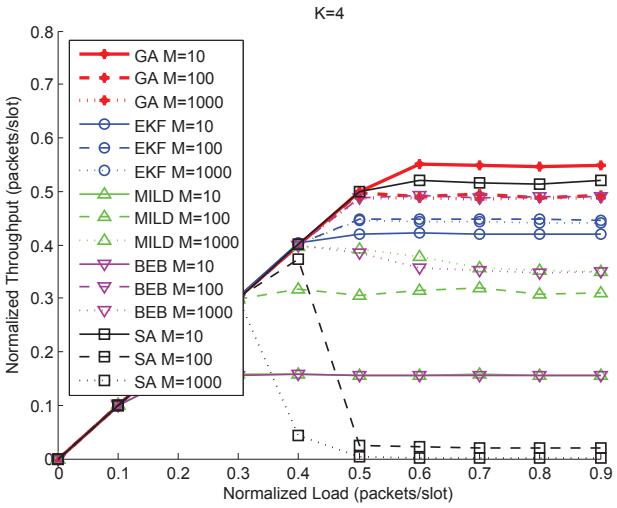

Figure 4.3: Throughput performance of the proposed EKF-based scheme compared to other schemes for $K = 4$.

In order to examine the throughput performance, we simulated a system of 100 users and a receiver with different values of K. Figure 4.5 shows the throughput, normalized by the value of K, compared to the peak of ALOHA and the ideal genie-aided case. It can be seen that the proposed method succeeds in achieving the peak of ALOHA, which was shown to increase asymptotically as $K \to \infty$. Results of

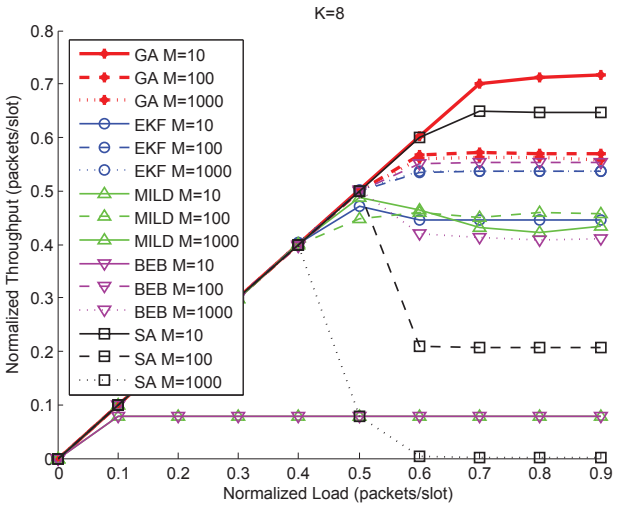

Figure 4.4: Throughput performance of the proposed EKF-based scheme compared to other schemes for $K = 8$.

simulating similar systems, where nodes are informed of the system state through a genie, are also shown for comparison.

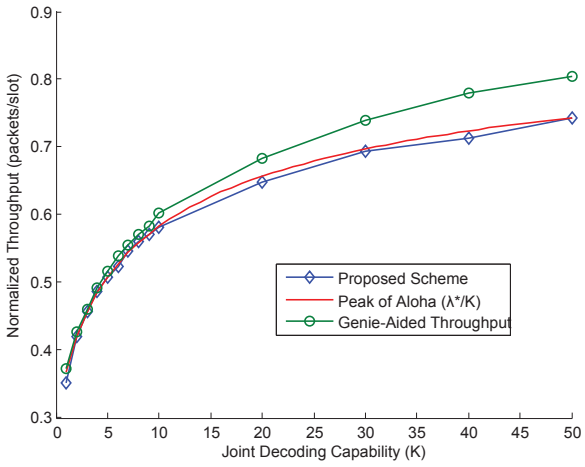

Figure 4.5: Comparison of the proposed scheme to peak of ALOHA and the ideal genie-aided scenario.

Delay performance of the proposed scheme is shown in Figure 4.6. We know that the delay increases unboundedly as λ approaches λ^*. Nevertheless, comparing λ/K (x-axis) in Figure 4.6 with λ^*/K (y-axis) in Figure 4.5, one can see that delay figures are below 50 slots when the normalized offered packet rate λ/K is about

10% below the maximum supportable load λ^*/K. The corresponding genie-aided packet delays are also plotted for comparison.

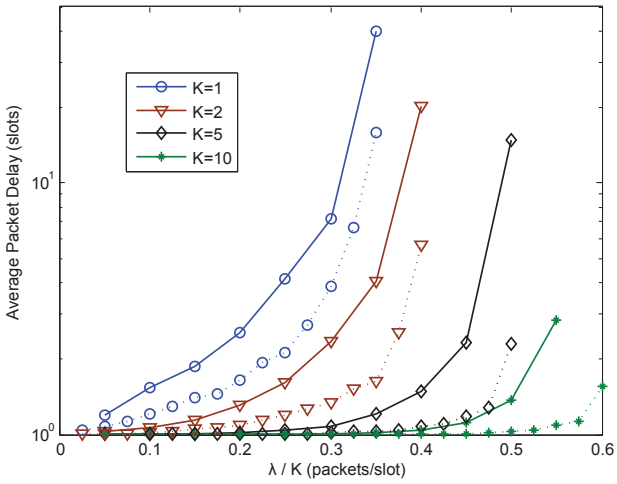

Figure 4.6: Average packet delay of the proposed scheme as a function of the offered packet rate. Dotted lines show the packet delay in the corresponding genie-aided scenarios.

4.4 Alternative Channel Models

4.4.1 Imperfect ACK Channel

Thus far, we have assumed that the ACK channel is "perfect" in the sense that nodes receive all the ACK packets. This assumption is not far from reality because short control packets such as ACK packets are normally transmitted with lower rates and, furthermore, ACK packets are transmitted on a broadcast channel where signals are less susceptible to interference. Nevertheless, no physical channel is perfect in practice and it should be instructive to examine scenarios with imperfect ACK channels.

In this section, we simply model ACK channel errors by a missing probability P_{miss} per packet. Since we assume that missing an ACK packet by a node is independent of other nodes, an imperfect ACK channel results in an asymmetry – nodes receive different pieces of information and, therefore, compute different estimates of the system state.

Another effect of missing ACK packets is that nodes may transmit duplicates of

successfully received packets, which reduces the effective throughput as the receiver recognizes and discards those duplicates. Indeed, the effective throughput equals $R(1 - P_{\mathsf{miss}})$ where R is the "raw" throughput of the system defined as the average number of successfully received packets including all the duplicates. In other words, missing ACK packets not only affects the raw throughput R through affecting the state estimate $\hat{N}$, but also reduces the effective throughput by a factor of $1 - P_{\mathsf{miss}}$.

Suppose that nodes simply skip computing a new estimate when an ACK packet is missed.[2] We simulated such a system with $M = 30$ nodes and different values of K and P_{miss}. Figure 4.7 shows the results. It can be seen that the decline of the effective system throughput is mostly due to the factor $1 - P_{\mathsf{miss}}$.

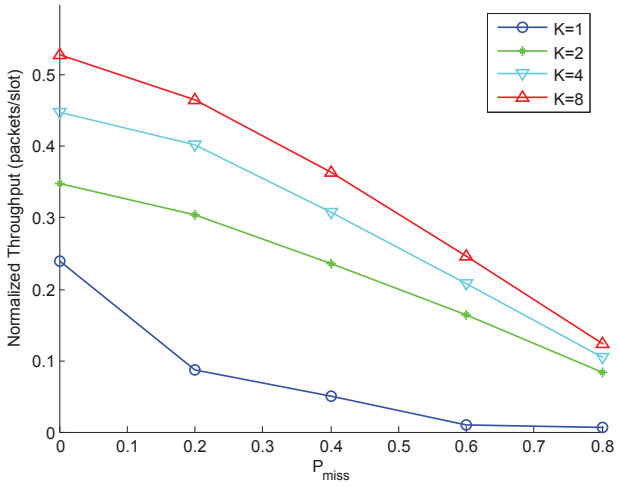

Figure 4.7: Throughput effect of missing ACK packets.

4.4.2 Delayed ACK Channel

In most terrestrial systems such as IEEE 802.11, when a node transmits a packet successfully, it receives an acknowledgement (ACK) packet almost immediately due to negligible propagation delays. This is the assumption we have used thus far in this chapter. In satellite systems or underwater acoustic networks, however, nodes have to wait for a non-negligible amount of time to receive the ACK packets. For example, the propagation delay from a ground station to a geosynchronous satellite

[2]Kalman filtering with intermittently missing observations is studied in the literature; e.g., see [55].

to another ground station is about 253 milliseconds. The round-trip time is twice this value, more than half a second, which is orders of magnitudes larger than slots of a few tens of microseconds.

If not taken into account, long round-trip times may cause instability in random access systems where nodes respond to the feedback they receive through reception acknowledgements and other control packets. It is instructive in this chapter, also, to examine the effect of large delays on the performance of our proposed scheme.

The problem of adaptive filtering with delayed measurements have been addressed in the literature; e.g., see [56] and references therein. However, the additional complexity imposed on the estimation system is justified and, more importantly, relevant only when the delay is relatively small. This is not the case in geosynchronous satellite systems.

We simulate the proposed random access system as follows. Nodes maintain two queues, a contention queue and a waiting queue. New packets enter the contention queue until they are transmitted. When a node transmits a packet, it may be willing to contend to transmit other packets. Therefore, it sends the current packet to the waiting queue for a duration of the round-trip time of the system. Then, if an ACK is received for the packet, it will be removed from the waiting queue; otherwise, the packet is moved back to the contention queue to contend for a retransmission. This process will continue until the reception of the packet is acknowledged. In this scenario, the system state N, by definition, is the number of nodes that are contending for transmission, i.e., have packets in their contention queues. In order to track the system state, we allow the nodes to use the information from delayed ACK packets and proceed with computation of the equations presented in Section 4.2 without considering the propagation delay.

Figures 4.8 and 4.9 show the throughput performance of the proposed MAC scheme in the presence of ACK delay. It can be seen that larger values of K give the random access system robustness against ACK delays. For example, it can be seen that a round-trip time of only 2 slots result in a significant drop in the throughput for the collision channel (Figure 4.8), while a similar decline of the throughput occurs, for the case of $K = 2$, at round-trip times that are orders of magnitude larger (Figure 4.9). These results suggest that using multiuser detection receivers

in satellite systems is a promising way to handle sporadic packet communications through random access methods.

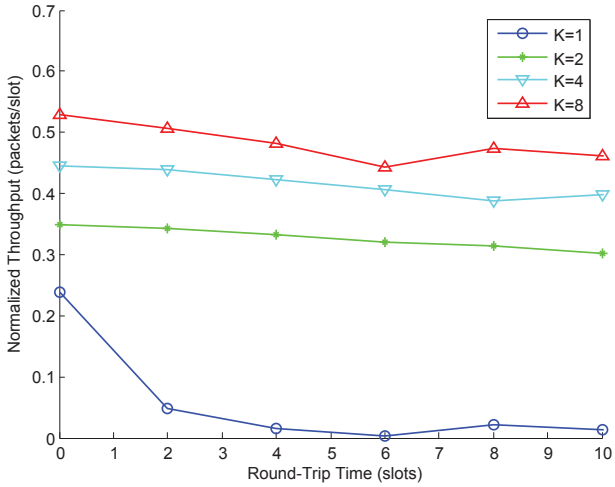

Figure 4.8: Throughput of the proposed scheme with delayed ACK packets.

4.5 Summary

In this chapter, we proposed a random access scheme that uses the theory of extended Kalman filters to track the system state. It was shown that the proposed scheme achieves the maximum of ALOHA without explicit knowledge of the traf-

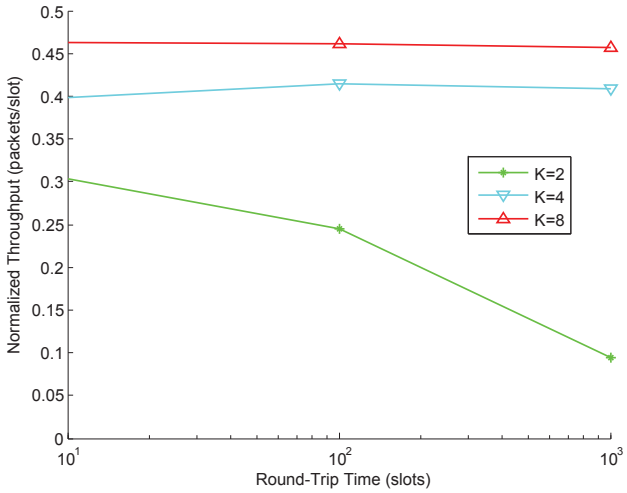

Figure 4.9: Throughput of the proposed scheme with delayed ACK packets.

fic load. This is accomplished through utilization of the limited feedback acquired through ACK messages from the receiver.

Average packet delay was shown to be reasonable as long as the offered packet rate maintains a safety margin to the maximum supportable packet rate. Particularly, delay figures remain below 50 slots when the normalized offered packet rate λ/K is about 10% below the maximum normalized supportable load λ^*/K.

Then, we examined two non-ideal cases: *i*) in the case that there is a nonzero probability that ACK packets are missed by nodes, it was shown that the dominant effect in decreasing the throughput is the extra replicas that nodes transmit while the proposed EKF-based state tracking system works properly; *ii*) when ACK packets are received with a delay, for example in the case of satellite systems, larger values of K results in more robustness against ACK delays, which suggests the use of random access systems and multiuser detectors for sporadic traffic in satellite communications.

Chapter 5

MPR Random Access on Satellite Links

The problem of controlling the access probability according to the partial state information received through ACK packets was addressed in the previous sections. In this section, we deviate from the models used thus far by adding a nonzero propagation delay to the round-trip time of the medium-access system. This model is particularly useful for satellite communication systems where long propagation delays introduce challenges at both PHY and MAC layers. In particular, random access traffic can obtain state information only with a delay that may have a significant impact on the system performance.

A recent trend in the field of random access on satellite links attempts to address the problem of long propagation delays by decreasing the dependence of medium access mechanisms on the ACK packets. In particular, it is suggested to transmit multiple replicas of the same packet in order to increase the link reliability in the forward path before delayed ACK packets are received. The main ideas of these random access schemes are *i*) providing diversity through repetitions and *ii*) improving the success probability further by allowing collision resolution aided by the successfully received replicas. While these papers focus on the collision model, we will examine repetition random access on the MPR channel in Section 5.1.[1]

Then, in Section 5.2, we will study the effect of propagation delay on the optimal access probability. Scaling laws for the throughput-optimal access versus delay-optimal access will be derived and will be compared to numerical results.

[1]This study was published in [57].

5.1 Repetition Aloha with MPR

As mentioned before, a recent research trend (e.g., [32, 33, 34]) has focused on improving the bandwidth efficiency through collision resolution by introducing transmission diversity and iterative interference cancelation. In the scheme proposed in [32], dubbed collision-resolution diversity slotted ALOHA (CRDSA), each packet is transmitted in two different randomly selected time slots. This, in part, increases the *physical* system load on the channel, but it provides diversity through the transmission of a redundant copy of each packet. Furthermore, this extra copy assists the receiver to resolve collisions through interference cancelation in the other slot.

For example, suppose the packet of user 1 is collided with the packet of user 2 in a slot s, but another copy of the same packet is received without collision in another slot s'. In this case, the receiver first decodes packet 1 received in s' and marks the user as "cleared." The received packet contains a pointer to s where the other copy of the same packet is transmitted. The receiver now reconstructs and subtracts the packet from the superposition of the two packets received at s and decodes the packet of user 2. This is illustrated in Figure 5.1. Here, retransmission of packet 1 not only provides diversity, but also helps resolving the collision with packet 2. With more users involved, more *iterations* may be needed to decode more packets. CRDSA is one of the random access schemes currently being adopted in the reverse (i.e., user-to-hub) link of two-way satellite communications standards [58].

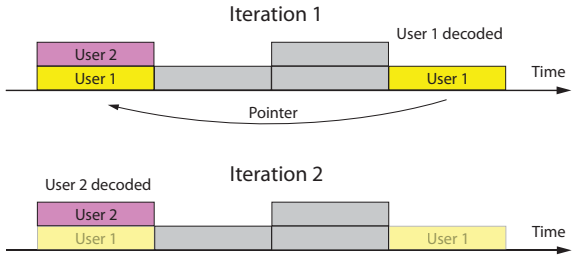

Figure 5.1: Illustration of CRDSA for two users.

The above scheme was generalized in [33] to allow users to have multiple transmissions of one packet. The number of retransmissions is drawn from a probability distribution that is optimized for best supportable load on the shared medium. Finally, [34] employs packet-level coding on graphs as follows: each packet is coded and

transmitted over multiple sub-packets, each sub-packet in one time slot, so that the receiver will be able to correct the "erased" (i.e., undecoded) sub-packets by using doubly-generalized low-density parity-check (DG-LDPC) decoding. The underlying idea of all these schemes is more or less similar: diversity through multiple transmissions per packets, followed by iterative collision resolution through cancelation of the interference contributed by packets that are decoded at previous iterations.

In the following, we examine the effect of MPR on irregular repetition slotted ALOHA (IRSA) introduced in [33].

5.1.1 System Model and Notations

Similarly to [33], m users each with one packet to transmit in a *frame* of n slots want to transmit their packets to a receiver. The transmission time of each packet is equal to the duration of one slot $T_S = T_F/n$, where T_F denotes frame duration. The goal is to maximize the *logical* system load $G = m/n$ through retransmission of packets.

Liva [33] studied the supportable system load of the above system on the collision channel. We extend the analysis and run simulations for the case of MPR channel.

This system can be modeled by a bipartite graph as illustrated in Figure 5.2; one part of the graph consists of m nodes $U = \{u_1, u_2, \cdots, u_m\}$ corresponding to the m users and the other part consists of n nodes $S = \{s_1, s_2, \cdots, s_n\}$ corresponding to the n slots in a frame. A node u_i is connected to s_j if user i transmits a copy of its packet in slot j. This graphical representation reveals the similarity of this collision resolution scheme to decoding low-density parity-check (LDPC) codes on the erasure channel.

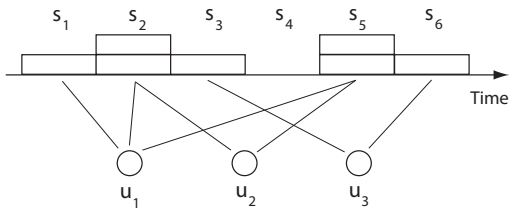

Figure 5.2: Bipartite graph representation of multiple access in a frame with $m = 3$ and $n = 6$. The resulting system load is $G = 3/6$.

In each frame, each user transmits its packet a random number of times r inde-

pendent of other users. r is drawn from a probability distribution that is common among users. We denote the probability of a user transmitting its packet $r \leq n$ times by Λ_r, so we have $\sum_r \Lambda_r = 1$. Once a user draws r from the distribution (at the beginning of the frame), it chooses r slots randomly uniformly to transmit its packet. $\Lambda = (\Lambda_1, \Lambda_2, \cdots)$ is, therefore, the degree distribution of the user nodes U: Λ_r is the probability of any node u_i having r edges.

Obviously, Λ_r is under full control of the system designer, which then determines the degree distribution $\Psi = (\Psi_1, \Psi_2, \cdots)$ of slot nodes. In the asymptotic regime where $m, n \to \infty$ and $G = m/n$ is fixed, Ψ follows a Poisson distribution with mean G times the average degree of a user node $\bar{r} = \sum_r r\Lambda_r$. That is,

$$\Psi_l = e^{-G\bar{r}} \frac{(G\bar{r})^l}{l!} \ . \tag{5.1}$$

Define the "edge-perspective" degree distribution $\lambda = (\lambda_1, \lambda_2, \cdots)$ as the probability distribution of an edge connected to a user node of degree r. We have

$$\lambda_r = \frac{r\Lambda_r}{\sum_k k\Lambda_k} \ . \tag{5.2}$$

Similarly, we have

$$\rho_l = \frac{l\Psi_l}{\sum_k k\Psi_k} \ , \tag{5.3}$$

for the edge-perspective degree distribution corresponding to slot nodes.

Furthermore, the following polynomials defined in [33] are used to facilitate mathematical representations:

$$\Lambda(x) \triangleq \sum_r \Lambda_r x^r \ , \tag{5.4}$$

$$\Psi(x) \triangleq \sum_l \Psi_l x^l \ , \tag{5.5}$$

$$\lambda(x) \triangleq \sum_r \lambda_r x^{r-1} \ , \tag{5.6}$$

$$\rho(x) \triangleq \sum_l \rho_l x^{l-1} \ . \tag{5.7}$$

It follows that $\bar{r} = \sum_r r\Lambda_r = \Lambda'(1)$ and $\Psi'(1) = G\Lambda'(1)$, where $g'(x)$ denotes the derivative of $g(x)$.

In the next section, we introduce and analyze the convergence of an iterative contention resolution scheme.

5.1.2 Convergence Analysis

The scheme proposed in [33] works as follows: The receiver starts by decoding packets that are not collided with any other packets. Once these packets are decoded, the receiver reconstructs and cancels those packets from the received signal of other slots where other replicas of the same packets have collided with other packets.[2] This possibly allows the receiver to resolve collisions between decoded packets and undecoded packets in other slots. Performing multiple iterations of these decoding and cancelation phases may allow the receiver to decode all the packets. Here, we generalize the model to the case where any combination of K or fewer colliding packets may be decoded jointly.

We are interested in analyzing the conditions under which the receiver is able to decode most packets successfully. We extend the analysis in [33] to obtain the probability that a random user is not yet decoded after some i iterations.

Let p_i and q_i denote, respectively, the probability that a randomly selected edge is not yet revealed (i.e., the corresponding user is still undecoded) after the decoding phase and subtracting phase of iteration i. An unrevealed edge connected to a degree-l slot node is revealed with probability $1-p$, after the first phase of an iteration, if the total number of unrevealed edges is less than or equal to K each with probability q. Hence,

$$1-p = \sum_{k=0}^{min(K,l)-1} \binom{l-1}{k} q^k (1-q)^{l-k-1} \ .$$

Averaging over the edge-perspective degree distribution of slot nodes ρ, we obtain

$$p_i = \sum_l \rho_l \left(1 - \sum_{k=0}^{min(K,l)-1} \binom{l-1}{k} q_{i-1}^k (1-q_{i-1})^{l-k-1} \right) \ .$$

It is then easy to show that

$$p_i = 1 - \sum_{k=0}^{K-1} \frac{q_{i-1}^{k-1} \rho^{(k)}(q_{i-1})}{k!} \ , \tag{5.8}$$

where $\rho(\cdot)$ is defined by (5.7), and $\rho^{(n)}(x)$ denotes the nth derivative of $\rho(x)$. It was shown in [33] that

$$\rho(x) = \frac{\Psi'(x)}{\Psi'(1)} = e^{-G\Lambda'(1)(1-x)} \ . \tag{5.9}$$

[2] In order for the receiver to know which packets are transmitted in which slots, each user may use a psuedo-random number generator whose initial seed is known to the receiver as well.

Hence, $\rho^{(n)}(x) = (G\Lambda'(1))^n \rho(x)$. Therefore,

$$\begin{aligned} p_i &= 1 - e^{-q_{i-1}G\Lambda'(1)} \sum_{k=0}^{K-1} \frac{(q_{i-1}G\Lambda'(1))^k}{k!} \\ &= P(K, q_{i-1}G\Lambda'(1)) \ , \end{aligned} \tag{5.10}$$

where $P(\cdot,\cdot)$ denotes the regularized lower incomplete gamma function.

In a second phase, an edge connected to a degree-r user node remains unknown with probability q only if none of the edges connected to that node are revealed with probability p; therefore, $q = p^{r-1}$. It then follows that

$$q_i = \sum_r \lambda_r p_i^{r-1} = \lambda(p_i) \ , \tag{5.11}$$

where $\lambda(\cdot)$ is defined by (5.6). Combining (5.10) and (5.11), we obtain

$$q_i = \lambda\left(P(K, q_{i-1}G\Lambda'(1))\right) \ . \tag{5.12}$$

Since q_i must be decreasing sequence, i.e. $q_{i-1} > q_i$ for all i, we have the following convergence condition:

$$q > \lambda\left(P(K, qG\Lambda'(1))\right) \text{ for } q > 0 \ , \tag{5.13}$$

where the iteration index is omitted for convenience.

For a given Λ, the expression $\lambda(P(K, qG\Lambda'(1)))$ is increasing in G, and therefore the above condition yields a load threshold G^* for which $G < G^*$ ensures convergence. Indeed, G^* is the supremum of network loads under which users' queues remain stable asymptotically as $m, n \to \infty$ (providing that sufficient time is given to the receiver to finish decoding all the users). Figure 5.3 illustrates the evolution of q_i for different values of G. Staircase curves are obtained by simulations with $n = 1000$, while the corresponding smooth curves show $f(q) \triangleq \lambda(P(K, qG\Lambda'(1)))$ that correspond to the boundaries of the staircase curves as $n \to \infty$. It can be observed that although the staircase curves do not perfectly follow what is predicted by $f(q)$, the function well predicts the qualitative behavior of q_i.

Since $f(q)$ is continuous and $f(0) = \lambda_1$, it is immediately concluded from the stability condition (5.13) that Λ_1 must be 0; that is, any distribution Λ that allows a nonzero probability of transmitting only one copy of packets is essentially unstable

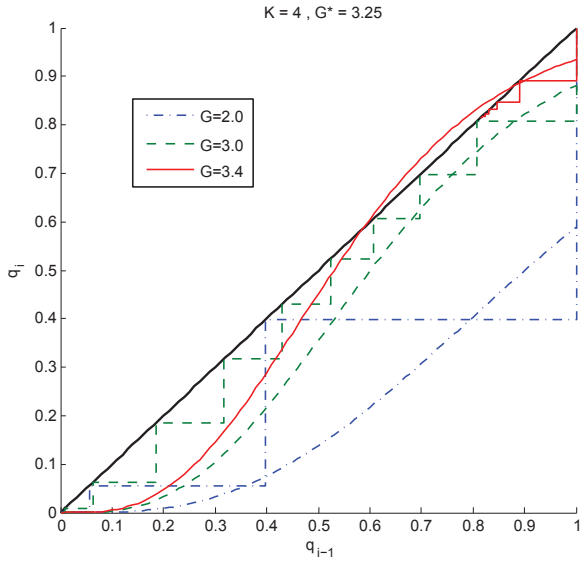

Figure 5.3: Evolution of the ratio of undecoded packets for a system with $K = 4$ and an arbitrary distribution $\Lambda = (0, 0.7, 0.3)$. For this system, $G^* = 3.25$. Therefore, $G = 2, 3$ result in stable systems while $G = 3.4$ is unstable. Staircase curves are obtained by simulations with $n = 1000$.

for any network load $G > 0$.[3] A trivial special case is ALOHA, which is equivalent to $\Lambda_1 = 1$ and $\Lambda_r = 0$ for all $r > 1$.

Numerical calculations suggest that the stability threshold G^* is obtained by

$$\begin{cases} q - f(q) = 0 \\ \frac{\partial}{\partial q}(q - f(q)) = 0 \end{cases} . \tag{5.14}$$

For $K = 1$, this reduces to the closed form [33, Eqn. (7)] since the above equations occur at $q = 0$. For $K > 1$, the above conditions are met at some $q > 0$, which can be obtained by numerical methods.

5.1.3 Numerical Results

In order to study the effect of K, we optimize G^* over $\Lambda = (\Lambda_1, \Lambda_2, \cdots, \Lambda_{r_{\max}})$ for an arbitrary, fixed maximum number of transmissions per packet $r_{\max}$. We assume that the MPR channel requires a K-fold increase in the bandwidth usage, hence the normalized load G^*/K is considered.

[3] All conclusions in this section, unless stated otherwise, are meant to address the asymptotic regime $m, n \to \infty$ and nontrivial cases where $m > K$.

Figure 5.4 shows the maximum normalized system load for different values of $r_{\max}$. It can be seen that fixing the maximum number of retransmissions favors systems with lower K in most cases. Indeed, in the examined scenarios, employing multiuser detection is beneficial only when the maximum number of transmissions is limited to $r_{\max} = 2$. In this particular case, the MPR capability as low as $K = 2, 3$ increases the normalized load by about 70%.

Nevertheless, results of Figure 5.4 hold for the asymptotic setting $n \to \infty$. For smaller values of n, larger K increases the probability of convergence and the number of decoded packets per frame. This is illustrated in Figure 5.5 for $r_{\max} = 8$ and fixed normalized load $G^*/K = 0.7$. Note that this load is within the range of supported loads for all the examined values of K. It is evident that larger K allows more users to be decoded per frame although the asymptotic analysis suggested larger supported normalized load by smaller K. For a frame of $n = 10$ slots, for example, a single-user detector can only decode an average of 60% of the users in a load of $G = 0.7$ although it was obtained by asymptotic analysis that $G^* = 0.94$. This ratio increases to 87% under the same normalized load $G^*/K = 0.7$ for $K = 5$.

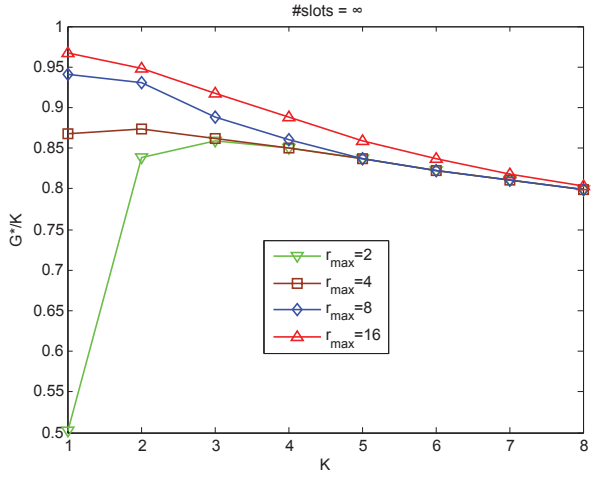

Figure 5.4: Maximum supportable system load for $n \to \infty$.

Next, we study the average number of transmissions per packet, which is a measure of transmission power. In the asymptotic regime, the average number of transmissions per packet is $\bar{r} = \Lambda'(1)$ shown in Figure 5.6. It can be seen that larger K allows smaller number of transmissions. Furthermore, for finite frame size n, it

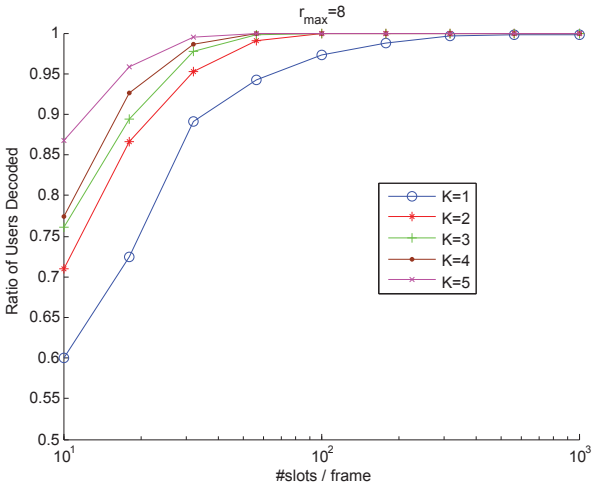

Figure 5.5: Ratio of decoded users per frame for finite n.

was observed that there is a nonzero probability that some packets will fail to be decoded. Suppose unsuccessful users attempt a retransmission in the next frame. This will increase the average number of transmissions per packet to $\Lambda'(1)/(1-\varepsilon)$, where ε denotes the probability of decoding failure per packet. This quantity is plotted in Figure 5.7 for different K. It is observed that for an implementation of 10 slots per frame, the number of transmissions per packet may be halved with K as low as 4.

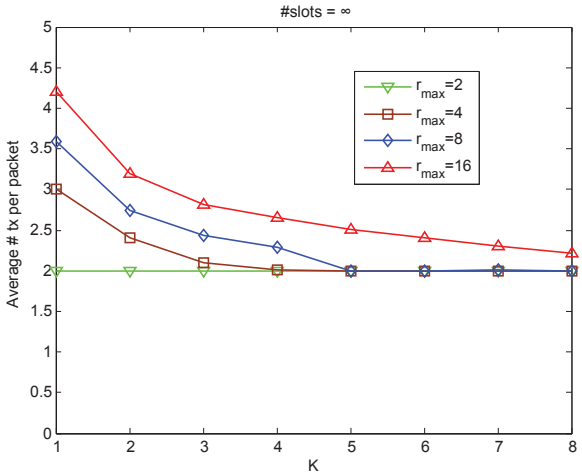

Figure 5.6: Average number of transmissions per packet for $n \to \infty$.

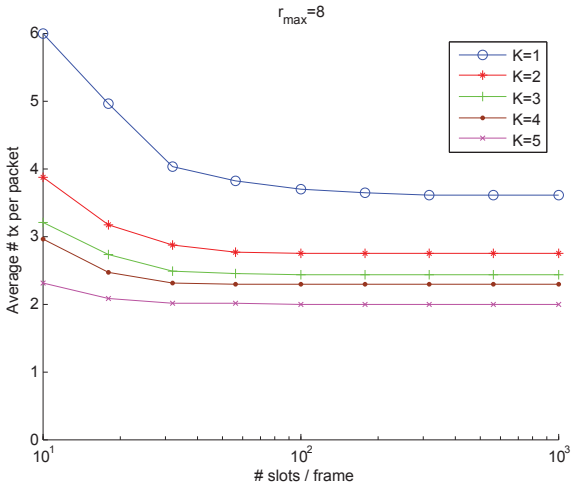

Figure 5.7: Average number of transmissions per packet for finite n.

5.2 Throughput-Delay Tradeoff over Satellite Links

We have, thus far, focused mainly on the throughput performance of random access systems. However, what is also of great importance is the "link delay," which is the delay a packet experiences from the moment it arrives at the MAC layer until it is delivered to the receiver. The link delay generally constitutes the waiting time in the MAC queue, the contention delays of transmissions and retransmissions, and the propagation time.

Similarly to the previous chapters, consider a random access system where each active node transmits a packet with some access probability (AP). Contending devices calculate the AP based on a measure of the current contention on the medium in order to maximize the throughput, minimize the link delay, and/or satisfy other quality-of-service requirements for the traffic. For example, in IEEE 802.11, the AP is the inverse of the "backoff window," which is adjusted according to the number of collisions that a packet has experienced as a measure of contention on the medium.

In most terrestrial systems such as IEEE 802.11 and the random access protocols we have studied thus far, when a node transmits a packet successfully, it receives an ACK packet almost immediately due to negligible propagation delays. The ACK packet informs the node that it can proceed to transmitting other packets in its queue. If the packet is not received, however, the node is notified immediately (due

to lack of an ACK packet) and continues to contend for retransmission of the same packet. In this system, maximizing the throughput is equivalent to minimizing the link delay because when a packet is received, it is removed from the queue immediately.

In satellite systems or underwater acoustic networks, however, nodes have to wait for a non-negligible amount of time to receive the ACK packets. For example, the propagation delay from a ground station to a geosynchronous satellite to another ground station is about 253 milliseconds. The round-trip time is twice this value, more than half a second, which is orders of magnitudes larger than slots of a few tens of microseconds. If nodes do not attempt retransmission of the same packet during this period, throughput-optimal and delay-optimal access probabilities are not necessarily equal; indeed, we will see that, as the ACK delay increases, the delay-optimal AP decreases compared to the throughput optimal AP because at each moment, the delay penalty of a reception failure becomes more and more significant compared to delaying the transmission at the transmitter. Decreasing the AP from the throughput-optimal value, however, results in underutilization of the shared channel. In this section, we study this throughput-delay tradeoff and how significant it becomes in scenarios of relatively large delay.

5.2.1 System Model

Consider a slotted random access system where N active nodes are contending to send packets to a common receiver. The propagation time in each direction is d slots resulting in a round-trip time of $2d$ slots as depicted in Figure 5.8.[4] Nodes may be of any type including mobile devices and very-small-aperture terminals (VSATs). For simplicity of analysis, we assume a symmetric scenario where all active nodes transmit with the same access probability p.

In the following sections, we study the expected system throughput R (packets per slot) and the expected delay per packet D (slots). We do not include the queuing delay in the definition of D, but only calculate the expected waiting time of a packet from the moment it starts contending for transmission until it is successfully received. This is equivalent to MAC queues of maximum length 1, a model

[4]The processing time and guard intervals are neglected in this model.

widely used in the literature that is applicable for remote-terminal applications and machine-to-machine types of traffic [30] where messages are too short (and delay-sensitive in some applications) to be fragmented and queued.

Notations We use the following notations for brevity:

- Since we study the access probability for the two cases of maximizing the expected throughput and minimizing the link delay, we use the superscripts $\cdot^*$ and $\cdot^\dagger$ to distinguish between throughput-optimal and delay-optimal quantities, respectively. For example, the access probability that maximizes the expected throughput is denoted by p^*, and the resulted expected throughput and link delay are denoted by R^* and D^*, respectively.

- To simplify the analysis, we often assume that the node population N is large and, hence, the number of transmissions per slot follows approximately a Poisson distribution. We use $\tilde{\cdot}$ to denote quantities corresponding to this assumption.

- We use the asymptotic order notation $g(x) = \Theta(f(x))$ to state that $g(x)$ (or $g(x)^{-1}$, resp.) grows asymptotically as fast as $f(x)$ (or $f(x)^{-1}$, resp.). The asymptotic orders derived in this section are tight.

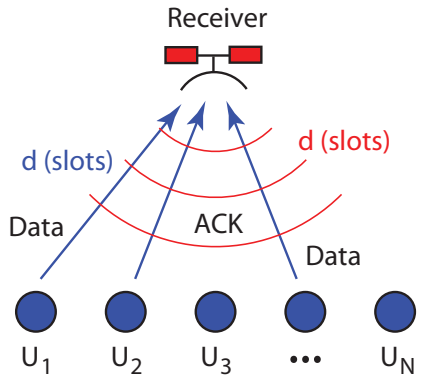

Figure 5.8: The medium access model used in this section.

5.2.2 Collision Channel

First, we study the collision channel. With N nodes contending to access the medium, each with probability p, the expected throughput is obtained by $R = Npq$, where the conditional success probability q is the probability that none of the contending nodes transmits simultaneously when a node transmits. Using Poisson approximation for large N, we have [3]

$$\tilde{R} = \lambda \exp(-\lambda) \quad \text{packets/slot} \ , \tag{5.15}$$

where $\lambda := Np$ is the average load offered to the medium. The expected throughput $\tilde{R}$ is maximized at $\lambda = 1$, i.e. $p^* = 1/N$, and we have $\tilde{R}^* = e^{-1}$ packets per slot [3].

Although the probability distribution of the packet delay takes a tedious form, the *expected* packet delay can be written in a short closed form. Recall that each packet is transmitted and retransmitted upon collision until received successfully. Since each node (re)transmits with probability p in each slot, there is a backoff period before each (re)transmission. Therefore, the expected duration of each backoff plus (re)transmission attempt is geometrically distributed with mean $1/p$. Then, it takes the signal d slots to travel through the link. If there is no collision, the packet will be received successfully and the transmitter will be notified d slots later; otherwise, the node attempts a retransmission of that packet after another backoff period. Note that the node does not need to remain idle during the $2d$ slots of round-trip time and is assumed to be contending for transmitting other packets in its queue. Since the success probability of each (re)transmission is q, the expected delay until the node receives ACK is $(1/p + 2d)/q$. Since the packet is received d slots before the ACK is received, the expected delay per packet follows

$$D = \frac{1/p + 2d}{q} - d \simeq \frac{1/p + 2d}{\exp(-Np)} - d = \tilde{D} \quad \text{slots} \ . \tag{5.16}$$

Figure 5.9 illustrates the timeline of transmitting a packet with one retransmission attempt.

To obtain the delay-optimal transmission probability $p^\dagger$, one should solve

$$\frac{\partial \tilde{D}}{\partial p} = 0 \ ,$$

which gives

$$2Ndp^{\dagger 2} + Np^\dagger - 1 = 0 \ . \tag{5.17}$$

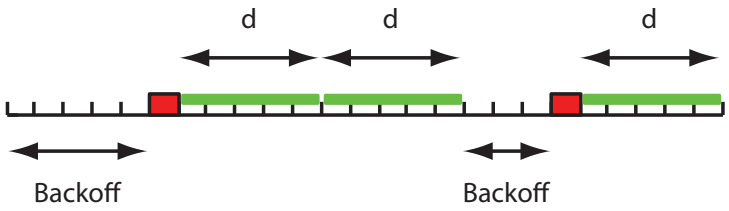

Figure 5.9: Timeline of transmitting a packet with one retransmission attempt.

If $d = 0$, we obtain $p^\dagger = 1/N = p^*$, which confirms that throughput-optimal and delay-optimal transmission probabilities are equal when ACK delay is zero. If $d > 0$, we obtain

$$p^\dagger = \frac{2}{\sqrt{N}\left(\sqrt{8d+N}+\sqrt{N}\right)} \ . \tag{5.18}$$

If $d \ll N$, we have $p^\dagger \simeq 1/N$. If $d \gg N$, we obtain

$$p^\dagger \simeq \frac{1}{\sqrt{2Nd}} \quad , \quad \lambda^\dagger = Np^\dagger \simeq \sqrt{\frac{N}{2d}} \ .$$

If $d \gg N$, the throughput loss ratio due to optimizing packet delay is

$$\frac{\tilde{R}^\dagger}{\tilde{R}^*} \simeq \sqrt{\frac{N}{2d}} \exp\left(1-\sqrt{\frac{N}{2d}}\right) \to \sqrt{\frac{N}{2d}} \ . \tag{5.19}$$

On the other hand, if throughput is optimized, packet delay will increase by the following ratio:

$$\frac{\tilde{D}^*}{\tilde{D}^\dagger} \simeq \frac{(N+2d)e-d}{(\sqrt{2Nd}+2d)\exp\left(\sqrt{N/(2d)}\right)-d} \to 2e-1 \ . \tag{5.20}$$

That is, the expected delay per packet may increase more than four folds in some scenarios. Similar closed-form equations can be derived also when d/N is finite and nonzero.

Figure 5.10 shows R and D versus $\lambda = Np$ for $N = 10$ and $d = 1000$ slots. It can be seen that the delay-optimal load is 0.072 of the throughput-optimal load, which is closely approximated by $\sqrt{N/(2d)} = 0.071$. For these parameter values, we obtain

$$\frac{R^\dagger}{R^*} = 0.17 \quad , \quad \frac{D^*}{D^\dagger} = 3.27 \ .$$

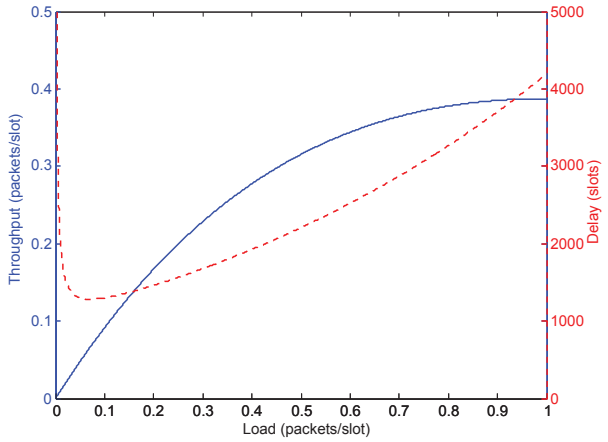

Figure 5.10: R (solid curve) and D (dashed curve) versus $\lambda = Np$ for $N = 10$ and $d = 1000$ slots.

5.2.3 Multipacket Reception Channel

For the MPR channel, we can obtain the success probability by

$$q \simeq \sum_{k=0}^{K-1} e^{-Np} \frac{(Np)^k}{k!} = Q(K, Np) \ , \tag{5.21}$$

for large N, where $Q(s,t) := \Gamma(s,t)/\Gamma(s)$ is the regularized upper incomplete gamma function. Therefore,

$$D = \frac{1/p + 2d}{q} - d \simeq \frac{1/p + 2d}{Q(K, Np)} - d = \tilde{D} \quad \text{slots} \ . \tag{5.22}$$

Let us first consider the special case of $K = 2$. We have

$$q \simeq \exp(-Np)(1 + Np) = Q(2, Np) \ ,$$

and, hence,

$$\tilde{D} = \frac{1/p + 2d}{\exp(-Np)(1 + Np)} - d \quad \text{slots} \ . \tag{5.23}$$

If $d = 0$, we obtain

$$\frac{\partial \tilde{D}}{\partial p} = 0 \quad \Rightarrow \quad p^\dagger = \frac{\sqrt{5} + 1}{2N} \ , \tag{5.24}$$

which, as expected, coincides with the throughput-optimal access probability p^* obtained by maximizing

$$\tilde{R} = Np \exp(-Np)(1 + Np) \ .$$

If $d \neq 0$, however, it follows that

$$p^\dagger = \frac{\sqrt[3]{A(N,d)}}{6dN^2} + \frac{6dN^3 + N^4}{6dN^2\sqrt[3]{A(N,d)}} - \frac{1}{6d} \ , \tag{5.25}$$

where

$$\begin{aligned} A(N,d) \ &\triangleq \ 54d^2N^4 - 9dN^5 - N^6 \\ &\quad +3\sqrt{3}\sqrt{108d^4N^8 - 44d^3N^9 - 5d^2N^{10}} \ . \end{aligned}$$

We are interested in the large-delay case $d \gg N$, where we have $A(N,d) = \Theta(d^2N^4)$ and, hence,

$$p^\dagger = \Theta\left(\frac{1}{\sqrt[3]{dN^2}}\right) \ .$$

Comparing the above with the results of the previous section, we can see that we have obtained a scaling improvement of the access probability, i.e. $\Theta(1/\sqrt[3]{dN^2})$ compared to $\Theta(1/\sqrt{dN})$ in (5.18), which is closer to the zero-delay scaling $\Theta(1/N)$. We expect to observe similar scaling improvements with $K > 2$. This is stated by the following proposition.

Proposition 2. *In an MPR system of capability K, the delay-optimal access probability scales as*

$$p^\dagger = \Theta\left(\frac{1}{\sqrt[K+1]{dN^K}}\right) \ . \tag{5.26}$$

Proof. The average delay per packet is given by (5.22). We have

$$\frac{\partial \tilde{D}}{\partial p} = \frac{\exp(-Np)(2dp+1)(Np)^K - \Gamma(K,Np)}{p^2\Gamma(K,Np)Q(K,Np)} \ .$$

Setting the above derivative to zero, we obtain

$$\exp(-Np)(2dp+1)(Np)^K = \Gamma(K,Np) \ ,$$

or, equivalently,

$$(2dp+1)(Np)^K = (K-1)!\sum_{k=0}^{K-1}\frac{(Np)^k}{k!} \ .$$

Note in the above that, since $d \gg N$, we must have $Np \ll 1$ in order for the left-hand side to remain bounded for bounded N. Therefore, for any K finite, the only significant term on the right-hand side is the constant term $(K-1)!$. That is, for constant K,

$$\Theta\left(dN^Kp^{K+1}\right) = \Theta\left(1\right) \ ,$$

which proves the proposition. □

It can be immediately concluded that the scaling of the delay-optimal access probability is closer to the throughput-optimal scaling $\Theta(1/N)$ for larger K. As a result, we have

$$\frac{\tilde{R}^\dagger}{\tilde{R}^*} = \Theta\left(\frac{p^\dagger}{p^*}\right) = \Theta\left(\sqrt[K+1]{\frac{N}{d}}\right) \ , \tag{5.27}$$

which means that, in the asymptotic sense, the throughput loss due to delay optimization is mitigated more significantly on a "stronger" MPR channel. This advantage is obtained at the cost of receiver complexity and channel bandwidth in order to implement the multiuser detector.

5.2.4 Repetition Random Access

We formerly assumed that a node does not attempt a retransmission during the round-trip time of $2d$ slots, which turns out to be the major source of long delays when $d \gg N$. Here, we study the throughput-delay tradeoff with repetition random access.

Let us consider the collision channel for simplicity. Suppose a node transmits a packet r times after each backoff period. The r slots should be random and nonconsecutive, obviously, to avoid repeated collisions with other nodes transmitting r replicas of their own packets. The required period for transmitting the r replicas, however, is assumed to be small compared to $1/p$ and d. This is a reasonable assumption because the required number of slots for some "sufficient" randomization of the slot selection depends on the number of contending nodes N, which is assumed to be much smaller than d and the associated optimal backoff period $1/p$. Therefore, we may simply assume that the equations of the previous section still hold, except that the new success probability that follows is

$$q = 1 - (1 - \exp(-Nrp))^r \ . \tag{5.28}$$

We have assumed, in the above equation, that the slot selection for transmitting r replicas is sufficiently randomized so that each of the r replicas may be received successfully independently of the others. Furthermore, the intensity of the offered traffic is assumed to be increased r times on average. Therefore, the expected packet delay follows

$$\tilde{D} \simeq \frac{1/p + 2d}{1 - (1 - \exp(-Nrp))^r} - d \quad \text{slots} \ . \tag{5.29}$$

Let us examine delay-optimal access under the new model. For $r = 2$, by solving

$$\frac{\partial}{\partial p}\left(\frac{1/p + 2d}{1-(1-\exp(-2Np))^2} - d\right) = 0 \ ,$$

we obtain

$$Np^\dagger\left(dp^\dagger + \frac{1}{2}\right) = \frac{1 - e^{-2Np^\dagger}/8}{1 - e^{-2Np^\dagger}} \ . \tag{5.30}$$

The left-hand side is $\Theta(dNp^2)$ while the right-hand side is $\Theta(1/(Np))$ given that $Np \ll 1$. Therefore, we have

$$p^\dagger = \Theta\left(\frac{1}{\sqrt[3]{dN^2}}\right) \ . \tag{5.31}$$

The following proposition generalizes the above result.

Proposition 3. *In a repetition random access system of r repetitions, the delay-optimal access probability scales as*

$$p^\dagger = \Theta\left(\frac{1}{\sqrt[r+1]{dN^r}}\right) \ . \tag{5.32}$$

Proof. When the expected packet delay is given by (5.29), by solving $\partial\tilde{D}/\partial p = 0$ we obtain

$$Nr^2p(2dp+1)e^{-Nrp}(1-e^{-Nrp})^{r-1} + (1-e^{-Nrp})^r = 1 \ .$$

Therefore, assuming $Np \ll 1$, we have

$$\Theta\left((Np)^r dp\right) + \Theta\left((Np)^r\right) = \Theta\left(1\right) \ ,$$

for constant r. Since we assume $d \gg N$, if $\Theta(dp) \leq \Theta(1)$, then $p = \Theta(1/N)$ contradicting $Np \ll 1$. Hence, we must have $\Theta(dp) > \Theta(1)$, which gives

$$\Theta\left(dN^r p^{r+1}\right) = \Theta\left(1\right) \ .$$

□

Comparing (5.32) to (5.26), we can conclude that repetition random access results in the same asymptotic scaling as using the MPR channel. Similarly,

$$\frac{\tilde{R}^\dagger}{\tilde{R}^*} = \Theta\left(\sqrt[r+1]{\frac{N}{d}}\right) \ . \tag{5.33}$$

This scaling improvement is obtained at the cost of higher energy spent per (re)transmission attempt. Although this higher energy increases the success probability that reduces the number of attempts, the average energy spent per packet may possibly be larger for larger r. Let E denote the transmission energy in one slot. The average energy per packet is E/q when only one replica is transmitted at each attempt. When transmitting r replicas, this quantity changes to

$$\frac{rE}{1-(1-\exp(-Nrp))^r} = \frac{rE}{1-(Nrp)^r+\Theta\left((Nrp)^{r+1}\right)} \ ,$$

which is on the order of r times

$$\frac{E}{1-\exp(-Np)} = \frac{E}{1-Nrp+\Theta\left((Nrp)^2\right)} \ ,$$

when $Np \ll 1$. In summary, it can be concluded that, since the success probability q is generally not small when $Np \ll 1$, transmitting r replicas of a packet at each attempt results in an r-fold increase in the transmission power as the cost of improving the scaling of $p^\dagger$.

Repetition random access on MPR channel The scaling results of Propositions 2 and 3 can be shown to be special cases of the more general form

$$p^\dagger = \Theta\left(\frac{1}{\sqrt[rK+1]{dN^{rK}}}\right) \ , \tag{5.34}$$

for r repetitions per attempt on a channel of MPR capability K, where r and K are finite constants.

5.2.5 Results

Unless examined numerically, the scaling results obtained so far do not show the effectiveness of the MPR channel in smoothing the throughput-delay compromise in practical scenarios. First, we examine Proposition 2. Figure 5.11 shows p^* (dashed curves) and $p^\dagger$ (solid curves) for $N = 10$ and different values of K when $d = 1, \cdots, 10^5$ slots. It can be seen that, for large values of the link delay, the $p^\dagger$ curves follow the dotted lines that show the scalings suggested by Proposition 2.

Figures 5.12 and 5.13 show the corresponding throughput loss $R^\dagger/R^*$ and delay increase $D^*/D^\dagger$, respectively. Values closer to 1 are favorable, which are clearly shown to be achieved by larger values of K.

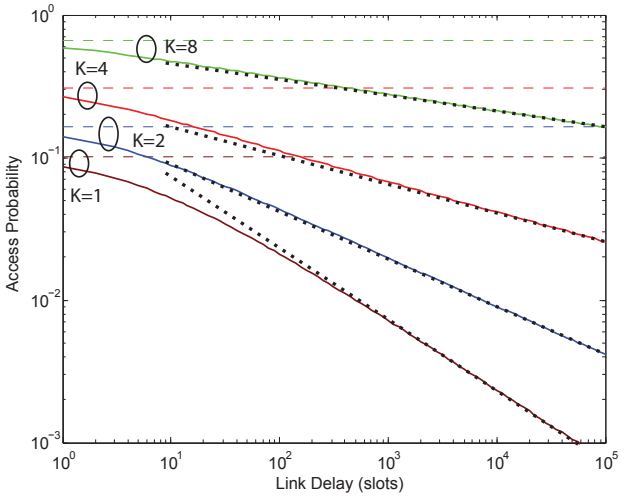

Figure 5.11: Optimal access probabilities for $N = 10$ and different values of K when $d = 1, \cdots, 10^5$ slots.

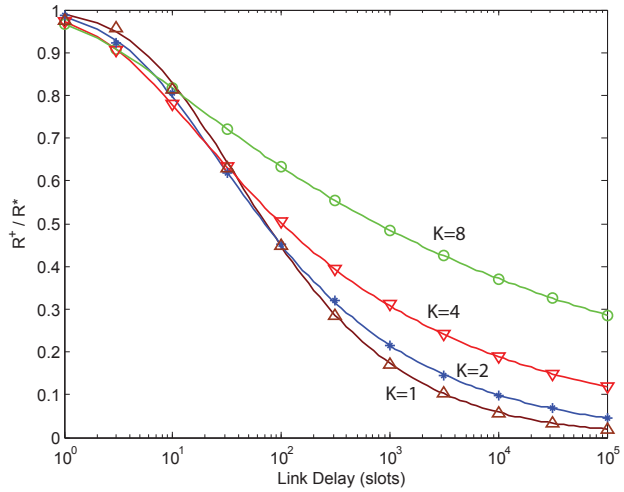

Figure 5.12: $R^\dagger/R^*$ for $N = 10$ and different values of K. Solid curves are plotted by numerical computations while the markers show simulation results.

Next, we examine Proposition 3. Figure 5.14 shows p^* (dashed curves) and $p^\dagger$ (solid curves) for $N = 10$ and different values of r when $d = 10^2, \cdots, 10^5$ slots. It can be seen that, for large values of the link delay, the $p^\dagger$ curves follow the dotted lines that show the scalings suggested by Proposition 3.

Figures 5.15 and 5.16 show the corresponding throughput loss $R^\dagger/R^*$ and delay increase $D^*/D^\dagger$, respectively. Again, values closer to 1 are favorable, which are shown to be achieved by larger values of r.

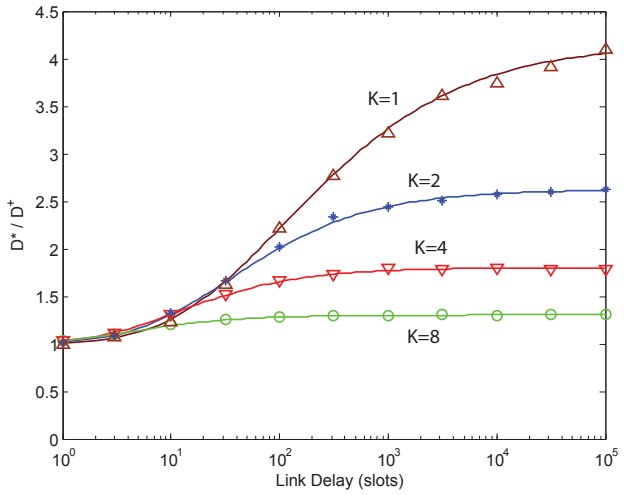

Figure 5.13: $D^*/D^\dagger$ for $N = 10$ and different values of K. Solid curves are plotted by numerical computations while the markers show simulation results.

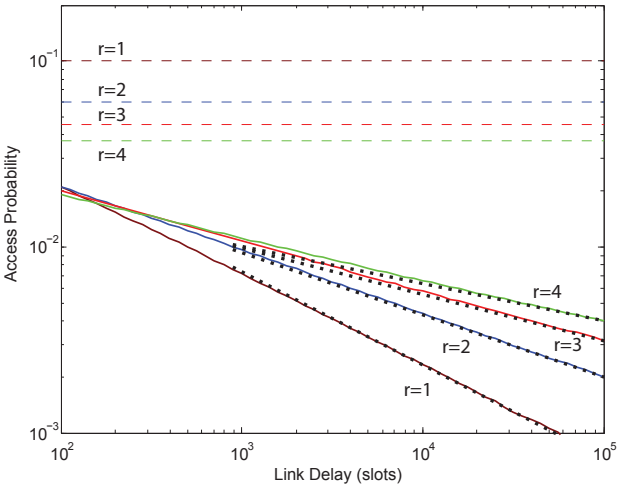

Figure 5.14: Optimal access probabilities for $N = 10$ and different values of r when $d = 10^2, \cdots, 10^5$ slots.

5.3 Summary

In this chapter, we generalized the asymptotic analysis of irregular repetition slotted ALOHA [33] to the case of MPR channel, by which we could compute optimal probability distribution of the number of transmissions per packet for $K > 1$. Results of simulations with finite population suggested that MPR allows larger loads on the system with practically low number of slots per frame, hence reducing total

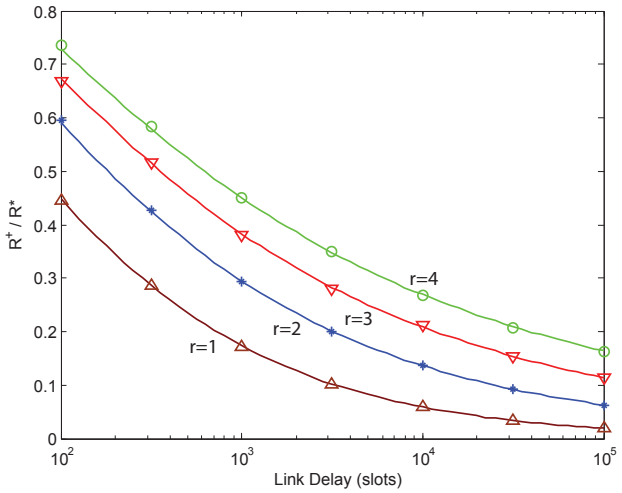

Figure 5.15: $R^{\dagger}/R^{*}$ for $N = 10$ and different values of r. Solid curves are plotted by numerical computations while the markers show simulation results.

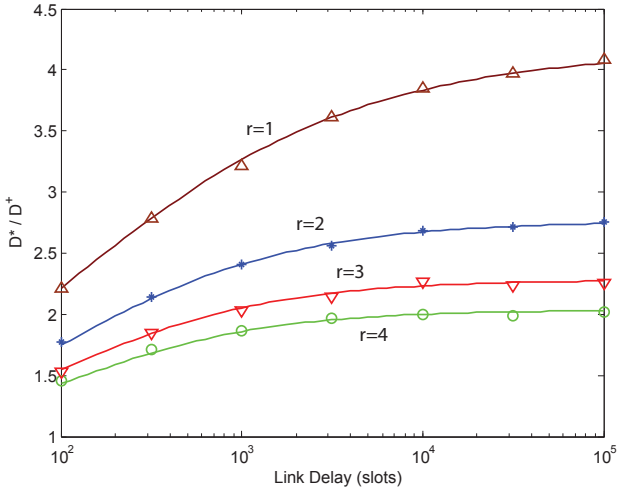

Figure 5.16: $D^{*}/D^{\dagger}$ for $N = 10$ and different values of r. Solid curves are plotted by numerical computations while the markers show simulation results.

transmission delay and power consumption. For example, for frames of 10 slots, it was shown that an average of 3 transmissions per packet is required for successful decoding when $K = 4$, while this average doubles on the collision channel. The ratio of the decoded users is also larger on the MPR channel, as a 45% increase in the ratio of decoded users can be obtained by using a multiuser detector of $K = 5$.

Next, throughput-delay tradeoff of random access over delayed links was analyzed and scaling laws were derived for the cases of collision channel and the MPR

channel as well as repetition random access. It was shown that MPR receivers improve the multiple access performance in the sense that the compromise between throughput and delay is smoothed when the receiver has joint-detection capabilities. In particular, we proved that if the receiver is capable of receiving K concurrent transmissions, the throughput-optimal and delay-optimal access probabilities scale as $\Theta(1/N)$ and $\Theta((dN^K)^{\frac{-1}{K+1}})$, respectively, which shows that the two scalings are closer for larger values of K. Similarly, with repetition random access where r replicas of a packet are transmitted, the delay-optimal access probability scales as $\Theta((dN^r)^{\frac{-1}{r+1}})$. These improvements are obtained at the cost of larger bandwidth, more complex signal processing and/or higher power consumption. This work motivates the use of multiuser detection schemes and/or repetition schemes in large-delay links such as user-to-hub random access in satellite communication systems.

Chapter 6

Conclusions and Future Work

Random access on the MPR channel was studied in this thesis. It was shown that random access, as a simple medium-access strategy, is potentially more efficient in terms of system throughput and average packet delay. MAC schemes were proposed for different scenarios and their performance was analyzed. The proposed schemes were shown to be effective in achieving the potential advantages of the MPR channel in the MAC layer. The contributions of this thesis are summarized in the following.

6.1 Summary of Contributions

6.1.1 Aloha Random Access on MPR Channel

We first studied the throughput performance of Aloha random access with Poisson arrivals. The main question to answer was how effectively multiuser detection improves the throughput of Aloha. A closed-form equation was derived that conveniently compares the throughput of Aloha as a function of the MPR capability K. The family of curves introduced through the variable parameter K highlights the throughput advantage of the MPR channel compared to the conventional collision channel and, in particular, shows that Aloha is asymptotically optimal as K goes to infinity providing that traffic control mechanisms are devised to maintain the average arrival below a threshold. The peak of the Aloha curve corresponding to a specific value of K provides a ballpark for the throughput that efficient random access schemes should target. This maximum was targeted in the next chapters and was used as the baseline for throughput comparisons. Bounds on the throughput performance were introduced and a closed-form approximation for the upper bound

corresponding to genie-aided random access with Poisson arrivals was proposed. Next, a method of computing the optimal access probability was presented for the cases where only partial information about the system state was available. This method was then employed for a scenario of Markovian incoming traffic.

6.1.2 MPR Random Access with Markovian Traffic

Throughput performance of MPR random access with Markovian traffic was examined. It was shown that, in a system of nodes offering Markovian traffic, burstiness of the packetized data traffic improves the predictability of the system state and increases the achievable throughput when only limited feedback is available. A Markov decision model was described for the system and the throughput-optimal access strategy was derived. Lower and upper bounds on the throughput were introduced that correspond, respectively, to no information and perfect information about the system state. It was observed that the throughput with fixed traffic load may vary between the introduced bounds as burstiness varies, starting from the lower bound for zero burstiness ($\gamma = 0$) and approaching the upper bound with extreme burstiness ($\gamma \to 1$). This may mean a 25% improvement of the MAC efficiency, depending on K and the number of nodes, by controlling the traffic burstiness. For a large system size, however, the gap between the two bounds vanishes and the feedback becomes unnecessary as long as the average traffic parameters are known. Finally, a closed-form equation for the average delay of the system was derived as a function of the throughput.

6.1.3 MPR Random Access with Adaptive Filtering

A MAC protocol using an extended Kalman filter was designed and shown to achieve the maximum of Aloha without explicit knowledge of traffic parameters. This was accomplished through utilization of the limited feedback acquired through ACK packets from the receiver. Average packet delay was shown to be reasonable as long as the offered packet rate maintains a safety margin to the maximum supportable packet rate. Particularly, delay figures remained below 50 slots when the normalized offered packet rate was about 10% below the maximum normalized supportable load.

Two unideal cases were then examined: i) in the case that there is a nonzero

probability that ACK packets are missed by nodes, it was shown that the dominant effect in decreasing the throughput is the extra replicas that nodes transmit while the proposed EKF-based state tracking system works properly; *ii*) when ACK packets are received with a delay, for example in the case of satellite systems, larger values of K result in more robustness against ACK delays, which motivates the use of random access systems and multiuser detectors for sporadic traffic over satellite links.

6.1.4 Repetition Random Access on MPR Channel

Irregular repetition slotted ALOHA [33] was generalized to the cases where the receiver was capable of decoding multiple colliding packets simultaneously. The asymptotic analysis of infinite user population (or equivalently, infinite number of slots per processed frame) was provided for this general case. Results of simulations with finite population suggested that multiuser detection allows larger loads on the system with practically low number of slots per frame, hence reducing total transmission delay and power consumption. It was also shown that the ratio of the decoded users was larger when multiuser detection was employed.

6.1.5 Throughput-Delay Tradeoff over Delayed Links

Throughput-delay tradeoff of random access over delayed links was analyzed and scaling laws were derived for the cases of collision channel and the MPR channel as well as repetition random access. It was shown that MPR receivers improve the multiple access performance in the sense that the compromise between throughput and delay is smoothed when the receiver has joint-detection capabilities. The improvements are obtained at the cost of larger bandwidth, more complex signal processing and/or higher power consumption. This work motivates the use of multiuser detection schemes and/or repetition schemes in large-delay links such as user-to-hub random access in satellite communication systems.

6.2 Future Work

There are several interesting problems related to this thesis. First of all, the majority of the body of research on multiuser detection has focused on coherent detection where channel states are known to the receiver. This assumption poses a challenge

in random access systems where we desire to require no coordination among users. As a result, pilot signals transmitted with packets may be distorted by multiuser interference. This subject is currently being studied for satellite uplink, e.g., see [59]. Furthermore, discrepancies between our MPR channel model and real-world wireless channels may result in effects not fully addressed in our study thus far. For example, asymmetry in channel conditions among nodes can lead to estimation errors and fairness issues where additional mechanisms will be required in order for the system to satisfy quality-of-service requirements of the users. We will study such effects in our future work.

We defined the rate/burstiness description of an on/off source, in Chapter 3, and related the achievable system throughput to burstiness of the packet generating sources. What that suggests, in essence, is that a larger entropy at the source introduces inherent limitations on the amount of throughput in a multiple-access system. This subject is related to Gallager's concept of "protocol information" [60], a type of information that is related to uncertainties about the packetized traffic and uses a portion of the shared capacity, hence limiting the capacity for useful "data information." An interesting question is whether there are analytical bounds on the multiple-access capacity as a function of the entropy of packet generating sources.

An interesting question on repetition random access regards scenarios where users are not cooperative in selecting the number of replicas they transmit per packet per frame. Selfish users define a game-theoretic framework for repetition random access where each user attempts to maximize the chance of having its packet decoded successfully by transmitting a larger number of replicas in one frame. What may happen as a result is overloading the channel that can result in a large number of decoding failures and, hence, low system throughput. The equilibrium point of a system of selfish users may be different from the optimum operating point obtained in Chapter 5 for cooperative users. It is an interesting problem to investigate how the equilibrium point of repetition random access depends on the cost of decoding failure as this is known about conventional random access [61].

Finally, we studied throughput-delay tradeoff in delayed-ACK random access systems. What has not been considered in that study is the queuing delay and how it affects the throughput-delay tradeoff. Also, it is useful to study how different

types of delay-sensitive and delay-insensitive traffic may coexist and share resources while they both benefit and meet the required quality-of-service demands.

Bibliography

[1] N. Abramson, "The aloha system–another alternative for computer communications," in *Proc. Fall Joint Comput. Conf.*, Houston, TX, Nov. 1970, pp. 281–285.

[2] F. A. Tobagi, "Analysis of a two-hop centralized packet radio network–part i: Slotted aloha," *IEEE Trans. Communications*, vol. COM-28, no. 2, pp. 196–207, Feb. 1980.

[3] N. Abramson, "The throughput of packet broadcasting channels," *IEEE Trans. Communications*, vol. COM-25, no. 1, pp. 117–128, Jan. 1977.

[4] Y.-C. Jenq, "On the stability of slotted aloha systems," *IEEE Trans. Communications*, vol. COM-28, no. 11, pp. 1936–1939, Nov. 1980.

[5] S. S. Lam and L. Kleinrock, "Packet switching in a multiaccess broadcast channel: Dynamic control procedures," *IEEE Trans. Communications*, vol. COM-23, no. 9, pp. 891–904, Sep. 1975.

[6] L. Kleinrock and S. S. Lam, "Packet switching in a multiaccess broadcast channel: Performance evaluation," *IEEE Trans. Communications*, vol. COM-23, no. 4, pp. 410–423, Apr. 1975.

[7] L. Kleinrock and F. A. Tobagi, "Packet switching in radio channels: Part i–carrier sense multiple-access modes and their throughput-delay characteristics," *IEEE Trans. Communications*, vol. COM-23, no. 12, pp. 1400–1416, Dec. 1975.

[8] Z. J. Haas and J. Deng, "On optimizing the backoff interval for random access schemes," *IEEE Trans. Communications*, vol. 51, no. 12, pp. 2081–2090, Dec. 2003.

[9] C. Ware, J. Judge, J. Chicharo, and E. Dutkiewicz, "Unfairness and capture behaviour in 802.11 adhoc networks," in *Proc. IEEE Intl. Conf. Comm. (ICC)*, New Orleans, LA, Jun. 2000, pp. 159–163.

[10] T. D. Lagkas, D. G. Stratogiannis, and P. Chatzimisios, "Modeling and performance analysis of an alternative to ieee 802.11e hybrid control function," *Telecom. Systems Jnl.*, Jun. 2011. [Online]. Available: http://dx.doi.org/10.1007/s11235-011-9477-5

[11] M. L. Molle and K. J. Christensen, "The effects of controlling capture on multimedia traffic for shared ethernet systems," *Telecom. Systems Jnl.*, vol. 9, no. 3-4, pp. 287–314, 1998.

[12] C. Schlegel and A. Grant, *Coordinated Multiuser Communications*. Dordrecht, Netherlands: Springer, 2006.

[13] S. Nagaraj, D. Truhachev, and C. Schlegel, "Analysis of a random channel access scheme with multi-packet reception," in *Proc. IEEE Globecom'08*, New Orleans, LA, Nov./Dec. 2008, pp. 1–5.

[14] S. Ghez, S. Verdú, and S. C. Schwartz, "Stability properties of slotted aloha with multipacket reception capability," *IEEE Trans. Automatic Control*, vol. 33, no. 7, pp. 640–649, Jul. 1988.

[15] ——, "Optimal decentralized control in the random access multipacket channel," *IEEE Trans. Automatic Control*, vol. 34, no. 11, pp. 1153–1163, Nov. 1989.

[16] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted aloha with multipacket reception," *IEEE Trans. Info. Theory*, vol. 51, no. 7, pp. 2636–2656, Jul. 2005.

[17] S. Verdú, *Multiuser Detection*, 1st ed. Cambridge, UK: Cambridge University Press, 1998.

[18] P. Mathys, "A class of codes for a t active users out of n multiple-access communication system," *IEEE Trans. Info. Theory*, vol. 36, no. 6, pp. 1206–1219, Nov. 1990.

[19] X. Wang and J. K. Tugnait, "A bit-map-assisted dynamic queue protocol for multiaccess wireless networks with multiple packet reception," *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2068–2081, Aug. 2003.

[20] Q. Zhao and L. Tong, "A multiqueue service room mac protocol for wireless networks with multipacket reception," *IEEE Trans. Networking*, vol. 11, no. 1, pp. 125–137, Feb. 2003.

[21] ——, "A dynamic queue protocol for multiaccess wireless networks with multipacket reception," *IEEE Trans. Wireless Comm.*, vol. 3, no. 6, pp. 2221–2231, Nov. 2004.

[22] W.-F. Yang, J.-Y. Wu, L.-C. Wang, and T.-S. Lee, "A multigroup priority queueing mac protocol for wireless networks with multipacket reception," *EURASIP Journal on Wireless Comm. and Net.*, vol. 2008, no. 5, pp. 1–12, Jan. 2008.

[23] J. Luo and A. Ephremides, "On the throughput, capacity, and stability regions of random multiple access," *IEEE Trans. Info. Theory*, vol. 52, no. 6, pp. 2593–2607, Jun. 2006.

[24] G. D. Celik, G. Zussman, W. F. Khan, and E. Modiano, "Mac for networks with multipacket reception capability and spatially distributed nodes," in *Proc. IEEE Infocom'08*, Phoenix, AZ, Apr. 2008, pp. 2110–2118.

[25] M. Coupechoux, T. Lestable, C. Bonnet, and V. Kumar, "Throughput of the multi-hop slotted aloha with multi-packet reception," in *Wireless On-Demand Network Systems*, ser. Lecture Notes in Computer Science. Berlin/Heidelberg, Germany: Springer, 2003, vol. 2928, pp. 239–243.

[26] J.-B. Seo and V. C. M. Leung, "Analysis of an exponential backoff algorithm for multipacket reception slotted aloha systems," in *Proc. Int'l Conf. Communications*, Cape Town, South Africa, May 2010, pp. 855–862.

[27] ——, "Design and analysis of cross-layer contention resolution algorithms for multi-packet reception slotted aloha systems," *IEEE Trans. Wireless Comm.*, vol. 10, no. 3, pp. 825–833, Mar. 2011.

[28] Y. H. Bae, B. D. Choi, and A. S. Alfa, "Achieving maximum throughput in random access protocols with multipacket reception," *IEEE Trans. Mobile Computing*, vol. 13, no. 3, pp. 497–511, Mar. 2014.

[29] D. S. Chan, T. Berger, and L. Tong, "Carrier sense multiple access communications on multipacket reception channels: Theory and applications to ieee 802.11 wireless networks," *IEEE Trans. Communications*, vol. 61, no. 1, pp. 266–278, Jan. 2013.

[30] F. Monsees, C. Bockelmann, D. Wubben, and A. Dekorsy, "Sparsity aware multiuser detection for machine to machine communication," in *Proc. Globecom Workshops*, Anaheim, CA, Dec. 2012, pp. 1706–1711.

[31] G. Bianchi, "Performance analysis of the ieee 802.11 distributed coordination function," *IEEE Jnl. Selected Areas in Comm.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[32] E. Casini, R. D. Gaudenzi, and O. del Rio Herrero, "Contention resolution diversity slotted aloha (crdsa): An enhanced random access scheme for satellite access packet networks," *IEEE Trans. Wireless Comm.*, vol. 6, no. 4, pp. 1408–1419, Apr. 2007.

[33] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted aloha," *IEEE Trans. Communications*, vol. 59, no. 2, pp. 477–487, Feb. 2011.

[34] M. Chiani, G. Liva, and E. Paolini, "The marriage between random access and codes on graphs: Coded slotted aloha," in *Proc. 7th Int'l Symp. Turbo Codes and Iterative Information Processing*, Gothenburg, Sweden, Aug. 2012, pp. 1–6.

[35] M. Ghanbarinejad, C. Schlegel, and P. Gburzynski, "Adaptive probabilistic medium access in mpr-capable ad-hoc wireless networks," in *Proc. IEEE Globecom'09*, Honolulu, HI, Nov./Dec. 2009, pp. 1–5.

[36] M. Ghanbarinejad and C. Schlegel, "Controlled random access with multipacket reception and traffic uncertainty," in *Proc. IEEE Globecom'10 Workshops*, Miami, FL, Dec. 2010, pp. 1238–1242.

[37] ——, "Analysis of controlled probabilistic access with multipacket reception," in *Proc. IEEE Globecom'11*, Houston, TX, Dec. 2011, pp. 1–5.

[38] H. Inaltekin, M. Chiang, H. V. Poor, and S. B. Wicker, "Selfish random access over wireless channels with multipacket reception," *IEEE JSAC*, vol. 30, no. 1, pp. 138–152, Jan. 2012.

[39] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Info. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[40] C. Schlegel and D. Truhachev, "Generalized modulation and iterative demodulation," in *Proc. Int'l Zurich Seminar on Comm. (IZS'08)*, Zurich, Switzerland, Mar. 2008, pp. 76–79.

[41] C. Schlegel, M. Burnashev, and D. Truhachev, "Generalized superposition modulation and iterative demodulation: A capacity investigation," *Jnl. Electrical and Computer Engineering*, vol. 2010, pp. 1–9, Jan. 2010.

[42] F. P. Fontan, M. Vazquez-Castro, C. E. Cabado, J. P. Garcia, and E. Kubista, "Statistical modeling of the lms channel," *IEEE Trans. Vehicular Technology*, vol. 50, no. 6, pp. 1549–1567, Nov. 2001.

[43] M. Ghanbarinejad and C. Schlegel, "Throughput-optimal distributed probabilistic medium-access in mpr-capable networks," in *Multiple Access Comm.*, ser. Lecture Notes in Computer Science, C. Sacchi *et al.*, Eds. Springer Berlin / Heidelberg, 2011, vol. 6886, pp. 75–86.

[44] ——, "Distributed probabilistic medium access with multipacket reception and markovian traffic," *Telecom. Systems Jnl.*, vol. 56, no. 2, pp. 311–321, Jun. 2014.

[45] A. Adas, "Traffic models in broadband networks," *IEEE Comm. Magazine*, vol. 35, no. 7, pp. 82–89, Jul. 1997.

[46] G. E. Monahan, "A survey of partially observable markov decision processes: Theory, models, and algorithms," *Management Science*, vol. 28, no. 1, pp. 1–16, Jan. 1982.

[47] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. Hoboken, NJ: John Wiley & Sons, 1994.

[48] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *The Annals of Mathematical Statistics*, vol. 41, no. 1, pp. 164–171, Feb. 1970.

[49] G. Bianchi and I. Tinnirello, "Kalman filter estimation of the number of competing terminals in an ieee 802.11 network," in *Proc. INFOCOM'03*, San Francisco, CA, Mar./Apr. 2003, pp. 844–852.

[50] M. Ghanbarinejad, C. Schlegel, and M. Khabbazian, "Random access with multipacket reception and adaptive filtering," in *Proc. Globecom'14*, Austin, TX, Dec. 2014.

[51] A. B. Carleial and M. E. Hellman, "Bistable behavior of aloha-type systems," *IEEE Trans. Communications*, vol. COM-23, no. 4, pp. 401–410, Apr. 1975.

[52] Y.-C. Jenq, "Optimal retransmission control of slotted aloha systems," *IEEE Trans. Communications*, vol. COM-29, no. 6, pp. 891–895, Jun. 1981.

[53] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "Macaw: A media access protocol for wireless lans," in *Proc. SIGCOMM'94*, London, UK, Aug./Sep. 1994, pp. 212–225.

[54] C. K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, 4th ed. Berlin, Germany: Springer-Verlag, 2009.

[55] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automatic Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[56] T. D. Larsen, N. A. Andersen, O. Ravn, and N. K. Poulsen, "Incorporation of time delayed measurements in a discrete-time kalman filter," in *Proc. IEEE Conf. Decision and Control*, Tampa, FL, Dec. 1998, pp. 3972–3977.

[57] M. Ghanbarinejad and C. Schlegel, "Irregular repetition slotted aloha with multiuser detection," in *Proc. Wireless On-Demand Network Systems and Services*, Banff, Canada, Mar. 2013, pp. 201–205.

[58] N. Celandroni *et al.*, "A survey of architectures and scenarios in satellite-based wireless sensor networks: system design aspects," *Int'l Jnl. Satellite Comm. Networks*, vol. 31, no. 1, pp. 1–38, Jan./Feb. 2013.

[59] P. Dickson and C. Schlegel, "Iterative demodulation and channel estimation for joint random access satellite communications," in *Proc. Int'l Symp. Turbo Codes and Iterative Info. Processing*, Bremen, Germany, Aug. 2014.

[60] R. G. Gallager, "Basic limits on protocol information in data communication networks," *IEEE Trans. Info. Theory*, vol. IT-22, no. 4, pp. 385–398, Jul. 1976.

[61] A. B. MacKenzie and S. B. Wicker, "Selfish users in aloha: a game-theoretic approach," in *Proc. Vehicular Technology Conference – Fall*, Atlantic City, NJ, Oct. 2001, pp. 1354–1357.